Revisiting Dynamic Regret of Strongly Adaptive Methods

Dheeraj Baby [1 2] **Hilaf Hasson** [2] **Bernie Wang** [2]

Abstract

We consider the framework of non-stationary Online Convex Optimization where a learner seeks to control its *dynamic regret* against an *arbitrary* sequence of comparators. When the loss functions are strongly convex or exp-concave, we demonstrate that Strongly Adaptive (SA) algorithms can be viewed as a principled way of controlling dynamic regret in terms of *path variation V_T of the comparator sequence*. Specifically, we show that SA algorithms enjoy $\tilde{O}(\sqrt{TV_T} \vee \log T)$ and $\tilde{O}(\sqrt{dTV_T} \vee d \log T)$ dynamic regret for strongly convex and exp-concave losses respectively *without* apriori knowledge of V_T, thus answering an open question in (Zhang et al., 2018b). The versatility of the principled approach is further demonstrated by the novel results in the setting of learning against bounded linear predictors and online regression with Gaussian kernels.

1. Introduction

Online Convex Optimization (OCO) is a powerful learning paradigm for real-time decision making. It has been applied in many influential applications such as portfolio selection, time series forecasting, and online recommendation systems to cite a few (Hazan et al., 2007; Koolen et al., 2015; Hazan, 2016). The OCO problem is modelled as an iterative game between a learner and an adversary that proceeds for T rounds as follows. At each time step t, the learner chooses a point $\boldsymbol{x}_t$ in a convex decision set $\mathcal{D}$. Then the adversary reveals a convex loss function $f_t : \mathcal{D} \to \mathbb{R}$. The most common way of measuring the performance of a learner is via its static regret, $R_T^s(z) := \sum_{t=1}^T (f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{z}))$, where $\boldsymbol{z}$ is termed as a fixed *comparator* in hindsight which can be any point in $\mathcal{D}$. For example $\boldsymbol{z}$ can be chosen as $\operatorname{argmin}_{\boldsymbol{x} \in \mathcal{D}} \sum_{t=1}^T f_t(\boldsymbol{x})$ with the knowledge of the entire sequence of loss functions. Learning is said to happen whenever the regret grows sub-linearly w.r.t. T. However, in the case of non-stationary environments such as stock market, one is often interested in matching the performance of a sequence of decisions in hindsight. In such circumstances, the notion of static regret fails to assess the performance of the learner. To better capture the non-stationarity, (Zinkevich, 2003) introduces the notion of *dynamic regret*:

$$R_T(\boldsymbol{z}_1, \dots, \boldsymbol{z}_T) := \sum_{t=1}^T (f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{z}_t)),$$

where $\boldsymbol{z}_1, \dots, \boldsymbol{z}_T$ is *any* sequence of comparators in $\mathcal{D}$. The degree of non-stationarity present in the comparator sequence is measured using the path variational defined as

$$V_T(\boldsymbol{z}_1, \dots, \boldsymbol{z}_T) := \sum_{t=2}^T \|\boldsymbol{z}_t - \boldsymbol{z}_{t-1}\|,$$

where $\|\cdot\|$ is the Euclidean norm. In what follows, we drop the arguments and represent the variation by V_T for brevity. The dynamic regret bounds are usually expressed as a function of T and V_T.

It is known that with convex loss functions the optimal dynamic regret is $O(\sqrt{T(1+V_T)})$ (Zhang et al., 2018a) which improves to $O(\sqrt{dTV_T} \vee d \log T)$ (Yuan & Lamperski, 2019), where d is the dimensionality of $\mathcal{D}$ and $(a \vee b) = \max\{a, b\}$, with additional curvature properties such as exp-concavity.

A parallel line of research (Hazan & Seshadhri, 2007; Daniely et al., 2015; Adamskiy et al., 2016) focus on developing algorithms whose static regret is controlled in any time interval. Specifically, (Daniely et al., 2015) develops the notion of Strongly Adaptive (SA) algorithms defined as:

Definition 1. *(Daniely et al., 2015) Let $[T] := \{1, \dots, T\}$. An algorithm is said to be Strongly Adaptive if for every continuous interval $I \subseteq [T]$, the static regret incurred by the algorithm is $O(poly(\log T) R^*(|I|))$, where $R^*(|I|)$ is the value of minimax static regret incurred in an interval of length $|I|$.*

(Zhang et al., 2018b) shows that SA algorithms incur dynamic regret of $\tilde{O}(T^{2/3} C_T^{1/3})$ [1] for convex losses and

[1]Department of Computer Science, University of California Santa Barbara, USA [2]Amazon Web Services, Palo Alto, California, USA. Correspondence to: Dheeraj Baby <dheeraj@ucsb.edu>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

[1] $\tilde{O}(\cdot)$ hides polynomial factors of $\log T$.

