A. Extended Literature Review

We review here additional prior work not detailed in the main paper.

A.1. General online learning

We recommend the monographs of Shalev-Shwartz (2012); Orabona (2019) and the textbook of Cesa-Bianchi & Lugosi (2006) for surveys of the field of online learning and Joulani et al. (2017); McMahan (2017) for widely applicable and modular analyses of online learning algorithms.

A.2. Online learning with optimism but without delay

Syrgkanis et al. (2015) analyzed optimistic FTRL and two-step variant of optimistic MD without delay. The work focuses on a particular form of optimism (using the last observed subgradient as a hint) and shows improved rates of convergence to correlated equilibria in multiplayer games. In the absence of delay, Steinhardt & Liang (2014) combined optimism and adaptivity to obtain improvements over standard optimistic regret bounds.

A.3. Online learning with delay but without optimism

Overview Joulani et al. (2013; 2016); McMahan & Streeter (2014) provide broad reviews of progress on delayed online learning.

Delayed stochastic optimization Recht et al. (2011); Agarwal & Duchi (2011); Nesterov (2012); Liu et al. (2014); Liu & Wright (2015); Sra et al. (2016) studied the effects of delay on stochastic optimization but do not treat the adversarial setting studied here.

FTRL-Prox vs. FTRL Joulani et al. (2016) analyzed the delayed feedback regret of the *FTRL-Prox* algorithm, which regularizes toward the last played iterate as in online mirror descent, but did not study the standard FTRL algorithms (sometimes called *FTRL-Centered*) analyzed in this work.

A.4. Self-tuned online learning without delay or optimism

In the absence of optimism and delay, de Rooij et al. (2014); Orabona & Pál (2015); Koolen et al. (2014) developed alternative variants of FTRL algorithms that self-tune their learning rates.

A.5. Online learning without delay for climate forecasting

Monteleoni et al. (2011) applied the Learn-α online learning algorithm of Monteleoni & Jaakkola (2004) to the task of ensembling climate models. The authors considered historical temperature data from 20 climate models and tracked the changing sequence of which model predicts best at any given time. In this context, the algorithm used was based on a set of generalized Hidden Markov Models, in which the identity of the current best model is the hidden variable and the updates are derived as Bayesian updates. This work was extended to take into account the influence of regional neighboring locations when performing updates (McQuade & Monteleoni, 2012). These initial results demonstrated the promise of applying online learning to climate model ensembling, but both methods rely on receiving feedback without delay.

B. Proof of Thm. 3: OFTRL regret

We will prove the following more general result for optimistic adaptive FTRL (OAFTRL)

$$\mathbf{w}_{t+1} = \operatorname{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{g}_{1:t} + \tilde{\mathbf{g}}_{t+1}, \mathbf{w}\rangle + \lambda_{t+1}\psi(\mathbf{w}), \qquad \text{(OAFTRL)}$$

from which Thm. 3 will follow with the choice $\lambda_t = \lambda$ for all $t \geq 1$.

Theorem 14 (OAFTRL regret)**.** *If ψ is nonnegative and $(\lambda_t)_{t\geq 1}$ is non-decreasing, then, $\forall \mathbf{u} \in \mathbf{W}$, the* OAFTRL *iterates $\mathbf{w}_t$ satisfy,*

$$\begin{aligned}\text{Regret}_T(\mathbf{u}) &\leq \lambda_T\psi(\mathbf{u}) + \textstyle\sum_{t=1}^T \delta_t \\ &\leq \lambda_T\psi(\mathbf{u}) + \textstyle\sum_{t=1}^T \min\left(\frac{1}{\lambda_t}\text{huber}(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t\|_*), \operatorname{diam}(\mathbf{W})\min(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t\|_*)\right)\end{aligned}$$

for

$$\delta_t \triangleq \min(F_{t+1}(\mathbf{w}_t, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t),\ \ \langle \mathbf{g}_t, \mathbf{w}_t - \bar{\mathbf{w}}_t \rangle,$$
$$F_{t+1}(\hat{\mathbf{w}}_t, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) + \langle \mathbf{g}_t, \mathbf{w}_t - \hat{\mathbf{w}}_t \rangle)_+ \quad \textit{with}$$
$$\bar{\mathbf{w}}_t \triangleq \mathrm{argmin}_{\mathbf{w}\in\mathbf{W}} F_{t+1}(\mathbf{w}, \lambda_t), \quad F_{t+1}(\mathbf{w}, \lambda_t) \triangleq \lambda_t \psi(\mathbf{w}) + \langle \mathbf{g}_{1:t}, \mathbf{w} \rangle, \quad \textit{and}$$
$$\hat{\mathbf{w}}_t \triangleq \mathrm{argmin}_{\mathbf{w}\in\mathbf{W}} \lambda_t \psi(\mathbf{w}) + \langle \mathbf{g}_{1:t} + \min(\tfrac{\|\mathbf{g}_t\|_*}{\|\tilde{\mathbf{g}}_t - \mathbf{g}_t\|_*}, 1)(\tilde{\mathbf{g}}_t - \mathbf{g}_t), \mathbf{w} \rangle.$$

Proof. Consider a sequence of arbitrary auxiliary subgradient hints $\tilde{\mathbf{g}}_1^*, \dots, \tilde{\mathbf{g}}_T^* \in \mathbb{R}^d$ and the auxiliary OAFTRL sequence

$$\mathbf{w}_{t+1}^* = \mathrm{argmin}_{\mathbf{w}^*\in\mathbf{W}} \langle \mathbf{g}_{1:t} + \tilde{\mathbf{g}}_{t+1}^*, \mathbf{w}^* \rangle + \lambda_{t+1}\psi(\mathbf{w}^*) \quad \text{for} \quad 0 \le t \le T \quad \text{with} \quad \tilde{\mathbf{g}}_{T+1}^* \triangleq \mathbf{0} \quad \text{and} \quad \lambda_{T+1} = \lambda_T. \quad (3)$$

Generalizing the forward regret decomposition of Joulani et al. (2017) and the prediction drift decomposition of Joulani et al. (2016), we will decompose the regret of our original $(\mathbf{w}_t)_{t=1}^T$ sequence into the regret of the auxiliary sequence $(\mathbf{w}_t^*)_{t=1}^T$ and the drift between $(\mathbf{w}_t)_{t=1}^T$ and $(\mathbf{w}_t^*)_{t=1}^T$.

For each time t, define the auxiliary optimistic objective function $\tilde{F}_t^*(\mathbf{w}) = F_t(\mathbf{w}) + \langle \tilde{\mathbf{g}}_t^*, \mathbf{w} \rangle$. Fixing any $\mathbf{u} \in \mathbf{W}$, we have the regret bound

$$\mathrm{Regret}_T(\mathbf{u}) = \textstyle\sum_{t=1}^T \ell_t(\mathbf{w}_t) - \ell_t(\mathbf{u}) \le \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{u} \rangle \quad \text{(since each } \ell_t \text{ is convex with } \mathbf{g}_t \in \partial \ell_t(\mathbf{w}_t))$$
$$= \underbrace{\textstyle\sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{w}_t^* \rangle}_{\text{drift}} + \underbrace{\textstyle\sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t^* - \mathbf{u} \rangle}_{\text{auxiliary regret}}.$$

To control the drift term we employ the following lemma, proved in App. B.1, which bounds the difference between two OAFTRL optimizers with different losses but common regularizers.

Lemma 15 (OAFTRL difference bound)**.** *The* OAFTRL *and auxiliary* OAFTRL *iterates* (3)*,* $\mathbf{w}_t$ *and* $\mathbf{w}_t^*$*, satisfy*

$$\|\mathbf{w}_t - \mathbf{w}_t^*\| \le \min(\tfrac{1}{\lambda_t}\|\tilde{\mathbf{g}}_t - \tilde{\mathbf{g}}_t^*\|_*, \mathrm{diam}(\mathbf{W})).$$

Letting $a = \mathrm{diam}(\mathbf{W}) \in \mathbb{R} \cup \{\infty\}$, we now bound each drift term summand using the Fenchel-Young inequality for dual norms and Lem. 15:

$$\langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{w}_t^* \rangle \le \|\mathbf{g}_t\|_* \|\mathbf{w}_t - \mathbf{w}_t^*\| \le \min\left(\tfrac{1}{\lambda_t}\|\mathbf{g}_t\|_* \|\tilde{\mathbf{g}}_t - \tilde{\mathbf{g}}_t^*\|_*, a\|\mathbf{g}_t\|_*\right).$$

To control the auxiliary regret, we begin by invoking the OAFTRL regret bound of Orabona (2019, proof of Thm. 7.28), the nonnegativity of ψ, and the assumption that $(\lambda_t)_{t\ge1}$ is non-decreasing:

$$\textstyle\sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t^* - \mathbf{u} \rangle \le \lambda_{T+1}\psi(\mathbf{u}) - \lambda_1 \psi(\mathbf{w}_1^*) + \sum_{t=1}^T F_{t+1}(\mathbf{w}_t^*, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) + (\lambda_t - \lambda_{t+1})\psi(\mathbf{w}_{t+1}^*)$$
$$\le \lambda_{T+1}\psi(\mathbf{u}) - \lambda_1 \psi(\mathbf{w}_1^*) + \textstyle\sum_{t=1}^T F_{t+1}(\mathbf{w}_t^*, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t).$$

We next bound the summands in this expression in two ways. Since $\mathbf{w}_t^*$ is the minimizer of $\tilde{F}_t^*$, we may apply the Fenchel-Young inequality for dual norms to conclude that

$$F_{t+1}(\mathbf{w}_t^*, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) = \tilde{F}_t^*(\mathbf{w}_t^*) + \langle \mathbf{w}_t^*, \mathbf{g}_t - \tilde{\mathbf{g}}_t^* \rangle - (\tilde{F}_t^*(\bar{\mathbf{w}}_t) + \langle \bar{\mathbf{w}}_t, \mathbf{g}_t - \tilde{\mathbf{g}}_t^* \rangle)$$
$$\le \langle \mathbf{w}_t^* - \bar{\mathbf{w}}_t, \mathbf{g}_t - \tilde{\mathbf{g}}_t^* \rangle \le \|\mathbf{w}_t^* - \bar{\mathbf{w}}_t\| \|\mathbf{g}_t - \tilde{\mathbf{g}}_t^*\|_* \le a\|\mathbf{g}_t - \tilde{\mathbf{g}}_t^*\|_*.$$

Moreover, by Orabona (2019, proof of Thm. 7.28) and the fact that $\bar{\mathbf{w}}_t$ minimizes $F_{t+1}(\cdot, \lambda_t)$ over $\mathbf{W}$,

$$F_{t+1}(\mathbf{w}_t^*, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) \le \tfrac{\|\mathbf{g}_t - \tilde{\mathbf{g}}_t^*\|_*^2}{2\lambda_t}.$$

Our collective bounds establish that

$$\delta_t(\tilde{\mathbf{g}}_t^*) \triangleq F_{t+1}(\mathbf{w}_t^*, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) + \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{w}_t^* \rangle$$
$$\le \min(\tfrac{1}{2\lambda_t}\|\mathbf{g}_t - \tilde{\mathbf{g}}_t^*\|_*^2, a\|\mathbf{g}_t - \tilde{\mathbf{g}}_t^*\|_*) + \min(\tfrac{1}{\lambda_t}\|\mathbf{g}_t\|_* \|\tilde{\mathbf{g}}_t - \tilde{\mathbf{g}}_t^*\|_*, a\|\mathbf{g}_t\|_*)$$
$$\le \tfrac{1}{2\lambda_t}\|\mathbf{g}_t - \tilde{\mathbf{g}}_t^*\|_*^2 + \tfrac{1}{\lambda_t}\|\mathbf{g}_t\|_* \|\tilde{\mathbf{g}}_t - \tilde{\mathbf{g}}_t^*\|_*.$$

To obtain an interpretable bound on regret, we will minimize the final expression over all convex combinations $\tilde{\mathbf{g}}_t^*$ of $\mathbf{g}_t$ and $\tilde{\mathbf{g}}_t$. The optimal choice is given by

$$\begin{aligned}
&\hat{\mathbf{g}}_t = \mathbf{g}_t + c_*(\tilde{\mathbf{g}}_t - \mathbf{g}_t) \quad \text{for} \\
&c_* \triangleq \min\left(\tfrac{\|\mathbf{g}_t\|_*}{\|\tilde{\mathbf{g}}_t - \mathbf{g}_t\|_*}, 1\right) = \operatorname*{argmin}_{c \le 1, \tilde{\mathbf{g}}_t^* = \mathbf{g}_t + c(\tilde{\mathbf{g}}_t - \mathbf{g}_t)} \tfrac{1}{2\lambda_t}\|\mathbf{g}_t - \tilde{\mathbf{g}}_t^*\|_*^2 + \tfrac{1}{\lambda_t}\|\mathbf{g}_t\|_*\|\tilde{\mathbf{g}}_t - \tilde{\mathbf{g}}_t^*\|_* \\
&\quad = \operatorname{argmin}_{c \le 1} \tfrac{c^2}{2\lambda_t}\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*^2 + \tfrac{1-c}{\lambda_t}\|\mathbf{g}_t\|_*\|\tilde{\mathbf{g}}_t - \mathbf{g}_t\|_*.
\end{aligned}$$

For this choice, we obtain the bound

$$\begin{aligned}
(\delta_t(\hat{\mathbf{g}}_t))_+ &\le \tfrac{1}{2\lambda_t}\|\mathbf{g}_t - \hat{\mathbf{g}}_t\|_*^2 + \tfrac{1}{\lambda_t}\|\mathbf{g}_t\|_*\|\hat{\mathbf{g}}_t - \tilde{\mathbf{g}}_t\|_* \\
&= \tfrac{c_*^2}{2\lambda_t}\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*^2 + \tfrac{1-c_*}{\lambda_t}\|\mathbf{g}_t\|_*\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_* \\
&= \tfrac{1}{2\lambda_t}\min(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t\|_*)^2 + \tfrac{1}{\lambda_t}\|\mathbf{g}_t\|_*(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_* - \|\mathbf{g}_t\|_*)_+ \\
&= \tfrac{1}{2\lambda_t}(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*^2 - (\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_* - \|\mathbf{g}_t\|_*)_+^2) \\
&= \tfrac{1}{\lambda_t}\text{huber}(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t\|_*)
\end{aligned}$$

and therefore

$$\delta_t = \min(\delta_t(\tilde{\mathbf{g}}_t), \delta_t(\mathbf{g}_t), \delta_t(\hat{\mathbf{g}}_t))_+ \le \min(\tfrac{1}{\lambda_t}\text{huber}(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t\|_*), a\min(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t\|_*)). \tag{4}$$

Since $\tilde{\mathbf{g}}_t^*$ is arbitrary, the advertised regret bounds follow as

$$\begin{aligned}
\text{Regret}_T(\mathbf{u}) &\le \inf_{\tilde{\mathbf{g}}_1^*, \dots, \tilde{\mathbf{g}}_T^* \in \mathbb{R}^d} \lambda_{T+1}\psi(\mathbf{u}) + \textstyle\sum_{t=1}^T \delta_t(\tilde{\mathbf{g}}_t^*) \\
&= \lambda_{T+1}\psi(\mathbf{u}) + \textstyle\sum_{t=1}^T \inf_{\tilde{\mathbf{g}}_t^* \in \mathbb{R}^d} \delta_t(\tilde{\mathbf{g}}_t^*) \\
&\le \lambda_{T+1}\psi(\mathbf{u}) + \textstyle\sum_{t=1}^T \min(\delta_t(\tilde{\mathbf{g}}_t), \delta_t(\mathbf{g}_t), \delta_t(\hat{\mathbf{g}}_t))_+.
\end{aligned}$$

□

B.1. Proof of Lem. 15: OAFTRL difference bound

Fix any time t, and define the optimistic objective function $\tilde{F}_t(\mathbf{w}) = \lambda_t\psi(\mathbf{w}) + \sum_{i=1}^{t-1}\langle \mathbf{g}_i, \mathbf{w}\rangle + \langle \tilde{\mathbf{g}}_t, \mathbf{w}\rangle$ and the auxiliary optimistic objective function $\tilde{F}_t^*(\mathbf{w}) = \lambda_t\psi(\mathbf{w}) + \sum_{i=1}^{t-1}\langle \mathbf{g}_i, \mathbf{w}\rangle + \langle \tilde{\mathbf{g}}_t^*, \mathbf{w}\rangle$ so that $\mathbf{w}_t \in \operatorname{argmin}_{\mathbf{w}\in\mathbf{W}} \tilde{F}_t(\mathbf{w})$ and $\mathbf{w}_t^* \in \operatorname{argmin}_{\mathbf{w}\in\mathbf{W}} \tilde{F}_t^*(\mathbf{w})$. We have

$$\begin{aligned}
\tilde{F}_t^*(\mathbf{w}_t) - \tilde{F}_t^*(\mathbf{w}_t^*) &\ge \tfrac{\lambda_t}{2}\|\mathbf{w}_t - \mathbf{w}_t^*\|^2 \quad \text{by the strong convexity of } \tilde{F}_t^* \text{ and} \\
\tilde{F}_t(\mathbf{w}_t^*) - \tilde{F}_t(\mathbf{w}_t) &\ge \tfrac{\lambda_t}{2}\|\mathbf{w}_t - \mathbf{w}_t^*\|^2 \quad \text{by the strong convexity of } \tilde{F}_t.
\end{aligned}$$

Summing the above inequalities and applying the Fenchel-Young inequality for dual norms, we obtain

$$\lambda_t\|\mathbf{w}_t - \mathbf{w}_t^*\|^2 \le \langle \tilde{\mathbf{g}}_t^* - \tilde{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{w}_t^*\rangle \le \|\tilde{\mathbf{g}}_t - \tilde{\mathbf{g}}_t^*\|_*\|\mathbf{w}_t - \mathbf{w}_t^*\|,$$

which yields the first half of our target bound after rearrangement. The second half follows from the definition of diameter, as $\|\mathbf{w}_t - \mathbf{w}_t^*\| \le \text{diam}(\mathbf{W})$.

C. Proof of Thm. 4: SOOMD regret

We will prove the following more general result for adaptive SOOMD (ASOOMD)

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{g}_t + \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t, \mathbf{w}\rangle + \lambda_{t+1}\mathcal{B}_\psi(\mathbf{w}, \mathbf{w}_t) \quad \text{with arbitrary} \quad \mathbf{w}_0 \quad \text{and} \quad \mathbf{g}_0 = \tilde{\mathbf{g}}_0 = \mathbf{0} \tag{ASOOMD}$$

from which Thm. 4 will follow with the choice $\lambda_t = \lambda$ for all $t \ge 1$.

