Online Linear Regression in Dynamic Environments via Discounting

Andrew Jacobsen [1] **Ashok Cutkosky** [2]

Abstract

We develop algorithms for online linear regression which achieve optimal static and dynamic regret guarantees *even in the complete absence of prior knowledge*. We present a novel analysis showing that a discounted variant of the Vovk-Azoury-Warmuth forecaster achieves dynamic regret of the form $R_T(\boldsymbol{u}) \le O\left(d\log(T) \vee \sqrt{dP_T^\gamma(\boldsymbol{u})T}\right)$, where $P_T^\gamma(\boldsymbol{u})$ is a measure of variability of the comparator sequence, and show that the discount factor achieving this result can be learned on-the-fly. We show that this result is optimal by providing a matching lower bound. We also extend our results to *strongly-adaptive* guarantees which hold over every sub-interval $[a,b] \subseteq [1,T]$ simultaneously.

1. Online Linear Regression

This paper presents new techniques and analyses for online linear regression, a variant of the classic least-squares regression problem tailored to streaming data (Azoury & Warmuth, 2001; Vovk, 2001; Orabona et al., 2015; Foster et al., 2016). Formally, consider T rounds of interaction between a learner and an environment, in which learner's objective is to accurately predict some observable target signal $y_t \in \mathbb{R}$ before it's revealed. On each round, a vector of *features* $x_t \in \mathbb{R}^d$ is first revealed, representing the context of the environment at the start of the round, and the learner predicts $\widehat{y}_t = \langle x_t, w_t\rangle$ by means of a weight vector $w_t \in \mathbb{R}^d$. The signal $y_t \in \mathbb{R}^d$ is then observed, and the learner incurs a loss proportional to the prediction error, $\ell_t(w_t) = \frac{1}{2}(y_t - \langle x_t, w_t\rangle)^2$. Since w_t is allowed to depend on x_t, this protocol is sometimes referred to as *improper* online regression, as the learner is able to make predictions outside of the class of linear models. Indeed, since x_t is revealed *before* the learner must make their prediction, it is always possible to make predictions $\widehat{y}_t = f_t(x_t)$ for any arbitrary transformation $f_t : \mathbb{R}^d \to \mathbb{R}$, for instance by setting $w_t = f_t(x_t)x_t/\left\|x_t\right\|^2$.

The classical measure of the learner's performance in this setting is *regret*, the cumulative prediction error relative to some fixed benchmark point $u \in \mathbb{R}^d$:

$$R_T(u) = \sum_{t=1}^T \ell_t(w_t) - \ell_t(u).$$

Notice that this performance measure can only properly reflect prediction accuracy when there exists a *fixed* $u \in \mathbb{R}^d$ which predicts well on average. For example, this may occur when when the (x_t, y_t) pairs are all generated *i.i.d.* from some well-behaved distribution. However, in many true streaming settings the data-generating distribution may change over time due to changes in the environment. *Dynamic* regret attempts to model such settings by comparing against a *sequence* of comparators $\boldsymbol{u} = (u_1, \dots, u_T)$:

$$R_T(\boldsymbol{u}) = \sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t).$$

Notice that dynamic regret captures the usual notion of regret (referred to as *static* regret) as a special case by setting $u_1 = \dots = u_T$. Our goal in this work is to make *favorable dynamic regret guarantees even in the complete absence of any prior knowledge of the underlying data-generating process*. Naturally, because such an algorithm leverages no prior knowledge, it necessarily must be adaptive to all problem-dependent quantities without requiring any instance-specific hyperparameter tuning.

Contributions. In this work we achieve the goal laid out above and develop the first algorithms for online linear regression that require no prior knowledge about the data stream, yet still make strong performance guarantees. In particular, our contributions are as follows:

- We show that even in the absence of any boundedness assumptions, a discounted variant of the VAW forecaster with a well-chosen discount factor achieves dynamic regret $R_T(\boldsymbol{u}) \le O\left(d\log(T) \vee \sqrt{dP_T^\gamma(\boldsymbol{u})T}\right)$, where $P_T^\gamma(\boldsymbol{u})$ is a measure of variability of the comparator sequence (*i.e.* the magnitude of $P_T^\gamma(\boldsymbol{u})$ is related to how drastically the comparator changes over time). We

[1]Department of Computing Science, University of Alberta, Edmonton, Canada [2]Department of Electrical and Computer Engineering, Boston University, Boston, Massachussetts. Correspondence to: Andrew Jacobsen <ajjacobs@ualberta.ca>.

Proceedings of the 41^{st} International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

also obtain *small-loss* guarantees of the form $R_T(\boldsymbol{u}) \leq O\Big(d\log(T) \vee \sqrt{dP_T^\gamma(\boldsymbol{u})\sum_{t=1}^T \ell_t(u_t)}\Big)$, so that the algorithm will automatically perform better on "easy" data where the comparator has low loss.

- We provide a matching lower bound of the form $R_T(\boldsymbol{u}) \geq \Omega\Big(d\log(T) \vee \sqrt{dTP_T^\gamma(\boldsymbol{u})}\Big)$, demonstrating optimality of the discounted VAW forecaster.
- We show that the discount factors required to obtain the results in the first point can be learned on-the-fly, leading to algorithms that make guarantees matching our lower bound. Moreover, we show how to extend our approach to achieve bounds of a similar form over *every sub-interval* $[a, b] \subseteq [1, T]$ *simultaneously*. These are the first strongly-adaptive guarantees have been achieved in the absence of all boundedness assumptions.

1.1. Related Works

Despite being a well-studied problem setting, there are no prior works which approach online linear regression with sufficient generality to be considered free from prior knowledge. The closest works to our own are Vovk (2001); Azoury & Warmuth (2001); Orabona et al. (2015); Mayo et al. (2022), each of which consider the same improper online learning setting as this work and present algorithms that can be run in an unbounded domain (hence requiring no prior knowledge about the comparator) and without any prior knowledge of the data stream. Yet these works provide guarantees that only hold for *static* regret—the *dynamic* regret of the algorithms in these works may be arbitrarily bad. In this sense, deploying any such algorithm implicitly requires rather strong prior knowledge: that the data-generating distribution is not changing over time.

A closely related problem setting which does account for potential non-stationarity is the classic *filtering* problem (Kalman, 1960; Simon, 2006; Kozdoba et al., 2019; Hazan & Singh, 2022). This problem setting assumes that the y_t are generated from a dynamical system of a specific form, and seeks to estimate the hidden state of the system. Thus, these works revolve around strong structural assumptions about the data-generating process from the outset. Similarly, there is a large literature on *adaptive* filtering which seeks to solve the filtering problem without *a priori* knowledge of the system (Kivinen et al., 2006; Hazan et al., 2017; 2018; Rashidinejad et al., 2020; Tsiamis & Pappas, 2022; Ghai et al., 2020), though these works still implicitly require prior knowledge that the underlying dynamical system is from some specific class, as any performance guarantees may otherwise fail to hold.

Alternatively, there are several related problem settings that one might hope to leverage results from, but these all inevitably require additional assumptions of some form to be applied to the online linear regression problem. For instance, many prior works develop algorithms for general online regression settings that capture linear regression as a special case (Orabona et al., 2015; Luo et al., 2016; Kotłowski, 2017; Kempka et al., 2019; Mhammedi & Koolen, 2020). Even more generally, one might hope to approach online linear regression via reduction to a more general online convex optimization setting (Zhang et al., 2018; Yuan & Lamperski, 2019; Zhao et al., 2020; Baby et al., 2021; Baby & Wang, 2021; Luo et al., 2022; Jacobsen & Cutkosky, 2022; Zhang et al., 2023; Zhao et al., 2024). Unfortunately, all of these works require additional boundedness assumptions on the losses such as Lipschitzness or exp-concavity, both of which require a bounded domain in the context of losses $\ell_t(w) = \frac{1}{2}(y_t - \langle x_t, w\rangle)^2$. Yet assuming a bounded domain amounts amounts to having strong prior knowledge that the comparator sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ lies entirely within some bounded subset $W \subset \mathbb{R}^d$, which must be known and accounted for *a priori* for the guarantees to hold.

One recent exception to the limitations mentioned above is the work of Jacobsen & Cutkosky (2023). They develop an approach that can be applied to any loss functions satisfying $\|\nabla\ell_t(w)\| \leq G_t + L_t\|w\|$ for some non-negative constants G_t and L_t, and hence could be applied in our setting for $G_t = |y_t|\,\|x_t\|$ and $L_t = \|x_t\|^2$. Their algorithm achieves a dynamic regret guarantee on the order of $O(M^{3/2}\sqrt{P_T T})$ where $M = \max_t \|u_t\|$ and $P_T = \sum_{t=2}^T \|u_t - u_{t-1}\|$. However, their approach fails to achieve logarithmic regret against a fixed comparator, and their approach requires prior knowledge of a $G_{\max} \geq G_t$ and $L_{\max} \geq L_t$ for all t. Moreover their approach requires $O(dT\log(T))$ per-round computation, making it inappropriate for many of the long-running problems where non-stationarity naturally emerges due to subtle changes in the environment over time.

1.2. Notations

We define $\ell_0(w) = \frac{\lambda}{2}\|w\|_2^2$, so that updates can be written purely in terms of losses ℓ_t. Given a positive definite matrix M, the weighted norm w.r.t M is $\|w\|_M = \sqrt{\langle w, Mw\rangle}$. For any sequence $a_1, a_2, \ldots$, we denote $a_{\max} = \max_t |a_t|$. Positive thresholding is denoted as $[\cdot]_+ = \max\{\cdot, 0\}$. The Bregman divergence *w.r.t.* a differentiable function ψ is $D_\psi(x|y) = \psi(x) - \psi(y) - \langle\nabla\psi(y), x - y\rangle$. We denote $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$, $[N] = \{1, \ldots, N\}$, $\mathbb{N} = \{0, 1, \ldots\}$ denotes the natural numbers, and $\mathbf{1}_N$ is the N-dimensional vector of ones. We use the short-hand $\text{Clip}_{[a,b]}(y) = (y \vee a) \wedge b$ and the compressed sum notations $g_{i:j} = \sum_{t=i}^j g_t$ and $\|g\|_{a:b}^2 = \sum_{t=a}^b \|g_t\|^2$. The N-dimensional simplex is denoted Δ_N. $O(\cdot)$ hides constant factors and $\widehat{O}(\cdot)$ hides constant and $\log\log$ factors.

2. The Vovk-Azoury-Warmuth Forecaster

In the context of *static* regret, it is well known that the optimal strategy in our improper online linear regression setting is the Vovk-Azoury-Warmuth (VAW) forecaster, discovered independently by Azoury & Warmuth (2001) and Vovk (2001). On each round, the standard VAW forecaster sets

$$w_t = \left(\lambda I + \sum_{s=1}^{t} x_s x_s^\top\right)^{-1} \sum_{s=1}^{t-1} y_s x_s. \tag{1}$$

The VAW forecaster is well-known for the following regret guarantee (Azoury & Warmuth, 2001; Vovk, 2001; Orabona et al., 2015).

Theorem 2.1. *For any $u \in \mathbb{R}^d$ and any sequences $(y_t)_{t=1}^T$ in $\mathbb{R}$ and $(x_t)_{t=1}^T$ in $\mathbb{R}^d$, the VAW forecaster guarantees*

$$R_T(u) \le \frac{\lambda}{2}\|u\|_2^2 + \frac{d \max_t y_t^2}{2} \log\left(1 + \frac{\sum_{t=1}^T \|x_t\|_2^2}{\lambda d}\right),$$

Let us briefly pause to appreciate some of the subtleties of this result, as it represents a very high standard of excellence in online learning. First, note that the result holds using *no prior knowledge about the data* — there are no underlying assumptions about how the features x_t or the targets y_t are distributed, the algorithm requires no specific statistics or bounds such as $|y_t| \le Y$ or $\|x_t\| \le X$, and the algorithm works in an unbounded domain — a relative rarity in adversarial settings. Yet despite this incredible degree of generality, the VAW forecaster boasts a strong *logarithmic* regret guarantee, which can be shown to be optimal up to constant factors (See, *e.g.*, Cesa-Bianchi & Lugosi (2006, Theorem 11.9)). Thus, the VAW forecaster achieves a harmony between theory and practice which is quite rare in online learning, requiring no problem-specific information or assumptions while still guaranteeing optimal regret.

However, a major caveat to the above discussion is that these favorable properties hold only within the context of *static* regret. The *dynamic* regret of the VAW forecaster can be arbitrarily bad in general. To see why, let us consider the simple case where $d = 1$ and $x_t = 1$ for all t. In this case, the VAW forecaster predicts $\widehat{y}_t = x_t w_t = (\lambda + t)^{-1} \sum_{s=1}^{t-1} y_s$, which approximates an empirical average of the targets observed up to round t. It is easy to see that any such prediction strategy can fail when competing with a time-varying comparator. For instance, if the first $T/2$ targets are -1 but the second half are $+1$, the VAW forecaster will quickly converge to predicting -1 in the first $T/2$ rounds, but will be unable to quickly adapt after the change in the latter $T/2$ rounds, leading to linear regret overall. In this sense, the VAW forecaster actually *implicitly requires quite strong prior knowledge* about the data: that it is, in some sense, *stationary*. Because of this, its predictions can *not* be trusted in the absence of prior knowledge, but rather only when the practitioner knows they are dealing with data that can be reasonably predicted using only a single fixed hypothesis $u \in \mathbb{R}^d$. In the next section, we will see that this issue can be alleviated by incorporating a suitable recency bias to the statistics of the VAW forecaster.

3. Dynamic Regret via Discounting

Despite making strong static regret guarantees, we saw in the previous section that the standard VAW forecaster may fail to attain low regret when competing against a time-varying comparator. Loosely speaking, the problem is that the VAW forecaster treats all time-steps as equally important. Indeed, it can be shown that VAW forecaster can be understood as updating

$$w_t = \arg\min_{w \in \mathbb{R}^d} \frac{1}{2}\|w\|_{\Lambda_t}^2 + \sum_{s=1}^{t-1} \ell_s(w),$$

where $\Lambda_t = \lambda I + x_t x_t^\top$.[1] The latter term $\sum_{s=1}^{t-1} \ell_s(w)$ forces the VAW forecaster to choose a w which balances all of the losses encountered so-far. Yet in dynamic scenarios, the losses that contain the most-relevant information for predicting y_t are typically the ones that have been observed the most recently. In order to more closely track these recently-observed losses, we make two modifications to the VAW forecaster. First, we incorporate a *forgetting* or *discount* factor γ in to the algorithm's statistics, placing less emphasis on losses observed far in the past. Second, we allow the update to additionally make use of a sequence of "predicted labels" or "hints" $\widetilde{y}_t$ that are available before we commit to $\widehat{y}_t$. Intuitively, we would like our algorithm to do better when $\widetilde{y}_t = y_t$. Later, we will provide some concrete ways to set $\widetilde{y}_t$ that yield strong regret bounds.

The variant of the VAW forecaster described above is provided concretely in Algorithm 1. Observe that by unrolling the recursions for θ_t and Σ_t, the update $w_t = \Sigma_t^{-1}\left[\widetilde{y}_t x_t + \gamma\theta_t\right]$ can be written in closed-form as

$$w_t = \left(\gamma^t \lambda I + \sum_{s=1}^{t} \gamma^{t-s} x_s x_s^\top\right)^{-1} \left[\widetilde{y}_t x_t + \gamma \sum_{s=1}^{t-1} \gamma^{t-1-s} y_s x_s\right].$$

By setting $\gamma = 1$ and $\widetilde{y}_t = 0$, the update precisely reduces to Equation (1), so the discounted VAW forecaster is a strict generalization of the standard VAW forecaster. Likewise, the following theorem shows that Algorithm 1 obtains a regret guarantee which captures Theorem 2.1 as a special case. Proof can be found in Appendix A.2.

[1] The equivalence to Equation (1) is readily checked via the first-order optimality condition, though this claim can also be derived as a special case of a more general claim Proposition A.1 provided in the appendix.

Algorithm 1: Discounted VAW Forecaster

Input $\lambda > 0$, $\gamma \in (0, 1]$
Initialize $w_1 = \mathbf{0}$, $\Sigma_0 = \lambda I$, $\theta_1 = \mathbf{0}$
for $t = 1 : T$ **do**
  Receive features $x_t \in \mathbb{R}^d$
  Set $\Sigma_t = x_t x_t^\top + \gamma\Sigma_{t-1}$, choose $\widetilde{y}_t \in \mathbb{R}$
  Update $w_t = \Sigma_t^{-1}\left[\widetilde{y}_t x_t + \gamma\theta_t\right]$

  Predict $\langle x_t, w_t\rangle$ and observe y_t
  Incur loss $\ell_t(w_t) = \frac{1}{2}(y_t - \langle x_t, w_t\rangle)^2$
  Set $\theta_{t+1} = y_t x_t + \gamma\theta_t$
end

Theorem 3.1. *Let $\lambda > 0$ and $\gamma \in (0, 1]$. Then for any sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ in $\mathbb{R}^d$, Algorithm 1 guarantees dynamic regret $R_T(\boldsymbol{u})$ bounded above by*

$$\frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{d}{2}\max_t (y_t - \widetilde{y}_t)^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right)$$
$$+ \gamma\sum_{t=1}^{T-1}\left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right] + \frac{d}{2}\log(1/\gamma)\sum_{t=1}^T (y_t - \widetilde{y}_t)^2$$

where $F_t^\gamma(w) = \gamma^t \frac{\lambda}{2}\|w\|_2^2 + \sum_{s=1}^t \gamma^{t-s}\ell_s(w)$.

The regret decomposition obtained in Theorem 3.1 is appealing for two reasons. First, it captures Theorem 2.1 as a special case: setting $\gamma = 1$, $\widetilde{y}_t = 0$, and $u_1 = \ldots = u_T = u$, the last two terms of the bound evaluate to zero, so the regret is bounded by $\frac{\lambda}{2}\|u\|_2^2 + \frac{d}{2}\max_t y_t^2 \log\left(1 + \frac{\sum_{t=1}^T \|x_t\|_2^2}{\lambda d}\right)$, which is precisely the guarantee promised by Theorem 2.1. Second, the decomposition displays a clean separation of concerns. The terms in the first line are the unavoidable penalties associated with *static* regret, which are of course also unavoidable here in the more general dynamic regret setting. In the second line, any penalties incurred as a result of a changing comparator sequence are captured entirely by the *variability term* $\gamma\sum_{t=1}^T F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)$, while the term $d\log(1/\gamma)\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2$ represents a *stability penalty* incurred due to discounting.

Intuitively, the terms in the second line represent a tracking/stability trade-off: against a volatile comparator sequence, we would ideally like to set the discount factor γ to be small to control the variability penalty, yet this will come at the expense of increasing the stability penalty $d\log(1/\gamma)\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2$. In its current form, however, this trade-off is still a bit mysterious. The variability term $\gamma\sum_{t=1}^{T-1} F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)$ is not necessarily monotonic as a function of γ nor is it necessarily positive, making it difficult to meaningfully analyze or understand how it relates to the stability penalty $\frac{d}{2}\log(1/\gamma)\sum_{t=1}^T (y_t - \widetilde{y}_t)^2$. If we instead consider a modest upper bound on these terms we can reveal a more explicit trade-off. We provide proof of a slightly more general statement of the following lemma in Appendix A.3.

Lemma 3.2. *(simplified) Let $\ell_0, \ell_1, \ldots, \ell_T$ be arbitrary non-negative functions, $\gamma \in (0, 1)$, and $F_t^\gamma(w) = \sum_{s=0}^t \gamma^{t-s}\ell_s(w)$. For all t, define*

$$\bar{d}_t^\gamma(u, v) = \sum_{s=0}^t \frac{\gamma^{t-s}}{\sum_{s'=0}^t \gamma^{t-s'}}\left[\ell_s(u) - \ell_s(v)\right]_+$$

and $P_T^\gamma(\boldsymbol{u}) = \sum_{t=1}^{T-1}\bar{d}_t^\gamma(u_{t+1}, u_t)$. Then for any $V_T \geq 0$,

$$\gamma\sum_{t=1}^{T-1}\left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right] + \log\left(\frac{1}{\gamma}\right)V_T$$
$$\leq \frac{\gamma}{1-\gamma}P_T^\gamma(\boldsymbol{u}) + \frac{1-\gamma}{\gamma}V_T$$

The lemma bounds the variability term $\gamma\sum_{t=1}^{T-1}\left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right]$ from Theorem 3.1 in terms of a new one $P_T^\gamma(\boldsymbol{u})$. To understand this new measure of variability, for each t let us first define a γ-exponentially-decaying distribution over time-steps $s \leq t$ as $p_t^\gamma(s) = \frac{\gamma^{t-s}}{\sum_{s'=0}^t \gamma^{t-s'}}$. Then, given γ we can express $P_T^\gamma(\boldsymbol{u})$ as

$$\begin{aligned} P_T^\gamma(\boldsymbol{u}) &= \sum_{t=1}^{T-1}\overbrace{\sum_{s=0}^t p_t^\gamma(s)\left[\ell_s(u_{t+1}) - \ell_s(u_t)\right]_+}^{\bar{d}_t^\gamma(u_{t+1}, u_t)} \\ &= \sum_{t=1}^{T-1}\mathbb{E}_{s\sim p_t^\gamma}\Big[\left(\ell_s(u_{t+1}) - \ell_s(u_t)\right)_+\Big], \end{aligned}$$

so each term of $P_T^\gamma(\boldsymbol{u})$ is a measure of how different the prediction errors of u_t and u_{t+1} are on average across "recent" losses. The quantity $P_T^\gamma(\boldsymbol{u})$ can also be naively related to the more common measure of variability — the path-length $P_T^{\|\cdot\|} = \sum_{t=1}^{T-1}\|u_t - u_{t+1}\|$ — as follows:

$$\begin{aligned} P_T^\gamma(\boldsymbol{u}) &\leq \sum_{t=1}^{T-1}\max_s \|\nabla\ell_s(u_{t+1})\|\|u_t - u_{t+1}\| \\ &\leq \max_{t,s}\|\nabla\ell_s(u_t)\| P_T^{\|\cdot\|} \leq O\left(\max_t \|u_t\| P_T^{\|\cdot\|}\right). \end{aligned}$$

Thus, $P_T^\gamma(\boldsymbol{u})$ is proportional to the usual path-length. Note that a multiplicative penalty of $\max_t\|u_t\|$ is the same worst-case penalty that appears in prior works, even in bounded domains (Zhang et al., 2018; Jacobsen & Cutkosky, 2022; Zhang et al., 2023; Zhao et al., 2024).