$\tilde{O}(\sqrt{TC_T})$ for strongly convex losses, where

$$C_T := \sum_{t=2}^{T} \sup_{\boldsymbol{x} \in \mathcal{D}} |f_t(\boldsymbol{x}) - f_{t-1}(\boldsymbol{x})|$$

to captures the non-stationarity of the problem in terms of the degree to which the sequence of losses changes over time. They further show that both results are optimal modulo poly logarithmic factors of T. For exp-concave losses, they derive a regret bound of $\tilde{O}(\sqrt{dTC_T})$. However, in (Zhang et al., 2018b), a question that was left open is whether it is possible to derive dynamic regret rates for SA methods that depend on the variational V_T.

In this paper, we answer this affirmatively for strongly convex (Theorem 3) and exp-concave (Theorem 6) losses. Specifically, we show that for SA methods,

$$R_T(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T) = \tilde{O}(\sqrt{TV_T(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T)} \vee \log T),$$

for strongly convex losses and

$$R_T(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T) = \tilde{O}(\sqrt{dTV_T(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T)} \vee d \log T),$$

for exp-concave losses.

This result immediately implies that SA algorithms can be seen as a unifying framework that allows one to control dynamic regret under different variationals (C_T and V_T) *simultaneously* whenever losses have curvature properties. Though this dynamic regret is attained by (Yuan & Lamperski, 2019) (without $\log T$ factors) by fundamentally different algorithms, our proof techniques are much simpler and shorter. Further, we demonstrate the versatility of this perspective by deriving new dynamic regret rates in various other interesting use cases where the results of (Yuan & Lamperski, 2019) do not apply (see Section 3). Every dynamic regret rate proposed in this paper are adaptive to V_T in the sense that the algorithms do not require the knowledge of V_T ahead of time.

Before ending this section, we summarize our key contributions below.

- We show that Strongly Adaptive (SA) algorithms are sufficient to guarantee the dynamic regret rates of $\tilde{O}(\sqrt{TV_T} \vee \log T)$ for strongly convex losses and $\tilde{O}(\sqrt{dTV_T} \vee d \log T)$ for exp-concave losses (see Theorems 3 and 6 respectively). Combined with the results of (Zhang et al., 2018b), we feature SA methods as a *unifying framework* for *simultaneously* controlling dynamic regret with variationals V_T and C_T.

- We demonstrate the versatility of this perspective by deriving several extensions (Theorems 9 and 10) where the results of (Yuan & Lamperski, 2019) don't apply. In particular, for competing against set of linear predictors that output bounded predictions as in (Luo et al., 2016), we show that SA methods enjoy dynamic regret rate that is *independent* of the diameter of the decision set. To the best of our knowledge this is the *first* time, dynamic regret rate has been proposed for such a benchmark set which is often more of practical interest than set of linear predictors with bounded L^2 norm.

The rest of the paper is organized as follows. We present the dynamic regret guarantees for strongly convex losses and exp-concave losses in Section 2. The extensions to competing against bounded linear predictors and online kernel regression is presented in Section 3. A section on preliminaries is provided in Appendix A

2. Dynamic regret for strongly convex and exp-concave losses

We start by showing that SA methods can serve as a principled way of achieving dynamic regret rates (up to log factors) of (Yuan & Lamperski, 2019). We assume that the loss functions are Lipschitz in the decision set. The main SA method we use in this paper is the Follow-the-Leading-History algorithm from (Hazan & Seshadhri, 2007). We provide a description of this algorithm in Appendix A for completeness.

Assumption 2. *The loss functions f_t satisfy $|f(\boldsymbol{x})-f(\boldsymbol{y})| \leq G\|\boldsymbol{x} - \boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$.*

2.1. Strongly convex losses

In this section we derive dynamic regret rates when the loss functions are H-strongly convex. We show that by appropriately instantiating the base learners in FLH, one can control the dynamic regret rates. The unspecified proofs are provided in the Appendix.