Theorem 16 (ASOOMD regret). *Fix any $\lambda_{T+1} \geq 0$. If each $(\lambda_{t+1} - \lambda_t)\psi$ is proper and differentiable, $\lambda_0 \triangleq 0$, and $\tilde{\mathbf{g}}_{T+1} \triangleq \mathbf{0}$, then, for all $\mathbf{u} \in \mathbf{W}$, the* ASOOMD *iterates $\mathbf{w}_t$ satisfy*

$$\begin{aligned}\text{Regret}_T(\mathbf{u}) \leq \textstyle\sum_{t=0}^{T}(\lambda_{t+1} - \lambda_t)\mathcal{B}_\psi(\mathbf{u}, \mathbf{w}_t) + \\ \textstyle\sum_{t=1}^{T} \min\left(\text{diam}(\mathbf{W})\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \frac{1}{\lambda_{t+1}}\text{huber}(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t + \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t\|_*)\right).\end{aligned}$$

Proof. Fix any $\mathbf{u} \in \mathbf{W}$, instantiate the notation of Joulani et al. (2017, Sec. 7.2), and consider the choices

- $r_1 = \lambda_2\psi$, $r_t = (\lambda_{t+1} - \lambda_t)\psi$ for $t \geq 2$, so that $r_{1:t} = \lambda_{t+1}\psi$ for $t \geq 1$,
- $q_t = \tilde{q}_t + \langle \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t, \cdot\rangle$ for $t \geq 0$,
- $\tilde{q}_0(\mathbf{w}) = \lambda_1\mathcal{B}_\psi(\mathbf{w}, \mathbf{w}_0)$ and $\tilde{q}_t \equiv 0$ for all $t \geq 1$,
- $p_1 \triangleq r_1 - q_0 = r_1 - \tilde{q}_0 - \langle \tilde{\mathbf{g}}_1 - \tilde{\mathbf{g}}_0, \cdot\rangle = \lambda_2\psi - \lambda_1\mathcal{B}_\psi(\cdot, \mathbf{w}_0) - \langle \tilde{\mathbf{g}}_1 - \tilde{\mathbf{g}}_0, \cdot\rangle$,
- $p_t \triangleq r_t - q_{t-1} = r_t - \tilde{q}_{t-1} - \langle \tilde{\mathbf{g}}_t - \tilde{\mathbf{g}}_{t-1}, \cdot\rangle = (\lambda_{t+1} - \lambda_t)\psi - \langle \tilde{\mathbf{g}}_t - \tilde{\mathbf{g}}_{t-1}, \cdot\rangle$ for all $t \geq 2$.

Since, for each t, $\delta_t = 0$ and ℓ_t is convex, the ADA-MD regret inequality of Joulani et al. (2017, Eq. (24)) and the choice $\tilde{\mathbf{g}}_{T+1} = 0$ imply that

$$\begin{aligned}
\text{Regret}_T(\mathbf{u}) &= \sum_{t=1}^{T}\ell_t(\mathbf{w}_t) - \sum_{t=1}^{T}\ell_t(\mathbf{u}) \\
&\leq -\sum_{t=1}^{T}\mathcal{B}_{\ell_t}(\mathbf{u}, \mathbf{w}_t) + \sum_{t=0}^{T} q_t(\mathbf{u}) - q_t(\mathbf{w}_{t+1}) + \sum_{t=1}^{T}\mathcal{B}_{p_t}(\mathbf{u}, \mathbf{w}_t) \\
&\quad - \sum_{t=1}^{T}\mathcal{B}_{r_{1:t}}(\mathbf{w}_{t+1}, \mathbf{w}_t) + \sum_{t=1}^{T}\langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{w}_{t+1}\rangle + \sum_{t=1}^{T}\delta_t \\
&\leq \lambda_1(\mathcal{B}_\psi(\mathbf{u}, \mathbf{w}_0) - \mathcal{B}_\psi(\mathbf{w}_1, \mathbf{w}_0)) + \sum_{t=0}^{T}\langle \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t, \mathbf{u} - \mathbf{w}_{t+1}\rangle \\
&\quad + \sum_{t=1}^{T}(\lambda_{t+1} - \lambda_t)\mathcal{B}_\psi(\mathbf{u}, \mathbf{w}_t) + \sum_{t=1}^{T}\langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{w}_{t+1}\rangle - \lambda_{t+1}\mathcal{B}_\psi(\mathbf{w}_{t+1}, \mathbf{w}_t) \\
&= \sum_{t=0}^{T}(\lambda_{t+1} - \lambda_t)\mathcal{B}_\psi(\mathbf{u}, \mathbf{w}_t) + \sum_{t=0}^{T}\langle \mathbf{g}_t - \tilde{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{w}_{t+1}\rangle - \lambda_{t+1}\mathcal{B}_\psi(\mathbf{w}_{t+1}, \mathbf{w}_t).
\end{aligned} \tag{5}$$

To obtain our advertised bound, we begin with the expression (5) and invoke the 1-strong convexity of ψ and the nonnegativity of $\mathcal{B}_{\lambda\psi}(\mathbf{w}_1, \mathbf{w}_0)$ to find

$$\begin{aligned}
\text{Regret}_T(\mathbf{u}) &\leq \textstyle\sum_{t=0}^{T}(\lambda_{t+1} - \lambda_t)\mathcal{B}_\psi(\mathbf{u}, \mathbf{w}_t) + \sum_{t=0}^{T}\langle \mathbf{g}_t - \tilde{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{w}_{t+1}\rangle - \lambda_{t+1}\mathcal{B}_\psi(\mathbf{w}_{t+1}, \mathbf{w}_t) \\
&\leq \textstyle\sum_{t=0}^{T}(\lambda_{t+1} - \lambda_t)\mathcal{B}_\psi(\mathbf{u}, \mathbf{w}_t) + \sum_{t=1}^{T}\langle \mathbf{g}_t - \tilde{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{w}_{t+1}\rangle - \frac{\lambda_{t+1}}{2}\|\mathbf{w}_t - \mathbf{w}_{t+1}\|^2.
\end{aligned} \tag{6}$$

We will bound the final sum in this expression using two lemmas. The first is a bound on the difference between subsequent ASOOMD iterates distilled from Joulani et al. (2016, proof of Prop. 2).

Lemma 17 (ASOOMD iterate bound (Joulani et al., 2016, proof of Prop. 2)). *If ψ is differentiable and 1-strongly convex with respect to $\|\cdot\|$, then the* ASOOMD *iterates satisfy*

$$\|\mathbf{w}_t - \mathbf{w}_{t+1}\| \leq \frac{1}{\lambda_{t+1}}\|\mathbf{g}_t + \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t\|_*.$$

The second, proved in App. C.1, is a general bound on $\langle \mathbf{g}, \mathbf{v}\rangle - \frac{\lambda}{2}\|\mathbf{v}\|^2$ under a norm constraint on $\mathbf{v}$.

Lemma 18 (Norm-constrained conjugate). *For any* $\mathbf{g} \in \mathbb{R}^d$ *and* $\lambda, c, b > 0$,

$$\sup_{\mathbf{v}\in\mathbb{R}^d:\|\mathbf{v}\|\leq\min(\frac{c}{\lambda},b)} \langle \mathbf{g}, \mathbf{v}\rangle - \frac{\lambda}{2}\|\mathbf{v}\|^2 = \frac{1}{\lambda}\min(\|\mathbf{g}\|_*, c, b\lambda)(\|\mathbf{g}\|_* - \frac{1}{2}\min(\|\mathbf{g}\|_*, c, b\lambda))$$
$$\leq \min(b\|\mathbf{g}\|_*, \frac{1}{\lambda}\min(\|\mathbf{g}\|_*, c)(\|\mathbf{g}\|_* - \frac{1}{2}\min(\|\mathbf{g}\|_*, c)))$$
$$= \min(b\|\mathbf{g}\|_*, \frac{1}{2\lambda}(\|\mathbf{g}\|_*^2 - (\|\mathbf{g}\|_* - \min(\|\mathbf{g}\|_*, c))^2))$$
$$= \min(b\|\mathbf{g}\|_*, \frac{1}{2\lambda}(\|\mathbf{g}\|_*^2 - (\|\mathbf{g}\|_* - c)_+^2))$$
$$\leq \min(\frac{1}{2\lambda}\|\mathbf{g}\|_*^2, \frac{1}{\lambda}c\|\mathbf{g}\|_*, b\|\mathbf{g}\|_*).$$

By Lems. 17 and 18 and the definition of $a \triangleq \operatorname{diam}(\mathbf{W})$, each summand in our regret bound (6) satisfies

$$\langle \mathbf{g}_t - \tilde{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{w}_{t+1}\rangle - \frac{\lambda_{t+1}}{2}\|\mathbf{w}_t - \mathbf{w}_{t+1}\|^2 \leq \sup_{\mathbf{v}\in\mathbb{R}^d:\|\mathbf{v}\|\leq\min(\frac{1}{\lambda_{t+1}}\|\mathbf{g}_t+\tilde{\mathbf{g}}_{t+1}-\tilde{\mathbf{g}}_t\|_*, a)} \langle \mathbf{g}_t - \tilde{\mathbf{g}}_t, \mathbf{v}\rangle - \frac{\lambda_{t+1}}{2}\|\mathbf{v}\|^2$$
$$= \min\left(a\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \frac{1}{2\lambda_{t+1}}(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*^2 - (\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_* - \|\mathbf{g}_t + \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t\|_*)_+^2)\right)$$

yielding the advertised result. □

C.1. Proof of Lem. 18: Norm-constrained conjugate

By the definition of the dual norm,

$$\sup_{\mathbf{v}\in\mathbb{R}^d:\|\mathbf{v}\|\leq\min(\frac{c}{\lambda},b)} \langle \mathbf{g}, \mathbf{v}\rangle - \frac{\lambda}{2}\|\mathbf{v}\|^2 = \sup_{a\leq\min(\frac{c}{\lambda},b)} \sup_{\mathbf{v}\in\mathbb{R}^d:\|\mathbf{v}\|\leq a} \langle \mathbf{g}, \mathbf{v}\rangle - \frac{\lambda}{2}a^2 = \sup_{a\leq\min(\frac{c}{\lambda},b)} a\|\mathbf{g}\|_* - \frac{\lambda}{2}a^2$$
$$= \frac{1}{\lambda}\min(\|\mathbf{g}\|_*, c, b\lambda)(\|\mathbf{g}\|_* - \frac{1}{2}\min(\|\mathbf{g}\|_*, c, b\lambda)) \leq \min(\frac{1}{\lambda}c\|\mathbf{g}\|_*, b\|\mathbf{g}\|_*).$$

We compare to the values of less constrained optimization problems to obtain the final inequalities:

$$\sup_{a\leq\min(\frac{c}{\lambda},b)} a\|\mathbf{g}\|_* - \frac{\lambda}{2}a^2 \leq \sup_{a\leq\frac{c}{\lambda}} a\|\mathbf{g}\|_* - \frac{\lambda}{2}a^2 = \frac{1}{\lambda}\min(\|\mathbf{g}\|_*, c)(\|\mathbf{g}\|_* - \frac{1}{2}\min(\|\mathbf{g}\|_*, c))$$
$$\leq \sup_{a>0} a\|\mathbf{g}\|_* - \frac{\lambda}{2}a^2 = \frac{1}{\lambda}\frac{1}{2}\|\mathbf{g}\|_*^2.$$

D. Proof of Lem. 8: DORM is ODAFTRL and DORM + is DOOMD

Our derivations will make use of several facts about ℓ^p norms, summarized in the next lemma.

Lemma 19 (ℓ^p norm facts). *For* $p \in (1, \infty)$, $\psi(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|_p^2$, *and any vectors* $\mathbf{w}, \mathbf{v} \in \mathbb{R}^d$ *and* $\tilde{\mathbf{w}}_0 \in \mathbb{R}_+^d$,

$$\nabla\psi(\mathbf{w}) = \nabla\frac{1}{2}\|\mathbf{w}\|_p^2 = \operatorname{sign}(\mathbf{w})|\mathbf{w}|^{p-1}/\|\mathbf{w}\|_p^{p-2} \tag{7}$$
$$\langle \mathbf{w}, \nabla\psi(\mathbf{w})\rangle = \|\mathbf{w}\|_p^2 = 2\psi(\mathbf{w})$$
$$\psi^*(\mathbf{v}) = \sup_{\mathbf{w}\in\mathbb{R}^d} \langle \mathbf{w}, \mathbf{v}\rangle - \psi(\mathbf{w}) = \frac{1}{2}\|\mathbf{v}\|_q^2 \quad \textit{for} \quad 1/q = 1 - 1/p \tag{8}$$
$$\nabla\psi^*(\mathbf{v}) = \operatorname{sign}(\mathbf{v})|\mathbf{v}|^{q-1}/\|\mathbf{v}\|_q^{q-2}$$
$$\psi_+^*(\mathbf{v}) = \sup_{\mathbf{w}\in\mathbb{R}_+^d} \langle \mathbf{w}, \mathbf{v}\rangle - \psi(\mathbf{w}) = \sup_{\mathbf{w}\in\mathbb{R}^d} \langle \mathbf{w}, (\mathbf{v})_+\rangle - \psi(\mathbf{w}) = \frac{1}{2}\|(\mathbf{v})_+\|_q^2$$
$$\nabla\psi_+^*(\mathbf{v}) = \operatorname*{argmax}_{\mathbf{w}\in\mathbb{R}_+^d}\langle \mathbf{w}, \mathbf{v}\rangle - \psi(\mathbf{w}) = \operatorname*{argmin}_{\mathbf{w}\in\mathbb{R}_+^d} \psi(\mathbf{w}) - \langle \mathbf{w}, \mathbf{v}\rangle = (\mathbf{v})_+^{q-1}/\|(\mathbf{v})_+\|_q^{q-2} \tag{9}$$
$$\min_{\tilde{\mathbf{w}}\in\mathbb{R}_+^d} \mathcal{B}_{\lambda\psi}(\tilde{\mathbf{w}}, \tilde{\mathbf{w}}_0) - \langle \mathbf{v}, \tilde{\mathbf{w}}\rangle = \lambda(\langle \tilde{\mathbf{w}}_0, \nabla\psi(\tilde{\mathbf{w}}_0)\rangle - \psi(\tilde{\mathbf{w}}_0) - \sup_{\tilde{\mathbf{w}}\in\mathbb{R}_+^d} \langle \tilde{\mathbf{w}}, \nabla\psi(\tilde{\mathbf{w}}_0) + \mathbf{v}/\lambda\rangle - \psi(\tilde{\mathbf{w}}))$$
$$= \lambda(\langle \tilde{\mathbf{w}}_0, \nabla\psi(\tilde{\mathbf{w}}_0)\rangle - \psi(\tilde{\mathbf{w}}_0) - \psi_+^*(\nabla\psi(\tilde{\mathbf{w}}_0) + \mathbf{v}/\lambda))$$
$$= \lambda(\psi(\tilde{\mathbf{w}}_0) - \psi_+^*(\nabla\psi(\tilde{\mathbf{w}}_0) + \mathbf{v}/\lambda))$$
$$= \lambda(\psi(\tilde{\mathbf{w}}_0) - \frac{1}{2}\|(\nabla\psi(\tilde{\mathbf{w}}_0) + \mathbf{v}/\lambda)_+\|_q^2)$$
$$= \lambda(\frac{1}{2}\|\tilde{\mathbf{w}}_0\|_p^2 - \frac{1}{2}\|(\tilde{\mathbf{w}}_0^{p-1}/\|\tilde{\mathbf{w}}_0\|_p^{p-2} + \mathbf{v}/\lambda)_+\|_q^2).$$

Proof. The fact (7) follows from the chain rule as

$$\begin{aligned}\nabla_j \tfrac{1}{2}\|\mathbf{w}\|_p^2 &= \tfrac{1}{2}\nabla_j(\|\mathbf{w}\|_p^p)^{2/p} = \tfrac{1}{p}(\|\mathbf{w}\|_p^p)^{(2/p)-1}\nabla_j\|\mathbf{w}\|_p^p = \tfrac{1}{p}\|\mathbf{w}\|_p^{2-p}\nabla_j \textstyle\sum_{j'=1}^d |\mathbf{w}_{j'}|^p \\ &= \tfrac{1}{p}\|\mathbf{w}\|_p^{2-p} p \operatorname{sign}(\mathbf{w}_j)|\mathbf{w}_j|^{p-1} = \operatorname{sign}(\mathbf{w}_j)|\mathbf{w}_j|^{p-1}/\|\mathbf{w}\|_p^{p-2}.\end{aligned}$$

The fact (8) follows from Lem. 18 as $\|\cdot\|_q$ is the dual norm of $\|\cdot\|_p$. □

We now prove each claim in turn.

D.1. DORM is ODAFTRL

Fix $p \in (1, 2]$, $\lambda > 0$, and $t \geq 0$. The ODAFTRL iterate with hint $-\mathbf{h}_{t+1}$, $\mathbf{W} \triangleq \mathbb{R}_+^d$, $\psi(\tilde{\mathbf{w}}) = \frac{1}{2}\|\tilde{\mathbf{w}}\|_p^2$, loss subgradients $\mathbf{g}_{1:t-D}^{\text{ODAFTRL}} = -\mathbf{r}_{1:t-D}$, and regularization parameter λ takes the form

$$\begin{aligned}&\operatorname*{argmin}_{\tilde{\mathbf{w}}\in\mathbb{R}_+^d} \lambda\psi(\tilde{\mathbf{w}}) - \langle \tilde{\mathbf{w}}, \mathbf{h}_{t+1} + \mathbf{r}_{1:t-D}\rangle \\ &\quad= \operatorname*{argmin}_{\tilde{\mathbf{w}}\in\mathbb{R}_+^d} \psi(\tilde{\mathbf{w}}) - \langle \tilde{\mathbf{w}}, (\mathbf{h}_{t+1} + \mathbf{r}_{1:t-D})/\lambda\rangle \\ &\quad= ((\mathbf{r}_{1:t-D} + \mathbf{h}_{t+1})/\lambda)_+^{q-1}/\|((\mathbf{r}_{1:t-D} + \mathbf{h}_{t+1})/\lambda)_+\|_q^{q-2} \quad \text{by (9)} \\ &\quad= ((\mathbf{r}_{1:t-D} + \mathbf{h}_{t+1})/\lambda)_+^{q-1}\|((\mathbf{r}_{1:t-D} + \mathbf{h}_{t+1})/\lambda)_+^{q-1}\|_p^{p-2} \quad \text{since } (p-1)(q-1) = 1 \\ &\quad= \tilde{\mathbf{w}}_{t+1}\|\tilde{\mathbf{w}}_{t+1}\|_p^{p-2}\end{aligned}$$

proving the claim.

D.2. DORM+ is DOOMD

Fix $p \in (1, 2]$ and $\lambda > 0$, and let $(\tilde{\mathbf{w}}_t)_{t\geq 0}$ denote the unnormalized iterates generated by DORM+ with hints $\mathbf{h}_t$, instantaneous regrets $\mathbf{r}_t$, regularization parameter λ, and hyperparameter q. For $p = q/(q-1)$, let $(\bar{\mathbf{w}}_t)_{t\geq 0}$ denote the sequence generated by DOOMD with $\bar{\mathbf{w}}_0 = \mathbf{0}$, hints $-\mathbf{h}_t$, $\mathbf{W} \triangleq \mathbb{R}_+^d$, $\psi(\bar{\mathbf{w}}) = \frac{1}{2}\|\bar{\mathbf{w}}\|_p^2$, loss subgradients $\mathbf{g}_t^{\text{DOOMD}} = -\mathbf{r}_t$, and regularization parameter λ. We proceed by induction to show that, for each t, $\bar{\mathbf{w}}_t = \tilde{\mathbf{w}}_t\|\tilde{\mathbf{w}}_t\|_p^{p-2}$.

Base case By assumption, $\bar{\mathbf{w}}_0 = \mathbf{0} = \tilde{\mathbf{w}}_0\|\tilde{\mathbf{w}}_0\|_p^{p-2}$, confirming the base case.