Letting $\eta = \frac{\gamma}{1-\gamma}$, Lemma 3.2 tells us that that latter terms of Theorem 3.1 are bounded by

$$\eta P_T^\gamma(\boldsymbol{u}) + \frac{d}{2\eta}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2,$$

a trade-off which can be optimized by choosing $\eta = \sqrt{\frac{\frac{d}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}{P_T^\gamma(\boldsymbol{u})}}$ to get

$$\eta P_T^\gamma(\boldsymbol{u}) + \frac{d}{2\eta}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2 = 2\sqrt{dP_T^\gamma(\boldsymbol{u})\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2}.$$

This is very promising; as we will see in Section 3.2, a penalty of this form is unavoidable in general. Plugging this choice of η back into $\eta = \frac{\gamma}{1-\gamma}$ and solving for γ, we find that the ideal choice of discount factor would be a $\gamma \in [0,1]$ satisfying

$$\gamma = \frac{\sqrt{\frac{d}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}}{\sqrt{\frac{d}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2} + \sqrt{P_T^\gamma(\boldsymbol{u})}}.$$

Notice in particular that γ appears on both sides of the expression, and solving for this γ explicitly is non-trivial in general. Nonetheless, the following theorem shows that a discount factor satisfying the above expression always exists, and if it could somehow be provided to the discounted VAW forecaster we would achieve dynamic regret matching the lower bound in Section 3.2. Proof can be found in Appendix A.5.

Theorem 3.3. *For any sequences $y_1, \dots, y_T$ and $\widetilde{y}_1, \dots, \widetilde{y}_T$ in $\mathbb{R}$ and any sequence $\boldsymbol{u} = (u_1, \dots, u_T)$ in $\mathbb{R}^d$, there is a discount factor $\gamma^* \in [0,1]$ satisfying*

$$\gamma^* = \frac{\sqrt{\frac{d}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}}{\sqrt{\frac{d}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2} + \sqrt{P_T^{\gamma^*}(\boldsymbol{u})}} \tag{2}$$

with which the regret of Algorithm 1 is bounded above by

$$R_T(\boldsymbol{u}) \le O\Bigg(d\max_t (y_t - \widetilde{y}_t)^2 \log(T) + \sqrt{dP_T^{\gamma^*}(\boldsymbol{u})\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}\Bigg)$$

While this result is promising, it is important to note that it still falls short of our desired goal of prior-knowledge-free learning. Indeed, it seems that we require *exceptionally strong* prior knowledge to choose the prescribed discount factor γ^* satisfying Equation (2). We will return to this issue in Section 4 to show that this discount factor can be learned on-the-fly, resulting in algorithms that *are* truly free of prior knowledge.

Interestingly, the discount factor γ^* in Theorem 3.3 can help to shed some light on the variability measure $P_T^{\gamma^*}(\boldsymbol{u})$. Observe from the relation in Equation (2) that γ^* can be near zero only when $P_T^{\gamma^*}(\boldsymbol{u})$ is very large relative to the stability penalty, and likewise, if γ^* is near 1 then $P_T^{\gamma^*}(\boldsymbol{u})$ must be inconsequentially small. In this sense, the $P_T^{\gamma^*}(\boldsymbol{u})$ corresponding to small γ^* can be regarded as the worst-case measures of variability. Yet as γ^* approaches zero, $P_T^{\gamma^*}(\boldsymbol{u})$ approaches $\sum_{t=1}^{t-1}[\ell_t(u_{t+1}) - \ell_t(u_t)]_+$, which can be naturally related other standard measures of variability. Indeed, this penalty is similar in spirit to the temporal variability $\sum_{t=1}^{T-1}|\ell_{t+1}(u_t) - \ell_t(u_t)|$ studied in works such as Campolongo & Orabona (2021); Besbes et al. (2015), and can be related to the path-length $\sum_{t=1}^{T-1}\|u_t - u_{t+1}\|$ via convexity of ℓ_t. In this sense, $P_T^{\gamma^*}(\boldsymbol{u})$ can be thought of as a relaxation of the more common measures of variability.

3.1. Small-loss Bounds via Self-confident Predictions

In the previous section, we saw that the discounted VAW forecaster can achieve regret scaling as $O\left(\sqrt{dP_T^{\gamma^*}(\boldsymbol{u})\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}\right)$, where $\widetilde{y}_t \in \mathbb{R}$ is an arbitrary "hint" available before observing the true y_t. One particularly interesting option is to use *the learner's own prediction* as a hint, $\widetilde{y}_t = \langle x_t, w_t\rangle$. The reasoning is that any learner achieving low dynamic regret must be predicting y_t reasonably well on average, so their own predictions would naturally make for reasonable predicted labels $\widetilde{y}_t$. Concretely, observe that by choosing $\widetilde{y}_t = \langle x_t, w_t\rangle$ we would have $\sum_{t=1}^T (y_t - \widetilde{y}_t)^2 = \sum_{t=1}^T (y_t - \langle x_t, w_t\rangle)^2 = 2\sum_{t=1}^T \ell_t(w_t)$, and hence for some $\gamma \in [0,1]$ the guarantee in Theorem 3.3 would scale as

$$R_T(\boldsymbol{u}) = \sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) \le \widetilde{O}\left(\sqrt{dP_T^\gamma(\boldsymbol{u})\sum_{t=1}^T \ell_t(w_t)}\right),$$

where the $\widetilde{O}(\cdot)$ hides the logarithmic factor. Now notice that $\sum_{t=1}^T \ell_t(w_t)$ appears on both sides of this inequality. Solving for $\sum_{t=1}^T \ell_t(w_t)$, one finds that this implies that $\sqrt{\sum_{t=1}^T \ell_t(w_t)} \le O\left(\sqrt{dP_T^\gamma(\boldsymbol{u})} + \sqrt{\sum_{t=1}^T \ell_t(u_t)}\right)$, so plugging this back into the regret bound we have

$$R_T(\boldsymbol{u}) \le \widetilde{O}\left(P_T^\gamma(\boldsymbol{u}) + \sqrt{P_T^\gamma(\boldsymbol{u})\sum_{t=1}^T \ell_t(u_t)}\right).$$

Bounds of this form, sometimes called *small-loss* or L^* bounds, are highly desirable because they naturally adapt to the total loss of the comparator sequence, potentially leading to lower regret than more naive hint choices such as $\widetilde{y}_t = y_{t-1}$ or $\widetilde{y}_t = 0$.

Unfortunately, the above argument does not quite go through because the now the logarithmic penalty in Theorem 3.3 scales as $O\left(d\max_t (y_t - \widetilde{y}_t)^2 \log(T)\right) = O\left(d\max_t \ell_t(w_t)\log(T)\right)$, and this $\max_t \ell_t(w_t)$ could be arbitrarily large. Fortunately, it turns out that this issue can be remedied by a simple trust-region argument. On each round, instead of directly using hints $\widetilde{y}_t = \langle x_t, w_t\rangle$, we can constrain these predictions to be close to some arbitrary reference point y_t^{Ref}. In particular, in Lemma D.1 we show by clipping the learner's predictions to a suitable interval centered at y_t^{Ref} we can guarantee $(y_t - \widetilde{y}_t)^2 \le O\left(\max_t (y_t - y_t^{\text{Ref}})^2 \wedge \ell_t(w_t)\right)$. This gives us the best-of-both-worlds: a similar self-bounding argument to above still yields a small-loss penalty $O\left(\sqrt{dP_T^\gamma(\boldsymbol{u})\sum_{t=1}^T \ell_t(u_t)}\right)$, while the logarithmic penalty can be bounded as $O\left(d\max_t (y_t - y_t^{\text{Ref}})^2 \log(T)\right) \le O(d \max_t y_t^2 \log(T))$ by setting $y_t^{\text{Ref}} = y_{t-1}$ or $y_t^{\text{Ref}} = 0$. The following theorem follows this above argument through, demonstrating that the discounted VAW forecaster can achieve small-loss bounds when using a well-chosen discount factor.

Theorem 3.4. *Let $y_t^{Ref} \in \mathbb{R}$ be an arbitrary reference point and let $\mathcal{B}_t = [y_t^{Ref} - M_t, y_t^{Ref} + M_t]$ for $M_t = \max_{s<t} \left|y_s - y_s^{Ref}\right|$. Suppose that we apply Algorithm 1 with hints $\widetilde{y}_t = \text{Clip}_{\mathcal{B}_t}(\langle x_t, w_t\rangle)$. Then for any sequence of losses $\ell_1, \ldots, \ell_T$ and any sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ in $\mathbb{R}^d$, there is a $\gamma^\circ \in [0,1]$ satisfying*

$$\gamma^\circ = \frac{\sqrt{d\sum_{t=1}^T \ell_t(u_t)}}{\sqrt{d\sum_{t=1}^T \ell_t(u_t)} + \sqrt{P_T^{\gamma^\circ}(\boldsymbol{u})}}. \tag{3}$$

Moreover, running Algorithm 1 with discount $\gamma^\circ \vee \gamma_{\min}$ for $\gamma_{\min} = \frac{2d}{2d+1}$ ensures regret bounded above by

$$R_T(\boldsymbol{u}) \le O\left(dP_T^{\gamma_{\min}}(\boldsymbol{u}) + d\max_t (y_t - y_t^{Ref})^2 \log(T) + \sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u})\sum_{t=1}^T \ell_t(u_t)}\right),$$

Notice that unlike the previous section, there are two different variability penalties, $P_T^{\gamma^\circ}(\boldsymbol{u})$ and $P_T^{\gamma_{\min}}(\boldsymbol{u})$. The first mirrors the measure encountered in the last section. The other, $P_T^{\gamma_{\min}}(\boldsymbol{u})$, is rather annoying; in high dimensions $\gamma_{\min} = \frac{2d}{2d+1}$ is generally quite large, so $P_T^{\gamma_{\min}}(\boldsymbol{u})$ may evaluate losses at irrelevant comparators that are far away in time. Nevertheless, notice that this term satisfies $P_T^{\gamma_{\min}}(\boldsymbol{u}) \le \sum_{t=1}^{T-1} \max_s [\ell_s(u_{t+1}) - \ell_s(u_t)]_+$, a penalty which we will show is unavoidable in general in Theorem 3.5.

3.2. Dimension-dependent Lower Bound

In this section, we show that the regret penalties observed in the previous sections are unavoidable without further assumptions. The following lower bound is proven in Appendix C.1.

Theorem 3.5. *For any $d, T \ge 1$ and $P, Y > 0$ such that $dP \le 2TY^2$, there is a sequence of losses $\ell_t(w) = \frac{1}{2}(y_t - \langle x_t, w\rangle)^2$ and a comparator sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ satisfying $\max_t |y_t| \le Y$ and $\sum_{t=1}^{T-1} \max_s [\ell_s(u_{t+1}) - \ell_s(u_t)]_+ \le P$ such that*

$$R_T(\boldsymbol{u}) \ge \Omega\left(dY^2 \log(T) + dP + \sqrt{dP\sum_{t=2}^T (y_t - y_{t-1})^2}\right).$$

The key observation is that there is always a sequence of losses such that $\sum_{t=1}^T \ell_t(u_t) = 0$ can be ensured using only T/d different comparators. Indeed, letting the features x_t cycle through the standard basis vectors, for any sub-interval $[s, s+d] \subseteq [1, T]$ we can choose a single $u \in \mathbb{R}^d$ such that $\langle x_t, u\rangle = y_t$ for each t in the interval. Then by sampling the y_t randomly from $\{-Y\sigma, Y\sigma\}$ for some $\sigma \in [0,1]$, we can ensure variability of at most $O(TY^2\sigma^2/d) \le P$ but regret of at least $\Omega(TY^2\sigma^2) \ge \Omega\left(\sqrt{dP\left[\sum_{t=1}^T (y_t - y_{t-1})^2 \vee dP\right]}\right)$.

Note that the condition $dP \le 2TY^2$ captures a natural restriction of the problem setting, in that for larger P the vacuous lower bound $R_T(\boldsymbol{u}) \ge \Omega(TY^2)$ can be constructed. Indeed, in the boundary case where $dP = 2TY^2$, Theorem 3.5 tells us that there is a sequence such that $R_T(\boldsymbol{u}) \ge \Omega\left(\sqrt{dP\mathcal{V}_T}\right) = \Omega(dP) = \Omega\left(TY^2\right)$. Yet this bound is achieved against *any* comparator sequence by the algorithm that naively predicts $\mathbf{0}$ on every round: $R_T(\boldsymbol{u}) = \sum_{t=1}^T \ell_t(\mathbf{0}) - \ell_t(u_t) \le \sum_{t=1}^T \frac{1}{2} y_t^2 \le \frac{1}{2}TY^2$. Hence, no lower bound can exceed $\frac{1}{2}TY^2$, so it is sufficient to consider comparator sequences with variability bounded by $P \le 2TY^2$.

If we instead consider a more restricted problem setting by assuming a bounded domain, then the losses $\ell_t(w) = \frac{1}{2}(y_t - \langle x_t, w\rangle)^2$ can be considered to be exp-concave. In this setting, Baby & Wang (2021) have shown a lower bound of

$$R_T(\boldsymbol{u}) \ge \Omega\left(Y^{4/3} d^{1/3} T^{1/3} C_T^{2/3}\right), \tag{4}$$

where $C_T = \sum_{t=1}^{T-1} \left\|u_t - u_{t-1}\right\|_1$. A natural question is whether similar results also hold in the unbounded setting, and how they compare to our lower bound in Theorem 3.5. Note that even in the exp-concave setting, the bound in Equation (4) is not necessarily tight. Indeed, Baby & Wang (2021) provide an algorithm which guarantees

$$R_T(\boldsymbol{u}) \le \widetilde{O}(Y^{4/3} d^{3.5} T^{1/3} C_T^{2/3}),$$

which does not match the lower bound w.r.t the dimension d. In contrast, our lower bound in Theorem 3.5 matches our upper bounds in all involved quantities (see Sections 3 and 4). Regardless, we also demonstrate in Appendix F.1 that the same $\widetilde{O}(Y^{4/3} d^{3.5} T^{1/3} C_T^{2/3})$ upper bound can be attained, even in unbounded domains, using the strongly-adaptive guarantees developed in Section 5.

4. Learning the Optimal Discount Factor

Recall that our goal from the outset has been to design algorithms that achieve favourable dynamic regret guarantees using *no prior knowledge*. To this end, we showed in Section 3 that the discounted VAW forecaster can achieve dynamic regret guarantees of the form $R_T(\boldsymbol{u}) \leq O\left(\sqrt{dP_T^\gamma(\boldsymbol{u})T} \vee d\log(T)\right)$ where $P_T^\gamma(\boldsymbol{u})$ is a certain measure of variability of the comparator sequence, and in Section 3.2 we showed that these penalties are unavoidable in general. However, these results hold under the assumption that the learner chooses discount rates satisfying special conditions (Equations (2) and (3)), either of which would require exceptionally strong prior knowlege to ensure. Indeed, the learner would need to know the future! In order to achieve our goal of learning in the complete absence of prior knowledge, we need to ensure that the learner can adequately guess or learn these ideal discount factors on-the-fly.

A common way to achieve runtime parameter-tuning of this sort would be to run many instances of the algorithm for different choices of γ in parallel, and combine the predictions using a suitable meta-algorithm. In particular, suppose we have a collection of algorithms $\mathcal{A}_1, \ldots, \mathcal{A}_N$ and on each round we can query each $\mathcal{A}_i$ for a prediction $y_t^{(i)} \in \mathbb{R}$. Moreover, suppose we have a meta-algorithm $\mathcal{A}_{\text{Meta}}$ which tells us how to combine these predictions by outputting a p_t from the N-dimensional simplex Δ_N. Then by predicting $\overline{y}_t = \sum_{i=1}^N p_{ti} y_t^{(i)}$, [2] for any benchmark sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ and any $j \in [N]$ we have

$$R_T(\boldsymbol{u}) = \sum_{t=1}^T \ell_t(\overline{y}_t) - \ell_t(u_t)$$
$$= \underbrace{\sum_{t=1}^T \ell_t(y_t^{(j)}) - \ell_t(u_t)}_{=:R_T^{\mathcal{A}_j}(\boldsymbol{u})} + \underbrace{\sum_{t=1}^T \ell_t(\overline{y}_t) - \ell_t(y_t^{(j)})}_{=:R_T^{\text{Meta}}(e_j)}$$

where the last line observes that $y_t^{(j)} = \langle x_t, w_t^{(j)} \rangle$. Hence, we may achieve our goal if we can ensure 1) that there is a $j \in [N]$ such that $\mathcal{A}_j$ uses a near-optimal discount factor γ_j, and 2) we can provide a meta-algorithm which guarantees low regret $R_T^{\text{Meta}}(e_j)$. We first investigate the latter point, and return to the former in Theorems 4.2 and 4.3.

The obvious approach to bounding the meta-algorithm's regret would be to observe that the losses $\ell_t(\overline{y}_t) = \frac{1}{2}(y_t - \overline{y}_t)^2$ are α_t-exp-concave for $\alpha_t = \frac{1}{2\max_i \ell_t(y_t^{(i)})}$ (Lemma D.2), which will allow us to apply an instance of the fixed-share algorithm (Cesa-Bianchi et al., 2012) to get:

$$R_T^{\text{Meta}}(e_j) \leq O\left(\frac{\log(NT)}{\alpha_{T+1}}\right) \leq O\Big(\max_{t,i} \ell_t(y_t^{(i)}) \log(NT)\Big),$$

as shown in Theorem E.1. However, just like in Section 3.1, the term $\max_{t,i} \ell_t(y_t^{(i)})$ is hard to quantify and could be be arbitrarily large in general. Fortunately the very same clipping trick used in Section 3.1 also works here: instead of having the meta-algorithm combine the *raw* predictions $y_t^{(i)}$, we can simply clip the predictions to a trust-region around a given reference point y_t^{Ref}. In Lemma D.3 we show that the clipping strategy detailed in Algorithm 2 incurs only an additional constant penalty in the regret. Then, using Lemma D.1, using these clipped predictions leads to

$$R_T^{\text{Meta}}(e_j) \leq O(\max_t (y_t - y_t^{\text{Ref}})^2 \log(NT)).$$

Note that a penalty of a similar order is already present in the regret of the VAW forecaster (*e.g.* Theorem 3.1) so this result will be sufficient for our purposes. Overall, the following theorem formalizes the argument described above. We provide a simplified statement here for brevity, but the full statement and its proof can be found in Appendix D.4.

Algorithm 2: Range-clipped Meta-algorithm

Input *Online learning algorithms* $\mathcal{A}_1, \ldots, \mathcal{A}_N$*, experts algorithm* $\mathcal{A}_{Meta}$ *over the simplex* Δ_N*.*
Initialize $\mathcal{A}_{Meta}, \mathcal{A}_1, \ldots, \mathcal{A}_N$*, and set* $M_1 = 0$
for $t = 1 : T$ **do**
  Receive features x_t
  Choose reference point y_t^{Ref}
  Define $\mathcal{B}_t = [y_t^{\text{Ref}} - M_t, y_t^{\text{Ref}} + M_t]$
  for $i = 1, \ldots, N$ **do**
    Send x_t to $\mathcal{A}_i$
    Get prediction $y_t^{(i)} = \langle x_t, w_t^{(i)} \rangle$ from $\mathcal{A}_i$
    Compute $\overline{y}_t^{(i)} = \text{Clip}_{\mathcal{B}_t}(y_t^{(i)})$
  end
  Get $p_t \in \Delta_N$ from $\mathcal{A}_{\text{Meta}}$
  Predict $\overline{y}_t = \sum_{i=1}^N p_{ti} \overline{y}_t^{(i)}$ and observe y_t
  Update $M_{t+1} = M_t \vee \left|y_t - y_t^{\text{Ref}}\right|$
  Send $\ell_t(w) = \frac{1}{2}(y_t - \langle x_t, w \rangle)^2$ to $\mathcal{A}_i$ $\forall i$
  Send $\ell_t(\overline{y}_t^{(1)}), \ldots, \ell_t(\overline{y}_t^{(N)})$ to $\mathcal{A}_{\text{Meta}}$
end

Theorem 4.1. *(simplified) Let* $\mathcal{A}_{Meta}$ *be the instance of fixed-share characterized in Theorem E.1. Then for any sequence* $\boldsymbol{u} = (u_1, \ldots, u_T)$ *in* $\mathbb{R}$ *and any* $j \in [N]$*, Algorithm 2 guarantees*

$$R_T(\boldsymbol{u}) \leq \widehat{O}\left(R_T^{\mathcal{A}_j}(\boldsymbol{u}) + \max_t \left(y_t - y_t^{Ref}\right)^2 \log(NT)\right),$$

where $\widehat{O}(\cdot)$ *hides* $\log\log$ *terms.*

[2] Recall from the introduction that because the features x_t are provided at the start of the round, we can work directly in the output space $\mathbb{R}$ if we so choose by setting $w_t = \overline{y}_t x_t / \|x_t\|^2$. Hence, given $\overline{y} \in \mathbb{R}$ we allow a slight abuse of notation by letting $\ell_t(\overline{y}) = \frac{1}{2}(y_t - \overline{y})^2$.

A similar target-clipping strategy was recently used by Mayo et al. (2022) to prove a static regret result for scale-free unconstrained online regression. Theorem 4.1 generalizes their approach by clipping to a trust-region of an arbitrary center $y_t^{\text{Ref}} \in \mathbb{R}$, and offers a somewhat streamlined argument which does not appeal to probabilistic notions such as mixibility.

Finally, with Theorem 4.1 in hand, we can achieve our desired result by running Algorithm 2 with the base algorithms $\mathcal{A}_i$ being instances of the discounted VAW forecaster with different discount factors γ. The following theorems show that for a well-chosen set of discount factors, we can make guarantees that match the bounds attained under oracle tuning of γ (Theorems 3.3 and 3.4), yet require no prior knowledge of any sort. Proofs can be found in Appendices D.5 and D.6 respectively.