Theorem 3. *Suppose the loss function f_t are H-strongly convex loss and satisfy Assumption 2. Running FLH with learning rate $\zeta = H/G^2$ and base learners as online gradient descent (OGD) with step size $\eta_t = 1/Ht$ results in a dynamic regret of $\tilde{O}\left(\sqrt{TV_T} \vee 1\right)$, where $\tilde{O}(\cdot)$ hides dependence on constants H, G and poly-logarithmic factors of T.*

Remark 4. *When compared with the $\tilde{O}(\sqrt{T(1+V_T)})$ dynamic rate for convex functions from (Zhang et al., 2018b), the rate we derived for strongly-convex or exp-concave losses (which have extra curvature properties beyond convexity) shows its benefit in the regime when we can represent $V_T = c/T^{g(T)}$, where c is a constant and $g : \mathbb{N} \to \mathbb{R}^+$. In this case, the regret bounds of $\tilde{O}(\sqrt{TV_T})$ obtained using curvature could be better than those obtained without using curvature. The regret bounds we derived grows as*

$\tilde{O}(1 \vee \sqrt{T^{1-g(T)}})$ as opposed to the larger rate $O(\sqrt{T})$ obtained without curvature. Such a regime is of interest in a practical application where the comparator sequence changes slowly. A concrete example would be online time-series forecasting of average ocean temperatures.

2.2. Exp-concave losses

In this section, we assume that the losses are α-exp-concave and the domain is bounded. Specifically:

Assumption 5. *There exists a constant D such that $\max_{\boldsymbol{x},\boldsymbol{y}\in\mathcal{D}}\|\boldsymbol{x}-\boldsymbol{y}\| \le D$.*

We have the following Theorem.

Theorem 6. *Suppose the losses f_t are α-exp-concave and satisfy Assumptions 2 and 5. Running FLH with learning rate $\zeta = \alpha$ and ONS as base learners results in a dynamic regret of $\tilde{O}\left(\sqrt{dTV} \vee d\right)$, where $\tilde{O}(\cdot)$ hides dependence on constants G, D, α and poly-logarithmic factors of T.*

We conclude this section by two remarks that are applicable to every dynamic regret guarantee presented throughout the paper.

Remark 7. *Let τ be the running time of OGD per round. The FLH procedure incurs a run-time of $O(\tau T)$ per round. This can be improved to $O(\tau \log T)$ by using the AFLH procedure of (Hazan & Seshadhri, 2007) at the cost of increasing the dynamic regret by a logarithmic factor in time horizon T.*

Remark 8. *The FLH procedure doesn't require to know an apriori bound on V_T ahead of time. Hence the dynamic regret in Theorems 3 and 6 is adaptive to the variation V_T.*

3. Extensions

In this section, we demonstrate the versatility of SA methods by deriving new dynamic regret guarantees in various interesting settings.

3.1. Dynamic regret against bounded linear predictors

Consider the following learning protocol:

- For $t = 1, \dots, T$:
  1. Adversary reveals a feature vector $\boldsymbol{v}_t \in \mathbb{R}^d$.
  2. Learner chooses $\boldsymbol{w}_t \in \mathbb{R}^d$ and predict $\boldsymbol{w}_t^T \boldsymbol{v}_t$.
  3. Adversary reveals a loss $f_t(\boldsymbol{w}) := \ell_t(\boldsymbol{w}^T \boldsymbol{v}_t)$.
  4. Learner suffers loss $\ell_t(\boldsymbol{w}_t^T \boldsymbol{v}_t)$.

Under the above protocol, most of the OCO algorithms typically minimize the regret against a set of benchmark weights (where each weight define a linear predictor) that is bounded in some norm (e.g., the Euclidean $\|\cdot\|_2$ norm). In this section, we follow the path in (Ross et al., 2013; Luo et al., 2016) and study dynamic regret against a set of weights that rather produce bounded predictions. Specifically, define $\mathcal{K}_t := \{\boldsymbol{w} : |\boldsymbol{w}^T \boldsymbol{v}_t| \le B\}$. We aim to compete with a benchmark of linear predictors:

$$\begin{aligned}\mathcal{K} &= \cap_{t=1}^T \mathcal{K}_t, \\ &= \{\boldsymbol{w} : \forall t \in [T], |\boldsymbol{w}^T \boldsymbol{v}_t| \le B\},\end{aligned}$$

which basically defines a set of weights that outputs predictions in $[-B, B]$ at the given feature set. As noted in (Luo et al., 2016), the benchmark set $\mathcal{K}$ can be much larger than an L2 norm ball. The set $\mathcal{K}$ is often more useful in practice than a set of weights with bounded norm since it is more easier to choose a reasonable interval of predictions rather than choosing a bound on perhaps non-interpretable norm of the weights. We have the following dynamic regret guarantee.