Inductive step Fix any $t \geq 0$ and assume that for each $s \leq t$, $\bar{\mathbf{w}}_s = \tilde{\mathbf{w}}_s\|\tilde{\mathbf{w}}_s\|_p^{p-2}$. Then, by the definition of DOOMD and our ℓ^p norm facts,

$$\begin{aligned}\bar{\mathbf{w}}_{t+1} &= \operatorname*{argmin}_{\bar{\mathbf{w}}\in\mathbb{R}_+^d} \langle -\mathbf{h}_{t+1} + \mathbf{h}_t - \mathbf{r}_{t-D}, \bar{\mathbf{w}}\rangle + \mathcal{B}_{\lambda\psi}(\bar{\mathbf{w}}, \bar{\mathbf{w}}_t) \\ &= \operatorname*{argmin}_{\bar{\mathbf{w}}\in\mathbb{R}_+^d} \lambda(\psi(\bar{\mathbf{w}}) - \psi(\bar{\mathbf{w}}_t) - \langle \bar{\mathbf{w}} - \bar{\mathbf{w}}_t, \nabla\psi(\bar{\mathbf{w}}_t)\rangle) + \langle -\mathbf{h}_{t+1} + \mathbf{h}_t - \mathbf{r}_{t-D}, \bar{\mathbf{w}}\rangle \\ &= \operatorname*{argmin}_{\bar{\mathbf{w}}\in\mathbb{R}_+^d} \psi(\bar{\mathbf{w}}) - \langle \bar{\mathbf{w}}, \nabla\psi(\bar{\mathbf{w}}_t) + (\mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})/\lambda\rangle \\ &= \operatorname*{argmin}_{\bar{\mathbf{w}}\in\mathbb{R}_+^d} \psi(\bar{\mathbf{w}}) - \langle \bar{\mathbf{w}}, \bar{\mathbf{w}}_t^{p-1}/\|\bar{\mathbf{w}}_t\|_p^{p-2} + (\mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})/\lambda\rangle \quad \text{by (7)} \\ &= \operatorname*{argmin}_{\bar{\mathbf{w}}\in\mathbb{R}_+^d} \psi(\bar{\mathbf{w}}) - \langle \bar{\mathbf{w}}, \tilde{\mathbf{w}}_t^{p-1} + (\mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})/\lambda\rangle \quad \text{by the inductive hypothesis} \\ &= (\tilde{\mathbf{w}}_t^{p-1} + (\mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})/\lambda)_+^{q-1}/\|(\tilde{\mathbf{w}}_t^{p-1} + (\mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})/\lambda)_+\|_q^{q-2} \quad \text{by (9)} \\ &= (\tilde{\mathbf{w}}_t^{p-1} + (\mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})/\lambda)_+^{q-1}\|(\tilde{\mathbf{w}}_t^{p-1} + (\mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})/\lambda)_+^{q-1}\|_p^{p-2} \quad \text{since } (p-1)(q-1) = 1 \\ &= \tilde{\mathbf{w}}_{t+1}\|\tilde{\mathbf{w}}_{t+1}\|_p^{p-2},\end{aligned}$$

completing the inductive step.

E. Proof of Lem. 7: DORM and DORM+ are independent of λ

We will prove the following more general result, from which the stated result follows immediately.

Lemma 20 (DORM and DORM+ are independent of λ)**.** *Consider either* DORM *or* DORM+ *plays* $\tilde{\mathbf{w}}_t$ *as a function of* $\lambda > 0$*, and suppose that for all time points* t*, the observed subgradient* $\mathbf{g}_t$ *and chosen hint* $\mathbf{h}_{t+1}$ *only depend on* λ *through* $(\mathbf{w}_s, \lambda^{q-1}\tilde{\mathbf{w}}_s, \mathbf{g}_{s-1}, \mathbf{h}_s)_{s\leq t}$ *and* $(\mathbf{w}_s, \lambda^{q-1}\tilde{\mathbf{w}}_s, \mathbf{g}_s, \mathbf{h}_s)_{s\leq t}$ *respectively. Then if* $\lambda^{q-1}\tilde{\mathbf{w}}_0$ *is independent of the choice of* $\lambda > 0$*, then so is* $\lambda^{q-1}\tilde{\mathbf{w}}_t$ *for all time points* t*. As a result,* $\mathbf{w}_t \propto \lambda^{q-1}\tilde{\mathbf{w}}_t$ *is also independent of the choice of* $\lambda > 0$ *at all time points.*

Proof. We prove each result by induction on t.

E.1. Scaled DORM iterates $\lambda^{q-1}\tilde{\mathbf{w}}_t$ are independent of λ

Base case By assumption, $\mathbf{h}_1$ is independent of the choice of $\lambda > 0$. Hence $\lambda^{q-1}\tilde{\mathbf{w}}_1 = (\mathbf{h}_1)_+^{q-1}$ is independent of $\lambda > 0$, confirming the base case.

Inductive step Fix any $t \geq 0$, suppose $\lambda^{q-1}\tilde{\mathbf{w}}_s$ is independent of the choice of $\lambda > 0$ for all $s \leq t$, and consider

$$\lambda^{q-1}\tilde{\mathbf{w}}_{t+1} = (\mathbf{r}_{1:t-D} + \mathbf{h}_{t+1})_+^{q-1}.$$

Since $\mathbf{r}_{1:t-D}$ depends on λ only through $\mathbf{w}_s$ and $\mathbf{g}_s$ for $s \leq t-D$, our λ dependence assumptions for $(\mathbf{g}_s, \mathbf{h}_{s+1})_{s\leq t}$; the fact that, for each s, $\mathbf{w}_s \propto \lambda^{q-1}\tilde{\mathbf{w}}_s$; and our inductive hypothesis together imply that $\lambda^{q-1}\tilde{\mathbf{w}}_{t+1}$ is independent of $\lambda > 0$.

E.2. Scaled DORM+ iterates $\lambda^{q-1}\tilde{\mathbf{w}}_t$ are independent of λ

Base case By assumption, $\lambda^{q-1}\tilde{\mathbf{w}}_0$ is independent of the choice of $\lambda > 0$, confirming the base case.

Inductive step Fix any $t \geq 0$ and suppose $\lambda^{q-1}\tilde{\mathbf{w}}_s$ is independent of the choice of $\lambda > 0$ for all $s \leq t$. Since $(p-1)(q-1) = 1$,

$$\lambda^{q-1}\tilde{\mathbf{w}}_{t+1} = (\lambda\tilde{\mathbf{w}}_t^{p-1} + \mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})_+^{q-1} = ((\lambda^{q-1}\tilde{\mathbf{w}}_t)^{p-1} + \mathbf{r}_{t-D} - \mathbf{h}_t + \mathbf{h}_{t+1})_+^{q-1}.$$

Since $\mathbf{r}_{t-D}$ depends on λ only through $\mathbf{w}_{t-D}$ and $\mathbf{g}_{t-D}$, our λ dependence assumptions for $(\mathbf{g}_s, \mathbf{h}_{s+1})_{s\leq t}$; the fact that, for each $s \leq t$, $\mathbf{w}_s \propto \lambda^{q-1}\tilde{\mathbf{w}}_s$; and our inductive hypothesis together imply that $\lambda^{q-1}\tilde{\mathbf{w}}_{t+1}$ is independent of $\lambda > 0$. □

F. Proof of Cor. 9: DORM and DORM+ regret

Fix any $\lambda > 0$ and $\mathbf{u} \in \triangle_{d-1}$, consider the unnormalized DORM or DORM+ iterates $\tilde{\mathbf{w}}_t$, and define $\bar{\mathbf{w}}_t = \tilde{\mathbf{w}}_t\|\tilde{\mathbf{w}}_t\|_p^{p-2}$ for each t. For either algorithm, we will bound our regret in terms of the surrogate losses

$$\hat{\ell}_t(\tilde{\mathbf{w}}) \triangleq -\langle \mathbf{r}_t, \tilde{\mathbf{w}}\rangle = \langle \mathbf{g}_t, \tilde{\mathbf{w}}\rangle - \langle \tilde{\mathbf{w}}, \mathbf{1}\rangle\langle \mathbf{g}_t, \mathbf{w}_t\rangle$$

defined for $\tilde{\mathbf{w}} \in \mathbb{R}_+^d$. Since $\hat{\ell}_t(\mathbf{u}) = \langle \mathbf{g}_t, \mathbf{u} - \mathbf{w}_t\rangle$, $\hat{\ell}_t(\bar{\mathbf{w}}_t) = 0$, and each ℓ_t is convex, we have

$$\text{Regret}_T(\mathbf{u}) = \textstyle\sum_{t=1}^T \ell_t(\mathbf{w}_t) - \ell_t(\mathbf{u}) \leq \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{u}\rangle = \sum_{t=1}^T \hat{\ell}_t(\bar{\mathbf{w}}_t) - \hat{\ell}_t(\mathbf{u}).$$

For DORM, Lem. 8 implies that $(\bar{\mathbf{w}}_t)_{t\geq 1}$ are ODFTRL iterates, so the ODFTRL regret bound (Thm. 5) and the fact that ψ is 1-strongly convex with respect to $\|\cdot\| = \sqrt{p-1}\|\cdot\|_p$ (see Shalev-Shwartz, 2007, Lemma 17) with $\|\cdot\|_* = \frac{1}{\sqrt{p-1}}\|\cdot\|_q$ imply

$$\text{Regret}_T(\mathbf{u}) \leq \tfrac{\lambda}{2}\|\mathbf{u}\|_p^2 + \tfrac{1}{\lambda(p-1)}\textstyle\sum_{t=1}^T \mathbf{b}_{t,q}.$$

Similarly, for DORM+, Lem. 8 implies that $(\bar{\mathbf{w}}_t)_{t\geq 0}$ are DOOMD iterates with $\bar{\mathbf{w}}_0 = \mathbf{0}$, so the DOOMD regret bound (Thm. 6) and the strong convexity of ψ yield

$$\text{Regret}_T(\mathbf{u}) \leq \mathcal{B}_{\frac{\lambda}{2}\|\cdot\|_p^2}(\mathbf{u}, \mathbf{0}) + \tfrac{1}{\lambda(p-1)}\textstyle\sum_{t=1}^T \mathbf{b}_{t,q} = \tfrac{\lambda}{2}\|\mathbf{u}\|_p^2 + \tfrac{1}{\lambda(p-1)}\textstyle\sum_{t=1}^T \mathbf{b}_{t,q}.$$

Since, by Lem. 7, the choice of λ does not impact the iterate sequences played by DORM and DORM+, we may take the infimum over $\lambda > 0$ in these regret bounds. The second advertised inequality comes from the identity $\frac{1}{p-1} = q - 1$ and the norm equivalence relations $\|\mathbf{v}\|_q \le d^{1/q}\|\mathbf{v}\|_\infty$ and $\|\mathbf{v}\|_p \le \|\mathbf{v}\|_1 = 1$ for $\mathbf{v} \in \mathbb{R}^d$, as shown in Lem. 21 below. The final claim follows as

$$\inf_{q'\ge 2} d^{2/q'}(q'-1) = \inf_{q'\ge 2} 2^{2\log_2(d)/q'}(q'-1) \le 2^{2\log_2(d)/(2\log_2(d))}(2\log_2(d)-1) = 2(2\log_2(d)-1)$$

since $d > 1$.

Lemma 21 (Equivalence of p-norms). *If $\mathbf{x} \in \mathbb{R}^n$ and $q > q' \ge 1$, then $\|\mathbf{x}\|_q \le \|\mathbf{x}\|_{q'} \le n^{(1/q'-1/q)}\|\mathbf{x}\|_q$.*

Proof. To show $\|\mathbf{x}\|_q \le \|\mathbf{x}\|_{q'}$ for $q > q' \ge 1$, suppose without loss of generality that $\|\mathbf{x}\|_{q'} = 1$. Then, $\|\mathbf{x}\|_q^q = \sum_{i=1}^n |x_i|^q \le \sum_{i=1}^n |x_i|^{q'} = \|\mathbf{x}\|_{q'}^{q'} = 1$. Hence $\|\mathbf{x}\|_q \le 1 = \|\mathbf{x}\|_{q'}$.

For the inequality $\|\mathbf{x}\|_{q'} \le n^{1/q'-1/q}\|\mathbf{x}\|_q$, applying Hölder's inequality yields

$$\|\mathbf{x}\|_{q'}^{q'} = \sum_{i=1}^n 1 \cdot |x_i|^{q'} \le (\sum_{i=1}^n 1)^{1-\frac{q'}{q}} (\sum_{i=1}^n |x_i|^q)^{\frac{q'}{q}} = n^{1-\frac{q'}{q}}\|\mathbf{x}\|_q^{q'},$$

so $\|\mathbf{x}\|_{q'} \le n^{1/q'-1/q}\|\mathbf{x}\|_q$. □

G. Proof of Thm. 10: ODAFTRL regret

Since ODAFTRL is an instance of OAFTRL with $\tilde{\mathbf{g}}_{t+1} = \mathbf{h}_{t+1} - \sum_{s=t-D+1}^t \mathbf{g}_s$, the ODAFTRL result follows immediately from the OAFTRL regret bound, Thm. 14.

H. Proof of Thm. 11: DUB Regret

Fix any $\mathbf{u} \in \mathbf{W}$. By Thm. 10, ODAFTRL admits the regret bound

$$\text{Regret}_T(\mathbf{u}) \le \lambda_T \psi(\mathbf{u}) + \sum_{t=1}^T \min(\tfrac{1}{\lambda_t}\mathbf{b}_{t,F}, \mathbf{a}_{t,F}).$$

To control the second term in this bound, we apply the following lemma proved in App. H.1.

Lemma 22 (DUB-style tuning bound). *Fix any $\alpha > 0$ and any non-negative sequences $(a_t)_{t=1}^T$, $(b_t)_{t=1}^T$. If*

$$\Delta_{t+1}^* \triangleq 2\max_{j\le t-D-1} a_{j-D+1:j} + \sqrt{\textstyle\sum_{i=1}^{t-D} a_i^2 + 2\alpha b_i} \le \alpha\lambda_{t+1} \quad \textit{for each} \quad t$$

then

$$\sum_{t=1}^T \min(b_t^2/\lambda_t, a_t) \le \Delta_{T+D+1}^* \le \alpha\lambda_{T+D+1}.$$

Since $\lambda_T \le \lambda_{T+D+1}$, the result now follows by setting $a_t = \mathbf{a}_{t,F}$ and $b_t = \mathbf{b}_{t,F}$, so that

$$\text{Regret}_T(\mathbf{u}) \le \lambda_T\psi(\mathbf{u}) + \alpha\lambda_{T+D+1} \le (\psi(\mathbf{u}) + \alpha)\lambda_{T+D+1}.$$

H.1. Proof of Lem. 22: DUB-style tuning bound

We prove the claim

$$\Delta_t \triangleq \sum_{i=1}^t \min(b_i/\lambda_i, a_i) \le \Delta_{t+D+1}^* \le \alpha\lambda_{t+D+1}$$

by induction on t.

Base case For $t \in [D+1]$,

$$\sum_{i=1}^t \min(b_i/\lambda_i, a_i) \le a_{1:t-1} + a_t \le 2\max_{j\le t-1} a_{j-D+1:j} + \sqrt{\textstyle\sum_{i=1}^t a_i^2 + 2\alpha b_i} = \Delta_{t+D+1}^* \le \alpha\lambda_{t+D+1}$$

confirming the base case.

Inductive step Now fix any $t+1 \geq D+2$ and suppose that

$$\Delta_i \leq \Delta^*_{i+D+1} \leq \alpha\lambda_{i+D+1}$$

for all $1 \leq i \leq t$. We apply this inductive hypothesis to deduce that, for each $0 \leq i \leq t$,

$$\begin{aligned}
\Delta_{i+1}^2 - \Delta_i^2 &= (\Delta_i + \min(b_{i+1}/\lambda_{i+1}, a_{i+1}))^2 - \Delta_i^2 = 2\Delta_i \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + \min(b_{i+1}/\lambda_{i+1}, a_{i+1})^2 \\
&= 2\Delta_{i-D} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + 2(\Delta_i - \Delta_{i-D}) \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + \min(b_{i+1}/\lambda_{i+1}, a_{i+1})^2 \\
&= 2\Delta_{i-D} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + 2\sum_{j=i-D+1}^{i} \min(b_j/\lambda_j, a_j) \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + \min(b_{i+1}/\lambda_{i+1}, a_{i+1})^2 \\
&\leq 2\alpha\lambda_{i+1} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + 2a_{i-D+1:i} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + a_{i+1}^2 \\
&\leq 2\alpha b_{i+1} + a_{i+1}^2 + 2a_{i-D+1:i} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}).
\end{aligned}$$

Now, we sum this inequality over $i = 0, \ldots, t$, to obtain

$$\begin{aligned}
\Delta_{t+1}^2 &\leq \textstyle\sum_{i=0}^{t}(2\alpha b_{i+1} + a_{i+1}^2) + 2\sum_{i=0}^{t} a_{i-D+1:i} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) \\
&= \textstyle\sum_{i=1}^{t+1}(2\alpha b_i + a_i^2) + 2\sum_{i=1}^{t+1} a_{i-D:i-1} \min(b_i/\lambda_i, a_i) \\
&\leq \textstyle\sum_{i=1}^{t+1}(a_i^2 + 2\alpha b_i) + 2\max_{j\leq t} a_{j-D+1:j} \sum_{i=1}^{t+1} \min(b_i/\lambda_i, a_i) \\
&= \textstyle\sum_{i=1}^{t+1}(a_i^2 + 2\alpha b_i) + 2\Delta_{t+1} \max_{j\leq t} a_{j-D+1:j}.
\end{aligned}$$

Solving this quadratic inequality and applying the triangle inequality, we have

$$\begin{aligned}
\Delta_{t+1} &\leq \max_{j\leq t} a_{j-D+1:j} + \tfrac{1}{2}\sqrt{(2\max_{j\leq t} a_{j-D+1:j})^2 + 4\textstyle\sum_{i=1}^{t+1} a_i^2 + 2\alpha b_i} \\
&\leq 2\max_{j\leq t} a_{j-D+1:j} + \sqrt{\textstyle\sum_{i=1}^{t+1} a_i^2 + 2\alpha b_i} = \Delta^*_{t+D+2} \leq \alpha\lambda_{t+D+2}.
\end{aligned}$$

I. Proof of Thm. 12: AdaHedgeD Regret

Fix any $\mathbf{u} \in \mathbf{W}$. Since the AdaHedgeD regularization sequence $(\lambda_t)_{t\geq 1}$ is non-decreasing, Thm. 14 gives the regret bound

$$\text{Regret}_T(\mathbf{u}) \leq \lambda_T\psi(\mathbf{u}) + \textstyle\sum_{t=1}^{T}\delta_t = \lambda_T\psi(\mathbf{u}) + \alpha\lambda_{T+D+1} \leq (\psi(\mathbf{u}) + \alpha)\lambda_{T+D+1},$$

and the proof of Thm. 14 gives the upper estimate (4):

$$\delta_t \leq \min\left(\frac{\mathbf{b}_{t,F}}{\lambda_t}, \mathbf{a}_{t,F}\right) \quad \text{for all} \quad t \in [T]. \tag{10}$$

Hence, it remains to bound λ_{T+D+1}. Since $\lambda_1 = \cdots = \lambda_{D+1} = 0$ and $\alpha(\lambda_{t+1} - \lambda_t) = \delta_{t-D}$ for $t \geq D+1$,

$$\begin{aligned}
\alpha\lambda_{T+D+1}^2 &= \textstyle\sum_{t=1}^{T+D} \alpha(\lambda_{t+1}^2 - \lambda_t^2) = \sum_{t=D+1}^{T+D}\big(\alpha(\lambda_{t+1} - \lambda_t)^2 + 2\alpha(\lambda_{t+1} - \lambda_t)\lambda_t\big) \\
&= \textstyle\sum_{t=1}^{T}\big(\delta_t^2/\alpha + 2\delta_t\lambda_{t+D}\big) \quad \text{by the definition of } \lambda_{t+1} \\
&= \textstyle\sum_{t=1}^{T}\big(\delta_t^2/\alpha + 2\delta_t\lambda_t + 2\delta_t(\lambda_{t+D} - \lambda_t)\big) \\
&\leq \textstyle\sum_{t=1}^{T}\big(\delta_t^2/\alpha + 2\delta_t\lambda_t + 2\delta_t \max_{t\in[T]}(\lambda_{t+D} - \lambda_t)\big) \\
&= \textstyle\sum_{t=1}^{T}\big(\delta_t^2/\alpha + 2\delta_t\lambda_t\big) + 2\lambda_{T+D+1}\max_{t\in[T]} \delta_{t-D:t-1} \\
&\leq \textstyle\sum_{t=1}^{T}\big(\mathbf{a}_{t,F}^2/\alpha + 2\mathbf{b}_{t,F}\big) + 2\lambda_{T+D+1}\max_{t\in[T]} \mathbf{a}_{t-D:t-1,F} \quad \text{by (10).}
\end{aligned}$$

Solving the above quadratic inequality for λ_{T+D+1} and applying the triangle inequality, we find

$$\begin{aligned}
\alpha\lambda_{T+D+1} &\leq \max_{t\in[T]} \mathbf{a}_{t-D:t-1,F} + \tfrac{1}{2}\sqrt{4(\max_{t\in[T]} \mathbf{a}_{t-D:t-1,F})^2 + 4\textstyle\sum_{t=1}^{T} \mathbf{a}_{t,F}^2 + 2\alpha\mathbf{b}_{t,F}} \\
&\leq 2\max_{t\in[T]} \mathbf{a}_{t-D:t-1,F} + \sqrt{\textstyle\sum_{t=1}^{T} \mathbf{a}_{t,F}^2 + 2\alpha\mathbf{b}_{t,F}}.
\end{aligned}$$

J. Proof of Thm. 13: Learning to hint regret

We begin by bounding the hinting problem regret. Since DORM+ is used for the hinting problem, the following result is an immediate corollary of Cor. 9.