Theorem 4.2. *Let $b > 1$, $\eta_{\min} = 2d$, $\eta_{\max} = dT$, and for all $i \in \mathbb{N}$ let $\eta_i = \eta_{\min} b^i \wedge \eta_{\max}$, and construct the set of discount factors $\mathcal{S}_\gamma = \left\{\gamma_i = \frac{\eta_i}{1+\eta_i} : i \in \mathbb{N}\right\} \cup \{0\}$. For any γ in $\mathcal{S}_\gamma$, let $\mathcal{A}_\gamma$ denote an instance of Algorithm 1 with discount γ.[3] Let $\mathcal{A}_{Meta}$ be an instance of the algorithm characterized in Theorem 4.1, and suppose we set $y_t^{Ref} = \widetilde{y}_t$ for all t. Then for any $\boldsymbol{u} = (u_1, \dots, u_T)$ in $\mathbb{R}^d$, Algorithm 2 guarantees*

$$R_T(\boldsymbol{u}) \le O\Bigg(d \max_t (y_t - y_t^{Ref})^2 \log(T) + b\sqrt{dP_T^{\gamma^*}(\boldsymbol{u}) \sum_{t=1}^T (y_t - \widetilde{y}_t)^2}\Bigg)$$

where $\gamma^* \in [0, 1]$ satisfies Equation (2).

Theorem 4.3. *Under the same conditions as Theorem 4.2, suppose each $\mathcal{A}_\gamma$ sets hints $\widetilde{y}_t = \overline{y}_t^\gamma = \text{Clip}_{\mathcal{B}_t}(\langle, x_t, w_t^\gamma \rangle)$, where $\mathcal{B}_t = [y_t^{Ref} - M_t, y_t^{Ref} + M_t]$ and $M_t = \max_{s<t} |y_s - y_s^{Ref}|$. Then for any $\boldsymbol{u} = (u_1, \dots, u_T)$ in $\mathbb{R}^d$, Algorithm 2 guarantees*

$$R_T(\boldsymbol{u}) \le O\Bigg(dP_T^{\gamma_{\min}}(\boldsymbol{u}) + d \max_t \left(y_t - y_t^{Ref}\right)^2 \log(T) + b\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)}\Bigg)$$

where $\gamma_{\min} = \frac{2d}{2d+1}$ and $\gamma^\circ \in [0, 1]$ satisfies Equation (3).

It is worth noting that Theorems 4.2 and 4.3 use knowledge of the horizon T to construct the set of experts. All of our results extend immediately to the unknown T setting as well via the standard doubling trick (Cesa-Bianchi & Lugosi, 2006), so for simplicity we treat T as part of the problem setting rather than a potentially unknown property of the data. An interesting direction for future development would be to construct the set of experts in a more on-the-fly way, so as to avoid using the doubling trick to adapt to unknown T.

5. Strongly-Adaptive Guarantees

While our original goal was only to achieve dynamic regret guarantees in the absence of prior knowledge, it turns out that we can actually achieve an even stronger result: dynamic regret guarantees that hold over *every sub-interal* $[a, b] \subseteq [1, T]$ *simultaneously*. To our knowledge, *strongly-adaptive* guarantees of this sort have previously only been achieved under various boundedness assumptions (Baby et al., 2021; Baby & Wang, 2022b;a; Jun et al., 2017; Cutkosky, 2020; Daniely et al., 2015).

The results can be derived using the results in the previous section. As shown in Appendix D.4, for any $[s, \tau] \subseteq [1, T]$, $\boldsymbol{u} = (u_s, \dots, u_\tau)$, and $\gamma \in \mathcal{S}_\gamma$, Algorithm 2 more generally guarantees that

$$R_{[s,\tau]}(\boldsymbol{u}) \le \widehat{O}\left(R_{[s,\tau]}^{\mathcal{A}_\gamma}(\boldsymbol{u}) + \max_t (y_t - y_t^{\text{Ref}})^2 \log(N\tau)\right),$$

where $R_{[s,\tau]}$ denotes the regret over sub-interval $[s, \tau] \subseteq [1, T]$. The only caveat is that the regret guarantees of the discounted VAW forecaster only hold when the algorithm *begins learning* on round s.[4] However, suppose that for each $s \in [1, T]$ and each $\gamma \in \mathcal{S}_\gamma$ we define an algorithm $\mathcal{A}_{\gamma,s}$ which uses discount γ but begins learning at time s. Then for any $[s, \tau]$ Lemma D.4 implies that there is a $\mathcal{A}_{\gamma,s}$ such that $R_{[s,\tau]}^{\mathcal{A}_{\gamma,s}}(\boldsymbol{u}) \le O(d \max_t (y_t - y_t^{\text{Ref}})^2 \log(\tau - s) + b\sqrt{dP_{[s,\tau]}^{\gamma^*}(\boldsymbol{u}) \sum_{t=s}^\tau (y_t - \widetilde{y}_t)^2})$. Plugging this back into the previous display and choosing $|\mathcal{S}_\gamma| \le O(\log(T))$, we have $N \le O(T \log(T))$ and an overall regret bound of

$$R_{[s,\tau]}(\boldsymbol{u}) \le \widehat{O}\Bigg(d \max_t (y_t - \widetilde{y}_t)^2 \log(T) + b\sqrt{dP_{[s,\tau]}^{\gamma^*}(\boldsymbol{u}) \sum_{t=s}^\tau (y_t - \widetilde{y}_t)^2}\Bigg).$$

This is the essence of the Follow the Leading History algorithm of Hazan & Seshadhri (2007; 2009).

While the above approach leads to a strongly-adaptive guarantee, it would be excessively expensive in general,

[3]For brevity, here we refer to an algorithm that directly predicts $\widetilde{y}_t$ on every round as being an instance of the discounted VAW forecaster with $\gamma = 0$. This terminology can be justified by Remark A.2, but for our purposes here it's sufficient to consider it convenient alias.

[4]More generally, it can be seen from the analysis that if the algorithm starts at time $t = 1$ and we try to bound the regret over $[s, \tau]$, then after telescoping the divergence terms we will end up with a non-trivial term $D_{\psi_s}(u_s | w_s)$ which is hard to quantify in general for $s > 1$ without further assumptions.

since we'd now have $O(T\log(T))$ total experts to update on every round. We may instead lower this to $O(\log^2(T))$ experts using the geometric covering intervals of Daniely et al. (2015); Veness et al. (2013). The idea is as follows: instead of initializing a new instance of each $\mathcal{A}_\gamma$ on every round $s \in [T]$, we will construct a set of intervals S such that any $[s,\tau] \subseteq [1,T]$ can be covered using only a small number of intervals from S. Then for each $\gamma \in \mathcal{S}_\gamma$ and each $I \in S$, we can define an instance of the discounted VAW forecaster $\mathcal{A}_{\gamma,I}$ which is run only during the interval I. The geometric covering intervals are constructed in such a way that 1) any round t can fall into at most $O(\log(T))$ of the intervals, and 2) any $[s,\tau] \subseteq [1,T]$ can be covered using only $O(\log(\tau - s))$ disjoint intervals from S. The first property ensures that there at most $O(\log^2(T))$ active experts on each round, while the second property implies that there is a disjoint set of intervals $I_1,\ldots,I_K$ such that $R_{[s,\tau]}(\boldsymbol{u}) = \sum_{i=1}^K R_{I_i}(\boldsymbol{u})$, so bounding each of these using a similar argument to the above followed by an application of Cauchy-Schwarz inequality yields

$$R_{[s,\tau]}(\boldsymbol{u}) \le \widehat{O}\Bigg(d\max_t (y_t - \widetilde{y}_t)^2\log^2(T) + b\sqrt{dP^{\gamma^*}_{[s,\tau]}(\boldsymbol{u})\sum_{t=s}^{\tau}(y_t - \widetilde{y}_t)^2}\Bigg),$$

where $P_{[s,\tau]}(\boldsymbol{u})$ is the total variability over the intervals and we've used $K\log(T) \le O(\log^2(T))$. Hence, overall the penalty we incur for using the geometric covering is a modest increase from $\log(T)$ to $K\log(T) \le O(\log^2(T))$ in the leading term. Likewise, a similar argument holds for our small-loss bounds. We provide a formal statement and proof of these results in Appendix F.

6. Conclusion

In this paper, we designed algorithms for online linear regression which achieve optimal dynamic regret guarantees, even in the absence of all prior knowledge. We developed a novel analysis of a discounted variant of the Vovk-Azoury-Warmuth forecaster, showing that it can guarantee dynamic regret of the form $R_T(\boldsymbol{u}) \le O\left(d\log(T) \vee \sqrt{dP^{\gamma}_T(\boldsymbol{u})T}\right)$ when equipped with an appropriate discount factor (Section 3). We also provided a matching lower bound, demonstrating that these penalties are unavoidable in general (Section 3.2). We then showed that the ideal discount factors can be learned on-the-fly, resulting in algorithms that can be applied with no prior knowledge yet still make optimal dynamic regret guarantees (Section 4) and strongly-adaptive guarantees (Section 5). These are the first algorithms for online linear regression that make meaningful guarantees without making assumptions of any kind on the underlying data.

An important direction for future work is to reduce the computational complexity of the algorithms. Similar to the traditional VAW forecaster, the approach developed here can be infeasible for very high-dimensional features, requiring roughly $O(d^2\log(T))$ computation every round. The d^2 factor likely can be reduced by extending our analysis to use modern sketching techniques (Luo et al., 2016), and the $\log(T)$ factor can possibly be reduced using similar techniques to the recent work of Lu & Hazan (2022).

Acknowledgements

We thank David Janz for valuable feedback on the initial draft of this work. AJ was supported by an NSERC CGS-D scholarship. AC acknowledges support from NSF grant no. 2211718.

Impact Statement

This paper presents theoretical work that improves online linear regression. We do not anticipate any significant negative societal consequences.

References

Azoury, K. S. and Warmuth, M. K. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine learning*, 43:211–246, 2001.

Baby, D. and Wang, Y.-X. Optimal dynamic regret in exp-concave online learning. In *Proceedings of Thirty Fourth Conference on Learning Theory*. PMLR, 2021.

Baby, D. and Wang, Y.-X. Optimal dynamic regret in lqr control. In *Advances in Neural Information Processing Systems*, 2022a.

Baby, D. and Wang, Y.-X. Optimal dynamic regret in proper online learning with strongly convex losses and beyond. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*. PMLR, 2022b.

Baby, D., Hasson, H., and Wang, Y. Dynamic regret for strongly adaptive methods and optimality of online krr, 2021.

Besbes, O., Gur, Y., and Zeevi, A. Non-stationary stochastic optimization. *Operations Research*, 2015.

Campolongo, N. and Orabona, F. A closer look at temporal variability in dynamic online learning, 2021.

Cesa-Bianchi, N. and Lugosi, G. *Prediction, learning, and games*. Cambridge university press, 2006.

Cesa-Bianchi, N., Gaillard, P., Lugosi, G., and Stoltz, G. Mirror descent meets fixed share (and feels no regret). *Advances in Neural Information Processing Systems*, 25, 2012.

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

Daniely, A., Gonen, A., and Shalev-Shwartz, S. Strongly adaptive online learning. In *Proceedings of the 32nd International Conference on Machine Learning*. PMLR, 2015.

Foster, D., Kale, S., and Karloff, H. Online sparse linear regression. In *29th Annual Conference on Learning Theory*. PMLR, 2016.

Gaillard, P., Gerchinovitz, S., Huard, M., and Stoltz, G. Uniform regret bounds over $\mathbb{R}^d$ for the sequential linear regression problem with the square loss. In *Proceedings of the 30th International Conference on Algorithmic Learning Theory*. PMLR, 2019.

Ghai, U., Lee, H., Singh, K., Zhang, C., and Zhang, Y. No-regret prediction in marginally stable systems. In *Proceedings of Thirty Third Conference on Learning Theory*. PMLR, 2020.

Hazan, E. Introduction to online convex optimization. *CoRR*, abs/1909.05207, 2019.

Hazan, E. and Seshadhri, C. Adaptive algorithms for online decision problems. In *Electronic colloquium on computational complexity (ECCC)*, number 088, 2007.

Hazan, E. and Seshadhri, C. Efficient learning algorithms for changing environments. In *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009.

Hazan, E. and Singh, K. Introduction to online nonstochastic control, 2022.

Hazan, E., Singh, K., and Zhang, C. Learning linear dynamical systems via spectral filtering. *Advances in Neural Information Processing Systems*, 2017.

Hazan, E., Lee, H., Singh, K., Zhang, C., and Zhang, Y. Spectral filtering for general linear dynamical systems. *Advances in Neural Information Processing Systems*, 2018.

Jacobsen, A. and Cutkosky, A. Parameter-free mirror descent. In *Proceedings of Thirty Fifth Conference on Learning Theory*. PMLR, 2022.

Jacobsen, A. and Cutkosky, A. Unconstrained online learning with unbounded losses. In *International Conference on Machine Learning (ICML)*. PMLR, 2023.

Jun, K.-S., Orabona, F., Wright, S., and Willett, R. Improved Strongly Adaptive Online Learning using Coin Betting. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*. PMLR, 2017.

Kalman, R. E. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 1960.

Kempka, M., Kotlowski, W., and Warmuth, M. K. Adaptive scale-invariant online algorithms for learning linear models. In *Proceedings of the 36th International Conference on Machine Learning*. PMLR, 2019.

Kivinen, J., Warmuth, M., and Hassibi, B. The p-norm generalization of the lms algorithm for adaptive filtering. *IEEE Transactions on Signal Processing*, 2006.

Kotłowski, W. Scale-invariant unconstrained online learning. In *Proceedings of the 28th International Conference on Algorithmic Learning Theory*. PMLR, 2017.

Kozdoba, M., Marecek, J., Tchrakian, T., and Mannor, S. On-line learning of linear dynamical systems: Exponential forgetting in kalman filters. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019.

Lu, Z. and Hazan, E. Efficient adaptive regret minimization, 2022.

Luo, H., Agarwal, A., Cesa-Bianchi, N., and Langford, J. Efficient second order online learning by sketching. In *Advances in Neural Information Processing Systems*, 2016.

Luo, H., Zhang, M., Zhao, P., and Zhou, Z.-H. Corralling a larger band of bandits: A case study on switching regret for linear bandits. In *Proceedings of Thirty Fifth Conference on Learning Theory*. PMLR, 2022.

Mayo, J. J., Hadiji, H., and van Erven, T. Scale-free unconstrained online learning for curved losses. In *Proceedings of Thirty Fifth Conference on Learning Theory*, 2022.

Mhammedi, Z. and Koolen, W. M. Lipschitz and comparator-norm adaptivity in online learning. In Abernethy, J. and Agarwal, S. (eds.), *Proceedings of Thirty Third Conference on Learning Theory*. PMLR, 2020.

Orabona, F., Crammer, K., and Cesa-Bianchi, N. A generalized online mirror descent with applications to classification and regression. *Mach. Learn.*, 2015.

Rashidinejad, P., Jiao, J., and Russell, S. Slip: Learning to predict in unknown dynamical systems with long-term memory. In *Advances in Neural Information Processing Systems*, 2020.

Simon, D. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

Tsiamis, A. and Pappas, G. J. Online learning of the kalman filter with logarithmic regret. *IEEE Transactions on Automatic Control*, 2022.

Veness, J., White, M., Bowling, M., and György, A. Partition tree weighting. In *2013 Data Compression Conference*, 2013.

Vovk, V. Competitive on-line statistics. *International Statistical Review*, 2001.

Yuan, J. and Lamperski, A. G. Trading-off static and dynamic regret in online least-squares and beyond. *CoRR*, abs/1909.03118, 2019.

Zhang, L., Lu, S., and Zhou, Z.-H. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems*, 2018.

Zhang, Z., Cutkosky, A., and Paschalidis, Y. Unconstrained dynamic regret via sparse coding. In *Advances in Neural Information Processing Systems*, 2023.

Zhao, P., Zhang, Y.-J., Zhang, L., and Zhou, Z.-H. Dynamic regret of convex and smooth functions. In *Advances in Neural Information Processing Systems*, 2020.

Zhao, P., Zhang, Y.-J., Zhang, L., and Zhou, Z.-H. Adaptivity and non-stationarity: Problem-dependent dynamic regret for online convex optimization. *Journal of Machine Learning Research*, 2024.

A. Proofs for Section 3 (Dynamic Regret via Discounting)

A.1. Equivalence to FTRL and Mirror Descent

We accomplish our analysis of the discounted VAW forecaster using the equivalence in the following proposition, proving both optimistic FTRL and and optimistic mirror descent interpretations of the discounted VAW forecaster. Equation (6) is perhaps the most natural interpretation of the update: it says that the discounted VAW forecaster chooses the w which minimizes the *discounted sum* $h_t(w)+\gamma\ell_{t-1}(w)+\gamma^2\ell_{t-2}(w)+\ldots$, thus placing greater emphasis on the most-recent losses and the hint function $h_t(w)$. However, it is not at all obvious how to analyze the dynamic regret of the discounted VAW forecaster when interpreted in this FTRL-like form. Rather, the key to our results in this work is to instead approach the analysis through the lens of the mirror descent update (Equation (7)). Interestingly, a similar mirror descent interpretation was used in the seminal work of Azoury & Warmuth (2001), though they did not account for an arbitrary $\widetilde{y}_t$ and they did not refer to the algorithm in terms of mirror descent.

Proposition A.1. *(Discounted VAW Forecaster) Let* $\gamma \in (0,1]$, $\lambda > 0$, $\widetilde{y}_1 = 0$, *and* $\widetilde{y}_t \in \mathbb{R}$ *for* $t > 1$. *Define* $h_t(w) = \frac{1}{2}(\widetilde{y}_t - \langle x_t, w\rangle)^2$ *and* $\ell_0(w) = \frac{\lambda}{2}\|w\|_2^2$. *Recursively define* $\Sigma_t = x_t x_t^\top + \gamma\Sigma_{t-1}$ *starting from* $\Sigma_0 = \lambda I$, *let* $\psi_t(w) = \frac{1}{2}\|w\|_{\Sigma_t}^2$ *and set* $w_1 = \arg\min_{w\in\mathbb{R}^d} \psi_1(w) = \mathbf{0}$. *Then the following are equivalent*

$$\Sigma_t^{-1}\Big[\widetilde{y}_t x_t + \gamma\sum_{s=1}^{t-1}\gamma^{t-1-s} y_s x_s\Big] \tag{5}$$

$$\arg\min_{w\in\mathbb{R}^d} h_t(w) + \gamma\sum_{s=0}^{t-1}\gamma^{t-1-s}\ell_s(s) \tag{6}$$

$$\arg\min_{w\in\mathbb{R}^d} (\gamma\ell_{t-1} - \gamma h_{t-1} + h_t)(w) + \gamma D_{\psi_{t-1}}(w|w_{t-1}) \tag{7}$$

Remark A.2. Note that with $\gamma = 0$, Equations (6) and (7) prescribe choosing any w_t satisfying $\langle w_t, x_t\rangle = \widetilde{y}_t$. The choice is not unique, but nevertheless it will often be convenient to refer to an algorithm which greedily predicts $\widetilde{y}_t$ on each round as an instance of Algorithm 1 with $\gamma = 0$.

Proof. The result follows by showing that Equations (6) and (7) are both equivalent to Equation (5). First consider the former, Equation (6). From the first-order optimality condition we have

$$\begin{aligned}
\mathbf{0} &= \nabla h_t(w_t) + \gamma\sum_{s=0}^{t-1}\gamma^{t-1-s}\nabla\ell_s(w_t)\\
&= -(\widetilde{y}_t - \langle x_t, w_t\rangle)x_t - \gamma\sum_{s=1}^{t-1}\gamma^{t-1-s}(y_s - \langle x_s, w_t\rangle)x_s + \gamma^t\lambda w_t,
\end{aligned}$$

where the last line recalls that we defined $\ell_0(w) = \frac{\lambda}{2}\|w\|_2^2$. Hence,

$$\begin{aligned}
\left(\gamma^t\lambda I + \sum_{s=1}^{t}\gamma^{t-s}x_s x_s^\top\right)w_t &= \widetilde{y}_t x_t + \sum_{s=1}^{t}\gamma^{t-s}y_s x_s\\
\implies w_t &= \left(\gamma^t\lambda I + \sum_{s=1}^{t}\gamma^{t-s}x_s x_s^\top\right)^{-1}\left[\widetilde{y}_t x_t + \gamma\sum_{s=1}^{t-1}\gamma^{t-1-s}y_s x_s\right]\\
&= \Sigma_t^{-1}\left[\widetilde{y}_t x_t + \gamma\sum_{s=1}^{t-1}\gamma^{t-1-s}y_s x_s\right],
\end{aligned}$$

where the last line can be seen by unrolling the recursion for Σ_t.

Likewise, consider Equation (7). From the first-order optimality condition $w_t = \arg\min_{w\in\mathbb{R}^d}(\gamma\ell_{t-1} - \gamma h_{t-1} + h_t)(w) + \gamma D_{\psi_{t-1}}(w|w_t)$, we have

$$\begin{aligned}
\mathbf{0} &= \gamma(\nabla\ell_{t-1}(w_t) - \nabla h_{t-1}(w_t)) + \nabla h_t(w_t) + \gamma\left[\nabla\psi_{t-1}(w_t) - \nabla\psi_{t-1}(w_{t-1})\right]\\
&= -\gamma y_{t-1}x_{t-1} + \gamma\widetilde{y}_{t-1}x_{t-1} - \widetilde{y}_t x_t + x_t x_t^\top w_t + \gamma\Sigma_{t-1}w_t - \gamma\Sigma_{t-1}w_{t-1}\\
&= -\gamma y_{t-1}x_{t-1} + \gamma\widetilde{y}_{t-1}x_{t-1} - \widetilde{y}_t x_t + \Sigma_t w_t - \gamma\Sigma_{t-1}w_{t-1},
\end{aligned}$$

where the last line observes that $\Sigma_t = x_t x_t^\top + \gamma\Sigma_{t-1}$ by construction. Hence, re-arranging we have

$$\Sigma_t w_t = \widetilde{y}_t x_t + \gamma y_{t-1}x_{t-1} - \gamma\widetilde{y}_{t-1}x_{t-1} + \gamma\Sigma_{t-1}w_{t-1}$$

and unrolling the recursion:

$$\begin{aligned}
&= \widetilde{y}_t x_t + \gamma y_{t-1}x_{t-1} - \gamma\widetilde{y}_{t-1}x_{t-1} + \gamma\left[\widetilde{y}_{t-1}x_{t-1} + \gamma y_{t-2}x_{t-2} - \gamma\widetilde{y}_{t-2}x_t + \gamma\Sigma_{t-2}w_{t-2}\right]\\
&= \widetilde{y}_t x_t + \gamma y_{t-1}x_{t-1} + \gamma^2 y_{t-2}x_{t-2} - \gamma^2\widetilde{y}_{t-2}x_{t-2} + \gamma^2\Sigma_{t-2}w_{t-2}\\
&= \dots\\
&= \widetilde{y}_t x_t - \gamma^{t-1}\widetilde{y}_1 x_1 + \gamma\sum_{s=1}^{t-1}\gamma^{t-1-s}y_s x_s\\
&= \widetilde{y}_t x_t + \gamma\sum_{s=1}^{t-1}\gamma^{t-1-s}y_s x_s,
\end{aligned}$$

for $\widetilde{y}_1 = 0$. Hence, applying Σ_t^{-1} to both sides we have

$$w_t = \Sigma_t^{-1}\left[\widetilde{y}_t x_t + \gamma\sum_{s=1}^{t-1}\gamma^{t-1-s}y_s x_s\right]$$

□

A.2. Proof of Theorem 3.1

Theorem 3.1. *Let $\lambda > 0$ and $\gamma \in (0,1]$. Then for any sequence $\boldsymbol{u} = (u_1,\dots,u_T)$ in $\mathbb{R}^d$, Algorithm 1 guarantees dynamic regret $R_T(\boldsymbol{u})$ bounded above by*

$$\begin{aligned}
&\frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{d}{2}\max_t(y_t-\widetilde{y}_t)^2\log\left(1+\frac{\sum_{t=1}^T\gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right)\\
&+\gamma\sum_{t=1}^{T-1}\left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right] + \frac{d}{2}\log(1/\gamma)\sum_{t=1}^T(y_t-\widetilde{y}_t)^2
\end{aligned}$$

where $F_t^\gamma(w) = \gamma^t\frac{\lambda}{2}\|w\|_2^2 + \sum_{s=1}^t\gamma^{t-s}\ell_s(w)$.