Theorem 9. *Suppose the losses f_t are α-exp-concave and satisfy Assumption 2. Further assume that ℓ_t are Lipschitz smooth. Running FLH with learning rate $\zeta = \alpha$ and invariant ONS algorithm from (Luo et al., 2016) as base learners results in a dynamic regret of $\tilde{O}\left(d\sqrt{TV} \vee d^2\right)$, where $\tilde{O}(\cdot)$ hides dependence on constants G, α and poly-logarithmic factors of T.*

The dynamic regret bounds of (Yuan & Lamperski, 2019) are derived under the assumption that the norm of the elements in the benchmark set is bounded by some known constant. Specifically, the dynamic regret bound of (Yuan & Lamperski, 2019) grows as $O(D\sqrt{dTV_T})$ where D is the maximum L^2 norm of a predictor in the benchmark set. With benchmark set being $\mathcal{K}$, this D can be prohibitively large. In this case the *diameter independent* regret guarantee in Theorem 9 can be much smaller. To the best of our knowledge this is the first time a diameter independent regret guarantee has been proposed for controlling the dynamic regret in terms of V_T when the losses are exp-concave.

3.2. Dynamic regret for regression against a function space

In this section, we derive dynamic regret guarantees for competing against a sequence of functions in an RKHS induced by the Gaussian kernel. We study a regression setup where the loss is measured using squared errors. Specifically we consider the protocol in Fig.1

Setup and notations. We represent each function in the RKHS $\mathcal{H}_k$ by a weight vector $\boldsymbol{w} \in \mathbb{R}^D$ where D can be possible infinite. Let $\boldsymbol{x} \in \mathbb{R}^d$. For a function $f_{\boldsymbol{w}}(\boldsymbol{x}) \in \mathcal{H}_k$, we have $f_{\boldsymbol{w}}(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{\phi}(\boldsymbol{x})$ where $\boldsymbol{\phi}(\boldsymbol{x}) \in \mathbb{R}^D$ is the feature embedding of the vector $\boldsymbol{x}$ induced by the Kernel function $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. We consider the gaussian

1. For time $t = 1, \ldots, T$:
   (a) Receive $\boldsymbol{x}_t \in \mathbb{R}^d$.
   (b) Learner predicts $\hat{y}_t \in \mathbb{R}$.
   (c) Adversary reveals a label $y_t \in [-B, B]$.
   (d) Player suffers a loss of $(y_t - \hat{y}_t)^2$.

Figure 1: *Interaction protocol with squared error losses.*

kernel where $k(\boldsymbol{x}, \boldsymbol{y}) = \exp(-\|\boldsymbol{x} - \boldsymbol{y}\|^2/(2\sigma^2))$ for some bandwidth parameter σ. The RKHS norm of the function $f_{\boldsymbol{w}}$ which corresponds to the weight vector $\boldsymbol{w}$ is denoted by $\|\boldsymbol{w}\| := \|f_{\boldsymbol{w}}\|_{\mathcal{H}_k}$. We denote the determinant of a matrix $\boldsymbol{A}$ by $|\boldsymbol{A}|$.

For a sequence of comparator functions $f_{\boldsymbol{w}_1}, \ldots, f_{\boldsymbol{w}_T} \in \mathcal{H}_k$, define the path variational as

$$V_T = \sum_{t=2}^{T} \|f_{\boldsymbol{w}_t} - f_{\boldsymbol{w}_{t-1}}\|_{\mathcal{H}_k}.$$