Corollary 23 (DORM+ hinting problem regret)**.** *With convex losses $l_t(\omega) = f_t(H_t\omega)$ and no meta-hints, the* DORM+ *hinting problem iterates ω_t satisfy, for each $v \in \triangle_{m-1}$,*

$$\mathrm{HintRegret}_T(v) \triangleq \sum_{t=1}^T l_t(\omega_t) - \sum_{t=1}^T l_t(v) \le \sqrt{\tfrac{m^{2/q}(q-1)}{2} \sum_{t=1}^T \beta_{t,\infty}} \quad \textit{for}$$

$$\beta_{t,\infty} = \begin{cases} \mathrm{huber}(\|\sum_{s=t-D}^t \rho_s\|_\infty, \|\rho_{t-D}\|_\infty), & \textit{for } t < T \\ \frac{1}{2}\|\sum_{s=t-D}^t \rho_s\|_\infty^2, & \textit{for } t = T \end{cases}$$

$$\textit{where} \quad \rho_t \triangleq \mathbf{1}\langle \gamma_t, \omega_t\rangle - \gamma_t \quad \textit{for} \quad \gamma_t \in \partial l_t(\omega_t) \quad \textit{is the } \text{instantaneous hinting problem regret.}$$

If, in addition, $q = \mathrm{argmin}_{q' \ge 2} m^{2/q'}(q'-1)$, then $\mathrm{HintRegret}_T(v) \le \sqrt{(2\log_2(m)-1)\sum_{t=1}^T \beta_{t,\infty}}$.

Our next lemma, proved in App. J.1, provides an interpretable bound for each $\beta_{t,\infty}$ term in terms of the hinting problem subgradients $(\gamma_t)_{t\ge 1}$.

Lemma 24 (Hinting problem subgradient regret bound)**.** *Under the notation and assumptions of Cor. 23,*

$$\beta_{t,\infty} \le \begin{cases} \mathrm{huber}(\xi_t, \zeta_t) & \textit{if } t < T \\ \frac{1}{2}\xi_t & \textit{if } t = T \end{cases}, \quad \textit{for}$$

$$\xi_t \triangleq 4(D+1)\sum_{s=t-D}^t \|\gamma_s\|_\infty^2 \quad \textit{and}$$

$$\zeta_t \triangleq 4\|\gamma_{t-D}\|_\infty \sum_{s=t-D}^t \|\gamma_s\|_\infty.$$

Now fix any $\mathbf{u} \in \mathbf{W}$. We invoke Assump. 1, Cor. 23, and Lem. 24 in turn to bound the base problem regret

$$\begin{aligned}
\mathrm{Regret}_T(\mathbf{u}) &= \textstyle\sum_{t=1}^T \ell_t(\mathbf{w}_t) - \ell_t(\mathbf{u}) \\
&\le C_0(\mathbf{u}) + C_1(\mathbf{u})\sqrt{\textstyle\sum_{t=1}^T f_t(\mathbf{h}_t(\omega_t))} \quad \text{by Assump. 1} \\
&\le C_0(\mathbf{u}) + C_1(\mathbf{u})\sqrt{\inf_{v\in\mathbf{V}} \textstyle\sum_{t=1}^T f_t(\mathbf{h}_t(v)) + \sqrt{(2\log_2(m)-1)\sum_{t=1}^T \beta_{t,\infty}}} \quad \text{by Cor. 23} \\
&\le C_0(\mathbf{u}) + C_1(\mathbf{u})\sqrt{\inf_{v\in\mathbf{V}} \textstyle\sum_{t=1}^T f_t(\mathbf{h}_t(v)) + \sqrt{(2\log_2(m)-1)(\frac{1}{2}\xi_T + \sum_{t=1}^{T-1} \mathrm{huber}(\xi_t, \zeta_t))}} \quad \text{by Lem. 24.}
\end{aligned}$$

The advertised bound now follows from the triangle inequality.

J.1. Proof of Lem. 24: Hinting problem subgradient regret bound

Fix any $t \in [T]$. The triangle inequality implies that

$$\|\rho_t\|_\infty = \|\gamma_t - \mathbf{1}\langle\omega_t, \gamma_t\rangle\|_\infty \le \|\gamma_t\|_\infty + |\langle\omega_t, \gamma_t\rangle| \le 2\|\gamma_t\|_\infty$$

since $\omega_t \in \triangle_{m-1}$. We repeatedly apply this finding in conjunction with Jensen's inequality to conclude

$$\begin{aligned}
\|\textstyle\sum_{s=t-D}^t \rho_s\|_\infty^2 &\le (D+1)\textstyle\sum_{s=t-D}^t \|\rho_s\|_\infty^2 \le 4(D+1)\sum_{s=t-D}^t \|\gamma_s\|_\infty^2 \quad \text{and} \\
\|\rho_{t-D}\|_\infty \|\textstyle\sum_{s=t-D}^t \rho_s\|_\infty &\le \|\rho_{t-D}\|_\infty \textstyle\sum_{s=t-D}^t \|\rho_s\|_\infty \le 4\|\gamma_{t-D}\|_\infty \sum_{s=t-D}^t \|\gamma_s\|_\infty.
\end{aligned}$$

K. Examples: Learning to Hint with DORM+ and AdaHedgeD

By Thm. 12, AdaHedgeD satisfies Assump. 1 with $f_t(\mathbf{h}_t) = \|\mathbf{r}_t\|_* \|\mathbf{h}_t - \sum_{s=t-D}^t \mathbf{r}_s\|_* \ge \frac{\mathbf{a}_{t,F}^2 + 2\alpha \mathbf{b}_{t,F}}{\mathrm{diam}(\mathbf{W})^2 + 2\alpha}$, $C_1(\mathbf{u}) = \sqrt{\mathrm{diam}(\mathbf{W})^2 + 2\alpha}$, and $C_0(\mathbf{u}) = 2\,\mathrm{diam}(\mathbf{W}) \max_{t\in[T]} \sum_{s=t-D}^{t-1} \|\mathbf{g}_s\|_*$.

By Cor. 9, DORM+ satisfies Assump. 1 with $f_t(\mathbf{h}) = \|\mathbf{r}_{t-D} + \mathbf{h}_{t+1} - \mathbf{h}_t\|_q \|\mathbf{h} - \sum_{s=t-D}^{t} \mathbf{r}_s\|_q$, $C_0(\mathbf{u}) = 0$, and $C_1(\mathbf{u}) = \sqrt{\frac{\|\mathbf{u}\|_p^2}{2(p-1)}}$.

These choices give rise to the hinting losses

$$l_t^{\text{DORM+}}(\omega) = \|\mathbf{r}_{t-D} + \mathbf{h}_{t+1} - \mathbf{h}_t\|_q \|H_t\omega - \textstyle\sum_{s=t-D}^{t} \mathbf{r}_s\|_q \quad \text{and} \tag{11}$$
$$l_t^{\text{AdaHedgeD}}(\omega) = \|\mathbf{g}_t\|_q \|H_t\omega - \textstyle\sum_{s=t-D}^{t} \mathbf{g}_s\|_q \quad \text{when} \quad \|\cdot\|_* = \|\cdot\|_q \quad \text{for} \quad q \in [1, \infty].$$

The following lemma, proved in App. K.1, identifies subgradients of these hinting losses.

Lemma 25 (Hinting loss subgradient)**.** *If $l_t(\omega) = \|\bar{\mathbf{g}}_t\|_q \|H_t\omega - \mathbf{v}_t\|_q$ for some $\bar{\mathbf{g}}_t, \mathbf{v}_t \in \mathbb{R}^d$ and $H_t \in \mathbb{R}^{d\times m}$, then*

$$\gamma_t = \begin{cases} \frac{\|\bar{\mathbf{g}}_t\|_q}{\|H_t\omega - \mathbf{v}_t\|_q^{q-1}} H_t^\top |H_t\omega - \mathbf{v}_t|^{q-1} \operatorname{sign}(H_t\omega - \mathbf{v}_t) & \text{if } q < \infty \\ \|\bar{\mathbf{g}}_t\|_\infty \operatorname{sign}(\mu) H_t^\top \mathbf{e}_k & \text{if } q = \infty \end{cases} \in \quad \partial l_t(\omega) \tag{12}$$

for $k = \operatorname{argmax}_{j\in[d]}(H_t\omega - \mathbf{v}_t)_j$ and $\mu = \max_{j\in[d]}(H_t\omega - \mathbf{v}_t)_j$.

Our next lemma, proved in App. K.2, bounds the ∞-norm of this hinting loss subgradient in terms of the base problem subgradients.

Lemma 26 (Hinting loss subgradient bound)**.** *Under the assumptions and notation of Lem. 25, the subgradient γ_t satisfies $\|\gamma_t\|_\infty \le d^{1/q} \|\bar{\mathbf{g}}_t\|_q \|H_t\|_\infty$ for $\|H_t\|_\infty$ the maximum absolute entry of H_t.*

K.1. Proof of Lem. 25: Hinting loss subgradient

The result follows immediately from the chain rule and the following lemma.

Lemma 27 (Subgradients of p-norms)**.** *Suppose $\mathbf{w} \in \mathbb{R}^d$ and $k \in \operatorname{argmax}_{j\in[d]} |\mathbf{w}_j|$. Then*

$$\partial\|\mathbf{w}\|_p \ni \begin{cases} \frac{|\mathbf{w}|^{p-1}}{\|\mathbf{w}\|_p^{p-1}} \operatorname{sign}(\mathbf{w}) & \text{if } \|\mathbf{w}\|_p \ne 0, p \in [1,\infty) \\ \mathbf{e}_k \operatorname{sign}(\mathbf{w}_k) & \text{if } \|\mathbf{w}\|_p \ne 0, p = \infty \\ \mathbf{0} & \text{if } \|\mathbf{w}\|_p = 0 \end{cases}.$$

Proof. Since $\mathbf{0}$ is a minimizer of $\|\cdot\|_p$, we have $\|\mathbf{u}\|_p \ge \|\mathbf{0}\|_p + \langle \mathbf{0}, \mathbf{u} - \mathbf{0}\rangle$ for any $\mathbf{u} \in \mathbb{R}^d$ and hence $\mathbf{0} \in \partial\|\mathbf{0}\|_p$.

For $p \in [1,\infty)$, by the chain rule, if $\|\mathbf{w}\|_p \ne \mathbf{0}$,

$$\begin{aligned} \partial_j \|\mathbf{w}\|_p &= \partial_j \big(\textstyle\sum_{k=1}^n |\mathbf{w}_k|^p \big)^{1/p} = \frac{1}{p} \big(\textstyle\sum_{k=1}^n |\mathbf{w}_k|^p \big)^{(1/p)-1} p |\mathbf{w}_j|^{p-1} \operatorname{sign}(\mathbf{w}_j) \\ &= \Big(\big(\textstyle\sum_{k=1}^n |\mathbf{w}_k|^p \big)^{1/p} \Big)^{-(p-1)} |\mathbf{w}_j|^{p-1} \operatorname{sign}(\mathbf{w}_j) \\ &= \Big(\frac{|\mathbf{w}_j|}{\|\mathbf{w}\|_p} \Big)^{p-1} \operatorname{sign}(\mathbf{w}_j). \end{aligned}$$

For $p = \infty$, we have that $\|\mathbf{w}\|_\infty = \max_{j\in[n]} |\mathbf{w}_j|$. By the Danskin-Bertsekas Theorem (Danskin, 2012) for subdifferentials, $\partial\|\mathbf{w}\|_\infty = \operatorname{conv}\{\cup \partial|\mathbf{w}_j| \quad \text{s.t.} \quad |\mathbf{w}_j| = \|\mathbf{w}\|_\infty\} = \operatorname{conv}\{\cup \operatorname{sign}(\mathbf{w}_j)\mathbf{e}_j \quad \text{s.t.} \quad |\mathbf{w}_j| = \|\mathbf{w}\|_\infty\}$, where conv is the convex hull operation. □

K.2. Proof of Lem. 26: Hinting loss subgradient bound

If $q \in [1,\infty)$, we have

$$\begin{aligned} \|\gamma_t\|_\infty &= \Big\| \frac{\|\bar{\mathbf{g}}_t\|_q}{\|H_t\omega - \sum_{s=t-D}^{t} \mathbf{g}_s\|_q^{q-1}} H_t^\top |H_t\omega - \textstyle\sum_{s=t-D}^{t} \mathbf{g}_s|^{q-1} \operatorname{sign}(H_t\omega - \textstyle\sum_{s=t-D}^{t} \mathbf{g}_s) \Big\|_\infty \\ &\le \frac{\|\bar{\mathbf{g}}_t\|_q \max_{j\in[d]} \|H_t\mathbf{e}_j\|_q}{\|H_t\omega - \sum_{s=t-D}^{t} \mathbf{g}_s\|_q^{q-1}} \|H_t\omega - \textstyle\sum_{s=t-D}^{t} \mathbf{g}_s\|_q^{q-1} \quad \text{by Hölder's inequality for } (q,p) \\ &\le d^{1/q} \|\bar{\mathbf{g}}_t\|_q \|H_t\|_\infty \quad \text{by Lem. 21.} \end{aligned}$$

If $q = \infty$, we have

$$\|\gamma_t\|_\infty = \left\| \|\bar{\mathbf{g}}_t\|_\infty \operatorname{sign}(\mu) H_t^\top \mathbf{e}_k \right\|_\infty = \mathbb{I}[\mu \neq 0] \|\bar{\mathbf{g}}_t\|_\infty \|H_t\|_\infty \leq d^{1/q} \|\bar{\mathbf{g}}_t\|_\infty \|H_t\|_\infty.$$

L. Experiment Details

L.1. Subseasonal Forecasting Application

We apply the online learning techniques developed in this paper to the problem of adaptive ensembling for subseasonal weather forecasting. Subseasonal forecasting is the problem predicting meteorological variables, often temperature and precipitation, 2-6 weeks in advance. These mid-range forecasts are critical for managing water resources and mitigating wildfires, droughts, floods, and other extreme weather events (Hwang et al., 2019). However, the subseasonal forecasting task is notoriously difficult due to the joint influences of short-term initial conditions and long-term boundary conditions (White et al., 2017).

To improve subseasonal weather forecasting capabilities, the US Department of Reclamation launched the Sub-Seasonal Climate Forecast Rodeo competition (Nowak et al., 2020), a yearlong real-time forecasting competition for the Western United States. Our experiments are based on Flaspohler et al. (2021), a snapshot of public subseasonal model forecasts including both physics-based and machine learning models. These models were developed for the subseasonal forecasting challenge and make semimonthly forecasts for the contest period (19 October 2019 – 29 September 2020).

To expand our evaluation beyond the subseasonal forecasting competition, we used the forecasts in Flaspohler et al. (2021) for analogous yearlong periods (26 semi-monthly dates starting from the last Wednesday in October) beginning in Oct. 2010 and ending in Sep. 2020. Throughout, we refer to the yearlong period beginning in Oct. 2010 – Sep. 2011 as the 2011 year and so on for each subsequent year. For each forecast date t, the models in Flaspohler et al. (2021) were trained only on data available at time t and model hyper-parameters were tuned to optimize average RMSE loss on the 3-year period preceding the forecast date t. For a few of the forecast dates, one or more models had missing forecasts; only dates for which all models have forecasts were used in evaluation.

L.2. Problem Definition

Denote the set of $d = 6$ input models $\{\mathcal{M}_1, \dots \mathcal{M}_d\}$ with labels: `llr` (Model1), `multillr` (Model2), `tuned_catboost` (Model3), `tuned_cfsv2` (Model4), `tuned_doy` (Model5) and `tuned_salient_fri` (Model6). On each semimonthly forecast date, each model $\mathcal{M}_i$ makes a prediction for each of two meteorological variables (cumulative precipitation and average temperature over 14 days) and two forecasting horizons (3-4 weeks and 5-6 weeks). For the 3-4 week and 5-6 horizons respectively, the forecaster experiences a delay of $D = 2$ and $D = 3$ forecasts. Each model makes a total of $T = 26$ semimonthly forecasts for these four tasks.

At each time t, each input model $\mathcal{M}_i$ produces a prediction at $G = 514$ gridpoints in the Western United States: $\mathbf{x}_{t,i}^c \in \mathbb{R}^G = \mathcal{M}_i(t)$ for task c at time t. Let $\mathbf{X}_t^c \in \mathbb{R}^{G \times d}$ be the matrix containing each input model's predictions as columns. The true meterological outcome for task c is $\mathbf{y}_t^c \in \mathbb{R}^G$. As online learning is performed for each task separately, we drop the task superscript c in the following.

