Online Learning with Optimism and Delay

Genevieve Flaspohler [1] **Francesco Orabona** [2] **Judah Cohen** [3] **Soukayna Mouatadid** [4] **Miruna Oprescu** [5] **Paulo Orenstein** [6] **Lester Mackey** [5]

Abstract

Inspired by the demands of real-time time-series forecasting, we develop and analyze optimistic online learning algorithms under delayed feedback. We present a novel "delay as optimism" analysis that reduces online learning under delay to optimistic online learning. This reduction enables optimal regret bounds for delayed online learning and exposes how side-information or optimistic "hints" can be used to combat the effects of delay. We use these theoretical tools to develop the first optimistic online learning algorithms that require no parameter tuning and have optimal regret guarantees under delay. These algorithms — DORM, DORM+, and AdaHedgeD— are robust and practical choices for real-world time-series forecasting. We conclude by benchmarking our algorithms on four subseasonal climate forecasting tasks, demonstrating low regret relative to state-of-the-art forecasting models.

1. Introduction

Online learning is a classical sequential decision-making paradigm in which a learner is pitted against a potentially adversarial environment (Orabona, 2019; Shalev-Shwartz, 2007). At time t, the learner must select a play $\mathbf{w}_t$ from some set of possible plays $\mathbf{W}$. The environment then reveals the loss function ℓ_t and the learner pays the cost $\ell_t(\mathbf{w}_t)$. The learner uses information collected in previous rounds to improve its plays in subsequent rounds. *Optimistic* online learners additionally make use of side-information or "hints" about expected future losses to improve their plays. Over a period of length T, the objective of the learner is to minimize *regret*, an objective that quantifies the performance gap between the learner and the best possible constant play in retrospect in some competitor set $\mathbf{U}$: $\text{Regret}_T = \sup_{u\in\mathbf{U}} \sum_{t=1}^{T} \ell_t(\mathbf{w}_t) - \ell_t(\mathbf{u})$. Adversarial online learning algorithms can provide robust performance in many complex real-world online prediction problems such as climate or financial forecasting.

In traditional online learning paradigms, the loss for round t is revealed to the learner immediately at the end of round t. However, many real-world applications produce delayed feedback, i.e., the loss for round t is not available until round $t + D$ for some delay period D. Several delayed algorithms are known to achieve optimal worst-case regret rates against adversarial loss sequences (Weinberger & Ordentlich, 2002; Joulani et al., 2013; McMahan & Streeter, 2014; Joulani et al., 2017), but each has its drawbacks when deployed for real applications with short horizons T. Some use only a small fraction of the data to train each learner (Weinberger & Ordentlich, 2002; Joulani et al., 2013); others rely on uniform upper bounds on future loss gradients to set their tuning parameters (McMahan & Streeter, 2014; Joulani et al., 2017). None leverage optimistic hints to improve performance when the delayed losses are partially predictable. The concurrent work of Hsieh et al. (2020) analyzes optimistic gradient descent under delay but relies on uniform bounds on future gradients that are often challenging to obtain and overly conservative in applications.

In this work, we aim to develop robust and practical algorithms for real-world delayed online learning. To this end, we introduce three novel algorithms — DORM, DORM+, and AdaHedgeD— that use every observation to train the learner, have no parameters to tune, exhibit optimal worst-case regret rates under delay, *and* enjoy improved performance when accurate hints for unobserved past and future losses are available. We begin by viewing delayed online learning as a special case of optimistic online learning and use this "delay as optimism" perspective to develop:

1. A formal reduction of delayed online learning to optimistic online learning (Lems. 1 and 2).
2. The first optimistic tuning-free and self-tuning algorithms with optimal regret guarantees under delay (DORM, DORM+, and AdaHedgeD).
3. A tightening of standard optimistic online learning regret bounds that reveals the robustness of optimistic al-

[1]Dept. of EECS, Massachusetts Institute of Technology [2]Dept. of ECE, Boston University [3]Atmospheric and Environmental Research [4]Dept. of CS, University of Toronto [5]Microsoft Research New England [6]Instituto de Matemática Pura e Aplicada. Correspondence to: Genevieve Flaspohler <geflaspo@mit.edu>.

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

gorithms to inaccurate hints (Thms. 3 and 4 of App. B).

4. The first general analysis of follow-the-regularized leader algorithms under delay (Thm. 5 of App. B and Thm. 18 of App. I).
5. The first analysis of delayed online mirror descent algorithms with optimistic hints (Thm. 6 of App. B).

We validate our algorithms on the problem of subseasonal forecasting in Sec. 5. Subseasonal forecasting — predicting precipitation and temperature 2-6 weeks in advance — is a crucial task for allocating water resources, managing wildfires, and preparing for other weather extremes (White et al., 2017). Several challenges emerge when applying existing online learning methods to subseasonal forecasting that our algorithms are equipped to manage. First, real-time subseasonal forecasting suffers from delayed feedback: multiple forecasts are issued before receiving feedback on the first. Second, the regret horizons are short: a common evaluation period for semimonthly forecasting is one year, resulting in 26 total forecasts. Third, self-tuned or tuning-free algorithms are essential for real-time, practical deployment. We demonstrate that our algorithms DORM, DORM+, and AdaHedgeD all produce strong performance and that DORM+ particularly achieves consistently low regret compared to the best forecasting models.