We are interested in controlling the dynamic regret,

$$\sum_{t=1}^{T} (\hat{y}_t - y_t)^2 - (f_{\boldsymbol{w}_t}(\boldsymbol{x}_t) - y_t)^2,$$

for a sequence of functions $f_{\boldsymbol{w}_t}$ that belong to the class of functions with bounded RKHS norm defined as $\mathcal{D} = \{\boldsymbol{w} : \|\boldsymbol{w}\| \leq B\}$. For any $\boldsymbol{w} \in \mathcal{D}$, since $|\boldsymbol{w}^T \phi(\boldsymbol{x}_t)| \leq \|\boldsymbol{w}\| \|\phi(\boldsymbol{x}_t)\| = \|\boldsymbol{w}\| \sqrt{k(\boldsymbol{x}_t, \boldsymbol{x}_t)} \leq B$ and $|y_t| \leq B$, we have that the losses $\ell_t(\boldsymbol{w}) := (\boldsymbol{w}^T \phi(\boldsymbol{x}_t) - y_t)^2$ are $1/(8B^2)$ exp-concave in the domain $\mathcal{D}$ (Hazan et al., 2007). Specifically, for all $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{D}$, we have

$$\ell_t(\boldsymbol{v}) \geq \ell_t(\boldsymbol{u}) + (\boldsymbol{v} - \boldsymbol{u})^T \nabla \ell_t(u) + \frac{\alpha}{2} \left((\boldsymbol{v} - \boldsymbol{u})^T \nabla \ell_t(u) \right)^2,$$

where $\alpha = 1/(8B^2)$.

The following theorem (proof deferred to Appendix) controls the dynamic regret in the above prediction framework.

Theorem 10. *Assume that the comparator function sequence obeys $\|f_{\boldsymbol{w}_t}\|_{\mathcal{H}_k} \leq B$ and labels obey $|y_t| \leq B$ for all $t \in [T]$. Running FLH with learning rate $\zeta = 1/(8B^2)$ and base learners as PKAWV from (Jézéquel et al., 2019) with parameter $\lambda = 1$ and basis functions that approximate Gaussian kernel yields a dynamic regret:*

$$\sum_{t=1}^{T} (\hat{y}_t - y_t)^2 - (f_{\boldsymbol{w}_t}(\boldsymbol{x}_t) - y_t)^2 \leq O\left((\log T)^{\frac{d+1}{2}} \sqrt{TV_T} \vee (\log T)^{\frac{d+1}{2}} \right).$$

By Theorem 4 of (Jézéquel et al., 2019), computational complexity of PKAWV run with the configurations in Theorem 10 is $O((\log T)^{2d})$ per round. Hence by Remark 7 the runtime of the strategy in Theorem 10 is $O(T(\log T)^{2d})$ per iteration, and improves to $O((\log T)^{2d+1})$ via AFLH.

4. Conclusion

In this work, we derived dynamic regret rates for SA methods when the loss functions have curvature properties such as strong convexity or exp-concavity. Combined with the work of (Zhang et al., 2018b), our results indicate that SA methods can be viewed as a unifying framework for controlling dynamic regret in an OCO setting. (Yuan & Lamperski, 2019) (see Proposition 1 there) establishes minimax optimality of $O(\sqrt{TV_T})$ rate for certain ranges of V_T. However it is an open question to establish lower bounds that holds for all values of V_T.

References

Adamskiy, D., Koolen, W. M., Chernov, A., and Vovk, V. A closer look at adaptive regret. *Journal of Machine Learning Research*, 2016.

Altschuler, J., Bach, F., Rudi, A., and Weed, J. Massively scalable sinkhorn distances via the nyström method. *ArXiv*, abs/1812.05189, 2019.

Calandriello, D., Lazaric, A., and Valko, M. Second-order kernel online convex optimization with adaptive sketching. Proceedings of Machine Learning Research, 2017.

Daniely, A., Gonen, A., and Shalev-Shwartz, S. Strongly adaptive online learning. In *International Conference on Machine Learning*, pp. 1405–1411, 2015.

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

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

Hazan, E., Agarwal, A., and Kale, S. Logarithmic regret algorithms for online convex optimization. 2007.

Jézéquel, R., Gaillard, P., and Rudi, A. Efficient online learning with kernels for adversarial large scale problems. In *NeurIPS*, 2019.

Koolen, W. M., Malek, A., Bartlett, P. L., and Abbasi, Y. Minimax time series prediction. In *Advances in Neural Information Processing Systems (NIPS'15)*, pp. 2557–2565. 2015.

Luo, H., Agarwal, A., Cesa-Bianchi, N., and Langford, J. Efficient second order online learning by sketching. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, 2016.

Ross, S., Mineiro, P., and Langford, J. Normalized online learning. *UAI*, 2013.

Yuan, J. and Lamperski, A. Trading-off static and dynamic regret in online least-squares and beyond. Proceedings of the AAAI Conference on Artificial Intelligence, 34, 2019.