At each timestep, the online learner makes a forecast prediction $\hat{\mathbf{y}}_t$ by playing $\mathbf{w}_t \in \mathbf{W} = \triangle_{d-1}$, corresponding to a convex combination of the individual models: $\hat{\mathbf{y}}_t = \mathbf{X}_t \mathbf{w}_t$. The learner then incurs a loss for the play $\mathbf{w}_t$ according to the root mean squared (RMSE) error over the geography of interest:

$$\ell_t(\mathbf{w}_t) = \frac{1}{\sqrt{G}} \|\mathbf{y}_t - \mathbf{X}_t \mathbf{w}_t\|_2,$$

$$\partial \ell_t(\mathbf{w}_t) \ni \mathbf{g}_t = \begin{cases} \frac{\mathbf{X}_t^\top (\mathbf{X}_t \mathbf{w}_t - \mathbf{y}_t)}{\sqrt{G} \|\mathbf{X}_t \mathbf{w}_t - \mathbf{y}_t\|_2} & \text{if} \quad \mathbf{X}_t \mathbf{w}_t - \mathbf{y}_t \neq \mathbf{0} \\ \mathbf{0} & \text{if} \quad \mathbf{X}_t \mathbf{w}_t - \mathbf{y}_t = \mathbf{0} \end{cases}$$

Our objective for the subseasonal forecasting application is to produce an adaptive ensemble forecast that competes with the best input model over the yearlong period. Hence, in our evaluation, we take the competitor set to be the set of individual models $\mathbf{U} = \{\mathbf{e}_i : i \in [d]\}$.

M. Extended Experimental Results

We present complete experimental results for the four experiments presented in the main paper (see Sec. 7).

M.1. Competing with the Best Input Model

Results for our three delayed online learning algorithms — DORM, DORM+, and AdaHedgeD— on the four subseasonal prediction tasks for the four optimism strategies described in Sec. 7 (`recent_g`, `prev_g`, `mean_g`, `none`) are presented below. Each table and figure shows the average RMSE loss and the annual regret versus the best input model in any given year respectively for each algorithm and task.

DORM+ is a competitive model for all three hinting strategies and under the `recent_g` hinting strategy achieves negative regret on all tasks except Temp. 5-6w. For the Temp. 5-6w task, no online learning model outperforms the best input model for any hinting strategy. For the precipitation tasks, the online learning algorithms presented achieve negative regret using all three hinting strategies for all four tasks. Within the subseasonal forecasting domain, precipitation is often considered a more challenging forecasting task than temperature (White et al., 2017). The gap between the best model and the worst model tends to be larger for precipitation than for temperature, and this could in part explain the strength of the online learning algorithms for these tasks.

Table 2: **Hint `recent_g`:** Average RMSE of the 2011-2020 semimonthly forecasts for online learning algorithms (left) and input models (right) over a 10-year evaluation period with the top-performing learners and input models bolded and blue. In each task, the online learners compare favorably with the best input model and learn to downweight the lower-performing candidates, like the worst models italicized in red.

RECENT_G	ADAHEDGED	DORM	DORM+	MODEL1	MODEL2	MODEL3	MODEL4	MODEL5	MODEL6
PRECIP. 3-4W	21.726	21.731	**21.675**	**21.973**	22.431	22.357	21.978	21.986	*23.344*
PRECIP. 5-6W	21.868	21.957	**21.838**	22.030	22.570	22.383	22.004	**21.993**	*23.257*
TEMP. 3-4W	2.273	2.259	**2.247**	**2.253**	2.352	2.394	2.277	2.319	*2.508*
TEMP. 5-6W	2.316	2.316	**2.303**	**2.270**	2.368	2.459	2.278	2.317	*2.569*

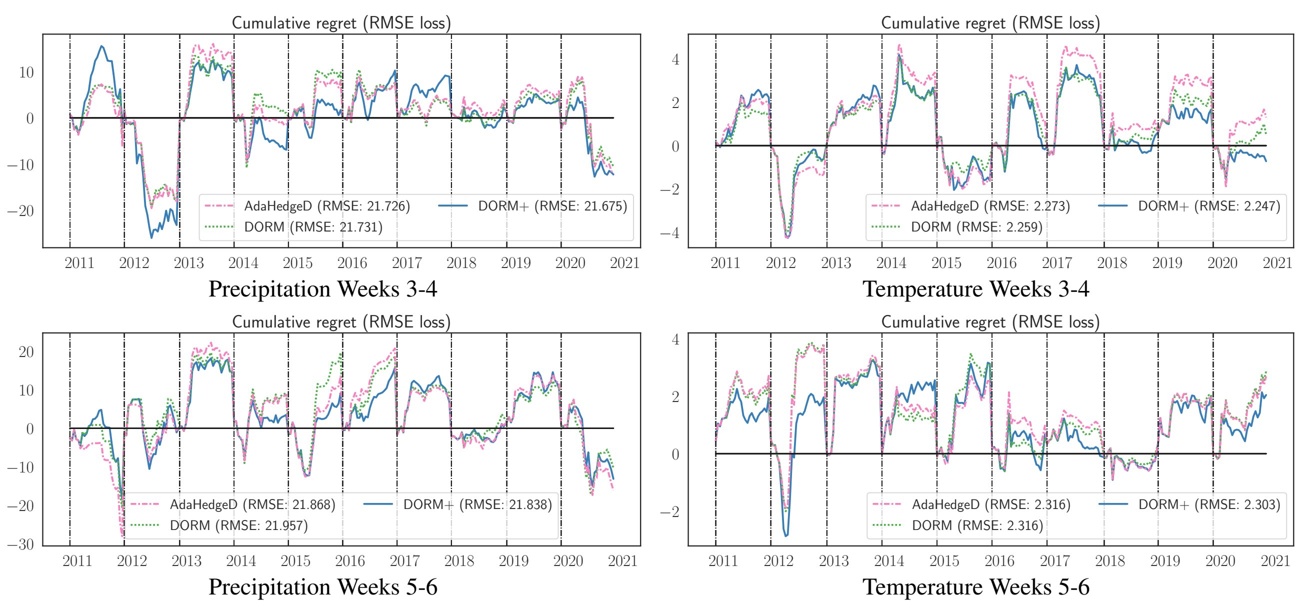


Figure 6: **Hint `recent_g`:** Yearly cumulative regret under RMSE loss for the three delayed online learning algorithms presented, over the 10-year evaluation period. The zero line corresponds to the performance of the best input model in a given year.

Table 3: **Hint prev_g:** Average RMSE of the 2010-2020 semimonthly forecasts for all four tasks over over a 10-year evaluation period.

PREV_G	ADAHEDGED	DORM	DORM+	MODEL1	MODEL2	MODEL3	MODEL4	MODEL5	MODEL6
PRECIP. 3-4W	21.760	21.777	**21.729**	**21.973**	22.431	22.357	21.978	21.986	*23.344*
PRECIP. 5-6W	21.943	21.964	**21.911**	22.030	22.570	22.383	22.004	**21.993**	*23.257*
TEMP. 3-4W	2.266	2.269	**2.250**	**2.253**	2.352	2.394	2.277	2.319	*2.508*
TEMP. 5-6W	2.306	2.307	**2.305**	**2.270**	2.368	2.459	2.278	2.317	*2.569*

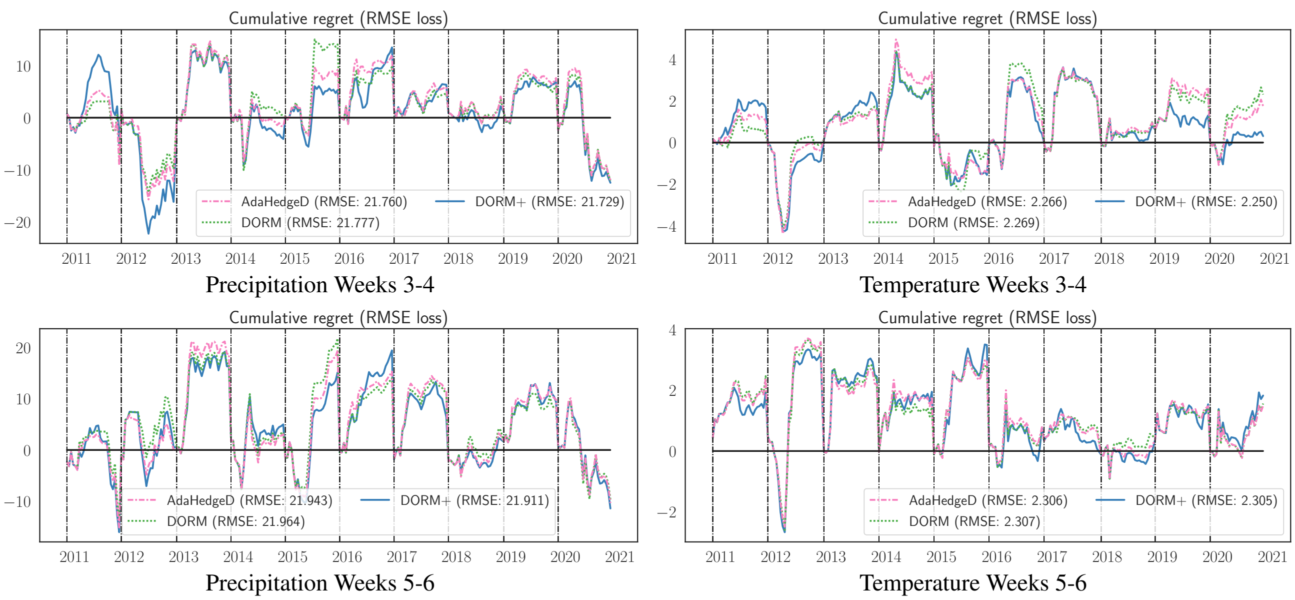


Figure 7: **Hint prev_g:** Yearly cumulative regret under RMSE loss for the three delayed online learning algorithms presented.

Table 4: **Hint mean_g:** Average RMSE of the 2010-2020 semimonthly forecasts for all four tasks over over a 10-year evaluation period.

MEAN_G	ADAHEDGED	DORM	DORM+	MODEL1	MODEL2	MODEL3	MODEL4	MODEL5	MODEL6
PRECIP. 3-4W	21.864	21.945	**21.830**	**21.973**	22.431	22.357	21.978	21.986	*23.344*
PRECIP. 5-6W	21.993	22.054	**21.946**	22.030	22.570	22.383	22.004	**21.993**	*23.257*
TEMP. 3-4W	2.273	2.277	**2.257**	**2.253**	2.352	2.394	2.277	2.319	*2.508*
TEMP. 5-6W	**2.311**	2.320	2.314	**2.270**	2.368	2.459	2.278	2.317	*2.569*

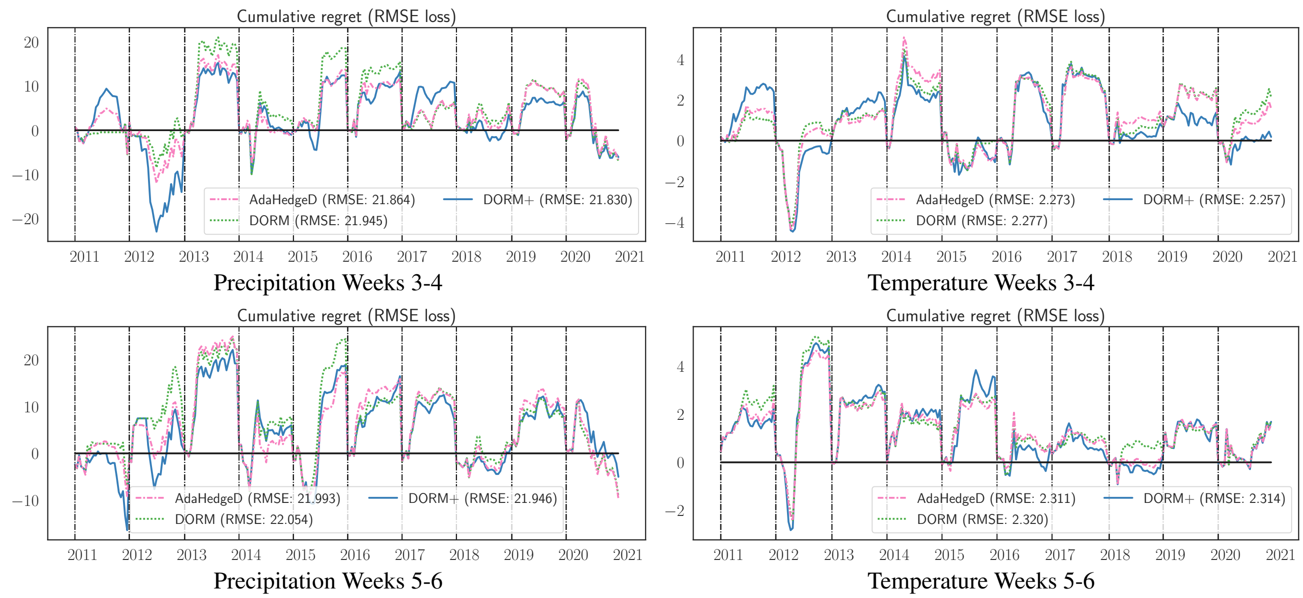


Figure 8: **Hint `mean_g`:** Yearly cumulative regret under RMSE loss for the three delayed online learning algorithms presented.

Table 5: **Hint `none`:** Average RMSE of the 2010-2020 semimonthly forecasts for all four tasks over over a 10-year evaluation period.

NONE	ADAHEDGED	DORM	DORM+	MODEL1	MODEL2	MODEL3	MODEL4	MODEL5	MODEL6
PRECIP. 3-4W	**21.760**	21.835	21.796	**21.973**	22.431	22.357	21.978	21.986	*23.344*
PRECIP. 5-6W	**21.860**	21.967	21.916	22.030	22.570	22.383	22.004	**21.993**	*23.257*
TEMP. 3-4W	2.266	2.272	**2.258**	**2.253**	2.352	2.394	2.277	2.319	*2.508*
TEMP. 5-6W	**2.296**	2.311	2.308	**2.270**	2.368	2.459	2.278	2.317	*2.569*

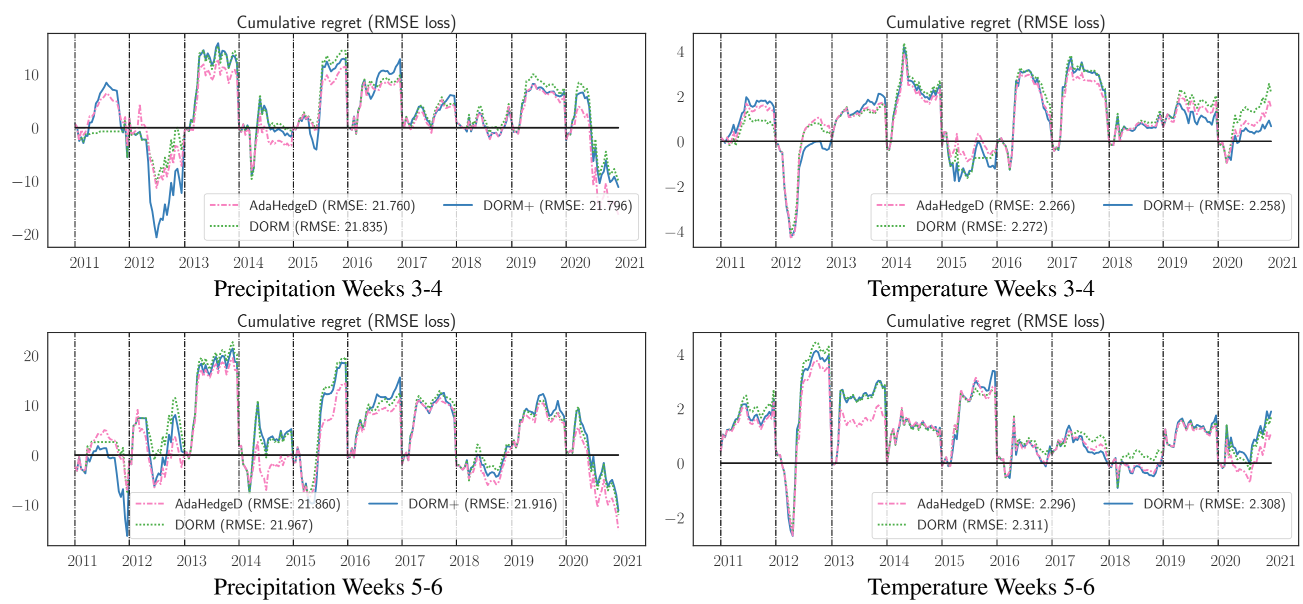


Figure 9: **Hint `none`:** Yearly cumulative regret under RMSE loss for the three delayed online learning algorithms presented.

M.2. Impact of Regularization

Results for three regularization strategies—AdaHedgeD, DORM+, and DUB—on all four subseasonal prediction as described in Sec. 7. Fig. 10 shows the annual regret versus the best input model in any given year for each algorithm and task, and Fig. 11 presents an example of the weights played by each algorithm in the final evaluation year, as well as the regularization weight used by each algorithm.

The under- and over-regularization of AdaHedgeD and DUB respectively compared with DORM+ is evident in all four tasks, both in the regret and weight plots. Due to the looseness of the regularization settings used in DUB, its plays can be seen to be very close to the uniform ensemble in all four tasks. For this subseasonal prediction problem, the uniform ensemble is competitive, especially for the 5-6 week horizons. However, in problems where the uniform ensemble has higher regret, this over-regularization property of DUB would be undesirable. The more adaptive plays of DORM+ and AdaHedgeD have the potential to better exploit heterogeneous performance among different input models.