Proof. Begin by applying the regret template provided by Lemma A.3:

$$R_T(\boldsymbol{u}) \le \sum_{t=1}^T D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) + \sum_{t=1}^T h_{t+1}(u_t) - h_t(u_t) + \frac{1}{2}\sum_{t=1}^T(y_t-\widetilde{y}_t)^2\|x_t\|_{\Sigma_t^{-1}}^2,$$

bound the first two summations using Lemma A.4:

$$\le \frac{\gamma\lambda}{2}\|u_1\|_2^2 + h_{T+1}(u_T) + \gamma\sum_{t=1}^{T-1}\left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right] + \frac{1}{2}\sum_{t=1}^T(y_t-\widetilde{y}_t)^2\|x_t\|_{\Sigma_t^{-1}}^2,$$

and apply a discounted variant of the log-determinant lemma (Lemma G.2) to bound the final summation:

$$\begin{aligned}
&\le \frac{\gamma\lambda}{2}\|u_1\|_2^2 + h_{T+1}(u_T) + \frac{d}{2}\max_t(y_t-\widetilde{y}_t)^2\log\left(1+\frac{\sum_{t=1}^T\gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right)\\
&\quad+\gamma\sum_{t=1}^{T-1}\left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right] + \frac{d}{2}\log(1/\gamma)\sum_{t=1}^T(y_t-\widetilde{y}_t)^2
\end{aligned}$$

Finally, since the regret does not depend on $h_{T+1}(\cdot)$ we may set $h_{T+1}(\cdot)\equiv 0$ in the analysis and hide constants to arrive at the stated bound. □

A.2.1. Proof of Lemma A.3

The following lemma provides the base regret decomposition that we use as a jumping-off point to prove Theorem 3.1. The result follows using mostly standard mirror descent analysis, though with a bit of additional care to handle issues related to the discounted regularizer.

Lemma A.3. *Let $\gamma \in (0,1]$. Then for any sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ in $\mathbb{R}^d$, Algorithm 1 guarantees*

$$\begin{aligned} R_T(\boldsymbol{u}) \le \sum_{t=1}^T D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) \\ + \sum_{t=1}^T h_{t+1}(u_t) - h_t(u_t) \\ + \sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2 \|x_t\|^2_{\Sigma_t^{-1}} \end{aligned}$$

Proof. We will proceed following a mirror-descent-based analysis, and thus begin by exposing the terms $(\gamma\ell_t - \gamma h_t + h_{t+1})(w_{t+1})$ observed in the mirror-descent interpretation of the update (Equation (7)):

$$\begin{aligned} R_T(\boldsymbol{u}) &= \sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) \\ &= \sum_{t=1}^T \gamma\left[\ell_t(w_t) - \ell_t(u_t)\right] + (1-\gamma)\sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) \\ &= \sum_{t=1}^T \gamma\left[(\ell_t - h_t)(w_t) - (\ell_t - h_t)(u_t)\right] + \sum_{t=1}^T \gamma h_t(w_t) - \gamma h_t(u_t) \\ &\quad + (1-\gamma)\sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) \\ &= \sum_{t=1}^T \gamma\left[(\ell_t - h_t)(w_{t+1}) - (\ell_t - h_t)(u_t)\right] + \sum_{t=1}^T \gamma h_t(w_t) - \gamma h_t(u_t) \\ &\quad + \gamma\sum_{t=1}^T (\ell_t - h_t)(w_t) - (\ell_t - h_t)(w_{t+1}) \\ &\quad + (1-\gamma)\sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) \\ &= \sum_{t=1}^T (\gamma\ell_t - \gamma h_t + h_{t+1})(w_{t+1}) - (\gamma\ell_t - \gamma h_t + h_{t+1})(u_t) \\ &\quad + \sum_{t=1}^T \gamma h_t(w_t) - h_{t+1}(w_{t+1}) + \sum_{t=1}^T h_{t+1}(u_t) - \gamma h_t(u_t) \\ &\quad + \gamma\sum_{t=1}^T (\ell_t - h_t)(w_t) - (\ell_t - h_t)(w_{t+1}) \\ &\quad + (1-\gamma)\sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) \end{aligned}$$

Re-arranging factors of γ from the second-line and observing that $\sum_{t=1}^{T} h_t(w_t) - h_{t+1}(w_{t+1}) = h_1(w_1) - h_{T+1}(w_{T+1})$:

$$
\begin{aligned}
&= \sum_{t=1}^{T} (\gamma\ell_t - \gamma h_t + h_{t+1})(w_{t+1}) - (\gamma\ell_t - \gamma h_t + h_{t+1})(u_t) \\
&\quad + \sum_{t=1}^{T} h_t(w_t) - h_{t+1}(w_{t+1}) + \sum_{t=1}^{T} -(1-\gamma)h_t(w_t) + (1-\gamma)h_t(u_t) + \sum_{t=1}^{T} h_{t+1}(u_t) - h_t(u_t) \\
&\quad + \gamma \sum_{t=1}^{T} (\ell_t - h_t)(w_t) - (\ell_t - h_t)(w_{t+1}) \\
&\quad + (1-\gamma) \sum_{t=1}^{T} \ell_t(w_t) - \ell_t(u_t) \\
&= \sum_{t=1}^{T} (\gamma\ell_t - \gamma h_t + h_{t+1})(w_{t+1}) - (\gamma\ell_t - \gamma h_t + h_{t+1})(u_t) \\
&\quad + h_1(w_1) - h_{T+1}(w_{T+1}) + \sum_{t=1}^{T} h_{t+1}(u_t) - h_t(u_t) \\
&\quad + \gamma \sum_{t=1}^{T} (\ell_t - h_t)(w_t) - (\ell_t - h_t)(w_{t+1}) \\
&\quad + (1-\gamma) \sum_{t=1}^{T} (\ell_t - h_t)(w_t) - (\ell_t - h_t)(u_t)
\end{aligned} \tag{8}
$$

Moreover, from the first-order optimality condition $w_{t+1} = \arg\min_{w \in \mathbb{R}^d} (\gamma\ell_t - \gamma h_t + h_{t+1})(w) + \gamma D_{\psi_t}(w|w_t)$, we have

$$
\langle \nabla(\gamma\ell_t - \gamma h_t + h_{t+1})(w_{t+1}) + \gamma\nabla\psi_t(w_{t+1}) - \gamma\nabla\psi_t(w_t), w_{t+1} - u_t \rangle \le 0
$$

so re-arranging:

$$
\begin{aligned}
\langle \nabla(\gamma\ell_t - \gamma h_t + h_{t+1})(w_{t+1}), w_{t+1} - u_t \rangle &\le \gamma \langle \nabla\psi_t(w_t) - \nabla\psi_t(w_{t+1}), w_{t+1} - u_t \rangle \\
&= \gamma D_{\psi_t}(u_t|w_t) - \gamma D_{\psi_t}(u_t|w_{t+1}) - \gamma D_{\psi_t}(w_{t+1}|w_t),
\end{aligned}
$$

where the last line uses the three-point relation for bregman divergences, $\langle \nabla f(w) - \nabla f(w'), w' - u \rangle = D_f(u|w) - D_f(u|w') - D_f(w'|w)$. Thus,

$$
\begin{aligned}
&\sum_{t=1}^{T} (\gamma\ell_t - \gamma h_t + h_{t+1})(w_{t+1}) - (\gamma\ell_t - \gamma h_t + h_{t+1})(u_t) \\
&\quad \overset{(a)}{=} \sum_{t=1}^{T} \langle \nabla(\gamma\ell_t - \gamma h_t + h_{t+1})(w_{t+1}), w_{t+1} - u_t \rangle - D_{\gamma\ell_t - \gamma h_t + h_{t+1}}(u_t|w_{t+1}) \\
&\quad \le \sum_{t=1}^{T} \gamma D_{\psi_t}(u_t|w_t) - \gamma D_{\psi_t}(u_t|w_{t+1}) - \gamma D_{\psi_t}(w_{t+1}|w_t) - D_{\gamma\ell_t - \gamma h_t + h_{t+1}}(u_t|w_{t+1}) \\
&\quad \overset{(b)}{=} \sum_{t=1}^{T} \gamma D_{\psi_t}(u_t|w_t) - \gamma D_{\psi_t}(u_t|w_{t+1}) - D_{h_{t+1}}(u_t|w_{t+1}) - \gamma D_{\psi_t}(w_{t+1}|w_t) \\
&\quad \overset{(c)}{=} \sum_{t=1}^{T} \gamma D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) - \gamma D_{\psi_t}(w_{t+1}|w_t) \\
&\quad = \sum_{t=1}^{T} D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) - (1-\gamma) D_{\psi_t}(u_t|w_t) - \gamma D_{\psi_t}(w_{t+1}|w_t),
\end{aligned}
$$

where (a) uses the definition of Bregman divergence to re-write $f(w) - f(u) = \langle \nabla f(w), w - u \rangle - D_f(u|w)$, (b) observes that $\gamma(\ell_t - h_t)(w) = \gamma\left(\frac{1}{2}y_t^2 - \frac{1}{2}\widetilde{y}_t^2 + (\widetilde{y}_t - y_t)\langle x_t, w \rangle\right)$, so $D_{\gamma\ell_t - \gamma h_t + h_{t+1}}(\cdot|\cdot) = D_{h_{t+1}}(\cdot|\cdot)$ due to the invariance of Bregman divergences to linear terms, and (c) recalls that $\Sigma_{t+1} = x_{t+1}x_{t+1}^\top + \gamma\Sigma_t$ so that overall we have:

$$\begin{aligned}\gamma D_{\psi_t}(u_t|w_{t+1}) + D_{h_{t+1}}(u_t|w_{t+1}) &= \frac{\gamma}{2}\left\|u_t - w_{t+1}\right\|_{\Sigma_t}^2 + \frac{1}{2}\left\langle x_{t+1}, u_t - w_{t+1}\right\rangle^2 \\ &= \frac{1}{2}\left\|u_t - w_{t+1}\right\|_{\Sigma_{t+1}}^2 \\ &= D_{\psi_{t+1}}(u_t|w_{t+1}).\end{aligned}$$

Plugging this back into Equation (8), we have

$$\begin{aligned}R_T(\boldsymbol{u}) \le \sum_{t=1}^T & D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) \\ &+ h_1(w_1) - h_{T+1}(w_{T+1}) + \sum_{t=1}^T h_{t+1}(u_t) - h_t(u_t) \\ &+ \gamma\sum_{t=1}^T (\ell_t - h_t)(w_t) - (\ell_t - h_t)(w_{t+1}) - D_{\psi_t}(w_{t+1}|w_t) \\ &+ (1-\gamma)\sum_{t=1}^T (\ell_t - h_t)(w_t) - (\ell_t - h_t)(u_t) - D_{\psi_t}(u_t|w_{t+1}).\end{aligned}$$

Finally, observe that for any $u, v \in \mathbb{R}^d$, $(\ell_t - h_t)(u) - (\ell_t - h_t)(v) = (\widetilde{y}_t - y_t)\left\langle x_t, u - v\right\rangle$, so an application of Fenchel-Young inequality yields

$$\begin{aligned}(\ell_t - h_t)(u) - (\ell_t - h_t)(v) - D_{\psi_t}(v|u) &= (\widetilde{y}_t - y_t)\left\langle x_t, u - v\right\rangle - \frac{1}{2}\left\|u - v\right\|_{\Sigma_t}^2 \\ &\le \frac{1}{2}(y_t - \widetilde{y}_t)^2\left\|x_t\right\|_{\Sigma_t^{-1}}^2.\end{aligned}$$

Applying this in the last two lines of the previous display yields

$$\begin{aligned}R_T(\boldsymbol{u}) \le \sum_{t=1}^T & D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) \\ & \underbrace{h_1(w_1) - h_{T+1}(w_{T+1})}_{\le 0} + \sum_{t=1}^T h_{t+1}(u_t) - h_t(u_t) \\ & \gamma\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2\left\|x_t\right\|_{\Sigma_t^{-1}}^2 + (1-\gamma)\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2\left\|x_t\right\|_{\Sigma_t^{-1}}^2 \\ \le \sum_{t=1}^T & D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) \\ &+ \sum_{t=1}^T h_{t+1}(u_t) - h_t(u_t) \\ &+ \sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2\left\|x_t\right\|_{\Sigma_t^{-1}}^2\end{aligned}$$

□

A.2.2. Proof of Lemma A.4

The following lemma bounds the sum of divergence terms. Intuitively, the goal here is to remove all instances of w_t from the analysis, since in an unbounded domain any terms depending on w_t will be hard to quantify and could be arbitrarily large in general. Lemma A.4 shows how get rid of the w_t-dependent terms left in the bound from Lemma A.3, such that only dependencies on the comparators u_t remain.

Lemma A.4. *Under the same conditions as Lemma A.3,*

$$\sum_{t=1}^{T} D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) + \sum_{t=1}^{T} h_{t+1}(u_t) - h_t(u_t) \leq \frac{\gamma\lambda}{2}\left\|u_1\right\|_2^2 + h_{T+1}(u_T) + \gamma\sum_{t=1}^{T-1} F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t).$$

where $F_t^\gamma(w) = \sum_{s=0}^t \gamma^{t-s}\ell_s(w)$.

Proof. Observe that by Lemma G.1 we have $D_{\ell_t}(u|v) = \frac{1}{2}\langle x_t, u-v\rangle^2 = D_{h_t}(u|v)$ for any $u, v \in W$. Hence, letting $F_t^\gamma(w) = \sum_{s=0}^t \gamma^{t-s}\ell_s(w)$ and $\widehat{F}_t^\gamma(w) = h_t(w) + \gamma F_{t-1}^\gamma(w)$, and recalling $\psi_t(w) = \frac{1}{2}\left\|w\right\|_{\Sigma_t}^2 = \frac{\gamma^t\lambda}{2}\left\|w\right\|_2^2 + \frac{1}{2}\sum_{s=1}^t \gamma^{t-s}\langle x_s, w\rangle^2$, we have $D_{\psi_t}(u|v) = D_{\widehat{F}_t^\gamma}(u|v)$ for any $u, v \in\in \mathbb{R}^d$. Thus:

$$\begin{aligned}
&\sum_{t=1}^{T} D_{\psi_t}(u_t|w_t) - D_{\psi_{t+1}}(u_t|w_{t+1}) \\
&\quad= D_{\psi_1}(u_1|w_1) - D_{\psi_{T+1}}(u_T|w_{T+1}) + \sum_{t=2}^{T} D_{\psi_t}(u_t|w_t) - D_{\psi_t}(u_{t-1}|w_t) \\
&\quad= D_{\psi_1}(u_1|w_1) - D_{\psi_{T+1}}(u_T|w_{T+1}) + \sum_{t=2}^{T} D_{\widehat{F}_t^\gamma}(u_t|w_t) - D_{\widehat{F}_t^\gamma}(u_{t-1}|w_t) \\
&\quad= D_{\psi_1}(u_1|w_1) - D_{\psi_{T+1}}(u_T|w_{T+1}) + \sum_{t=2}^{T} \widehat{F}_t^\gamma(u_t) - \widehat{F}_t^\gamma(u_{t-1}) - \left\langle \nabla\widehat{F}_t^\gamma(w_t), u_t - u_{t-1}\right\rangle.
\end{aligned}$$

Moreover, by Proposition A.1 we have

$$w_t = \operatorname*{arg\,min}_{w\in\mathbb{R}^d}\ h_t(w) + \gamma\sum_{s=0}^{t-1}\gamma^{t-1-s}\ell_s(w) = \operatorname*{arg\,min}_{w\in\mathbb{R}^d}\ \widehat{F}_t^\gamma(w),$$

hence by convexity of $\widehat{F}_t^\gamma$ and the first-order optimality condition we have $\nabla\widehat{F}_t^\gamma(w_t) = \mathbf{0}$, so overall we have

$$\begin{aligned}
&\sum_{t=1}^{T} D_{\psi_t}(u_t|w_t) - D_{\psi_t}(u_t|w_{t+1}) + \sum_{t=1}^{T} h_{t+1}(u_t) - h_t(u_t) \\
&\quad= D_{\psi_1}(u_1|w_1) - D_{\psi_{T+1}}(u_T|w_{T+1}) + \sum_{t=2}^{T} \widehat{F}_t^\gamma(u_t) - \widehat{F}_t^\gamma(u_{t-1}) + \sum_{t=1}^{T} h_{t+1}(u_t) - h_t(u_t) \\
&\quad= D_{\psi_1}(u_1|w_1) - D_{\psi_{T+1}}(u_T|w_{T+1}) + \sum_{t=2}^{T}\left[h_t(u_t) - h_t(u_{t-1}) + \gamma F_{t-1}^\gamma(u_t) - \gamma F_{t-1}^\gamma(u_{t-1})\right] + \sum_{t=1}^{T} h_{t+1}(u_t) - h_t(u_t) \\
&\quad= D_{\psi_1}(u_1|w_1) - D_{\psi_{T+1}}(u_T|w_{T+1}) + \gamma\sum_{t=1}^{T-1} F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t) + \sum_{t=2}^{T} h_{t+1}(u_t) - h_t(u_{t-1}) + h_2(u_1) - h_1(u_1) \\
&\quad= D_{\psi_1}(u_1|w_1) - D_{\psi_{T+1}}(u_T|w_{T+1}) + h_{T+1}(u_T) - h_1(u_1) + \gamma\sum_{t=1}^{T-1} F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t).
\end{aligned}$$

Finally, observe that with $w_1 = \mathbf{0}$ and $\widetilde{y}_1 = 0$ we have

$$D_{\psi_1}(u_1|w_1) = \psi_1(u_1) - \psi_1(\mathbf{0}) - \langle\nabla\psi_1(\mathbf{0}), u_1\rangle = h_1(u_1) + \gamma\ell_0(u_1) = h_1(u_1) + \frac{\gamma\lambda}{2}\left\|u_1\right\|_2^2$$

so we can express the bound as the bound as

$$\begin{aligned}
&\sum_{t=1}^{T} D_{\psi_t}(u_t|w_t) - D_{\psi_t}(u_t|w_{t+1}) + \sum_{t=1}^{T} h_{t+1}(u_t) - h_t(u_t) \\
&\qquad\leq \frac{\gamma\lambda}{2}\left\|u_1\right\|_2^2 + h_{T+1}(u_T) + \gamma\sum_{t=1}^{T-1} F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t).
\end{aligned}$$

□

A.3. Proof of Lemma 3.2

The following lemma bounds the variability and stability terms from Theorem 3.1 to expose a more explicit trade-off in terms of the discount factor γ.

Lemma 3.2. Let $\ell_0, \ell_1, \ldots, \ell_T$ be arbitrary non-negative functions, $0 < \gamma \le \beta < 1$, and $F_t^\gamma(w) = \sum_{s=0}^t \gamma^{t-s} \ell_s(w)$. For all t, define

$$\bar{d}_t^\beta(u, v) = \frac{1}{\sum_{s=0}^t \beta^{t-s}} \sum_{s=0}^t \beta^{t-s} \left[\ell_s(u) - \ell_s(v)\right]_+$$

and let $P_T^\beta(\boldsymbol{u}) = \sum_{t=1}^{T-1} \bar{d}_t^\beta(u_{t+1}, u_t)$. Then for any $V_T \ge 0$,

$$\gamma \sum_{t=1}^{T-1} \left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right] + \log\left(\frac{1}{\gamma}\right) V_T \le \frac{\beta}{1-\beta} P_T^\beta(\boldsymbol{u}) + \frac{1-\gamma}{\gamma} V_T$$

Proof. The first summation can be bounded as

$$\begin{aligned}
\gamma \sum_{t=1}^{T-1} \left[F_t^\gamma(u_t) - F_t^\gamma(u_{t-1})\right] &= \gamma \sum_{t=1}^{T-1} \sum_{s=0}^t \gamma^{t-s} \left[\ell_s(u_{t+1}) - \ell_s(u_t)\right] \\
&\le \gamma \sum_{t=1}^{T-1} \sum_{s=0}^t \gamma^{t-s} \left[\ell_s(u_{t+1}) - \ell_s(u_t)\right]_+ \\
&\le \beta \sum_{t=1}^{T-1} \sum_{s=0}^t \frac{\sum_{s'=0}^t \beta^{t-s'}}{\sum_{s'=0}^t \beta^{t-s'}} \beta^{t-s} \left[\ell_s(u_{t+1}) - \ell_s(u_t)\right]_+ \\
&\le \frac{\beta}{1-\beta} \sum_{t=1}^{T-1} \sum_{s=0}^t \frac{\beta^{t-s}}{\sum_{s'=0}^t \beta^{t-s'}} \left[\ell_s(u_{t+1}) - \ell_s(u_t)\right]_+ \\
&= \frac{\beta}{1-\beta} P_T^\beta(\boldsymbol{u}),
\end{aligned}$$

where the last inequality uses $\sum_{s=0}^t \beta^{t-s} = \frac{1-\beta^{t+1}}{1-\beta} \le \frac{1}{1-\beta}$. Using this along with the elementary inequality $\log(x) \le x - 1$, for any $V_T \ge 0$ we have

$$\begin{aligned}
\gamma \sum_{t=1}^{T-1} \left[F_t^\gamma(u_t) - F_t^\gamma(u_{t-1})\right] + \log\left(\frac{1}{\gamma}\right) V_T &\le \frac{\beta}{1-\beta} P_T^\beta(\boldsymbol{u}) + \left(\frac{1}{\gamma} - 1\right) V_T \\
&= \frac{\beta}{1-\beta} P_T^\beta(\boldsymbol{u}) + \frac{1-\gamma}{\gamma} V_T
\end{aligned}$$

□

A.4. Existence of a Good Discount Factor

The following lemma establishes the existence of a discount factor that will lead to favorable tuning of the γ-dependent terms in Lemma 3.2.