Zhang, L., Lu, S., and Zhou, Z.-H. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems (NeurIPS-18)*, pp. 1323–1333, 2018a.

Zhang, L., Yang, T., Zhou, Z.-H., et al. Dynamic regret of strongly adaptive methods. In *International Conference on Machine Learning (ICML-18)*, pp. 5877–5886, 2018b.

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

A. Preliminaries

The results in this paper hold for general Strongly Adaptive algorithms, but for concreteness we will phrase them in terms of a particular algorithm called Follow-the-Leading-History (FLH) (Hazan & Seshadhri, 2007).

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

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} base 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 2: FLH algorithm

We recall that a function f_t is said to be H-strongly convex in the domain in the domain $\mathcal{D}$ if it satisfies

$$f_t(\boldsymbol{y}) \geq f_t(\boldsymbol{x}) + (\boldsymbol{y} - \boldsymbol{x})^T \nabla f_t(\boldsymbol{x}) + \frac{H}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2, \tag{1}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{D}$. Further, f_t is said to be α-exp-concave if the last term in Eq.(1) is replaced by $\frac{\alpha}{2}\left((\boldsymbol{y} - \boldsymbol{x})^T \nabla f_t(\boldsymbol{x})\right)^2$.

FLH enjoys the following guarantee against any base learner.

Proposition 11. *(Hazan & Seshadhri, 2007) Suppose the loss functions are exp-concave with parameter α. For any interval $I = [r, s]$ in time, the algorithm FLH with learning rate $\zeta = \alpha$ gives $O(\alpha^{-1}(\log r + \log|I|))$ regret against the base learner in hindsight.*

For the case of exp-concave losses, one can maintain base learners $E^1, \ldots, E^T$ in Fig.2 as ONS algorithms that start at time points $1, \ldots, T$. Since each ONS instance achieves an $O(d \log T)$ static regret, Proposition 11 implies that the corresponding FLH with $\zeta = \alpha$ attains $O(d \log T)$ static regret in any interval.

Losses that are H-strongly convex and G-Lipschitz are known to be $O(H/G^2)$ exp-concave (Hazan et al., 2007). Further OGD attains $O(\log T)$ static regret. Hence FLH with OGD base learners and $\zeta = H/G^2$ can yield an $O(\log T)$ static regret in any interval when the loss functions are H-strongly convex. In Definition 1 if we restrict to minimax optimality wrt to interval length, these observations give rise to the following proposition.

Proposition 12. *FLH algorithm in Fig.2 with base learners as OGD and ONS are Strongly Adaptive when the losses are strongly convex and exp-concave respectively.*

B. Proofs for Section 2

We start with some useful lemmas for proving this theorem. In Lemma 13, we divide the time horizon into various bins such that the path variation of the comparator sequence incurred within these bins is at-most a quantity that will be tuned later. In Lemma 14, for each bin, we bound the dynamic regret by the sum of static regret against the first comparator point within the bin and a term that captures the drift of the remaining sequence of comparator points from the first point.

Lemma 13. *Let $\tilde{V} > 0$ be a constant. There exists a partitioning $\mathcal{P}$ of the sequence $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_t$ into M bins viz $\{[i_s, i_e]\}_{i=1}^M\}$ such that:*

1. $M := |\mathcal{P}| = O\left(\max\{V_T/\tilde{V}, 1\}\right)$.
2. *For all* $[i_s, i_e] \in \mathcal{P}$ *with* $i_s < i_e$, $\sum_{j=i_s+1}^{i_e} \|\boldsymbol{z}_j - \boldsymbol{z}_{j-1}\| \leq \tilde{V}$.

Proof. For an interval $[s,e]$ define $V_{s \to e} = \sum_{t=s+1}^{e} \|\boldsymbol{z}_t - \boldsymbol{z}_{t-1}\|$ which is the path variation incurred by the comparator within the interval $[s,e]$. Define $V_{s \to s} = 0$. Consider the following partitioning scheme.

1. Inputs: $\tilde{V} > 0$, $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T$.
2. Initialize $\mathcal{P} \leftarrow \Phi$ and $s \leftarrow 1$.
3. For time $t = 1, \ldots, T$:
   (a) If $V_{s \to t} \geq \tilde{V}$:
      i. Add $[s, t-1]$ to $\mathcal{P}$.
      ii. Set $s \leftarrow t$.