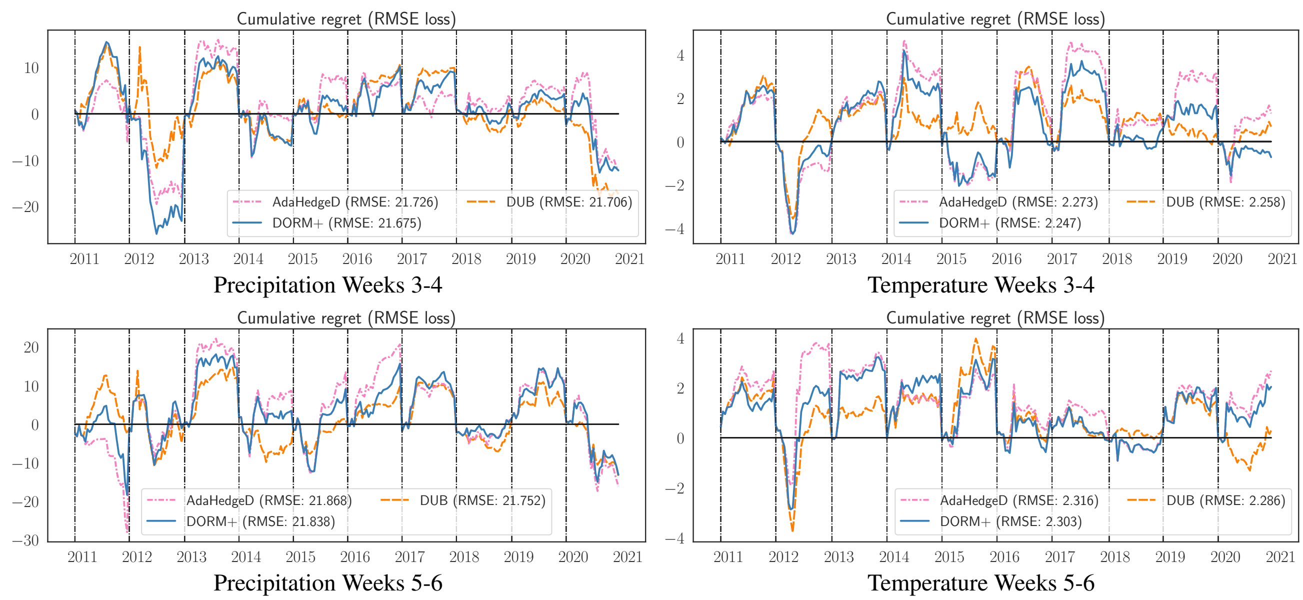
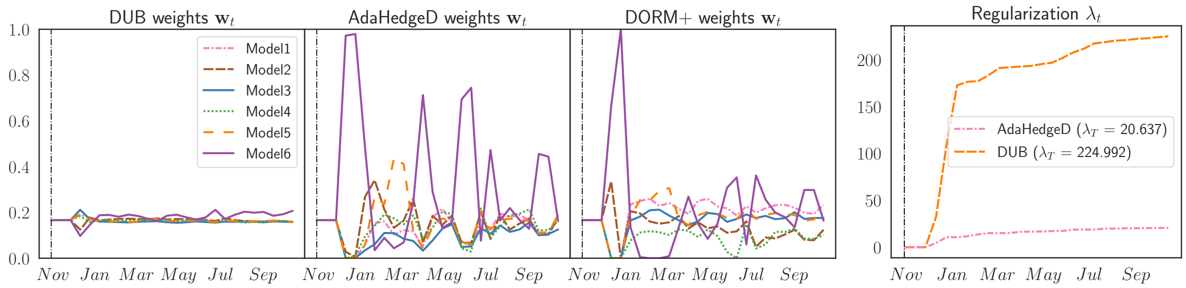
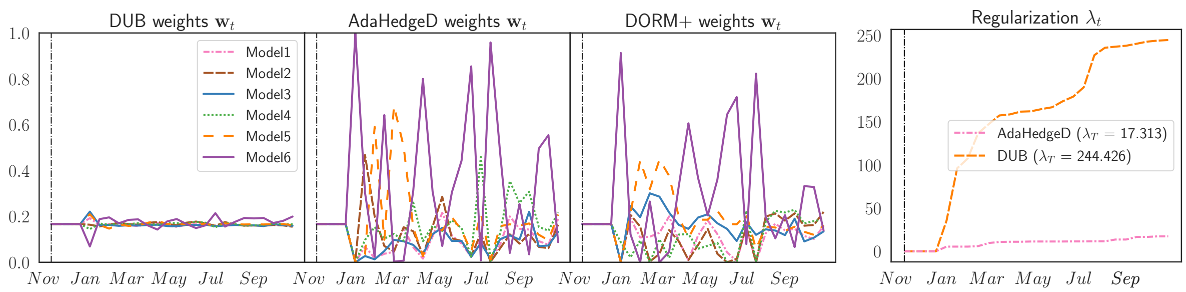


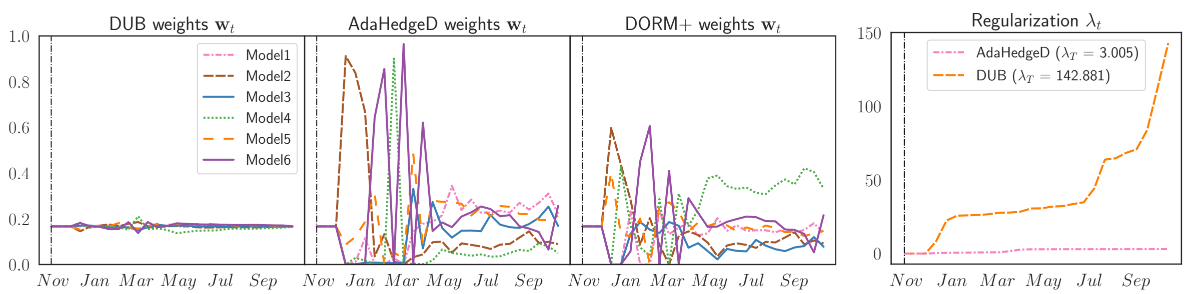
Figure 10: **Overall regret:** Yearly cumulative regret under the RMSE loss for the three regularization algorithms presented.



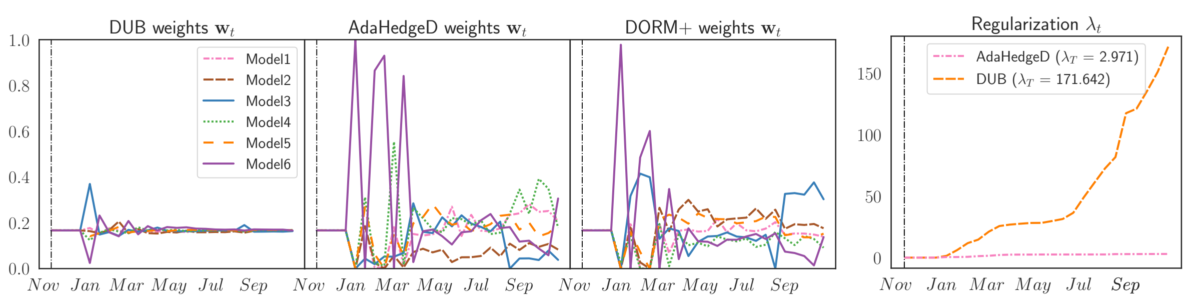
(a) Precipitation Weeks 3-4



(b) Precipitation Weeks 5-6



(c) Temperature Weeks 3-4



(d) Temperature Weeks 5-6

Figure 11: **Impact of regularization:** The plays $\mathbf{w}_t$ of online learning algorithms used to combine the input models for all four tasks in the 2020 evaluation year. The weights of DUB and AdaHedgeD appear respectively over and under regularized compared to DORM+ due to their selection of regularization strength λ_t (right).

M.3. To Replicate or Not to Replicate

We compare the performance of replicated and non-replicated variants of our DORM+ algorithm as in Sec. 7. Both algorithms perform well, but in all tasks, DORM+ outperforms replicated DORM+ (in which $D+1$ independent copies of DORM+ make staggered predictions). Fig. 12 provides an example of the weight plots produced by the replication strategy for all for tasks.

The replicated algorithms only have the opportunity to learn from $T/(D+1)$ plays. For the 3-4 week horizons tasks $D=2$ and for the 5-6 week horizons tasks $D=3$. Because our forecasting horizons are short ($T=26$), further limiting the feedback available to each online learner via replication could be detrimental to practical model performance.

Table 6: **Replication RMSE:** Average RMSE of the 2010-2020 semimonthly forecasts for four tasks over over a 10-year evaluation period for replicated versus standard DORM+.

	DORM+	Replicated DORM+	Model1	Model2	Model3	Model4	Model5	Model6
Precip. 3-4w	**21.675**	21.720	**21.973**	22.431	22.357	21.978	21.986	*23.344*
Precip. 5-6w	**21.838**	21.851	22.030	22.570	22.383	22.004	**21.993**	*23.257*
Temp. 3-4w	**2.247**	2.249	**2.253**	2.352	2.394	2.277	2.319	*2.508*
Temp. 5-6w	**2.303**	2.315	**2.270**	2.368	2.459	2.278	2.317	*2.569*

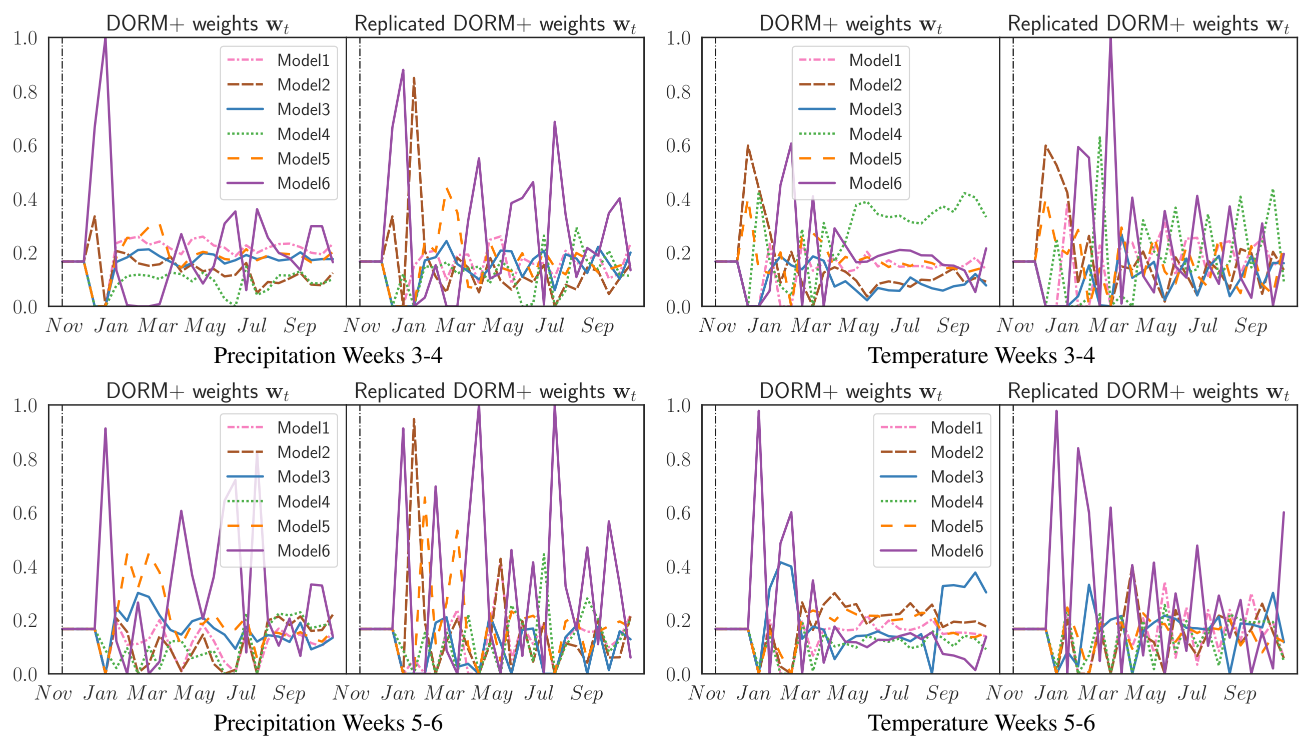


Figure 12: **Replication weights**: The plays $\mathbf{w}_t$ of DORM+ and replicated DORM+ for all four tasks in the final evaluation year.

M.4. Learning to Hint

We examine the effect of optimism on the DORM+ algorithms and the ability of our "learning to hint" strategy to recover the performance of the best optimism strategy in retrospect as described in Sec. 7. We use DORM+ as the meta-algorithm for hint learning to produce the `learned` optimism strategy that plays a convex combination of the three constant hinters.

As reported in the main text, the regret of the base algorithm using the learned hinting strategy generally falls between the worst and the best hinting strategy for any given year. Because the best hinting strategy for any given year is unknown *a priori*, the adaptivity of the hint learner is useful practically. Currently, the hint learner is only optimizing a loose upper bound on base problem regret. Deriving loss functions for hint learning that more accurately quantify the effect of the hinter on base model regret is an important next step in achieving negative regret for online hinting algorithms.

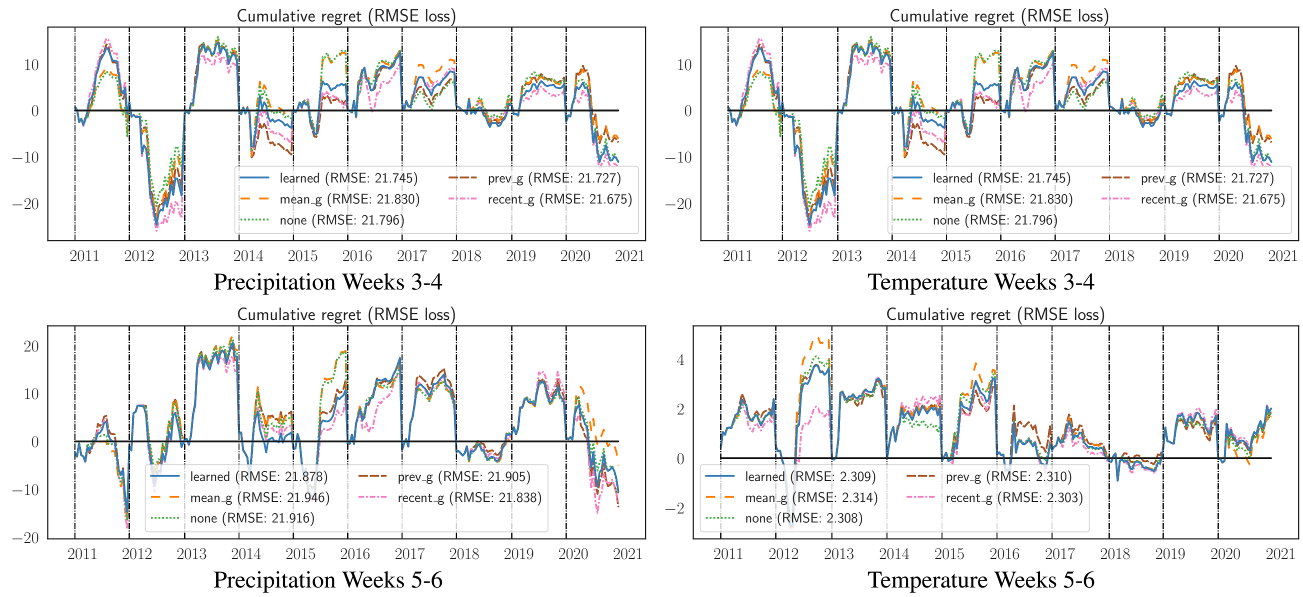


Figure 13: **Overall regret:** Yearly cumulative regret under the RMSE loss for DORM+ using the three constant hinting strategies presented and the learned hinter, over the 10-year evaluation period.

M.5. Impact of Different Forms of Optimism

The regret analysis presented in this work suggest that optimistic strategies under delay can benefit from hinting at both the "past" $\mathbf{g}_{t-D:t-1}$ missing losses and the "future" unobserved loss $\mathbf{g}_t$. To study the impact of different forms of optimism on DORM+, we provide a `recent_g` hint for either only the missing **future** loss $\mathbf{g}_t$, only the missing **past** losses $\mathbf{g}_{t-D:t-1}$, or both **past and future** losses (the strategy used in this paper) $\mathbf{g}_{t-D:t}$. Inspired by the recommendation of an anonymous reviewer, we also test two hint settings that only hint at the future unobserved loss but multiply the weight of that hint by **2D+1** or **3D+1**, effectively increasing the importance of the future hint in the online learning optimization. Fig. 14 presents the experimental results.

In this experiment, all settings of optimism improve upon the non-optimistic algorithm, and, for all tasks, providing hints for missing future losses outperforms hinting at missing past losses. For all tasks save Temp. 5-6w, hinting at both missing past and future losses yields a further improvement. The 2D+1 and 3D+1 settings demonstrate that, for some tasks, increasing the magnitude of the optimistic hint can further improve performance in line with the online gradient descent predictions of Hsieh et al. (2020, Thm. 13).

N. Algorithmic Details

N.1. ODAFTRL with AdaHedgeD and DUB tuning

The AdaHedgeD and DUB algorithms presented in the experiments are implementations of ODAFTRL with a negative entropy regularizer $\psi(\mathbf{w}) = \sum_{j=1}^{d} \mathbf{w}_j \ln \mathbf{w}_j + \ln d$, which is 1-strongly convex with respect to the norm $\|\cdot\|_1$ (Shalev-

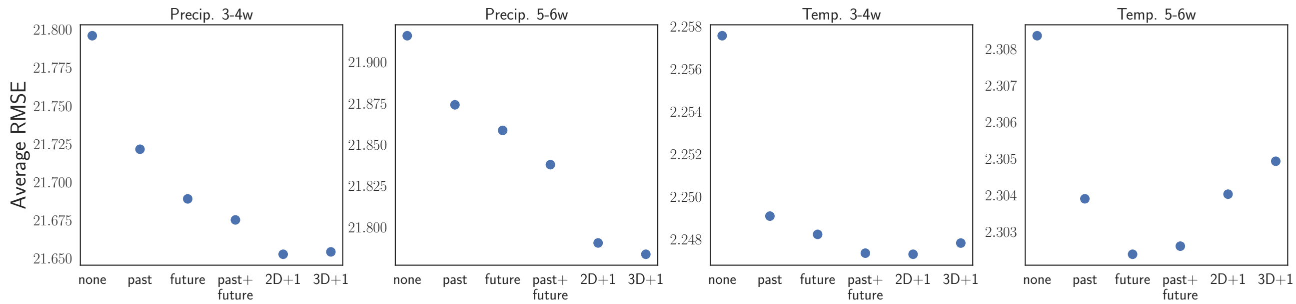


Figure 14: DORM+ average RMSE as in Table 1 as a function of optimism strategy; see App. M.5 for details.

Shwartz, 2007, Lemma 16) with dual norm $\|\cdot\|_\infty$. Each algorithm optimizes over the simplex and competes with the simplex: $\mathbf{W} = \mathbf{U} = \triangle_{d-1}$. We choose $\alpha = \sup_{\mathbf{u}\in\mathbf{U}} \psi(\mathbf{u}) = \ln(d)$. In the following, define $\psi_t \triangleq \lambda_t\psi$ for $\lambda_t \geq 0$. Our derivations of the update equations for AdaHedgeD and DUB make use of the following properties of the negative entropy regularizer, proved in App. N.4.

Lemma 28 (Negative entropy properties)**.** *The negative entropy regularizer $\psi(\mathbf{w}) = \sum_{j=1}^d \mathbf{w}_j \ln \mathbf{w}_j + \ln d$ with $\psi_t = \lambda_t\psi$ for $\lambda_t \geq 0$ satisfies the following properties on the simplex $\mathbf{W} = \triangle_{d-1}$.*

$$\psi^*_{\mathbf{W}}(\theta) \triangleq \sup_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{w}, \theta\rangle - \psi(\mathbf{w}) = \ln\left(\textstyle\sum_{j=1}^d \exp(\theta_j)\right) - \ln d,$$

$$(\lambda\psi)^*_{\mathbf{W}}(\theta) \triangleq \sup_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{w}, \theta\rangle - \lambda\psi(\mathbf{w}) = \begin{cases} \lambda\psi^*_{\mathbf{W}}(\theta/\lambda) = \lambda\ln(\sum_{j=1}^d \exp(\theta_j/\lambda)) - \lambda\ln d, & \textit{if } \lambda > 0 \\ \max_{j\in[d]} \theta_j & \textit{if } \lambda = 0 \end{cases},$$

$$\mathbf{w}^*(\theta, \lambda) \triangleq \begin{cases} \frac{\exp(\theta/\lambda)}{\sum_{j=1}^d \exp(\theta_j/\lambda)} & \textit{if } \lambda > 0 \\ \frac{\mathbb{I}[\theta = \max_j \theta_j]}{\sum_{k\in[d]} \mathbb{I}[\theta_k = \max_j \theta_j]} & \textit{if } \lambda = 0 \end{cases} \in \operatorname{argmin}_{\mathbf{w}\in\mathbf{W}} \lambda\psi(\mathbf{w}) - \langle \mathbf{w}, \theta\rangle \subseteq \partial(\lambda\psi)^*_{\mathbf{W}}(\theta).$$

Our next corollary concerning optimal ODAFTRL objectives follows directly from Lem. 28.