Lemma A.5. *Let $\ell_0, \ell_1, \ldots$ be arbitrary non-negative functions, $V_T \ge 0$, denote $\bar{d}_t^\gamma(u, v) = \frac{\sum_{s=0}^t \gamma^{t-s}[\ell_s(u) - \ell_s(v)]_+}{\sum_{s=0}^t \gamma^{t-s}}$ for $\gamma \in [0, 1]$, and let $P_T^\gamma(\boldsymbol{u}) = \sum_{t=1}^{T-1} \bar{d}_t^\gamma(u_{t+1}, u_t)$. Then there is a $\gamma^* \in [0, 1]$ such that*

$$\gamma^* = \frac{\sqrt{V_T}}{\sqrt{V_T} + \sqrt{P_T^{\gamma^*}(\boldsymbol{u})}}.$$

Proof. First, notice that that any such γ with the stated property must be in $[0, 1]$ since

$$0 \le \frac{\sqrt{V_T}}{\sqrt{V_T} + \sqrt{P_T^\gamma(\boldsymbol{u})}} \le \frac{\sqrt{V_T}}{\sqrt{V_T}} = 1.$$

Next, observe that the condition can be equivalently expressed as follows:

$$\begin{aligned}
&\gamma = \frac{\sqrt{V_T}}{\sqrt{V_T} + \sqrt{P_T^\gamma(\boldsymbol{u})}} \\
\iff & \sqrt{V_T}(1-\gamma) = \gamma\sqrt{P_T^\gamma(\boldsymbol{u})} \\
&= \gamma\sqrt{\sum_{t=1}^{T-1}\sum_{s=0}^{t} \frac{\gamma^{t-s}}{\sum_{s=0}^{t}\gamma^{t-s}}\left[\ell_s(u_{t+1}) - \ell_s(u_t)\right]_+} \\
&= \gamma\sqrt{\sum_{t=1}^{T-1}\sum_{s=0}^{t} \frac{\gamma^{t-s}}{1-\gamma^{t+1}}(1-\gamma)\left[\ell_s(u_{t+1}) - \ell_s(u_t)\right]_+} \\
\iff & \sqrt{V_T(1-\gamma)} = \gamma\sqrt{\sum_{t=1}^{T-1}\sum_{s=0}^{t} \frac{\gamma^{t-s}}{1-\gamma^{t+1}}\left[\ell_s(u_{t+1}) - \ell_s(u_t)\right]_+}.
\end{aligned}$$

The quantity on the LHS begins at $\sqrt{V_T}$ (for $\gamma = 0$) and then decreases to 0 as a function of γ. Likewise, the RHS begins at 0 (for $\gamma = 0$) and increases as a function of γ, approaching ∞ as $\gamma \to 1$. Hence, there must be some $\gamma \in [0,1]$ at which the two lines cross, and hence a $\gamma \in [0,1]$ which satisfies the above relation, so there is a $\gamma \in [0,1]$ such that

$$\gamma = \frac{\sqrt{V_T}}{\sqrt{V_T} + \sqrt{P_T^\gamma(\boldsymbol{u})}}.$$

□

A.5. Proof of Theorem 3.3

Now combining everything we've seen in the previous sections, we can easily prove the following bound for the discounted VAW forecaster under *oracle tuning* of the discount factor.

Theorem 3.3. *For any sequences $y_1, \dots, y_T$ and $\widetilde{y}_1, \dots, \widetilde{y}_T$ in $\mathbb{R}$ and any sequence $\boldsymbol{u} = (u_1, \dots, u_T)$ in $\mathbb{R}^d$, there is a discount factor $\gamma^* \in [0,1]$ satisfying*

$$\gamma^* = \frac{\sqrt{\frac{d}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}}{\sqrt{\frac{d}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2} + \sqrt{P_T^{\gamma^*}(\boldsymbol{u})}} \tag{2}$$

with which the regret of Algorithm 1 is bounded above by

$$\begin{aligned}
R_T(\boldsymbol{u}) \le O\Bigg(& d\max_t (y_t - \widetilde{y}_t)^2 \log(T) \\
&+ \sqrt{dP_T^{\gamma^*}(\boldsymbol{u})\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}\Bigg)
\end{aligned}$$

Proof. Lemma A.5 shows that for any sequence $\boldsymbol{u} = (u_1, \dots, u_T)$, there is a $\gamma^* \in [0,1]$ such that

$$\gamma^* = \frac{\sqrt{d\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2}}{\sqrt{d\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2} + \sqrt{P_T^{\gamma^*}(\boldsymbol{u})}},$$

so choosing $\gamma = \gamma^*$ and applying Theorem 3.1, we have

$$\begin{aligned} R_T(\boldsymbol{u}) &\leq \frac{\gamma^*\lambda}{2}\left\|u_1\right\|_2^2 + \frac{d}{2}\max_t(y_t - \widetilde{y}_t)^2\log\left(1 + \frac{\sum_{t=1}^T \left\|x_t\right\|_2^2}{\lambda d}\right) \\ &\quad + \gamma^*\sum_{t=1}^{T-1}\left[F_t^{\gamma^*}(u_{t+1}) - F_t^{\gamma^*}(u_t)\right] + \frac{d}{2}\log\left(1/\gamma^*\right)\sum_{t=1}^T(y_t - \widetilde{y}_t)^2 \\ &\overset{(*)}{\leq} \frac{\lambda}{2}\left\|u_1\right\|_2^2 + \frac{d}{2}\max_t(y_t - \widetilde{y}_t)^2\log\left(1 + \frac{\sum_{t=1}^T \left\|x_t\right\|_2^2}{\lambda d}\right) \\ &\quad + \frac{\gamma^*}{1-\gamma^*}P_T^{\gamma^*}(\boldsymbol{u}) + \frac{1-\gamma^*}{\gamma^*}\frac{d}{2}\sum_{t=1}^T(y_t - \widetilde{y}_t)^2 \\ &= \frac{\lambda}{2}\left\|u_1\right\|_2^2 + \frac{d}{2}\max_t(y_t - \widetilde{y}_t)^2\log\left(1 + \frac{\sum_{t=1}^T \left\|x_t\right\|_2^2}{\lambda d}\right) + \sqrt{2dP_T^{\gamma^*}(\boldsymbol{u})\sum_{t=1}^T(y_t - \widetilde{y}_t)^2} \end{aligned}$$

where $(*)$ uses Lemma 3.2 (with $\beta = \gamma = \gamma^*$). The stated result follows by hiding lower-order terms. □

B. Proofs for Section 3.1 (Small-loss Bounds via Self-confident Predictions)

B.1. Proof of Theorem 3.4

We split the proof of Theorem 3.4 into two parts. The following lemma, proven in Appendix B.1.1, first derives an initial regret template that does most of the heavy lifting. We will later re-use this template in the proof of Theorem 4.3 to avoid repeating the argument. The high-level intuition is that choosing hints $\widetilde{y}_t \approx \langle x_t, w_t\rangle$ leads to $\sum_{t=1}^T(y_t - \widetilde{y}_t)^2 \approx \sum_{t=1}^T \ell_t(w_t)$, which leads to a self-bounding argument that lets us replace $\sum_{t=1}^T(y_t - \widetilde{y}_t)^2$ with $\sum_{t=1}^T \ell_t(u_t)$ in the regret bound. We defer proof of the lemma to the next subsection, Appendix B.1.1.

Lemma B.1. *Let $y_t^{Ref} \in \mathbb{R}$ be an arbitrary reference point, available at the start of round t, and let $\mathcal{B}_t = \left\{y \in \mathbb{R} : y_t^{Ref} - M_t \leq y \leq y_t^{Ref} + M_t\right\}$ for $M_t = \max_{s<t}\left|y_s - y_s^{Ref}\right|$. Suppose that we apply Algorithm 1 with hints $\widetilde{y}_t = \overline{y}_t := \text{Clip}_{\mathcal{B}_t}(\langle x_t, w_t\rangle)$. Then for any sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ in $\mathbb{R}^d$ and any $\gamma, \beta \in [0,1]$ such that $\beta \geq \gamma \geq \gamma_{\min} = \frac{2d}{2d+1}$,*

$$\begin{aligned} R_T(\boldsymbol{u}) \leq \gamma\lambda\left\|u_1\right\|_2^2 + 4d\max_t(y_t - y_t^{Ref})^2\log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\left\|x_t\right\|_2^2}{\lambda d}\right) \\ + 2\frac{\beta}{1-\beta}P_T^{\beta}(\boldsymbol{u}) + \frac{1-\gamma}{\gamma}2d\sum_{t=1}^T \ell_t(u_t) \end{aligned}$$

Now using this template, Theorem 3.4 is easily proven by plugging in the stated discount factor $\gamma = \gamma^\circ \vee \gamma_{\min}$

Theorem 3.4. *Let $y_t^{Ref} \in \mathbb{R}$ be an arbitrary reference point and let $\mathcal{B}_t = \left[y_t^{Ref} - M_t, y_t^{Ref} + M_t\right]$ for $M_t = \max_{s<t}\left|y_s - y_s^{Ref}\right|$. Suppose that we apply Algorithm 1 with hints $\widetilde{y}_t = \text{Clip}_{\mathcal{B}_t}(\langle x_t, w_t\rangle)$. Then for any sequence of losses $\ell_1, \ldots, \ell_T$ and any sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ in $\mathbb{R}^d$, there is a $\gamma^\circ \in [0,1]$ satisfying*

$$\gamma^\circ = \frac{\sqrt{d\sum_{t=1}^T \ell_t(u_t)}}{\sqrt{d\sum_{t=1}^T \ell_t(u_t)} + \sqrt{P_T^{\gamma^\circ}(\boldsymbol{u})}}. \tag{3}$$

Moreover, running Algorithm 1 with discount $\gamma^\circ \vee \gamma_{\min}$ for $\gamma_{\min} = \frac{2d}{2d+1}$ ensures regret bounded above by

$$\begin{aligned} R_T(\boldsymbol{u}) \leq O\Bigg(&dP_T^{\gamma_{\min}}(\boldsymbol{u}) + d\max_t(y_t - y_t^{Ref})^2\log(T) \\ &+ \sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u})\sum_{t=1}^T \ell_t(u_t)}\Bigg), \end{aligned}$$

Proof. By Lemma B.1 (with $\beta = \gamma$), for any $\gamma \geq \gamma_{\min} = \frac{2d}{2d+1}$, we have

$$R_T(\boldsymbol{u}) \leq \gamma\lambda \|u_1\|_2^2 + 4d \max_t (y_t - y_t^{\text{Ref}})^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{\lambda d}\right) + 2\frac{\gamma}{1-\gamma} P_T^\gamma(\boldsymbol{u}) + \frac{1-\gamma}{\gamma} 2d \sum_{t=1}^T \ell_t(u_t).$$

Now by Lemma A.5, there is a $\gamma^\circ \in [0,1]$ satisfying $\gamma^\circ = \frac{\sqrt{d\sum_{t=1}^T \ell_t(u_t)}}{\sqrt{d\sum_{t=1}^T \ell_t(u_t)} + \sqrt{P_T^{\gamma^\circ}(\boldsymbol{u})}}$. If $\gamma^\circ \geq \gamma_{\min}$, then for $\gamma = \gamma^\circ \vee \gamma_{\min}$, the terms in the second line reduce to

$$2\frac{\gamma^\circ}{1-\gamma^\circ} P_T^{\gamma^\circ}(\boldsymbol{u}) + \frac{1-\gamma^\circ}{\gamma^\circ} 2d \sum_{t=1}^T \ell_t(u_t) = 4\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)},$$

and otherwise for $\gamma^\circ \leq \gamma_{\min}$ we have

$$\begin{aligned} 2\frac{\gamma_{\min}}{1-\gamma_{\min}} P_T^{\gamma_{\min}}(\boldsymbol{u}) + \frac{1-\gamma_{\min}}{\gamma_{\min}} 2d \sum_{t=1}^T \ell_t(u_t) &\leq 2\frac{\gamma_{\min}}{1-\gamma_{\min}} P_T^{\gamma_{\min}}(\boldsymbol{u}) + \frac{1-\gamma^\circ}{\gamma^\circ} 2d \sum_{t=1}^T \ell_t(u_t) \\ &\leq 4dP_T^{\gamma_{\min}}(\boldsymbol{u}) + 2\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)}, \end{aligned}$$

so combining these two bounds and plugging back into the regret bound above, we have

$$\begin{aligned} R_T(\boldsymbol{u}) &\leq \gamma\lambda \|u_1\|_2^2 + 4d \max_t (y_t - y_t^{\text{Ref}})^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{\lambda d}\right) \\ &\quad + 4dP_T^{\gamma_{\min}}(\boldsymbol{u}) + 4\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)} \\ &\leq O\left(dP_T^{\gamma_{\min}}(\boldsymbol{u}) + d \max_t (y_t - y_t^{\text{Ref}})^2 \log(T) + \sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)}.\right) \end{aligned}$$

□

B.1.1. Proof of Lemma B.1

Lemma B.1. *Let $y_t^{Ref} \in \mathbb{R}$ be an arbitrary reference point, available at the start of round t, and let $\mathcal{B}_t = \{y \in \mathbb{R} : y_t^{Ref} - M_t \leq y \leq y_t^{Ref} + M_t\}$ for $M_t = \max_{s<t} |y_s - y_s^{Ref}|$. Suppose that we apply Algorithm 1 with hints $\widetilde{y}_t = \overline{y}_t :=$ $\text{Clip}_{\mathcal{B}_t}(\langle x_t, w_t\rangle)$. Then for any sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ in $\mathbb{R}^d$ and any $\gamma, \beta \in [0,1]$ such that $\beta \geq \gamma \geq \gamma_{\min} = \frac{2d}{2d+1}$,*

$$R_T(\boldsymbol{u}) \leq \gamma\lambda \|u_1\|_2^2 + 4d \max_t (y_t - y_t^{Ref})^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{\lambda d}\right) + 2\frac{\beta}{1-\beta} P_T^\beta(\boldsymbol{u}) + \frac{1-\gamma}{\gamma} 2d \sum_{t=1}^T \ell_t(u_t)$$

Proof. Applying Theorem 3.1 followed by Lemma 3.2, for any $\gamma \in (0,1]$ and $\beta \geq \gamma$ we have

$$\begin{aligned}
R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) &\leq \frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{d}{2}\max_t(y_t - \overline{y}_t)^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right) \\
&\quad + \gamma\sum_{t=1}^{T-1}\left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right] + \frac{d}{2}\log(1/\gamma)\sum_{t=1}^T (y_t - \overline{y}_t)^2 \\
&\leq \frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{d}{2}\max_t(y_t - \overline{y}_t)^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right) \\
&\quad + \frac{\beta}{1-\beta}P_T^\beta(\boldsymbol{u}) + \frac{1-\gamma}{\gamma}\frac{d}{2}\sum_{t=1}^T (y_t - \overline{y}_t)^2,
\end{aligned}$$

Using Lemma D.1 we have

$$\sum_{t=1}^T (y_t - \overline{y}_t)^2 \leq \sum_{t=1}^T \left[M_{t+1}^2 - M_t^2 + 2\ell_t(w_t)\right] \leq M_{T+1}^2 + 2\sum_{t=1}^T \ell_t(w_t),$$

so for any $\gamma \geq \frac{2d}{2d+1}$, we have

$$\begin{aligned}
\frac{1-\gamma}{\gamma}\frac{d}{2}\sum_{t=1}^T (y_t - \overline{y}_t)^2 &\leq \frac{1-\gamma}{\gamma} d\left[\frac{1}{2}M_{T+1}^2 + \sum_{t=1}^T \ell_t(w_t)\right] \\
&= \frac{1-\gamma}{\gamma} d\left[\frac{1}{2}M_{T+1}^2 + \sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) + \sum_{t=1}^T \ell_t(u_t)\right] \\
&\leq \frac{1}{4}M_{T+1}^2 + \frac{1}{2}\sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) + \frac{1-\gamma}{\gamma} d\sum_{t=1}^T \ell_t(u_t),
\end{aligned}$$

where the final inequality uses $\gamma \geq \frac{2d}{2d+1} \implies \frac{1-\gamma}{\gamma} \leq \frac{1}{2d}$ and bounds $\frac{1-\gamma}{\gamma} d\sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) \leq \frac{1}{2}\sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t)$ (assuming $\sum_{t=1}^T \ell_t(w_t) - \ell_t(u_t) \geq 0$, which can be assumed without loss of generality since otherwise the stated bound trivially holds). Plugging this back into the regret bound and re-arranging terms, we have

$$\begin{aligned}
R_T(\boldsymbol{u}) &\leq \frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{d}{2}\max_t(y_t - \overline{y}_t)^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right) \\
&\quad + \frac{\gamma}{1-\gamma}P_T^\gamma(\boldsymbol{u}) + \frac{1}{2}R_T(\boldsymbol{u}) + \frac{1-\gamma}{\gamma}\sum_{t=1}^T \ell_t(u_t) \\
\implies R_T(\boldsymbol{u}) &\leq \gamma\lambda\|u_1\|_2^2 + 4d\max_t(y_t - y_t^{\text{Ref}})^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right) \\
&\quad + 2\frac{\beta}{1-\beta}P_T^\beta(\boldsymbol{u}) + \frac{1-\gamma}{\gamma}2d\sum_{t=1}^T \ell_t(u_t),
\end{aligned}$$

where we've bounded $\max_t(y_t - \overline{y}_t)^2 \leq 4M_{T+1} = 4\max_t(y_t - y_t^{\text{Ref}})^2$ using Lemma D.1. □

C. Proofs for Section 3.2 (Dimension-dependent Lower Bound)

C.1. Proof of Theorem 3.5

Theorem 3.5. *For any $d, T \geq 1$ and $P, Y > 0$ such that $dP \leq 2TY^2$, there is a sequence of losses $\ell_t(w) = \frac{1}{2}(y_t - \langle x_t, w\rangle)^2$ and a comparator sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ satisfying $\max_t |y_t| \leq Y$ and $\sum_{t=1}^{T-1}\max_s\left[\ell_s(u_{t+1}) - \ell_s(u_t)\right]_+ \leq P$ such that*

$$R_T(\boldsymbol{u}) \geq \Omega\left(dY^2\log(T) + dP + \sqrt{dP\sum_{t=2}^T (y_t - y_{t-1})^2}\right).$$

Proof. First notice that the trivial comparator sequence with $u_1 = \ldots = u_T$ always satisfies $\sum_{t=2}^{T} \max_s [\ell_s(u_{t+1}) - \ell_s(u_t)]_+ = 0 \leq P$, so we can always lower-bound the dynamic regret using the well-known lower bound for the static regret in this setting (see, *e.g.*, Vovk (2001); Gaillard et al. (2019); Mayo et al. (2022)). In particular, for any $u \in W$ we have

$$\sup_{y_1,\ldots,y_T} R_T(u) \geq \Omega\left(dY^2 \log(T)\right) \tag{9}$$

Next, let $\sigma \in [0,1]$ and let $\sigma_1, \ldots, \sigma_t$ be a sequence of iid random variables drawn uniformly from $\{-\sigma, \sigma\}$, and let $y_t = Y\sigma_t$. Choose feature vectors x_t which cycle through the standard basis vectors (*e.g.* define $\iota(t) = t \pmod d + 1$ and let $x_t = e_{\iota(t)}$). Now observe that the comparator sequence can always exactly fit a sequence $y_1, \ldots, y_T$ by setting u_t to satisfy $\langle x_t, u_t \rangle = u_{t,\iota(t)} = y_t$. In particular, by letting $\widetilde{u}_1 = (y_1, \ldots, y_d)$, $\widetilde{u}_2 = (y_{d+1}, \ldots, y_{2d}), \ldots, \widetilde{u}_{\lceil T/d \rceil} = (y_{\lceil T/d \rceil+1}, \ldots, y_T, 0, 0, \ldots)$ we can set $u_t = \widetilde{u}_{\lceil t/d \rceil}$ to guarantee $\langle x_t, u_t \rangle = y_t$ on all rounds, while only changing the comparator $\lceil T/d \rceil$ times at most. From this, we have the following initial bound on the regret:

$$\begin{aligned}
\sup_{y_1,\ldots,y_T} R_T(\boldsymbol{u}) &\geq \mathbb{E}_{\boldsymbol{y}}\left[\sum_{t=1}^{T} \ell_t(w_t) - \ell_t(u_t)\right] \\
&\geq \mathbb{E}_{\boldsymbol{y}}\left[\frac{1}{2}y_t^2 + \frac{1}{2}\langle x_t, w_t \rangle^2 + y_t \langle x_t, w_t \rangle\right] \\
&\geq \frac{1}{2}\sigma^2 Y^2 T,
\end{aligned} \tag{10}$$

where the last line uses $y_t^2 = Y^2\sigma^2$ and $\mathbb{E}[y_t] = 0$. Moreover, since the comparator changes only every d rounds, the variability is bounded as

$$\sum_{t=1}^{T-1} \max_s [\ell_s(u_{t+1}) - \ell_s(u_t)]_+ \leq \sum_{i=1}^{\lceil T/d \rceil - 1} \max_s [\ell_s(\widetilde{u}_{i+1}) - \ell_s(\widetilde{u}_i)]_+ .$$

Observe that $\ell_s(\widetilde{u}_{i+1}) - \ell_s(\widetilde{u}_i)$ can only be positive when $\langle x_s, \widetilde{u}_i \rangle = y_s$ and $\langle x_s, \widetilde{u}_{i+1} \rangle = -y_s$, hence

$$\begin{aligned}
\sum_{t=1}^{T-1} \max_s [\ell_s(u_{t+1}) - \ell_s(u_t)]_+ &\leq \sum_{i=1}^{\lceil T/d \rceil - 1} \max_s [\ell_s(\widetilde{u}_{i+1}) - \ell_s(\widetilde{u}_i)]_+ \\
&\leq \sum_{i=1}^{\lceil T/d \rceil - 1} \frac{1}{2}(y_s - (-y_s))^2 \\
&\leq \frac{2TY^2}{d}\sigma^2.
\end{aligned}$$

Hence, setting $\sigma = \sqrt{\frac{dP}{2TY^2}} \leq 1$ ensures $\sum_{t=1}^{T-1} \max_s [\ell_s(u_{t+1}) - \ell_s(u_t)]_+ \leq \frac{2TY^2}{d}\sigma^2 \leq P$, and the regret is bounded below by

$$\sup_{y_1,\ldots,y_T} R_T(\boldsymbol{u}) \geq \frac{1}{2}\sigma^2 Y^2 T = \frac{1}{4}dP,$$

which we can further lower bound as:

$$\begin{aligned}
&= \frac{1}{4}\sqrt{dP \cdot dP} \geq \frac{1}{4}\sqrt{dP \cdot d \sum_{t=1}^{T-1} \max_s [\ell_s(u_{t+1}) - \ell_s(u_t)]_+} \\
&\geq \frac{1}{4}\sqrt{dP \sum_{t=1}^{T-1} [\ell_t(u_{t+1}) - \ell_t(u_t)]_+} = \Omega\left(\sqrt{dP \sum_{t=2}^{T} \frac{1}{2}(y_t - y_{t-1})^2}\right).
\end{aligned} \tag{11}$$

Taken together with Equation (9), we have

$$R_T(\boldsymbol{u}) \geq \Omega\left(dY^2 \log(T) \vee \sqrt{dP\mathcal{V}_T}\right)$$

where $\mathcal{V}_T = dP \vee \sum_{t=2}^{T} \frac{1}{2}(y_t - y_{t-1})^2$.