Let $M := |\mathcal{P}|$. Let's enumerate the intervals in $\mathcal{P}$ by $[i_s, i_e]$ with $i = 1, \ldots, M$. Assume that $M > 1$. We have,

$$\begin{aligned} V_T &\geq \sum_{i=1}^{M-1} V_{i_s \to i_e + 1}, \\ &\geq (M-1)\tilde{V}, \end{aligned}$$

where the last line follows from Steps 3(a,i,ii) of the partitioning scheme. Now rearranging yields the lemma. □

Lemma 14. *Assume that the losses are* α*-exp-concave. Let* $\boldsymbol{x}_t$ *be the predictions made by FLH with learning rate set as* $\zeta = \alpha$*. Let* $R(L)$ *be the static regret incurred by the base learners in an interval of length* L*. Let* $\mathcal{P}$ *be the partition of* $[T]$ *produced in Lemma 13. Represent each element in the partition by* $[i_s, i_e]$, $i = 1, \ldots, M$*. Then we have,*

$$\sum_{t=1}^{T} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{z}_t) \leq \tilde{O}\left(\inf_{\tilde{V} : \frac{V_T}{V} \geq 1} \left(V_T/\tilde{V} + GT\tilde{V} + \sum_{i=1}^{M} \sum_{t=i_s}^{i_e} R(i_e - i_s + 1)\right) \vee R(T)\right) \tag{2}$$

Proof. Consider a bin $[i_s, i_e] \in \mathcal{P}$ where $i \in [M]$. Let $\boldsymbol{x}_t$ be the predictions of the FLH with learning rate $\zeta = \alpha$. Let $\boldsymbol{p}_t$ be the predictions made by the base learner that wakes at time i_s. Due to Theorem 3.2 of (Hazan & Seshadhri, 2007), we have

$$\sum_{t=i_s}^{i_e} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{z}_t) \leq \sum_{t=i_s}^{i_e} f_t(\boldsymbol{p}_t) - f_t(\boldsymbol{z}_t) + O(\log T).$$

We have,

$$\begin{aligned} \sum_{t=i_s}^{i_e} f_t(\boldsymbol{p}_t) - f_t(\boldsymbol{z}_t) &= \sum_{t=i_s}^{i_e} f_t(\boldsymbol{p}_t) - f_t(\boldsymbol{z}_{i_s}) + \sum_{t=i_s}^{i_e} f_t(\boldsymbol{z}_{i_s}) - f_t(\boldsymbol{z}_t), \\ &\leq_{(a)} R(i_e - i_s + 1) + \sum_{t=i_s}^{i_e} f_t(\boldsymbol{z}_{i_s}) - f_t(\boldsymbol{z}_t), \\ &\leq_{(b)} R(i_e - i_s + 1) + G\tilde{V}(i_e - i_s + 1), \end{aligned}$$

where line (a) follows from static regret guarantee of the base learner and line (b) is due to Assumption 2 and the fact that $\sum_{j=i_s+1}^{i_e} \|\boldsymbol{z}_j - \boldsymbol{z}_{j-1}\| \leq \tilde{V}$. due to the partitioning scheme in Lemma 13.

Hence summing across all bins yields,

$$\sum_{t=1}^{T} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{z}_t) \leq \tilde{O}\left(V_T/\tilde{V} + GT\tilde{V} + \sum_{i=1}^{M} \sum_{t=i_s}^{i_e} R(i_e - i_s + 1)\right),$$

whenever $V_T/\tilde{V} \geq 1$ due to Lemma 13. Taking an infimium across such $\tilde{V}$ and including the static regret case $V_T = 0$ now yields the Lemma. □

Proof. **of Theorem 3**. We assume the notations in Lemma 14. An H-strongly convex loss is H/G^2 exp-concave in the decision set $\mathcal{D}$ (Hazan et al., 2007). Further, when the losses are strongly convex, from Theorem 1 of (Hazan et al., 2007) we have $R(T) = O(\log T)$ for OGD with step size $\eta_t = 1/Ht$. Hence by Lemmas 13 and 14 we have,

$$\begin{aligned}
\sum_{t=1}^{T} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{z}_t) &\leq \tilde{O}\left(V_T/\tilde{V} + GT\tilde{V} + \sum_{i=1}^{M}\sum_{t=i_s}^{i_e} R(i_e - i_s + 1)\right), \\
&\leq \tilde{O}\left(V_T/\tilde{V} + GT\tilde{V} + \sum_{i=1}^{M} \log T\right), \\
&\leq \tilde{O}\left(V_T/\tilde{V} + GT\tilde{V}\right), \qquad (3)
\end{aligned}$$

whenever $V_T/\tilde{V} \geq 1$.