Corollary 29 (Optimal ODAFTRL objectives)**.** *Instantiate the notation of Lem. 28, and define the functions $F_t(\mathbf{w}, \lambda) \triangleq \lambda\psi(\mathbf{w}) + \langle \mathbf{g}_{1:t-1}, \mathbf{w}\rangle$ for $\mathbf{w} \in \mathbf{W}$. Then*

$$-(\lambda\psi)^*_{\mathbf{W}}(-(\mathbf{g}_{1:t-1} + \mathbf{h})) = \inf_{\mathbf{w}\in\mathbf{W}} F_t(\mathbf{w}, \lambda) + \langle \mathbf{h}, \mathbf{w}\rangle \quad \textit{and}$$
$$\mathbf{w}^*(-(\mathbf{g}_{1:t-1} + \mathbf{h}), \lambda) = \operatorname{argmin}_{\mathbf{w}\in\mathbf{W}} F_t(\mathbf{w}, \lambda) + \langle \mathbf{h}, \mathbf{w}\rangle.$$

Using Lem. 28 and Cor. 29, we can derive an expression, proved in App. N.5, for the AdaHedgeD δ_t updates.

Proposition 30 (AdaHedgeD δ_t)**.** *Instantiate the notation of Thm. 12, and define the auxiliary hint vector*

$$\hat{\mathbf{h}}_t \triangleq \mathbf{g}_{t-D:t} + \sigma_t(\mathbf{h}_t - \mathbf{g}_{t-D:t}) \quad \textit{for} \quad \sigma_t \triangleq \min\left(\frac{\|\mathbf{g}_t\|_*}{\|\mathbf{h}_t - \mathbf{g}_{t-D:t}\|_*}, 1\right) \tag{13}$$

along with the scalars

$$c_* = \max_{j:\mathbf{w}_{t,j}\neq 0} \mathbf{h}_{t,j} - \mathbf{g}_{t-D:t,j} \quad \textit{and} \quad \hat{c}_* = \max_{j:\hat{\mathbf{w}}_{t,j}\neq 0} \hat{\mathbf{h}}_{t,j} - \mathbf{g}_{t-D:t,j}$$

for

$$\bar{\mathbf{w}}_t = \operatorname{argmin}_{\mathbf{w}\in\mathbf{W}} F_{t+1}(\mathbf{w}, \lambda_t) = \frac{\exp(-\mathbf{g}_{1:t}/\lambda_t)}{\sum_{j=1}^d \exp(-\mathbf{g}_{1:t,j}/\lambda_t)} \quad \textit{and}$$
$$\hat{\mathbf{w}}_t = \operatorname{argmin}_{\mathbf{w}\in\mathbf{W}} F_{t+1}(\mathbf{w}, \lambda_t) + \langle \hat{\mathbf{h}}_t - \mathbf{g}_{t-D:t}, \mathbf{w}\rangle = \frac{\exp(-(\mathbf{g}_{1:t-D-1} + \hat{\mathbf{h}}_t)/\lambda_t)}{\sum_{j=1}^d \exp(-(\mathbf{g}_{1:t-D-1,j} + \hat{\mathbf{h}}_{t,j})/\lambda_t)}$$

by Cor. 29. If $\lambda_t > 0$,

$$\begin{aligned}
\delta_t &= \min(\delta_t^{(1)}, \delta_t^{(2)}, \delta_t^{(3)})_+ \quad \textit{for} \\
\delta_t^{(1)} &= F_{t+1}(\mathbf{w}_t, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) \\
&= \lambda_t \ln(\textstyle\sum_{j\in[d]} \mathbf{w}_{t,j} \exp((\mathbf{h}_{t,j} - \mathbf{g}_{t-D:t,j})/\lambda_t)) + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle \\
&= \lambda_t \ln(\textstyle\sum_{j\in[d]} \mathbf{w}_{t,j} \exp((\mathbf{h}_{t,j} - \mathbf{g}_{t-D:t,j} - c_*)/\lambda_t)) + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle + c_*, \\
\delta_t^{(2)} &= \langle \mathbf{g}_t, \mathbf{w}_t - \bar{\mathbf{w}}_t \rangle, \quad \textit{and} \\
\delta_t^{(3)} &= F_{t+1}(\hat{\mathbf{w}}_t, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) + \langle \mathbf{g}_t, \mathbf{w}_t - \hat{\mathbf{w}}_t \rangle \\
&= \lambda_t \ln(\textstyle\sum_{j\in[d]} \hat{\mathbf{w}}_{t,j} \exp((\hat{\mathbf{h}}_{t,j} - \mathbf{g}_{t-D:t,j})/\lambda_t)) + \langle \mathbf{g}_{t-D:t} - \hat{\mathbf{h}}_t, \hat{\mathbf{w}}_t \rangle + \langle \mathbf{g}_t, \mathbf{w}_t - \hat{\mathbf{w}}_t \rangle \\
&= \lambda_t \ln(\textstyle\sum_{j\in[d]} \hat{\mathbf{w}}_{t,j} \exp((\hat{\mathbf{h}}_{t,j} - \mathbf{g}_{t-D:t,j} - \hat{c}_*)/\lambda_t)) + \langle \mathbf{g}_{t-D:t} - \hat{\mathbf{h}}_t, \hat{\mathbf{w}}_t \rangle + \hat{c}_* + \langle \mathbf{g}_t, \mathbf{w}_t - \hat{\mathbf{w}}_t \rangle.
\end{aligned}$$

If $\lambda_t = 0$,

$$\begin{aligned}
\delta_t &= \min(\delta_t^{(1)}, \delta_t^{(2)}, \delta_t^{(3)})_+ \quad \textit{for} \\
\delta_t^{(1)} &= \langle \mathbf{g}_{1:t}, \mathbf{w}_t \rangle - \min_{j\in[d]} \mathbf{g}_{1:t,j}, \\
\delta_t^{(2)} &= \langle \mathbf{g}_t, \mathbf{w}_t - \bar{\mathbf{w}}_t \rangle, \quad \textit{and} \\
\delta_t^{(3)} &= \langle \mathbf{g}_{1:t}, \hat{\mathbf{w}}_t \rangle - \min_{j\in[d]} \mathbf{g}_{1:t,j} + \langle \mathbf{g}_t, \mathbf{w}_t - \hat{\mathbf{w}}_t \rangle.
\end{aligned}$$

Leveraging these results, we present the pseudocode for the AdaHedgeD and DUB instantiations of ODAFTRL in Algorithm 1.

N.2. DORM and DORM+

The DORM and DORM+ algorithms presented in the experiments are implementations of ODAFTRL and DOOMD respectively that play iterates in $\mathbf{W} \triangleq \triangle_{d-1}$ using the default value $\lambda = 1$. Both algorithms use a p-norm regularizer $\psi = \frac{1}{2}\|\cdot\|_p^2$, which is 1-strongly convex with respect to $\|\cdot\| = \sqrt{p-1}\|\cdot\|_p$ (see Shalev-Shwartz, 2007, Lemma 17) with $\|\cdot\|_* = \frac{1}{\sqrt{p-1}}\|\cdot\|_q$. For the paper experiments, we choose the optimal value $q = \inf_{q'\geq 2} d^{2/q'}(q'-1)$ to obtain $\ln(d)$ scaling in the algorithm regret; for $d = 6$, $p = q = 2$. The update equations for each algorithm are given in the main text by DORM and DORM+ respectively. The optimistic hinters provide delayed gradient hints $\tilde{\mathbf{g}}_t$, which are then used to compute regret gradient hints $\tilde{\mathbf{r}}_t$, where $\tilde{\mathbf{r}}_t = \langle \tilde{\mathbf{g}}_t, \mathbf{w}_t \rangle - \tilde{\mathbf{g}}_t$ and $\mathbf{h}_t = \sum_{s=t-D}^{t-1} \tilde{\mathbf{r}}_s + \langle \tilde{\mathbf{g}}_t, \mathbf{w}_{t-1} \rangle - \tilde{\mathbf{g}}_t$.

N.3. Adaptive Hinting

For the adaptive hinting experiments, we use the DORM+ as both the base and hint learner. For the hint learner with DORM base algorithm, the hint loss function is given by (11) with $q = 2$. The plays of the online hinter ω_t are used to generate the hints $\mathbf{h}_t$ for the base algorithm using the hint matrix $H_t \in \mathbb{R}^{d\times m}$. The j-th column of H_t contains hinter j's predictions for the cumulative missing regret subgradients $\mathbf{r}_{t-D:t}$. The final hint for the base learner is $\mathbf{h}_t = H_t\omega_t$. Psuedo-code for the adaptive hinter is given in Algorithm 2.

N.4. Proof of Lem. 28: Negative entropy properties

The expression of the Fenchel conjugate for $\lambda > 0$ is derived by solving an appropriate constrained convex optimization problem for $\mathbf{w} = \triangle_{d-1}$, as shown in Orabona (2019, Section 6.6). The value of $\mathbf{w}^*(\theta, \lambda) \in \partial(\lambda\psi)^*_{\mathbf{W}}(\theta)$ uses the properties of the Fenchel conjugate (Rockafellar, 1970; Orabona, 2019, Theorem 5.5) and is shown in Orabona (2019, Theorem 6.6).

Algorithm 1 ODAFTRL with $\mathbf{W} = \triangle_{d-1}$, $\psi(\mathbf{w}) = \sum_{j=1}^{d} \mathbf{w}_j \ln \mathbf{w}_j + \ln(d)$, delay $D \geq 0$, and tuning strategy `tuning`

Parameter $\alpha = \sup_{\mathbf{u} \in \triangle_{d-1}} \psi(\mathbf{u}) = \ln(d)$
Initial regularization weight: $\lambda_0 = 0$
if `tuning` is DUB **then**
  Initial regularization sum: $\Delta_0 = 0$
  Initial maximum: $\mathbf{a}^{\max} = 0$
end if
Initial subgradient sum: $\mathbf{g}_{1:1} = \mathbf{0} \in \mathbb{R}^d$
Dummy losses and iterates: $\mathbf{g}_{-D} = \cdots = \mathbf{g}_0 = \mathbf{0} \in \mathbb{R}^d$, $\mathbf{w}_{-D} = \cdots = \mathbf{w}_0 = \mathbf{0} \in \mathbb{R}^d$
for $t = 1, \ldots, T$ **do**
  Receive hint $\mathbf{h}_t \in \mathbb{R}^d$
  Output $\mathbf{w}_t = \operatorname{argmin}_{\mathbf{w} \in \mathbf{W}} F_{t-D}(\mathbf{w}, \lambda_t) + \langle \mathbf{h}_t, \mathbf{w} \rangle$ as in Cor. 29
  Receive $\mathbf{g}_{t-D} \in \mathbb{R}^d$ and pay $\langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D} \rangle$
  Update subgradient sum $\mathbf{g}_{1:t-D} = \mathbf{g}_{1:t-D-1} + \mathbf{g}_{t-D}$
  if `tuning` is AdaHedgeD **then**
    Compute the auxiliary play $\bar{\mathbf{w}}_{t-D} = \operatorname{argmin}_{\mathbf{w} \in \mathbf{W}} F_{t-D+1}(\mathbf{w}, \lambda_{t-D})$ as in Cor. 29
    Compute the auxiliary regret term $\delta_{t-D}^{(1)} = F_{t-D+1}(\mathbf{w}_{t-D}, \lambda_{t-D}) - F_{t-D+1}(\bar{\mathbf{w}}_{t-D}, \lambda_{t-D})$ as in Prop. 30
    Compute the drift term $\delta_{t-D}^{(2)} = \langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D} - \bar{\mathbf{w}}_{t-D} \rangle$
    Compute the auxiliary hint (13) $\hat{\mathbf{h}}_{t-D} \triangleq \mathbf{g}_{t-2D:t-D} + \min(\frac{\|\mathbf{g}_{t-D}\|_*}{\|\mathbf{h}_{t-D} - \mathbf{g}_{t-2D:t-D}\|_*}, 1)(\mathbf{h}_{t-D} - \mathbf{g}_{t-2D:t-D})$
    Compute the auxiliary play $\hat{\mathbf{w}}_{t-D} = \operatorname{argmin}_{\mathbf{w} \in \mathbf{W}} F_{t-D+1}(\mathbf{w}, \lambda_{t-D}) + \langle \hat{\mathbf{h}}_{t-D} - \mathbf{g}_{t-2D:t-D}, \mathbf{w} \rangle$ as in Cor. 29
    Compute the regret term $\delta_{t-D}^{(3)} = F_{t-D+1}(\hat{\mathbf{w}}_{t-D}, \lambda_{t-D}) - F_{t-D+1}(\bar{\mathbf{w}}_{t-D}, \lambda_{t-D}) + \langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D} - \hat{\mathbf{w}}_{t-D} \rangle$ as in Prop. 30
    Update $\lambda_{t+1} = \lambda_t + \frac{1}{\alpha} \min(\delta_{t-D}^{(1)}, \delta_{t-D}^{(2)}, \delta_{t-D}^{(3)})_+$ as in (2)
  else if `tuning` is DUB **then**
    Compute $\mathbf{a}_{t-D,F} = 2 \min\left(\|\mathbf{g}_{t-D}\|_\infty, \|\mathbf{h}_{t-D} - \sum_{s=t-2D}^{t-D} \mathbf{g}_s\|_\infty\right)$ as in (1)
    Compute $\mathbf{b}_{t-D,F} = \frac{1}{2}\|\mathbf{h}_{t-D} - \sum_{s=t-2D}^{t-D} \mathbf{g}_s\|_\infty^2 - \frac{1}{2}(\|\mathbf{h}_{t-D} - \sum_{s=t-2D}^{t-D} \mathbf{g}_s\|_\infty - \|\mathbf{g}_{t-D}\|_\infty)_+^2$ as in (1)
    Update $\Delta_{t+1} = \Delta_t + \mathbf{a}_{t-D,F}^2 + 2\alpha \mathbf{b}_{t-D,F}$
    Update maximum $\mathbf{a}^{\max} = \max(\mathbf{a}^{\max}, \mathbf{a}_{t-2D:t-D-1,F})$
    Update $\lambda_{t+1} = \frac{1}{\alpha}(2\mathbf{a}^{\max} + \sqrt{\Delta_{t+1}})$ as in DUB
  end if
end for

N.5. Proof of Prop. 30: AdaHedgeD δ_t

First suppose $\lambda_t > 0$. The first term in the $\min$ of AdaHedgeD's δ_t setting is derived as follows:

$$
\begin{aligned}
\delta_t^{(1)} &\triangleq F_{t+1}(\mathbf{w}_t, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) \quad \text{by definition (2)} \\
&= F_{t-D}(\mathbf{w}_t, \lambda_t) + \langle \mathbf{h}_t, \mathbf{w}_t \rangle + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle - \inf_{\mathbf{w} \in \mathbf{W}} F_{t+1}(\mathbf{w}, \lambda_t) \quad \text{by definition of } \bar{\mathbf{w}}_t \\
&= F_{t-D}(\mathbf{w}_t, \lambda_t) + \langle \mathbf{h}_t, \mathbf{w}_t \rangle + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle + \lambda_t \psi_{\mathbf{W}}^*(-\mathbf{g}_{1:t}/\lambda_t) \quad \text{by Cor. 29} \\
&= \lambda_t \psi_{\mathbf{W}}^*(-\mathbf{g}_{1:t}/\lambda_t) - \lambda_t \psi_{\mathbf{W}}^*((-\mathbf{h}_t - \mathbf{g}_{1:t-D-1})/\lambda_t) + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle \\
&\quad \text{because } \mathbf{w}_t \in \operatorname{argmin}_{\mathbf{w} \in \mathbf{W}} F_{t-D}(\mathbf{w}_t, \lambda_t) + \langle \mathbf{h}_t, \mathbf{w}_t \rangle \\
&= \lambda_t (\ln(\textstyle\sum_{j=1}^d \exp(-\mathbf{g}_{1:t,j}/\lambda_t)) - \lambda_t(\ln(\textstyle\sum_{j=1}^d \exp((-\mathbf{g}_{1:t-D-1,j} - \mathbf{h}_{t,j})/\lambda_t)) + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle \quad \text{by Lem. 28} \\
&= \lambda_t \ln\left(\textstyle\sum_{j=1}^d \frac{\exp(-\mathbf{g}_{1:t,j}/\lambda_t)}{\sum_{j=1}^d \exp((-\mathbf{g}_{1:t-D-1,j} - \mathbf{h}_{t,j})/\lambda_t)}\right) + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle \\
&= \lambda_t \ln\left(\textstyle\sum_{j=1}^d \frac{\exp((-\mathbf{g}_{1:t-D-1,j} - \mathbf{h}_{t,j})/\lambda_t) \exp((\mathbf{h}_{t,j} - \mathbf{g}_{t-D:t,j})/\lambda_t)}{\sum_{j=1}^d \exp((-\mathbf{g}_{1:t-D-1,j} - \mathbf{h}_{t,j})/\lambda_t)}\right) + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle \\
&= \lambda_t \ln\left(\textstyle\sum_{j=1}^d \mathbf{w}_{t,j} \exp((\mathbf{h}_{t,j} - \mathbf{g}_{t-D:t,j})/\lambda_t)\right) + \langle \mathbf{g}_{t-D:t} - \mathbf{h}_t, \mathbf{w}_t \rangle \quad \text{by the expression for } \mathbf{w}_t \text{ in Cor. 29.}
\end{aligned}
$$

The expression for the third term in the $\min$ of AdaHedgeD's δ_t setting follows from identical reasoning.

Algorithm 2 Learning to hint with DORM+ (q=2) hint learner, DORM+ base learner, and delay $D \geq 0$

Subgradient vector: $\mathbf{g}_{-D}, \cdots \mathbf{g}_0 = \mathbf{0} \in \mathbb{R}^d$
Meta-subgradient vector: $\gamma_{-D}, \cdots \gamma_0 = \mathbf{0} \in \mathbb{R}^m$
Initial instantaneous regret: $\mathbf{r}_{-D} = \mathbf{0} \in \mathbb{R}^d$
Initial instantaneous meta-regret: $\rho_{-D} = \mathbf{0} \in \mathbb{R}^m$
Initial hint $\mathbf{h}_0 = \mathbf{0} \in \mathbb{R}^d$
Initial orthant meta-vector: $\tilde{\omega}_0 = \mathbf{0} \in \mathbb{R}^m$
for $t = 1, \dots, T$ **do**
  // `Update online hinter using DORM+ with` $q = 2$
  Find optimal unnormalized hint combination vector $\tilde{\omega}_t = \max(\mathbf{0}, \tilde{\omega}_{t-1} + \rho_{t-D-1})$
  Normalize: $\omega_t = \begin{cases} \mathbf{1}/m & \text{if } \tilde{\omega}_t = \mathbf{0} \\ \tilde{\omega}_t / \langle \mathbf{1}, \tilde{\omega}_t \rangle & \text{otherwise} \end{cases}$
  Receive hint matrix: $H_t \in \mathbb{R}^{d \times m}$ in which each column is a hint for $\sum_{s=t-D}^{t} \mathbf{r}_s$
  Output hint $\mathbf{h}_t = H_t \omega_t$
  // `Update DORM+ base learner and get next play`
  Output $\mathbf{w}_t = \text{DORM+}(\mathbf{g}_{t-D-1}, \mathbf{h}_t)$
  Receive $\mathbf{g}_{t-D} \in \mathbb{R}^d$ and pay $\langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D} \rangle$
  Compute instantaneous regret $\mathbf{r}_{t-D} = \mathbf{1}\langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D} \rangle - \mathbf{g}_{t-D}$
  Compute hint meta-subgradient $\gamma_{t-D} \in \partial l_{t-D}(\omega_{t-D}) \in \mathbb{R}^m$ as in (12)
  Compute instantaneous hint regret $\rho_{t-D} = \mathbf{1}\langle \gamma_{t-D}, \omega_{t-D} \rangle - \gamma_{t-D}$
end for

Now suppose $\lambda_t = 0$. We have

$$\begin{aligned} \delta_t^{(1)} &\triangleq F_{t+1}(\mathbf{w}_t, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t) \quad \text{by definition (2)} \\ &= \langle \mathbf{g}_{1:t}, \mathbf{w}_t \rangle - \inf_{\mathbf{w} \in \mathbf{W}} F_{t+1}(\mathbf{w}, \lambda_t) \quad \text{by definition of } \bar{\mathbf{w}}_t \\ &= \langle \mathbf{g}_{1:t}, \mathbf{w}_t \rangle - \min_{j \in [d]} \mathbf{g}_{1:t,j} \quad \text{by Cor. 29.} \end{aligned}$$

Identical reasoning yields the advertised expression for the third term.