□

D. Proofs for Section 4 (Learning the Optimal Discount Factor)

D.1. Proof of Lemma D.1

The following lemma shows that by clipping our predictions to some crude "trust-region", the loss of the clipped prediction is at worst portional to the maximal deviation of the true y_t from the trust region. Intuitively, we can think of y^{Ref} as being some data-dependent but already-observed quantity, such as y_{t-1}.

Lemma D.1. *Define* $M_t = \max_{s<t} \left|y_s - y_s^{Ref}\right|$, $\mathcal{B}_t = \left\{x \in \mathbb{R} : y_t^{Ref} - M_t \le x \le y_t^{Ref} + M_t\right\}$, *and let* $\overline{y}_t = \text{Clip}_{\mathcal{B}_t}(\langle x_t, w_t\rangle)$ *for some* $w_t \in \mathbb{R}^d$. *Then for any* t *we have*

$$(y_t - \overline{y}_t)^2 \le \min\left\{4M_{t+1}^2, 2\ell_t(w_t) + M_{t+1}^2 - M_t^2\right\}$$

Proof. First, observe that we always have

$$(y_t - \overline{y}_t)^2 = \left(y_t - y_t^{\text{Ref}} + y_t^{\text{Ref}} - \overline{y}_t\right)^2 \le 2\left(y_t - y_t^{\text{Ref}}\right)^2 + 2\left(y_t^{\text{Ref}} - \overline{y}_t\right)^2 \le 2M_{t+1}^2 + 2M_t^2 \le 4M_{t+1}^2.$$

Next, observe that if $\langle x_t, w_t\rangle = \overline{y}_t$, then we trivially have $(y_t - \overline{y}_t)^2 = (y_t - \langle x_t, w_t\rangle)^2 = 2\ell_t(w_t)$. Otherwise, when $\langle x_t, w_t\rangle \ne \overline{y}_t$, we have clipped $\overline{y}_t$ to be a distance of M_t away from y_t^{Ref} and there are two cases to consider. If $\text{Sign}\left(\overline{y}_t - y_t^{\text{Ref}}\right) \ne \text{Sign}\left(y_t - y_t^{\text{Ref}}\right)$, then the clipping operation $\overline{y}_t = \text{Clip}_{\mathcal{B}_t}(\langle x_t, w_t\rangle)$ moves us closer to y_t, hence $|y_t - \overline{y}_t| \le |y_t - \langle x_t, w_t\rangle|$. If $\text{Sign}\left(\overline{y}_t - y_t^{\text{Ref}}\right) = \text{Sign}\left(y_t - y_t^{\text{Ref}}\right)$, then we precisely have $|y_t - \overline{y}_t| = M_{t+1} - M_t$ when $y_t \notin \mathcal{B}_t$ and $|y_t - \overline{y}_t| \le |y_t - \langle x_t, w_t\rangle|$ when $y_t \in \mathcal{B}_t$. Hence, combining these cases we have

$$(y_t - \overline{y}_t)^2 \le (y_t - \langle w_t, x_t\rangle)^2 + (M_{t+1} - M_t)^2 \le 2\ell_t(w_t) + M_{t+1}^2 - M_t^2,$$

where we have used $(u - l)^2 \le u^2 - l^2$ for $u \ge l \ge 0$. Hence, combining with the first display we have

$$(y_t - \overline{y}_t)^2 \le \min\left\{4M_{t+1}^2, M_{t+1}^2 - M_t^2 + 2\ell_t(w_t)\right\}.$$

□

D.2. Proof of Lemma D.2

The following lemma shows the following important property of the meta-learner's losses: they are α_t-exp-concave with $\alpha_t = \frac{1}{2\max_i \ell_t(y_t^{(i)})}$ in the domain $\widehat{\mathcal{Y}}_t = \left\{y = \sum_{i=1}^N p_i y_t^{(i)} : \sum_{i=1}^N p_i = 1\right\}$.

Lemma D.2. *Let* $y^{(1)}, \dots, y^{(N)}$ *be arbitrary real numbers and let* $\widehat{\mathcal{Y}}_t = \left\{\overline{y} = \sum_{i=1}^N p_i y^{(i)} : p \in \mathbb{R}_{\ge 0}^N, \sum_{i=1}^N p_i = 1\right\}$. *Then* $\ell_t(\overline{y}) = \frac{1}{2}(y_t - \overline{y})^2$ *is* α_t*-Exp-Concave on* $\widehat{\mathcal{Y}}_t$ *for* $\alpha_t \le \frac{1}{2\max_i \ell_t(y^{(i)})}$.

Proof. Letting $f_t(\overline{y}) = \exp(-\alpha_t \ell_t(\overline{y}))$ we have for any $\overline{y} \in \widehat{\mathcal{Y}}_t$:

$$\begin{aligned}
f_t'(\overline{y}) &= \left[\exp\left(-\frac{\alpha_t}{2}(y_t - \overline{y})^2\right)\right]' = \exp\left(-\frac{\alpha_t}{2}(y_t - \overline{y})^2\right)\alpha_t(y_t - \overline{y}) \\
f_t''(\overline{y}) &= \exp\left(-\frac{\alpha_t}{2}(y_t - \overline{y})^2\right)\left[\alpha_t^2(y_t - \overline{y})^2 - \alpha_t\right] \\
&= \exp\left(-\frac{\alpha_t}{2}(y_t - \overline{y})^2\right)\left[2\alpha_t^2\ell_t(\overline{y}) - \alpha_t\right]
\end{aligned}$$

Hence for $\alpha_t \le \frac{1}{2\max_i \ell_t(y^{(i)})}$ we have

$$f_t''(\overline{y}) \le \exp\left(-\frac{\alpha_t}{2}(y_t - \overline{y})^2\right)\alpha_t\left[2\alpha_t\ell_t(\overline{y}) - 1\right] \le 0$$

so $f_t(\overline{y}) = \exp(-\alpha_t\ell_t(\overline{y}))$ is concave and ℓ_t is α_t-Exp-Concave over $\widehat{\mathcal{Y}}$ for $\alpha_t \le \frac{1}{2\max_i \ell_t(y^{(i)})}$. □

D.3. Regret of the Range-Clipped Meta-Algorithm

In this section we prove a simple result showing that the range-clipping reduction described by Algorithm 2 incurs only an constant additional penalty. This lemma will be used to do most of the heavy-lifting in proving Theorem 4.1, which simply applies the following lemma and then chooses a specific meta-algorithm for $\mathcal{A}_{\text{Meta}}$.

Lemma D.3. *For any $[a,b] \subseteq [1,T]$, sequence $\boldsymbol{u} = (u_a, \dots, u_b)$ in $\mathbb{R}$, and $j \in [N]$, Algorithm 2 guarantees*

$$R_{[a,b]}(\boldsymbol{u}) \leq \frac{1}{2}\max_t (y_t - y_t^{Ref})^2 + R_{[a,b]}^{\mathcal{A}_j}(\boldsymbol{u}) + R_{[a,b]}^{Meta}(e_j),$$

where $R_{[a,b]}^{\mathcal{A}_j}(\boldsymbol{u}) = \sum_{t=a}^b \ell_t(w_t^{(j)}) - \ell_t(u_t)$ is the dynamic regret $\mathcal{A}_j$ and $R_{[a,b]}^{Meta}(e_j) = \sum_{t=a}^b \ell_t(\overline{y}_t) - \ell_t(\overline{y}_t^{(j)})$.

Proof. For ease of notation we let $y_t^{(i)} = \left\langle x_t, w_t^{(i)} \right\rangle$, where $w_t^{(i)}$ is the output of algorithm $\mathcal{A}_i$, and slightly abuse notation by writing $\ell_t(y) = \frac{1}{2}(y_t - y)^2$ for $y \in \mathbb{R}$. Hence, we may write $\ell_t(w_t) \equiv \ell_t(y_t^{(i)})$ interchangeably. Note that this equivalence is valid in the improper online learning setting since the features are observed *before* the learner makes a prediction, as discussed in the introduction.

Now, for for any $j \in [N]$ we have

$$\begin{aligned}
R_{[a,b]}(\boldsymbol{u}) &= \sum_{t=a}^b \ell_t(\overline{y}_t) - \ell_t(u_t) \\
&= \sum_{t=a}^b \ell_t(w_t^{(j)}) - \ell_t(u_t) + \sum_{t=a}^b \ell_t(\overline{y}_t) - \ell_t(w_t^{(j)}) \\
&= R_{[a,b]}^{\mathcal{A}_j}(\boldsymbol{u}) + \sum_{t=a}^b \ell_t(\overline{y}_t) - \ell_t\left(y_t^{(j)}\right),
\end{aligned}$$

where we have observed $y_t^{(j)} = \left\langle x_t, w_t^{(j)} \right\rangle$. Observe that by Lemma D.1 we have

$$\begin{aligned}
\ell_t(y_t^{(j)}) &\geq \frac{1}{2}M_t^2 - \frac{1}{2}M_{t+1}^2 + \frac{1}{2}(y_t - \overline{y}_t^{(j)})^2 \\
&= \frac{1}{2}M_t^2 - \frac{1}{2}M_{t+1}^2 + \ell_t\left(\overline{y}_t^{(j)}\right),
\end{aligned}$$

where $M_t = \max_{s<t} \left| y_s - y_s^{\text{Ref}} \right|$. Hence,

$$\begin{aligned}
R_{[a,b]}(\boldsymbol{u}) &\leq R_{[a,b]}^{\mathcal{A}_j}(\boldsymbol{u}) + \sum_{t=a}^b \ell_t(\overline{y}_t) - \ell_t\left(y_t^{(j)}\right) \\
&\leq R_{[a,b]}^{\mathcal{A}_j}(\boldsymbol{u}) + \sum_{t=a}^b \ell_t(\overline{y}_t) - \ell_t\left(\overline{y}_t^{(j)}\right) \\
&\quad + \sum_{t=a}^b \frac{1}{2}M_{t+1}^2 - \frac{1}{2}M_t^2 \\
&\leq \frac{1}{2}M_{b+1}^2 + R_{[a,b]}^{\mathcal{A}_j}(\boldsymbol{u}) + \underbrace{\sum_{t=a}^b \ell_t(\overline{y}_t) - \ell_t\left(\overline{y}_t^{(j)}\right)}_{=:R_{[a,b]}^{\text{Meta}}(e_j)}
\end{aligned}$$

□

D.4. Proof of Theorem 4.1

Theorem 4.1. Let $\mathcal{A}_{\text{Meta}}$ be an instance of Algorithm 3 with $\alpha_t = \frac{1}{2\max_{t,i}\ell_t(\overline{y}_t^{(i)})}$, $\beta_{t+1} = \frac{1}{(e+t)\log^2(e+t)+1}$ and $p_1 = \mathbf{1}_N/N$. Then for any sequence $\boldsymbol{u} = (u_1,\dots,u_T)$ in $\mathbb{R}$ and any $j \in [N]$, Algorithm 2 guarantees

$$R_{[a,b]}(\boldsymbol{u}) \le O\left(R^{\mathcal{A}_j}_{[a,b]}(\boldsymbol{u}) + \max_t (y_t - \widetilde{y}_t)^2 \log\left(Nb\log^2(b)\right)\right),$$

where $R_{[a,b]}$ denotes regret over the sub-interval $[a,b]$.

Proof. The proof follows almost immediately using the regret guarantee of the range-clipped meta-algorithm (Lemma D.3), from which we have

$$R_{[a,b]}(\boldsymbol{u}) \le \frac{1}{2}\max_t (y_t - \widetilde{y}_t)^2 + R^{\mathcal{A}_j}_{[a,b]}(\boldsymbol{u}) + R^{\text{Meta}}_{[a,b]}(e_j).$$

Now applying the guarantee of an appropriate instance of the fixed-share algorithm (Theorem E.1 with $\alpha_t = \frac{1}{2\max_{t,i}\ell_t(\overline{y}_t^{(i)})}$, $\beta_t = \frac{1}{(e+t)\log^2(e+t)+1}$, and $p_1 = \mathbf{1}_N/N$), we have

$$\begin{aligned}
R^{\text{Meta}}_{[a,b]}(e_j) &\le \frac{1}{\alpha_{b+1}}\left[2\log\left(\frac{1}{\beta_{b+1}p_{1j}}\right) + 1\right] \\
&\le \max_{t,i}\ell_t(\overline{y}_t^{(i)})\left[2\log\left(((e+b)\log^2(e+b)+1)N\right)+1\right] \\
&\le O\left(\max_t (y_t - \widetilde{y}_t)^2 \log\left(b\log^2(b)N\right)\right),
\end{aligned}$$

where the last line applies Lemma D.1 and hides constants. All together, we have

$$R_{[a,b]}(\boldsymbol{u}) \le O\left(R^{\mathcal{A}_j}_{[a,b]}(\boldsymbol{u}) + \max_t (y_t - \widetilde{y}_t)^2 \log\left(Nb\log^2(b)\right)\right).$$

□

D.5. Proof of Theorem 4.2

The proof of Theorem 4.2 follows by applying Theorem 4.1, and then showing that there exists a $\mathcal{A}_\gamma$ which attains the desired bound. We first provide proof of the latter claim in Lemma D.4 for the sake of modularity. In particular, we will also re-use this result to argue strongly-adaptive guarantees in Section 5. Proof of Theorem 4.2 is then easily proven at the end of this section.

Lemma D.4. *Let $b > 1$, $\eta_{\min} = 2d$, $\eta_{\max} = dT$, and define $\mathcal{S}_\eta = \left\{\eta_i = \eta_{\min}b^i \wedge \eta_{\max} : i = 0,1,\dots\right\}$ and $\mathcal{S}_\gamma = \left\{\gamma_i = \frac{\eta_i}{1+\eta_i} : i = 0,1,\dots\right\} \cup \{0\}$. For any γ in $\mathcal{S}_\gamma$, let $\mathcal{A}_\gamma$ denote an instance of Algorithm 1 with discount γ. Then for any $\boldsymbol{u} = (u_1,\dots,u_T)$ in $\mathbb{R}^d$, there is a $\gamma^* \in [0,1]$ satisfying $\gamma^* = \frac{\sqrt{d\sum_{t=1}^T \frac{1}{2}(y_t-\widetilde{y}_t)^2}}{\sqrt{d\sum_{t=1}^T \frac{1}{2}(y_t-\widetilde{y}_t)^2} + \sqrt{P_T^{\gamma^*}(\boldsymbol{u})}}$ and a $\gamma \in \mathcal{S}_\gamma$ such that*

$$R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) \le O\left(d\max_t (y_t - \widetilde{y}_t)^2 \log(T) + b\sqrt{dP_T^{\gamma^*}(\boldsymbol{u})\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}\right).$$

Proof. Denote $V_T = \frac{d}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2$. By Lemma A.5, there exists a $\gamma^* \in [0,1]$ such that

$$\gamma^* = \frac{\sqrt{V_T}}{\sqrt{V_T} + \sqrt{P_T^{\gamma^*}(\boldsymbol{u})}}.$$

Throughout the proof it will be convenient to work in terms of the related quantity $\eta^* = \frac{\gamma^*}{1-\gamma^*} = \sqrt{\frac{V_T}{P_T^{\gamma}(\boldsymbol{u})}}$. Let us first suppose that $0 \le \eta^* \le \eta_{\min}$. In this case, we have

$$\eta^* = \sqrt{\frac{V_T}{P_T^{\gamma^*}(\boldsymbol{u})}} \le \eta_{\min} \implies \sqrt{\frac{1}{2}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2} \le \eta_{\min}\sqrt{\frac{1}{d}P_T^{\gamma}(\boldsymbol{u})}.$$

Consider the algorithm $\mathcal{A}_0$ with $\gamma = 0$: in this case we have $w_t = \arg\min_{w \in \mathbb{R}^d} h_t(w)$, so $\langle x_t, w_t \rangle = \widetilde{y}_t$ and the regret is trivially

$$\begin{aligned}\sum_{t=1}^T \ell_t(w_t^{\mathcal{A}_0}) - \ell_t(u_t) &\leq \sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2 \\ &= \sqrt{\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2 \sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2} \\ &\leq \frac{\eta_{\min}}{\sqrt{d}} \sqrt{P_T^{\gamma^*}(\boldsymbol{u}) \sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2} \\ &= 2\sqrt{V_T P_T^{\gamma^*}(\boldsymbol{u})} \end{aligned} \tag{12}$$

for $\eta_{\min} = 2d$.

Otherwise, for $\eta^* \geq \eta_{\min}$, using Theorem 3.1 we have that for any $\gamma \in \mathcal{S}_\gamma$,

$$\begin{aligned} R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) &\leq \frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{d}{2}\max_t (y_t - \widetilde{y}_t)^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right) \\ &\quad + \gamma \sum_{t=1}^{T-1}\left[F_t^\gamma(u_{t+1}) - F_t^\gamma(u_t)\right] + \log(1/\gamma) V_T \\ &\overset{(*)}{\leq} \frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{d}{2}\max_t (y_t - \widetilde{y}_t)^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right) \\ &\quad + \eta^* P_T^{\gamma^*}(\boldsymbol{u}) + \frac{V_T}{\eta} \end{aligned}$$

where $(*)$ observes that $\eta_{\min} = \frac{\gamma_{\min}}{1-\gamma_{\min}} \leq \eta^* = \frac{\gamma^*}{1-\gamma^*} \implies \gamma_{\min} \leq \gamma^*$ and applies Lemma 3.2 (with $\beta = \gamma^*$) and substitutes $\eta = \frac{\gamma}{1-\gamma}$. If $\eta^* \geq \eta_{\max}$ then choosing $\eta = \eta_{\max} = dT$ yields

$$\frac{V_T}{\eta} = \frac{d}{2dT}\sum_{t=1}^T (y_t - \widetilde{y}_t)^2 \leq \frac{1}{2}\max_t (y_t - \widetilde{y}_t)^2,$$

and otherwise, there is an η_k in $\mathcal{S}_\eta$ such that $\eta_k \leq \eta^* \leq b\eta_k$, so choosing $\eta = \eta_k$ yields

$$\frac{V_T}{\eta_k} \leq b\frac{V_T}{\eta^*} = b\sqrt{P_T^{\gamma^*}(\boldsymbol{u}) V_T}$$

Hence, overall we have that there is a $\gamma \in \mathcal{S}_\gamma$ such that

$$\begin{aligned} R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) &\leq \frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{1}{2}\max_t (y_t - \widetilde{y}_t)^2 \left[d\log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right) \vee 1\right] + \eta^* P_T^{\gamma^*}(\boldsymbol{u}) + b\frac{V_T}{\eta^*} \\ &= \frac{\gamma\lambda}{2}\|u_1\|_2^2 + \frac{1}{2}\max_t (y_t - \widetilde{y}_t)^2 \left[d\log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right) \vee 1\right] + (b+1)\sqrt{V_T P_T^{\gamma^*}(\boldsymbol{u})} \\ &\leq O\left(d\max_t (y_t - \widetilde{y}_t)^2 \log(T) \vee b\sqrt{dP_T^{\gamma^*}(\boldsymbol{u})\sum_{t=1}^T (y_t - \widetilde{y}_t)^2}\right). \end{aligned}$$

□

With the previous lemma in hand, the proof of Theorem 4.2 follows easily. The theorem is re-stated for convenience.

Theorem 4.2. *Let $b > 1$, $\eta_{\min} = 2d$, $\eta_{\max} = dT$, and for all $i \in \mathbb{N}$ let $\eta_i = \eta_{\min} b^i \wedge \eta_{\max}$, and construct the set of discount factors $\mathcal{S}_\gamma = \left\{\gamma_i = \frac{\eta_i}{1+\eta_i} : i \in \mathbb{N}\right\} \cup \{0\}$. For any γ in $\mathcal{S}_\gamma$, let $\mathcal{A}_\gamma$ denote an instance of Algorithm 1 with discount γ.*[5] *Let $\mathcal{A}_{Meta}$ be an instance of the algorithm characterized in Theorem 4.1, and suppose we set $y_t^{Ref} = \widetilde{y}_t$ for all t. Then for any $\boldsymbol{u} = (u_1, \ldots, u_T)$ in $\mathbb{R}^d$, Algorithm 2 guarantees*

$$R_T(\boldsymbol{u}) \le O\Bigg(d \max_t (y_t - y_t^{Ref})^2 \log(T) + b\sqrt{dP_T^{\gamma^*}(\boldsymbol{u}) \sum_{t=1}^T (y_t - \widetilde{y}_t)^2}\Bigg)$$

where $\gamma^* \in [0, 1]$ satisfies Equation (2).