Assume that $V_T \geq 1/T$. In this setting, if we choose $\tilde{V} = \sqrt{V/T}$, we have $V/\tilde{V} \geq 1$. Plugging this value to Eq. (3) yields a dynamic regret of $\tilde{O}(\sqrt{TV})$.

When $V_T = O(1/T)$, then we have,

$$\begin{aligned}
\sum_{t=1}^{T} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{z}_t) &= \sum_{t=1}^{T} f_t(\boldsymbol{x}_t) - f_t(\boldsymbol{z}_1) + \sum_{t=1}^{T} f_t(\boldsymbol{z}_1) - f_t(\boldsymbol{z}_t), \\
&\leq_{(a)} O(\log T) + GTV_T, \\
&\leq_{(b)} O(\log T),
\end{aligned}$$

where line (a) is by strong adaptivity of FLH and Lispchitzness of f_t. Line (b) is by the assumption $V_T = O(1/T)$. Combining both cases now yields the theorem. □

Proof sketch of Theorem 6. Theorem 2 of (Hazan et al., 2007), provides $O(d \log T)$ static regret for ONS under Assumptions 2 and 5. Theorem 6 follows by plugging in this static regret guarantee in the arguments of the proof of Theorem 3

C. Proofs for Section 3

Proof sketch of Theorem 9. Theorem 4 of (Luo et al., 2016), provides $O(d^2 \log T)$ static regret for a variant of ONS when ℓ_t are Lipschitz and f_t are exp-concave. Theorem 9 follows by plugging in this static regret guarantee in the arguments of the proof of Theorem 3,

Proof. **of Theorem 10** By Theorem 4 of (Jézéquel et al., 2019), if PKAWV algorithm is run with parameter $\lambda > 0$ and appropriately chosen basis, then we incur a regret:

$$\sum_{t=1}^{T} (\hat{y}_t - y_t)^2 - (f_{\boldsymbol{w}}(\boldsymbol{x}_t) - y_t)^2 \leq \lambda \|f_{\boldsymbol{w}}\|_{\mathcal{H}_k}^2 + \frac{3B^2}{2} \log \left|\boldsymbol{I} + \lambda^{-1}\boldsymbol{K}\right|,$$

where $\boldsymbol{K} \in \mathbb{R}^T \times \mathbb{R}^T$ is the kernel evaluation matrix with $\boldsymbol{K}_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$. When $k(\boldsymbol{x}, \boldsymbol{x}) \leq 1$, Lemma 3 of (Calandriello et al., 2017) implies:

$$\log \left|\boldsymbol{I} + \lambda^{-1}\boldsymbol{K}\right| \leq d_{eff}(\lambda)(1 + \log(1 + T/\lambda)),$$

where the effective dimension is defined as $d_{eff}(\lambda) = \mathrm{Tr}\left(\boldsymbol{K}(\boldsymbol{K} + \lambda\boldsymbol{I})^{-1}\right)$. It is known from (Altschuler et al., 2019) that for Gaussian kernels and covariates $\boldsymbol{x} \in \mathbb{R}^d$ with $k(\boldsymbol{x}, \boldsymbol{x}) \leq 1$, we have $d_{eff}(\lambda) = O\left(\left(\log \frac{T}{\lambda}\right)^d\right)$.

Hence whenever the comparator functions have bounded RKHS norm, by choosing $\lambda = 1$ we have the static regret bounded as

$$\sum_{t=1}^{T} (\hat{y}_t - y_t)^2 - (f_{\boldsymbol{w}}(\boldsymbol{x}_t) - y_t)^2 \leq O\left((\log T)^{d+1}\right).$$

Let $\ell_t(\boldsymbol{w}) = (f_{\boldsymbol{w}}(\boldsymbol{x}_t) - y_t)^2$. Since $\|\nabla \ell_t(\boldsymbol{w})\| \leq 2B^2$, we have that $\ell_t(\boldsymbol{w})$ is Lipschitz smooth in $\mathcal{D}$. Now plugging the above static regret bound and $G = 2B^2$ into Eq.(2) and following similar steps as in the proof of Theorem 3, we obtain the stated final regret bound.

$\square$