O. Extension to Variable and Unbounded Delays

In this section we detail how our main results generalize to the case of variable and potentially unbounded delays. For each time t, we define $\text{last}(t)$ as the largest index s for which $\mathbf{g}_{1:s}$ is observable at time t (that is, available for constructing $\mathbf{w}_t$) and $\text{first}(t)$ as the first time s at which $\mathbf{g}_{1:t}$ is observable at time s (that is, available for constructing $\mathbf{w}_s$).

O.1. Regret of DOOMD with variable delays

Consider the DOOMD variable-delay generalization

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w} \in \mathbf{W}} \langle \mathbf{g}_{\text{last}(t)+1:\text{last}(t+1)} + \mathbf{h}_{t+1} - \mathbf{h}_t, \mathbf{w} \rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_t) \quad \text{with} \quad \mathbf{h}_0 \triangleq \mathbf{0} \quad \text{and arbitrary} \quad \mathbf{w}_0. \tag{DOOMD with variable delays}$$

We first note that DOOMD with variable delays is an instance of SOOMD respectively with a “bad” choice of optimistic hint $\tilde{\mathbf{g}}_{t+1}$ that deletes the unobserved loss subgradients $\mathbf{g}_{\text{last}(t+1)+1:t}$.

Lemma 31 (DOOMD with variable delays is SOOMD with a bad hint)**.** *DOOMD with variable delays is SOOMD with* $\tilde{\mathbf{g}}_{t+1} = \tilde{\mathbf{g}}_t + \mathbf{g}_{\text{last}(t)+1:\text{last}(t+1)} - \mathbf{g}_t + \mathbf{h}_{t+1} - \mathbf{h}_t = \mathbf{h}_{t+1} + \sum_{s=1}^{t} \mathbf{g}_{\text{last}(s)+1:\text{last}(s+1)} - \mathbf{g}_s. = \mathbf{h}_{t+1} - \mathbf{g}_{\text{last}(t+1)+1:t}$.

The following result now follows immediately from Thm. 4 and Lem. 31.

Theorem 32 (Regret of DOOMD with variable delays)**.** *If ψ is differentiable and $\mathbf{h}_{T+1} \triangleq \mathbf{g}_{\text{last}(T+1)+1:T}$, then, for all*

$\mathbf{u} \in \mathbf{W}$, *the* DOOMD with variable delays *iterates* $\mathbf{w}_t$ *satisfy*

$$\begin{aligned}\text{Regret}_T(\mathbf{u}) &\le \mathcal{B}_{\lambda\psi}(\mathbf{u}, \mathbf{w}_0) + \tfrac{1}{\lambda}\textstyle\sum_{t=1}^T \mathbf{b}_{t,O}^2, \quad \textit{for} \\ \mathbf{b}_{t,O}^2 &\triangleq \text{huber}(\|\mathbf{h}_t - \textstyle\sum_{s=\text{last}(t)+1}^t \mathbf{g}_s\|_*, \|\mathbf{g}_{\text{last}(t)+1:\text{last}(t+1)} + \mathbf{h}_{t+1} - \mathbf{h}_t\|_*).\end{aligned}$$

O.2. Regret of ODAFTRL with variable delays

Consider the ODAFTRL variable-delay generalization

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{g}_{1:\text{last}(t+1)} + \mathbf{h}_{t+1}, \mathbf{w}\rangle + \lambda_{t+1}\psi(\mathbf{w}). \qquad \text{(ODAFTRL with variable delays)}$$

Since ODAFTRL with variable delays is an instance of OAFTRL with $\tilde{\mathbf{g}}_{t+1} = \mathbf{h}_{t+1} - \sum_{s=\text{last}(t+1)+1}^t \mathbf{g}_s$, the following result follows immediately from the OAFTRL regret bound, Thm. 14.

Theorem 33 (Regret of ODAFTRL with variable delays)**.** *If ψ is nonnegative and λ_t is non-decreasing in t, then, $\forall \mathbf{u} \in \mathbf{W}$, the* ODAFTRL with variable delays *iterates* $\mathbf{w}_t$ *satisfy*

$$\begin{aligned}\text{Regret}_T(\mathbf{u}) &\le \lambda_T\psi(\mathbf{u}) + \textstyle\sum_{t=1}^T \min(\frac{\mathbf{b}_{t,F}}{\lambda_t}, \mathbf{a}_{t,F}) \quad \textit{with} \\ \mathbf{b}_{t,F} &\triangleq \text{huber}(\|\mathbf{h}_t - \textstyle\sum_{s=\text{last}(t)+1}^t \mathbf{g}_s\|_*, \|\mathbf{g}_t\|_*) \quad \textit{and} \\ \mathbf{a}_{t,F} &\triangleq \text{diam}(\mathbf{W}) \min\big(\|\mathbf{h}_t - \textstyle\sum_{s=\text{last}(t)+1}^t \mathbf{g}_s\|, \|\mathbf{g}_t\|_*\big).\end{aligned} \tag{14}$$

O.3. Regret of DUB with variable delays

Consider the DUB variable-delay generalization

$$\alpha\lambda_{t+1} = 2 \max_{j\le\text{last}(t+1)-1} \mathbf{a}_{\text{last}(j+1)+1:j,F} + \sqrt{\textstyle\sum_{i=1}^{\text{last}(t+1)} \mathbf{a}_{i,F}^2 + 2\alpha\mathbf{b}_{i,F}}. \qquad \text{(DUB with variable delays)}$$

Theorem 34 (Regret of DUB with variable delays)**.** *Fix $\alpha > 0$, and, for $\mathbf{a}_{t,F}, \mathbf{b}_{t,F}$ as in* (14)*, consider the* DUB with variable delays *sequence. If ψ is nonnegative, then, for all $\mathbf{u} \in \mathbf{W}$, the* ODAFTRL with variable delays *iterates* $\mathbf{w}_t$ *satisfy*

$$\begin{aligned}&\text{Regret}_T(\mathbf{u}) \le \big(\tfrac{\psi(\mathbf{u})}{\alpha} + 1\big) \\ &\big(2\max_{t\in[T]} \mathbf{a}_{\text{last}(t)+1:t-1,F} + \sqrt{\textstyle\sum_{t=1}^T \mathbf{a}_{t,F}^2 + 2\alpha\mathbf{b}_{t,F}}\big)\end{aligned}$$

Proof. Fix any $\mathbf{u} \in \mathbf{W}$. By Thm. 33, ODAFTRL with variable delays admits the regret bound

$$\text{Regret}_T(\mathbf{u}) \le \lambda_T\psi(\mathbf{u}) + \textstyle\sum_{t=1}^T \min(\frac{1}{\lambda_t}\mathbf{b}_{t,F}, \mathbf{a}_{t,F}).$$

To control the second term in this bound, we apply the following lemma proved in App. H.1.

Lemma 35 (DUB with variable delays-style tuning bound)**.** *Fix any $\alpha > 0$ and any non-negative sequences $(a_t)_{t=1}^T$, $(b_t)_{t=1}^T$. If $(\lambda_t)_{t\ge1}$ is non-decreasing and*

$$\Delta_{t+1}^* \triangleq 2\max_{j\le\text{last}(t+1)-1} a_{\text{last}(j+1)+1:j} + \sqrt{\textstyle\sum_{i=1}^{\text{last}(t+1)} a_i^2 + 2\alpha b_i} \le \alpha\lambda_{t+1} \quad \textit{for each} \quad t$$

then

$$\textstyle\sum_{t=1}^T \min(b_t/\lambda_t, a_t) \le \Delta_{\text{first}(T)}^* \le \alpha\lambda_{\text{first}(T)}.$$

□

Since $T \le \text{first}(T)$, $\lambda_T \le \lambda_{\text{first}(T)}$, and $\text{last}(\text{first}(T)) = T$, the result now follows by setting $a_t = \mathbf{a}_{t,F}$ and $b_t = \mathbf{b}_{t,F}$, so that

$$\text{Regret}_T(\mathbf{u}) \le \lambda_T\psi(\mathbf{u}) + \alpha\lambda_{\text{first}(T)} \le (\psi(\mathbf{u}) + \alpha)\lambda_{\text{first}(T)}.$$

O.4. Proof of Lem. 35: DUB with variable delays-style tuning bound

We prove the claim

$$\Delta_t \triangleq \sum_{i=1}^{t} \min(b_i/\lambda_i, a_i) \le \Delta^*_{\text{first}(t)} \le \alpha\lambda_{\text{first}(t)}$$

by induction on t.

Base case For $t = 1$, since $\text{last}(\text{first}(t)) \ge t$, we have

$$\begin{aligned}\textstyle\sum_{i=1}^{t} \min(b_i/\lambda_i, a_i) &\le a_1 \le 2\max_{j\le t-1} a_{\text{last}(j+1)+1:j} + \sqrt{\textstyle\sum_{i=1}^{t} a_i^2 + 2\alpha b_i} \\ &\le 2\max_{j\le \text{last}(\text{first}(t))-1} a_{\text{last}(j+1)+1:j} + \sqrt{\textstyle\sum_{i=1}^{\text{last}(\text{first}(t))} a_i^2 + 2\alpha b_i} = \Delta^*_{\text{first}(t)} \le \alpha\lambda_{\text{first}(t)}\end{aligned}$$

confirming the base case.

Inductive step Now fix any $t+1 \ge 2$ and suppose that

$$\Delta_i \le \Delta^*_{\text{first}(i)} \le \alpha\lambda_{\text{first}(i)}$$

for all $1 \le i \le t$. Since $\text{first}(\text{last}(i+1)) \le i+1$ and λ_s is non-decreasing in s, we apply this inductive hypothesis to deduce that, for each $0 \le i \le t$,

$$\begin{aligned}\Delta_{i+1}^2 - \Delta_i^2 &= (\Delta_i + \min(b_{i+1}/\lambda_{i+1}, a_{i+1}))^2 - \Delta_i^2 = 2\Delta_i \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + \min(b_{i+1}/\lambda_{i+1}, a_{i+1})^2 \\ &= 2\Delta_{\text{last}(i+1)} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + 2(\Delta_i - \Delta_{\text{last}(i+1)}) \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + \min(b_{i+1}/\lambda_{i+1}, a_{i+1})^2 \\ &= 2\Delta_{\text{last}(i+1)} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + 2\sum_{j=\text{last}(i+1)+1}^{i} \min(b_j/\lambda_j, a_j) \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + \min(b_{i+1}/\lambda_{i+1}, a_{i+1})^2 \\ &\le 2\alpha\lambda_{\text{first}(\text{last}(i+1))} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + 2a_{\text{last}(i+1)+1:i} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + a_{i+1}^2 \\ &\le 2\alpha\lambda_{i+1} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + 2a_{\text{last}(i+1)+1:i} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) + a_{i+1}^2 \\ &\le 2\alpha b_{i+1} + a_{i+1}^2 + 2a_{\text{last}(i+1)+1:i} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}).\end{aligned}$$

Now, we sum this inequality over $i = 0, \dots, t$, to obtain

$$\begin{aligned}\Delta_{t+1}^2 &\le \textstyle\sum_{i=0}^{t} (2\alpha b_{i+1} + a_{i+1}^2) + 2\sum_{i=0}^{t} a_{\text{last}(i+1)+1:i} \min(b_{i+1}/\lambda_{i+1}, a_{i+1}) \\ &= \textstyle\sum_{i=1}^{t+1} (2\alpha b_i + a_i^2) + 2\sum_{i=1}^{t+1} a_{\text{last}(i+1):i-1} \min(b_i/\lambda_i, a_i) \\ &\le \textstyle\sum_{i=1}^{t+1} (a_i^2 + 2\alpha b_i) + 2\max_{j\le t} a_{\text{last}(j+1)+1:j} \sum_{i=1}^{t+1} \min(b_i/\lambda_i, a_i) \\ &= \textstyle\sum_{i=1}^{t+1} (a_i^2 + 2\alpha b_i) + 2\Delta_{t+1} \max_{j\le t} a_{\text{last}(j+1)+1:j}.\end{aligned}$$

We now solve this quadratic inequality, apply the triangle inequality, and invoke the relation $\text{last}(\text{first}(t+1)) \ge t+1$ to conclude that

$$\begin{aligned}\Delta_{t+1} &\le \max_{j\le t} a_{\text{last}(j+1)+1:j} + \tfrac{1}{2}\sqrt{(2\max_{j\le t} a_{\text{last}(j+1)+1:j})^2 + 4\textstyle\sum_{i=1}^{t+1} a_i^2 + 2\alpha b_i} \\ &\le 2\max_{j\le t} a_{\text{last}(j+1)+1:j} + \sqrt{\textstyle\sum_{i=1}^{t+1} a_i^2 + 2\alpha b_i} \\ &\le 2\max_{j\le \text{last}(\text{first}(t+1))-1} a_{\text{last}(j+1)+1:j} + \sqrt{\textstyle\sum_{i=1}^{\text{last}(\text{first}(t+1))} a_i^2 + 2\alpha b_i} = \Delta^*_{\text{first}(t+1)} \le \alpha\lambda_{\text{first}(t+1)}.\end{aligned}$$

O.5. Regret of AdaHedgeD with variable delays

Consider the AdaHedgeD variable-delay generalization

$$\lambda_{t+1} = \tfrac{1}{\alpha}\textstyle\sum_{s=1}^{\text{last}(t+1)} \delta_s \quad \text{for} \quad \delta_t \quad \text{defined in (2).} \qquad \text{(AdaHedgeD with variable delays)}$$

Theorem 36 (Regret of AdaHedgeD with variable delays)**.** *Fix $\alpha > 0$, and consider the* AdaHedgeD with variable delays *sequence. If ψ is nonnegative, then, for all $\mathbf{u} \in \mathbf{W}$, the* ODAFTRL with variable delays *iterates satisfy*

$$\begin{aligned}&\mathrm{Regret}_T(\mathbf{u}) \le \big(\tfrac{\psi(\mathbf{u})}{\alpha} + 1\big)\\ &\big(2\max_{t\in[T]} \mathbf{a}_{\mathrm{last}(t+1)+1:t,F} + \sqrt{\textstyle\sum_{t=1}^T \mathbf{a}_{t,F}^2 + 2\alpha \mathbf{b}_{t,F}}\big).\end{aligned}$$

Proof. Fix any $\mathbf{u} \in \mathbf{W}$, and for each t, define $\lambda'_{t+1} = \frac{1}{\alpha}\sum_{s=1}^t \delta_s$ so that $\alpha(\lambda'_{t+1} - \lambda'_t) = \delta_t$. Since the AdaHedgeD with variable delays regularization sequence $(\lambda_t)_{t\ge1}$ is non-decreasing, last$(T) \le T$, and hence $\lambda_T \le \lambda'_{T+1}$, Thm. 14 gives the regret bound

$$\mathrm{Regret}_T(\mathbf{u}) \le \lambda_T\psi(\mathbf{u}) + \textstyle\sum_{t=1}^T \delta_t \le \lambda_T\psi(\mathbf{u}) + \alpha\lambda'_{T+1} \le (\psi(\mathbf{u}) + \alpha)\lambda'_{T+1}$$

and the proof of Thm. 14 gives the upper estimate (4):

$$\delta_t \le \min\left(\tfrac{\mathbf{b}_{t,F}}{\lambda_t}, \mathbf{a}_{t,F}\right) \quad \text{for all} \quad t \in [T]. \tag{15}$$

Hence, it remains to bound λ'_{T+1}. We have

$$\begin{aligned}\alpha{\lambda'_{T+1}}^2 &= \textstyle\sum_{t=1}^T \alpha({\lambda'_{t+1}}^2 - {\lambda'_t}^2) = \sum_{t=1}^T \big(\alpha(\lambda'_{t+1} - \lambda'_t)^2 + 2\alpha(\lambda'_{t+1} - \lambda'_t)\lambda'_t\big)\\ &= \textstyle\sum_{t=1}^T \big(\delta_t^2/\alpha + 2\delta_t\lambda'_t\big) \quad \text{by the definition of } \lambda'_{t+1}\\ &= \textstyle\sum_{t=1}^T \big(\delta_t^2/\alpha + 2\delta_t\lambda_t + 2\delta_t(\lambda'_t - \lambda_t)\big)\\ &\le \textstyle\sum_{t=1}^T \big(\delta_t^2/\alpha + 2\delta_t\lambda_t + 2\delta_t \max_{t\in[T]}(\lambda'_t - \lambda_t)\big)\\ &= \textstyle\sum_{t=1}^T \big(\delta_t^2/\alpha + 2\delta_t\lambda_t\big) + 2\alpha\lambda'_{T+1}\max_{t\in[T]}(\lambda'_t - \lambda_t)\\ &= \textstyle\sum_{t=1}^T \big(\delta_t^2/\alpha + 2\delta_t\lambda_t\big) + 2\lambda'_{T+1}\max_{t\in[T]}\delta_{\mathrm{last}(t+1)+1:t}\\ &\le \textstyle\sum_{t=1}^T \big(\mathbf{a}_{t,F}^2/\alpha + 2\mathbf{b}_{t,F}\big) + 2\lambda'_{T+1}\max_{t\in[T]}\mathbf{a}_{\mathrm{last}(t+1)+1:t,F} \quad \text{by (15).}\end{aligned}$$

Solving the above quadratic inequality for λ'_{T+1} and applying the triangle inequality, we find

$$\begin{aligned}\alpha\lambda'_{T+1} &\le \max_{t\in[T]}\mathbf{a}_{\mathrm{last}(t+1)+1:t,F} + \tfrac{1}{2}\sqrt{4(\max_{t\in[T]}\mathbf{a}_{\mathrm{last}(t+1)+1:t,F})^2 + 4\textstyle\sum_{t=1}^T \mathbf{a}_{t,F}^2 + 2\alpha\mathbf{b}_{t,F}}\\ &\le 2\max_{t\in[T]}\mathbf{a}_{\mathrm{last}(t+1)+1:t,F} + \sqrt{\textstyle\sum_{t=1}^T \mathbf{a}_{t,F}^2 + 2\alpha\mathbf{b}_{t,F}}.\end{aligned}$$

□