Proof. Applying Theorem 4.1, for any sequence $\boldsymbol{u} = (u_1, \ldots, u_T)$ in $\mathbb{R}^d$ and any $\gamma \in \mathcal{S}_\gamma$ we have

$$\begin{aligned} R_T(\boldsymbol{u}) &\le \widehat{O}\left(R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) + \max_t (y_t - y_t^{\text{Ref}})^2 \log(NT)\right) \\ &\le \widehat{O}\left(R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) + \max_t (y_t - y_t^{\text{Ref}})^2 \log(T)\right), \end{aligned} \tag{13}$$

where the last line uses $N = |\mathcal{S}_\gamma| = \log_b(\eta_{\max}/\eta_{\min}) \le O(\log_b(T))$, then hides $\log(\log)$ factors. Finally, by Lemma D.4, there is indeed a $\gamma^* \in [0, 1]$ satisfying $\gamma^* = \frac{\sqrt{d\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2}}{\sqrt{d\sum_{t=1}^T \frac{1}{2}(y_t - \widetilde{y}_t)^2} + \sqrt{P_T^{\gamma^*}(\boldsymbol{u})}}$ and a $\gamma \in \mathcal{S}_\gamma$ such that

$$R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) \le O\left(d \max_t (y_t - \widetilde{y}_t)^2 \log(T) + b\sqrt{dP_T^{\gamma^*}(\boldsymbol{u}) \sum_{t=1}^T (y_t - \widetilde{y}_t)^2}\right).$$

Plugging this back into Equation (13) and choosing $y_t^{\text{Ref}} = \widetilde{y}_t$ proves the result. □

D.6. Proof of Theorem 4.3

As in Appendix D.5, the proof of Theorem 4.3 follows by applying Theorem 4.1 and then showing that there is a $\mathcal{A}_\gamma$ attaining the desired regret bound. We first provide proof of the latter claim in Lemma D.5 for the sake of modularity, so that we can use it when arguing strongly-adaptive guarantees in Section 5. Proof of Theorem 4.3 is proven at the end of this section.

Lemma D.5. *Under the same conditions as Lemma D.4, suppose each $\mathcal{A}_\gamma$ sets hints $\widetilde{y}_t = \overline{y}_t^\gamma = \text{Clip}_{\mathcal{B}_t}(\langle x_t, w_t^\gamma \rangle)$, where $\mathcal{B}_t = [y_t^{Ref} - M_t, y_t^{Ref} + M_t]$ and $M_t = \max_{s<t} \left|y_s - y_s^{Ref}\right|$. Then for any $\boldsymbol{u} = (u_1, \ldots, u_T)$ in W, there is a $\gamma^\circ \in [0, 1]$ satisfying $\gamma^\circ = \frac{\sqrt{d\sum_{t=1}^T \ell_t(u_t)}}{\sqrt{d\sum_{t=1}^T \ell_t(u_t)} + \sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u})}}$ and a $\gamma \in \mathcal{S}_\gamma$ such that*

$$R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) \le O\Bigg(dP_T^{\gamma_{\min}}(\boldsymbol{u}) + d \max_t \left(y_t - y_t^{Ref}\right)^2 \log(T) + b\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)}\Bigg),$$

where $\gamma^{\min} = \min\{\gamma \in \mathcal{S}_\gamma\} = \frac{2d}{2d+1}$.

[5]For brevity, here we refer to an algorithm that directly predicts $\widetilde{y}_t$ on every round as being an instance of the discounted VAW forecaster with $\gamma = 0$. This terminology can be justified by Remark A.2, but for our purposes here it's sufficient to consider it convenient alias.

Proof. Using Lemma B.1, for any $\boldsymbol{u} = (u_1, \dots, u_T)$, $\gamma \in (0,1)$, and $\beta \geq \gamma \geq \gamma_{\min} = \frac{2d}{2d+1}$, we have

$$R_T(\boldsymbol{u}) \leq \gamma\lambda \|u_1\|_2^2 + 4d \max_t (y_t - y_t^{\text{Ref}})^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{\lambda d}\right)$$

$$+ 2\frac{\beta}{1-\beta} P_T^\beta(\boldsymbol{u}) + \frac{1-\gamma}{\gamma} 2d \sum_{t=1}^T \ell_t(u_t),$$

We will proceed by showing that there is a β and γ that suitably balances the summations in the last line. To this end, recall that by Lemma A.5, there is a γ° satisfying

$$\gamma^\circ = \frac{\sqrt{d\sum_{t=1}^T \ell_t(u_t)}}{\sqrt{d\sum_{t=1}^T \ell_t(u_t)} + \sqrt{P_T^{\gamma^\circ}(\boldsymbol{u})}}$$

Denote $\eta = \frac{\gamma}{1-\gamma}$ and $\eta^\circ = \frac{\gamma^\circ}{1-\gamma^\circ} = \sqrt{\frac{d\sum_{t=1}^T \ell_t(u_t)}{P_T^{\gamma^\circ}(\boldsymbol{u})}}$. If $\eta^\circ \geq \eta_{\max} = \frac{\gamma_{\max}}{1-\gamma_{\max}}$, then we can take $\beta = \gamma^\circ$ and $\gamma = \gamma_{\max}$ to get

$$\begin{aligned}
\frac{\beta}{1-\beta} P_T^\beta(\boldsymbol{u}) + \frac{\gamma}{1-\gamma} d \sum_{t=1}^T \ell_t(u_t) &= \eta^\circ P_T^{\gamma^\circ}(\boldsymbol{u}) + \frac{d\sum_{t=1}^T \ell_t(u_t)}{\eta_{\max}} \\
&= \sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)} + \frac{d\sum_{t=1}^T \ell_t(u_t)}{\eta_{\max}} \\
&\leq \sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)} + \max_t \ell_t(u_t),
\end{aligned}$$

where the last line recalls $\eta_{\max} = dT$. Otherwise, if $\eta^\circ \leq \eta_{\min} = \frac{\gamma_{\min}}{1-\gamma_{\min}} = 2d$, then taking $\beta = \gamma = \gamma_{\min}$ yields

$$\begin{aligned}
\eta_{\min} P_T^{\gamma_{\min}}(\boldsymbol{u}) + \frac{d\sum_{t=1}^T \ell_t(u_t)}{\eta_{\min}} &\leq \eta_{\min} P_T^{\gamma_{\min}}(\boldsymbol{u}) + \frac{d\sum_{t=1}^T \ell_t(u_t)}{\eta^\circ} \\
&= 2dP_T^{\gamma_{\min}}(\boldsymbol{u}) + \sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)}.
\end{aligned}$$

Lastly, if $\eta_{\min} \leq \eta^\circ \leq \eta_{\max}$, there is a $\eta_k = \frac{\gamma_k}{1-\gamma_k} \in \mathcal{S}_\eta$ such that $\eta_k \leq \eta^\circ \leq b\eta_k$, so choosing $\beta = \gamma^\circ$ and $\gamma = \gamma_k$ yields

$$\begin{aligned}
\eta^\circ P_T^{\gamma^\circ}(\boldsymbol{u}) + \frac{d\sum_{t=1}^T \ell_t(u_t)}{\eta_k} &\leq \eta^\circ P_T^{\gamma^\circ}(\boldsymbol{u}) + b\frac{d\sum_{t=1}^T \ell_t(u_t)}{\eta^\circ} \\
&= (b+1)\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)}
\end{aligned}$$

Combining the three cases, we have

$$2\frac{\beta}{1-\beta} P_T^\beta(\boldsymbol{u}) + \frac{1-\gamma}{\gamma} 2d \sum_{t=1}^T \ell_t(u_t) \leq 4dP_T^{\gamma_{\min}}(\boldsymbol{u}) + 2\max_t \ell_t(u_t) + 2(b+1)\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)}$$

Hence, overall the regret can be bound as

$$\begin{aligned}
R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) &\leq \gamma\lambda \|u_1\|_2^2 + d \max_t (y_t - \overline{y}_t^\gamma)^2 \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{\lambda d}\right) \\
&\quad + 4dP_T^{\gamma_{\min}}(\boldsymbol{u}) + 2\max_t \ell_t(u_t) + 2(b+1)\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)} \\
&\leq O\left(dP_T^{\gamma_{\min}}(\boldsymbol{u}) + d\max_t (y_t - y_t^{\text{Ref}})^2 \log(T) + b\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u}) \sum_{t=1}^T \ell_t(u_t)}\right),
\end{aligned}$$

where we've applied Lemma D.1 to bound $\max_t(y_t - \overline{y}_t^\gamma)^2 \le 4M_{T+1}^2 = 4\max_t(y_t - y_t^{\text{Ref}})^2$. Plugging this back into Equation (14) proves the stated bound. □

Now the proof of Theorem 4.3 follows by composing Theorem 4.1 and Lemma D.5. The theorem is restated below for convenience.

Theorem 4.3. *Under the same conditions as Theorem 4.2, suppose each $\mathcal{A}_\gamma$ sets hints $\widetilde{y}_t = \overline{y}_t^\gamma = \text{Clip}_{\mathcal{B}_t}(\langle, x_t, w_t^\gamma\rangle)$, where $\mathcal{B}_t = [y_t^{Ref} - M_t, y_t^{Ref} + M_t]$ and $M_t = \max_{s<t}\left|y_s - y_s^{Ref}\right|$. Then for any $\boldsymbol{u} = (u_1, \dots, u_T)$ in $\mathbb{R}^d$, Algorithm 2 guarantees*

$$R_T(\boldsymbol{u}) \le O\Bigg(dP_T^{\gamma_{\min}}(\boldsymbol{u}) + d\max_t\left(y_t - y_t^{Ref}\right)^2\log(T) + b\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u})\sum_{t=1}^T \ell_t(u_t)}\Bigg)$$

where $\gamma_{\min} = \frac{2d}{2d+1}$ and $\gamma^\circ \in [0,1]$ satisfies Equation (3).

Proof. As in the proof of Theorem 4.2, we apply Theorem 4.1, from which it follows that for any $\boldsymbol{u} = (u_1, \dots, u_T)$ in $\mathbb{R}^d$ and any $\gamma \in \mathcal{S}_\gamma$, the dynamic regret is bounded as

$$\begin{aligned} R_T(\boldsymbol{u}) &\le \widehat{O}\left(R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) + \max_t(y_t - y_t^{\text{Ref}})^2\log(NT)\right) \\ &\le \widehat{O}\left(R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) + \max_t(y_t - y_t^{\text{Ref}})^2\log(T)\right), \end{aligned} \tag{14}$$

where the last line uses $N = |\mathcal{S}_\gamma| = \log_b(\eta_{\max}/\eta_{\min}) \le O(\log_b(T))$, then hides $\log(\log)$ factors. And using Lemma D.5, for any $\boldsymbol{u} = (u_1, \dots, u_T)$ there is a $\gamma^\circ \in [0,1]$ satisfying $\gamma^\circ = \frac{\sqrt{d\sum_{t=1}^T \ell_t(u_t)}}{\sqrt{d\sum_{t=1}^T \ell_t(u_t)} + \sqrt{P_T^\circ(\boldsymbol{u})}}$ and a $\gamma \in \mathcal{S}_\gamma$ such that

$$R_T^{\mathcal{A}_\gamma}(\boldsymbol{u}) \le O\left(dP_T^{\gamma_{\min}}(\boldsymbol{u}) + d\max_t(y_t - y_t^{\text{Ref}})^2\log(T) + b\sqrt{dP_T^{\gamma^\circ}(\boldsymbol{u})\sum_{t=1}^T \ell_t(u_t)}\right),$$

Plugging this back into Equation (14) completes the proof. □

E. Adaptive Fixed-share

Algorithm 3: Adaptive Fixed-Share

Input *Experts* $\mathcal{A}_1, \dots, \mathcal{A}_N$, $p_1 \in \Delta_N$
for $t = 1 : T$ **do**
  Get $y_t^{(i)}$ from $\mathcal{A}_i$ for all i
  Play $\overline{y}_t = \sum_{i=1}^N p_{ti} y_t^{(i)}$
  Observe loss $\ell_t(y) = \frac{1}{2}(y_t - y)^2$ and let $\ell_{ti} = \ell_t(y_t^{(i)})$ for all i
  Let $q_{t+1,i} = \frac{p_{ti}\exp(-\alpha_t \ell_{ti})}{\sum_{j=1}^N p_{tj}\exp(-\alpha_t \ell_{tj})}$ for all i
  Choose β_{t+1} and set $p_{t+1} = (1 - \beta_{t+1})q_{t+1} + \beta_{t+1}p_1$
end

In this section, we provide for completeness analysis related to the fixed-share algorithm (Cesa-Bianchi et al., 2012) with time-varying modulus. The following is a modest generalization of the analysis of Hazan (2019, Theorem 10.3). Throughout this section we assume that the losses $\ell_t : \widehat{\mathcal{Y}} \to \mathbb{R}$ are exp-concave in their domain.

Theorem E.1. *For all t let ℓ_t be an α_t-Exp-Concave function and assume that $\alpha_t \ge \alpha_{t+1}$ for all t. For all t, set $\beta_t \le \frac{1}{(e+t)\log^2(e+t)+1}$. Then for any $j \in [N]$ and any $[a,b] \subseteq [1,T]$, Algorithm 3 guarantees*

$$\sum_{t=a}^b \ell_t(\overline{y}_t) - \ell_t(y_t^{(j)}) \le \frac{1}{\alpha_{b+1}}\left[2\log\left(\frac{1}{\beta_{b+1}p_{1j}}\right) + 1\right]$$

Proof. The heavy lifting is done mostly using Lemma E.2, after which the proof follows by choosing the sequence of mixing parameters β_t. Applying Lemma E.2 and observing the telescoping sum, we have

$$\begin{aligned}
\sum_{t=a}^{b} \ell_t(\overline{y}_t) - \ell_t\left(y_t^{(j)}\right) &\leq \sum_{t=a}^{b} \frac{1}{\alpha_t}\log\left(\frac{1}{p_{tj}}\right) - \frac{1}{\alpha_{t+1}}\log\left(\frac{1}{p_{t+1,j}}\right) \\
&+ \sum_{t=a}^{b} \frac{1}{\alpha_t}\log\left(\frac{1}{1-\beta_{t+1}}\right) \\
&+ \sum_{t=a}^{b} \left|\frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t}\right| \log\left(\frac{1}{\beta_{t+1}p_{1j}}\right) \\
&= \frac{1}{\alpha_a}\log\left(\frac{1}{p_{aj}}\right) - \frac{1}{\alpha_{b+1}}\log\left(\frac{1}{p_{b+1,j}}\right) \\
&+ \sum_{t=a}^{b} \frac{1}{\alpha_t}\log\left(\frac{1}{1-\beta_{t+1}}\right) \\
&+ \sum_{t=a}^{b} \left|\frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t}\right| \log\left(\frac{1}{\beta_{t+1}p_{1j}}\right).
\end{aligned}$$

Now observe that with $\beta_{t+1} \leq \frac{1}{(e+t)\log^2(e+t)+1}$, using the elementary inequality $\log(1+y) \leq y$ we have

$$\log\left(\frac{1}{1-\beta_{t+1}}\right) = \log\left(1 + \frac{\beta_{t+1}}{1-\beta_{t+1}}\right) \leq \frac{\beta_{t+1}}{1-\beta_{t+1}} = \frac{1}{(e+t)\log^2(e+t)}$$

so for non-increasing α_t we have

$$\begin{aligned}
\sum_{t=a}^{b} \frac{1}{\alpha_t}\log\left(\frac{1}{1-\beta_{t+1}}\right) &\leq \sum_{t=a}^{b} \frac{1}{\alpha_t}\frac{1}{(e+t)\log^2(e+t)} \\
&\leq \frac{1}{\alpha_b}\sum_{t=a}^{b} \frac{1}{(e+t)\log^2(e+t)} \\
&\leq \frac{1}{\alpha_b}\int_e^{e+b} \frac{1}{y\log^2 y}dy \\
&= \frac{1}{\alpha_b}\frac{-1}{\log(y)}\Bigg|_{y=e}^{e+b} \leq \frac{1}{\alpha_b}
\end{aligned}$$

and similarly,

$$\begin{aligned}
\sum_{t=a}^{b} \left|\frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t}\right| \log\left(\frac{1}{\beta_{t+1}p_{1j}}\right) &\leq \log\left(\frac{1}{\beta_{b+1}p_{1j}}\right)\sum_{t=a}^{b} \frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t} \\
&\leq \frac{1}{\alpha_{b+1}}\log\left(\frac{1}{\beta_{b+1}p_{1j}}\right),
\end{aligned}$$

so overall we have

$$\begin{aligned}
\sum_{t=a}^{b} \ell_t(\overline{y}_t) - \ell_t\left(y_t^{(j)}\right) &\leq \frac{1}{\alpha_a}\log\left(\frac{1}{p_{aj}}\right) - \frac{1}{\alpha_{b+1}}\log\left(\frac{1}{p_{b+1,j}}\right) + \frac{\log\left(\frac{1}{\beta_{b+1}p_{1j}}\right)+1}{\alpha_{b+1}} \\
&= \frac{1}{\alpha_a}\log\left(\frac{1}{p_{aj}}\right) + \frac{\log\left(\frac{p_{b+1,j}}{\beta_{b+1}p_{1j}}\right)+1}{\alpha_{b+1}} \\
&\leq \frac{1}{\alpha_{b+1}}\log\left(\frac{1}{(1-\beta_a)q_{aj}+\beta_a p_{1j}}\right) + \frac{\log\left(\frac{p_{b+1,j}}{\beta_{b+1}p_{1j}}\right)+1}{\alpha_{b+1}} \\
&\leq \frac{1}{\alpha_{b+1}}\left[2\log\left(\frac{1}{\beta_{b+1}p_{1j}}\right)+1\right] \qquad \leq
\end{aligned}$$

$\square$

E.1. Proof of Lemma E.2

The following provides an initial one-step bound to work from, which we use in the proof of Theorem E.1.

Lemma E.2. *For all t let ℓ_t be an α_t-Exp-Concave function. Then for any $j \in [N]$, Algorithm 3 guarantees*

$$\begin{aligned}
\ell_t(\overline{y}_t) - \ell_t(y_t^{(j)}) \le\ & \frac{1}{\alpha_t}\log\left(\frac{1}{p_{tj}}\right) - \frac{1}{\alpha_{t+1}}\log\left(\frac{1}{p_{t+1,j}}\right) \\
&+ \frac{1}{\alpha_t}\log\left(\frac{1}{1-\beta_{t+1}}\right) \\
&+ \left|\frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t}\right|\log\left(\frac{1}{\beta_{t+1}p_{1j}}\right)
\end{aligned}$$

Proof. By α_t-Exp-Concavity of ℓ_t, we have that $y \mapsto \exp\left(-\alpha_t\ell_t(y)\right)$ is concave. Hence, applying Jensen's inequality:

$$\exp\left(-\alpha_t\ell_t(\overline{y}_t)\right) \ge \sum_{i=1}^N p_{ti}\exp\left(-\alpha_t\ell_t\left(y_t^{(i)}\right)\right) = \sum_{i=1}^N p_{ti}\exp\left(-\alpha_t\ell_{ti}\right)$$

and taking the natural logarithm of both sides we have

$$\begin{aligned}
-\alpha_t\ell_t(\overline{y}_t) &\ge \log\left(\sum_{i=1}^N p_{ti}\exp\left(-\alpha_t\ell_{ti}\right)\right) \\
\ell_t(\overline{y}_t) &\le -\frac{1}{\alpha_t}\log\left(\sum_{i=1}^N p_{ti}\exp\left(-\alpha_t\ell_{ti}\right)\right).
\end{aligned}$$

Hence, for any $j \in [N]$ we have

$$\begin{aligned}
\ell_t(\overline{y}_t) - \ell_t\left(y_t^{(j)}\right) &\le -\frac{1}{\alpha_t}\log\left(\sum_{i=1}^N p_{ti}\exp\left(-\alpha_t\ell_{ti}\right)\right) - \ell_{tj} \\
&= -\frac{1}{\alpha_t}\log\left(\sum_{i=1}^N p_{ti}\exp\left(-\alpha_t\ell_{ti}\right)\right) + \frac{1}{\alpha_t}\log\left(\exp\left(-\alpha_t\ell_{tj}\right)\right) \\
&= \frac{1}{\alpha_t}\log\left(\frac{\exp\left(-\alpha_t\ell_{tj}\right)}{\sum_{i=1}^N p_{ti}\exp\left(-\alpha_t\ell_{ti}\right)}\right) \\
&= \frac{1}{\alpha_t}\log\left(\frac{p_{tj}\exp\left(-\alpha_t\ell_{tj}\right)}{p_{tj}\sum_{i=1}^N p_{ti}\exp\left(-\alpha_t\ell_{ti}\right)}\right) \\
&= \frac{1}{\alpha_t}\left[\log\left(\frac{q_{t+1,j}}{p_{tj}}\right)\right] \\
&= \frac{1}{\alpha_t}\left[\log\left(\frac{1}{p_{tj}}\right) - \log\left(\frac{1}{q_{t+1,j}}\right)\right].
\end{aligned}$$

Adding and subtracting $\frac{1}{\alpha_{t+1}}\log\left(\frac{1}{p_{t+1,j}}\right)$,

$$
\begin{aligned}
\ell_t(\overline{y}_t) - \ell_t\left(y_t^{(j)}\right) &\leq \frac{1}{\alpha_t}\log\left(\frac{1}{p_{tj}}\right) - \frac{1}{\alpha_{t+1}}\log\left(\frac{1}{p_{t+1,j}}\right) \\
&\quad + \frac{1}{\alpha_{t+1}}\log\left(\frac{1}{p_{t+1,j}}\right) - \frac{1}{\alpha_t}\log\left(\frac{1}{q_{t+1,j}}\right) \\
&= \frac{1}{\alpha_t}\log\left(\frac{1}{p_{tj}}\right) - \frac{1}{\alpha_{t+1}}\log\left(\frac{1}{p_{t+1,j}}\right) \\
&\quad + \underbrace{\frac{1}{\alpha_t}\log\left(\frac{1}{p_{t+1,j}}\right) - \frac{1}{\alpha_t}\log\left(\frac{1}{q_{t+1,j}}\right)}_{\log(q_{t+1,j}/p_{t+1,j})/\alpha_t} \\
&\quad + \left[\frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t}\right]\log\left(\frac{1}{p_{t+1,j}}\right)
\end{aligned}
$$

recalling $p_{t+1,j} = (1-\beta_{t+1})q_{t+1,j} + \beta_{t+1}p_{1j}$,

$$
\begin{aligned}
&= \frac{1}{\alpha_t}\log\left(\frac{1}{p_{tj}}\right) - \frac{1}{\alpha_{t+1}}\log\left(\frac{1}{p_{t+1,j}}\right) \\
&\quad + \frac{1}{\alpha_t}\log\left(\frac{q_{t+1,j}}{(1-\beta_{t+1})q_{t+1,j} + \beta_{t+1}p_{1j}}\right) \\
&\quad + \left[\frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t}\right]\log\left(\frac{1}{(1-\beta_{t+1})q_{t+1,j} + \beta_{t+1}p_{1j}}\right) \\
&\leq \frac{1}{\alpha_t}\log\left(\frac{1}{p_{tj}}\right) - \frac{1}{\alpha_{t+1}}\log\left(\frac{1}{p_{t+1,j}}\right) \\
&\quad + \frac{1}{\alpha_t}\log\left(\frac{1}{1-\beta_{t+1}}\right) \\
&\quad + \left|\frac{1}{\alpha_{t+1}} - \frac{1}{\alpha_t}\right|\log\left(\frac{1}{\beta_{t+1}p_{1j}}\right)
\end{aligned}
$$

□

F. Strongly-Adaptive Guarantees

In this section we provide a formal statement of the result sketched in Section 5. The result follows easily from the results in Section 4, after borrowing the geometric covering intervals from Daniely et al. (2015).

Theorem F.1. *Let $\mathcal{S}_\gamma$ be the set of discount factors defined in Theorem 4.2, let S denote a set of geometric covering intervals over $[1,T]$, and for each $\gamma \in \mathcal{S}_\gamma$ and $I \in S$, let $\mathcal{A}_{\gamma,I}$ be an instance of Algorithm 1 using discount γ and applied during interval I (and predicts y_t^{Ref} for $t \notin I$). Let $\mathcal{A}_{Meta}$ be an instance of the meta-algorithm characterized in Theorem 4.1. Then for any $[s,\tau] \subseteq [1,T]$, there is a set of disjoint intervals $I_1,\dots,I_K$ in S such that $\cup_{i=1}^K I_i = [s,\tau]$, and moreover, for any $\boldsymbol{u} = (u_s,\dots,u_\tau)$ Algorithm 2 with $y_t^{Ref} = \widetilde{y}_t$ guarantees*

$$R_{[s,\tau]}(\boldsymbol{u}) \leq \widehat{O}\left(d\max_t (y_t - y_t^{Ref})^2 \log^2(T) + b\sqrt{dP_{[s,\tau]}^{\gamma^*}(\boldsymbol{u}) \sum_{t\in[s,\tau]} (y_t - \widetilde{y}_t)^2}\right)$$

where $P_{[s,\tau]}^{\gamma^*}(\boldsymbol{u}) = \sum_{i=1}^K P_{I_i}^{\gamma_i^*}(\boldsymbol{u})$ and each $\gamma_i^* \in [0,1]$ satisfies $\gamma_i^* = \frac{\sqrt{\frac{d}{2}\sum_{t\in I_i}(y_t-\widetilde{y}_t)^2}}{\sqrt{\frac{d}{2}\sum_{t\in I_i}(y_t-\widetilde{y}_t)^2}+\sqrt{P_{I_i}^{\gamma_i^*}(\boldsymbol{u})}}$.

If we instead suppose each $\mathcal{A}_{\gamma,I}$ sets hints as in Theorem 4.3, then for any $\boldsymbol{u} = (u_s,\dots,u_\tau)$ Algorithm 2 guarantees

$$R_{[s,\tau]}(\boldsymbol{u}) \leq \widehat{O}\left(dP_{[s,\tau]}^{\gamma_{\min}}(\boldsymbol{u}) + d\max_t (y_t - y_t^{Ref})^2 \log^2(T) + b\sqrt{dP_{[s,\tau]}^{\gamma^\circ}(\boldsymbol{u}) \sum_{t\in[s,\tau]} \ell_t(u_t)}\right)$$

where $P_{[s,\tau]}^{\gamma^\circ}(\boldsymbol{u}) = \sum_{i=1}^K P_{I_i}^{\gamma_i^\circ}(\boldsymbol{u})$ and each $\gamma_i^\circ \in [0,1]$ satisfies $\gamma_i^\circ = \frac{\sqrt{d\sum_{t\in I_i}\ell_t(u_t)}}{\sqrt{d\sum_{t\in I_i}\ell_t(u_t)}+\sqrt{P_{I_i}^{\gamma_i^\circ}(\boldsymbol{u})}}$.

Proof. For any $[s,\tau] \subseteq [1,T]$, Daniely et al. (2015, Lemma 1.2) shows that there exists a disjoint set of intervals $I_1,\dots,I_K$ in S such that $\cup_{i=1}^K I_i = [s,\tau]$ and $K \leq O(\log(\tau - s))$. Hence, we can decompose $\sum_{i=1}^K R_{I_i}(\boldsymbol{u})$, so applying Theorem 4.1 to each of these sub-intervals, for any $\gamma \in \mathcal{S}_\gamma$ we have:

$$\begin{aligned} R_{[s,\tau]}(\boldsymbol{u}) = \sum_{i=1}^K R_{I_i}(\boldsymbol{u}) &\leq \sum_{i=1}^K \widehat{O}\left(R_{I_i}^{\mathcal{A}_{\gamma,I_i}}(\boldsymbol{u}) + \max_t (y_t - \widetilde{y}_t)^2 \log(N|I_i|)\right) \\ &\leq \widehat{O}\left(\sum_{i=1}^K R_{I_i}^{\mathcal{A}_{\gamma,I_i}}(\boldsymbol{u}) + K\max_t (y_t - \widetilde{y}_t)^2 \log(N(\tau - s))\right) \\ &\leq \widehat{O}\left(\sum_{i=1}^K R_{I_i}^{\mathcal{A}_{\gamma,I_i}}(\boldsymbol{u}) + \max_t (y_t - \widetilde{y}_t)^2 \log^2(T)\right), \end{aligned} \tag{15}$$

where $\widehat{O}(\cdot)$ hides $\log(\log)$ factors and the last line bounds $K \leq O(\log(\tau - s)) \leq O(\log(T))$ and $N \leq O(T\log(T))$. The bound on N can be seen from the fact that $|\mathcal{S}_\gamma| \leq O(\log(T))$, and from the fact that S is constructed as $S = \cup_{i=1}^{\lfloor \log(T) \rfloor} S_i$ where $S_i = \{[k2^i, (k+1)2^i - 1] : k = 0,1,\dots\}$, from which it is easily seen that $|S| \leq O(T)$ by observing that each S_i has at most $T/2^i$ intervals, hence summing them all up yields $|S| = \sum_{i=1}^{\lfloor \log(T) \rfloor} |S_i| \leq O(T)$.

Now for any interval I_i, Lemma D.4 shows that there is a $\gamma_i^* \in [0,1]$ satisfying $\gamma_i^* = \frac{\sqrt{d\sum_{t\in I_i}\frac{1}{2}(y_t-\widetilde{y}_t)^2}}{\sqrt{d\sum_{t\in I_i}\frac{1}{2}(y_t-\widetilde{y}_t)^2}+\sqrt{P_{I_i}^{\gamma_i^*}(\boldsymbol{u})}}$ and a $\gamma \in \mathcal{S}_\gamma$ such that

$$R_{I_i}^{\mathcal{A}_{\gamma,I_i}}(\boldsymbol{u}) \leq O\left(d\max_t (y_t - \widetilde{y}_t)^2 \log(|I_i|) + b\sqrt{dP_{I_i}^{\gamma_i^*}(\boldsymbol{u}) \sum_{t\in I_i} (y_t - \widetilde{y}_t)^2}\right)$$

so summing these up and applying Cauchy-Schwarz inequlity leads to

$$\begin{aligned} \sum_{i=1}^K R_{I_i}^{\mathcal{A}_{\gamma,I_i}}(\boldsymbol{u}) &\leq O\left(Kd\max_t (y_t - \widetilde{y}_t)^2 \log(|I_i|) + \sum_{i=1}^K b\sqrt{dP_{I_i}^{\gamma_i^*}(\boldsymbol{u}) \sum_{t\in I_i} (y_t - \widetilde{y}_t)^2}\right) \\ &\leq O\left(d\max_t (y_t - \widetilde{y}_t)^2 \log^2(\tau - s) + b\sqrt{dP_{[s,\tau]}^{\gamma^*}(\boldsymbol{u}) \sum_{t\in[s,\tau]} (y_t - \widetilde{y}_t)^2}\right) \end{aligned}$$

where we've defined $P^{\gamma^*}_{[s,\tau]}(\boldsymbol{u}) = \sum_{i=1}^K P^{\gamma_i^*}_{I_i}(\boldsymbol{u})$. Plugging this back into Equation (15), overall we may bound:

$$R_{[s,\tau]}(\boldsymbol{u}) \le \widehat{O}\left(d\max_t (y_t - y_t^{\text{Ref}})^2 \log^2(T) + b\sqrt{dP^{\gamma^*}_{[s,\tau]}(\boldsymbol{u})\sum_{t\in[s,\tau]}(y_t - \widetilde{y}_t)^2}\right)$$

where we've chosen $\widetilde{y}_t = y_t^{\text{Ref}}$ for simplicity.

An identical argument holds for the second statement: for any interval I_i, Lemma D.5 shows that there is a $\gamma_i^\circ \in [0,1]$ satisfying $\gamma_i^\circ = \frac{\sqrt{d\sum_{t\in I_i}\ell_t(u_t)}}{\sqrt{d\sum_{t\in I_i}\ell_t(u_t)}+\sqrt{P^{\gamma_i^\circ}_{I_i}(\boldsymbol{u})}}$ and a $\gamma \in \mathcal{S}_\gamma$ such that

$$R^{\mathcal{A}_{\gamma,I_i}}_{I_i}(\boldsymbol{u}) \le O\left(dP^{\gamma_{\min}}_{I_i}(\boldsymbol{u}) + d\max_t (y_t - y_t^{\text{Ref}})^2 \log(|I_i|) + b\sqrt{dP^{\gamma_i^\circ}_{I_i}(\boldsymbol{u})\sum_{t\in I_i}\ell_t(u_t)}\right)$$

so summing these up and applying Cauchy-Schwarz inequality again leads to

$$\begin{aligned}\sum_{i=1}^K R^{\mathcal{A}_{\gamma,I_i}}_{I_i}(\boldsymbol{u}) &\le O\left(dP^{\gamma_{\min}}_{[s,\tau]}(\boldsymbol{u}) + Kd\max_t (y_t - \widetilde{y}_t)^2 \log(|I_i|) + \sum_{i=1}^K b\sqrt{dP^{\gamma^\circ_i}_{I_i}(\boldsymbol{u})\sum_{t\in I_i}\ell_t(u_t)}\right)\\ &\le O\left(dP^{\gamma_{\min}}_{[s,\tau]}(\boldsymbol{u}) + d\max_t (y_t - \widetilde{y}_t)^2 \log^2(\tau - s) + b\sqrt{dP^{\gamma^\circ}_{[s,\tau]}(\boldsymbol{u})\sum_{t\in[s,\tau]}\ell_t(u_t)}\right)\end{aligned}$$

where we've defined $P^{\gamma^*}_{[s,\tau]}(\boldsymbol{u}) = \sum_{i=1}^K P^{\gamma_i^*}_{I_i}(\boldsymbol{u})$, so plugging this back into Equation (15), overall we may bound:

$$R_{[s,\tau]}(\boldsymbol{u}) \le \widehat{O}\left(dP^{\gamma_{\min}}_{[s,\tau]}(\boldsymbol{u}) + d\max_t (y_t - y_t^{\text{Ref}})^2 \log^2(T) + b\sqrt{dP^{\gamma^\circ}_{[s,\tau]}(\boldsymbol{u})\sum_{t\in[s,\tau]}\ell_t(u_t)}\right),$$

where we've defined $P^{\gamma^\circ}_{[s,\tau]} = \sum_{i=1}^K P^{\gamma_i^\circ}_{I_i}(\boldsymbol{u})$.

□

F.1. Matching the Exp-concave Guarantee in Unbounded Domains

Recall from Section 3.2 that in the Exp-concave setting, the algorithm of Baby & Wang (2021) achieves a dynamic regret bound of the form $R_T(\boldsymbol{u}) \le \widetilde{O}\left(T^{1/3}C_T^{2/3}\right)$ for $C_T = \sum_{t=1}^{T-1}\|u_t - u_{t-1}\|_1$. Our strongly-adaptive guarantees in Theorem F.1 show that a bound of this form can be achieved even in the unbounded domain setting. To see why, note that the essential intuition of Baby & Wang (2021) is that if we have access to a *strongly-adaptive* algorithm guaranteeing $R_{[a,b]}(u) \le O(\log(b-a))$ *static* regret on all intervals $[a,b] \subseteq [1,T]$, then to attain the desired bound up to log terms it suffices to show that there *exists* a set of intervals $\{I_1,\ldots,I_N\}$ partitioning $[1,T]$ such that $N \le T^{1/3}C_T^{2/3}$ and that the dynamic regret is bounded by the static regrets over the partition, leading to regret matching $O(T^{1/3}C_T^{2/3})$ up to logarithmic terms.

Our strongly-adaptive guarantee in Theorem F.1 actually achieves a stronger guarantee than is necessary to invoke the above argument, by guaranteeing $O\big(\log(b-a) \vee \sqrt{dP^{\gamma}_{[a,b]}(\boldsymbol{u})|b-a|}\big)$ *dynamic* regret on every interval $[a,b]$, and hence as a special case we have $O(\log(b-a))$ static regret on each interval as well. A similar partitioning argument then provides an analogous $T^{1/3}C_T^{2/3}$ bound, even in unbounded domains. If this is surprising, note that the exp-concave (and hence bounded domain) restriction is only really used to provide an algorithm which achieves logarithmic static regret, not to construct the essential partition. In the online linear regression setting, we do not need exp-concavity to guarantee logarithmic static regret — the VAW forecaster can provide the necessary guarantee even in an unbounded domain.

G. Supporting Lemmas

The following provides a useful relation between the squared loss and its Bregman divergence.

Lemma G.1. *Let $\ell_t(w) = \frac{1}{2}(y_t - \langle x_t, w_t\rangle)^2$. Then for any $u, w \in W$,*

$$D_{\ell_t}(u|w) = \frac{1}{2}\langle x_t, u - w\rangle^2$$

Proof. By definition of Bregman divergence, we have:

$$D_{\ell_t}(u|w) = \ell_t(u) - \ell_t(w) - \langle \nabla \ell_t(w), u - w\rangle .$$

Expanding the definition of ℓ_t, we have

$$\begin{aligned}
\ell_t(u) - \ell_t(w) &= \frac{1}{2}(y_t - \langle x_t, u\rangle)^2 - \frac{1}{2}(y_t - \langle x_t, w\rangle)^2 \\
&= \frac{1}{2}y_t^2 + \frac{1}{2}\langle x_t, u\rangle^2 - y_t\langle x_t, u\rangle - \frac{1}{2}y_t^2 - \frac{1}{2}\langle x_t, w\rangle^2 + y_t\langle x_t, w\rangle \\
&= \frac{1}{2}\langle x_t, u\rangle^2 - \frac{1}{2}\langle x_t, w\rangle^2 + y_t\langle x_t, w - u\rangle .
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
-\langle \nabla \ell_t(w), u - w\rangle &= \langle (y_t - \langle x_t, w\rangle)x_t, u - w\rangle \\
&= -y_t\langle x_t, w - u\rangle + \langle x_t, w\rangle^2 - \langle x_t, w\rangle\langle x_t, u\rangle ,
\end{aligned}$$

so combining with the previous display we have

$$\begin{aligned}
\ell_t(u) - \ell_t(w) - \langle \nabla \ell_t(w), u - w\rangle &= \frac{1}{2}\langle x_t, u\rangle^2 - \frac{1}{2}\langle x_t, w\rangle^2 + y_t\langle x_t, w - u\rangle \\
&\quad - y_t\langle x_t, w - u\rangle + \langle x_t, w\rangle^2 - \langle x_t, w\rangle\langle x_t, u\rangle \\
&= \frac{1}{2}\langle x_t, u\rangle^2 + \frac{1}{2}\langle x_t, w\rangle^2 - \langle x_t, w\rangle\langle x_t, u\rangle \\
&= \frac{1}{2}(\langle x_t, u\rangle - \langle x_t, w\rangle)^2 \\
&= \frac{1}{2}\langle x_t, u - w\rangle^2 .
\end{aligned}$$

□

The following provides a discounted version of the log-determinant lemma.

Lemma G.2. *Let $\gamma \in (0, 1]$, $\lambda > 0$, $x_t \in \mathbb{R}^d$, and define $M_0 = \lambda I$ and $M_t = x_t x_t^\top + \gamma M_{t-1}$ for each $t > 0$. Then for any sequence $\Delta_1, \Delta_2, \ldots$ in $\mathbb{R}$,*

$$\sum_{t=1}^{T} \Delta_t^2 \|x_t\|_{M_t^{-1}}^2 \le d\log(1/\gamma)\Delta_{1:T}^2 + \max_t \Delta_t^2 d\log\left(1 + \frac{\sum_{t=1}^{T}\gamma^{T-t}\|x_t\|_2^2}{\lambda d}\right)$$

Proof. By definition we have $M_t = x_t x_t^\top + \gamma M_{t-1}$, so re-arranging and taking the determinant of both sides we have

$$\begin{aligned}
\text{Det}(\gamma M_{t-1}) = \text{Det}\left(M_t - x_t x_t^\top\right) &= \text{Det}(M_t)\,\text{Det}\left(I - M_t^{-\frac{1}{2}} x_t x_t^\top M_t^{-\frac{1}{2}}\right) \\
&= \text{Det}(M_t)\left(1 - \|x_t\|_{M_t^{-1}}^2\right)
\end{aligned}$$

where the last line uses the fact that $\text{Det}\left(I - yy^\top\right) = 1 - \|y\|_2^2$. Re-arranging, using $\text{Det}\left(\gamma M_{t-1}\right) = \gamma^d \text{Det}\left(M_{t-1}\right)$, and using the fact that $1 - x \le -\log(x)$ we have

$$\begin{aligned}
\sum_{t=1}^T \Delta_t^2 \|x_t\|_{M_t^{-1}}^2 &= \sum_{t=1}^T \Delta_t^2 \left[1 - \frac{\gamma^d \text{Det}\left(M_{t-1}\right)}{\text{Det}\left(M_t\right)}\right] \\
&\le \sum_{t=1}^T \Delta_t^2 \log\left(\frac{\text{Det}\left(M_t\right)}{\gamma^d \text{Det}\left(M_{t-1}\right)}\right) \\
&= \sum_{t=1}^T \Delta_t^2 d \log\left(1/\gamma\right) + \sum_{t=1}^T \Delta_t^2 \log\left(\frac{\text{Det}\left(M_t\right)}{\text{Det}\left(M_{t-1}\right)}\right) \\
&\le d \log\left(1/\gamma\right) \Delta_{1:T}^2 + \max_t \Delta_t^2 \log\left(\prod_{t=1}^T \frac{\text{Det}\left(M_t\right)}{\text{Det}\left(M_{t-1}\right)}\right) \\
&= d \log\left(1/\gamma\right) \Delta_{1:T}^2 + \max_t \Delta_t^2 \log\left(\frac{\text{Det}\left(M_T\right)}{\text{Det}\left(M_0\right)}\right).
\end{aligned}$$

Observe that $\text{Det}\left(M_0\right) = \text{Det}\left(\lambda I\right) = \lambda^d$, and using AM-GM inequality we have

$$\begin{aligned}
\text{Det}\left(M_T\right) &\le \left(\frac{\text{Tr}\left(M_t\right)}{d}\right)^d = \left(\frac{\text{Tr}\left(\lambda \gamma^T I + \sum_{t=1}^T \gamma^{T-t} x_t x_t^\top\right)}{d}\right)^d \\
&= \left(\frac{d\lambda\gamma^T + \sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{d}\right)^d,
\end{aligned}$$

Hence $\frac{\text{Det}(M_T)}{\text{Det}(M_0)} \le \left(\frac{d\lambda\gamma^T + \sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{d\lambda}\right)^d$, so overall we have

$$\begin{aligned}
\sum_{t=1}^T \Delta_t^2 \|x_t\|_{M_t^{-1}}^2 &\le d \log\left(1/\gamma\right) \Delta_{1:T}^2 + \max_t \Delta_t^2 \log\left(\left(\frac{d\lambda\gamma^T + \sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{\lambda d}\right)^d\right) \\
&= d \log\left(1/\gamma\right) \Delta_{1:T}^2 + \max_t \Delta_t^2 d \log\left(\frac{d\lambda\gamma^T + \sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{\lambda d}\right) \\
&\le d \log\left(1/\gamma\right) \Delta_{1:T}^2 + \max_t \Delta_t^2 d \log\left(1 + \frac{\sum_{t=1}^T \gamma^{T-t} \|x_t\|_2^2}{\lambda d}\right)
\end{aligned}$$

□

Note that the Lemma G.2 also immediately gives us the usual log determinant lemma as a special case where $\gamma = 1$:

Lemma G.3. *Let $\lambda > 0$, $x_t \in \mathbb{R}^d$, and define Let $M_0 = \lambda I$ and $M_t = x_t x_t^\top + M_{t-1}$ for each $t > 0$. Then for any sequence $\Delta_1, \Delta_2, \ldots$ in $\mathbb{R}$,*

$$\sum_{t=1}^T \Delta_t^2 \|x_t\|_{M_t^{-1}}^2 \le d \max_t \Delta_t^2 \log\left(1 + \frac{\sum_{t=1}^T \|x_t\|_2^2}{\lambda d}\right)$$

The following lemma is common in adaptive online learning and provided for completeness.

Lemma G.4. *Let $a_1, \ldots, a_T$ be arbitrary non-negative numbers in $\mathbb{R}$. Then*

$$\sqrt{\sum_{t=1}^T a_t} \le \sum_{t=1}^T \frac{a_t}{\sqrt{\sum_{s=1}^t a_s}} \le 2\sqrt{\sum_{t=1}^T a_t}$$

Proof. By concavity of $x \mapsto \sqrt{x}$, we have

$$\sqrt{a_{1:t}} - \sqrt{a_{1:t-1}} \geq \frac{a_t}{2\sqrt{a_{1:t}}},$$

so summing over t and observing the resulting telescoping sum yields

$$\sum_{t=1}^{T} \frac{a_t}{\sqrt{a_{1:t}}} \leq 2 \sum_{t=1}^{T} \sqrt{a_{1:t}} - \sqrt{a_{1:t-1}} = 2\sqrt{a_{1:T}}.$$

For the lower bound, observe that

$$\sum_{t=1}^{T} \frac{a_t}{\sqrt{a_{1:t}}} \geq \sum_{t=1}^{T} \frac{a_t}{\sqrt{a_{1:T}}} = \frac{a_{1:T}}{\sqrt{a_{1:T}}} = \sqrt{a_{1:T}}$$

□