Open-source Python code implementing DORM, DORM+ and AdaHedgeD and recreating our subseasonal forecasting experiments is available at `redacted`.

Notation For integers a, b, we use the shorthand $\mathbf{g}_{a:b} \triangleq \sum_{i=a}^{b} \mathbf{g}_i$. We say a function f is proper if it is somewhere finite and never $-\infty$. We let $\partial f(\mathbf{w}) = \{\mathbf{g} \in \mathbb{R}^d : f(\mathbf{u}) \geq f(\mathbf{w}) + \langle \mathbf{g}, \mathbf{u} - \mathbf{w} \rangle, \ \forall \mathbf{u} \in \mathbb{R}^d\}$ denote the set of *subgradients* of f at $\mathbf{w} \in \mathbb{R}^d$ and say f is *μ-strongly convex* over a convex set $\mathbf{W} \subseteq \operatorname{int} \operatorname{dom} f$ with respect to $\|\cdot\|$ if $\forall \mathbf{w}, \mathbf{u} \in \mathbf{W}$ and $\mathbf{g} \in \partial f(\mathbf{w})$, we have $f(\mathbf{u}) \geq f(\mathbf{w}) + \langle \mathbf{g}, \mathbf{u} - \mathbf{w} \rangle + \frac{\mu}{2}\|\mathbf{w} - \mathbf{u}\|^2$. For differentiable ψ, we define the Bregman divergence $\mathcal{B}_\psi(\mathbf{w}, \mathbf{u}) \triangleq \psi(\mathbf{w}) - \psi(\mathbf{u}) - \langle \nabla \psi(\mathbf{u}), \mathbf{w} - \mathbf{u} \rangle$. We define $\operatorname{diam}(\mathbf{W}) = \inf_{\mathbf{w}, \mathbf{w}' \in \mathbf{W}} \|\mathbf{w} - \mathbf{w}'\|$ and $(r)_+ \triangleq \max(r, 0)$.

2. Online Learning with Optimism and Delay

Standard online learning algorithms, such as follow the regularized leader (FTRL) and online mirror descent (OMD) achieve optimal worst-case regret against adversarial loss sequences (Orabona, 2019). However, many loss sequences encountered in applications are not truly adversarial. *Optimistic* online learning algorithms aim to achieve improved performance when loss sequences are partially predictable, while maintaining robustness to adversarial sequences (see, e.g., Rakhlin & Sridharan, 2013b; Steinhardt & Liang, 2014; Kamalaruban, 2016; Chiang et al., 2012). In many formulations of optimistic online learning, the learner is provided with a pseudo-loss $\tilde{\ell}_t$ at the start of round t that represents a guess for the true, unknown loss at time t. The online learner can incorporate this hint into its learning process before making play $\mathbf{w}_t$. When loss feedback is delayed by D time steps, the learner observes the losses $\{\ell_s\}_{s=1}^{t-D-1}$ and the optimistic pseudolosses $\{\tilde{\ell}_s\}_{s=1}^{t}$ before playing $\mathbf{w}_t$.

In the remainder of the text, we use the following notation for the subdifferential of the online learning loss and optimistic pseudoloss respectively: $\mathbf{g}_t \in \partial \ell_t(\mathbf{w}_t)$, $\tilde{\mathbf{g}}_t \in \partial \tilde{\ell}_t(\mathbf{w}_{t-1})$.

In the delayed and optimistic setting, we propose counterparts of standard FTRL and OMD online learning algorithms, which we call *optimistic delayed FTRL (*ODFTRL*)* and *delayed optimistic online mirror descent (*DOOMD*)* respectively. These algorithms produce iterates $\mathbf{w}_t$ satisfying,

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w} \in \mathbf{W}} \langle \mathbf{g}_{1:t-D} + \mathbf{h}_{t+1}, \mathbf{w} \rangle + \lambda \psi(\mathbf{w}) \qquad \text{(ODFTRL)}$$

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w} \in \mathbf{W}} \langle \mathbf{g}_{t-D} + \mathbf{h}_{t+1} - \mathbf{h}_t, \mathbf{w} \rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_t)$$

$$\text{with} \quad \mathbf{h}_0 \triangleq \mathbf{0} \quad \text{and arbitrary} \quad \mathbf{w}_0. \qquad \text{(DOOMD)}$$

for constant delay period D, regularization parameter λ, and optimistic hint vector $\mathbf{h}_t = \sum_{s=t-D}^{t} \tilde{\mathbf{g}}_s$, representing our best guess of the summed gradients of missing delayed and future losses.

2.1. Delay as Optimism

A first key insight of this paper is that, for ODFTRL and DOOMD,

> *Learning with delay is a special case of learning with optimism.*

In particular, ODFTRL and DOOMD are instances of optimistic FTRL (OFTRL) and single-step optimistic OMD (SOOMD) respectively with a particularly "bad" choice of optimistic hint $\tilde{\mathbf{g}}_{t+1}$ that deletes the unobserved loss subgradients $\mathbf{g}_{t-D+1:t}$.

Lemma 1 (ODFTRL is OFTRL with a bad hint)**.** *ODFTRL is OFTRL with* $\tilde{\mathbf{g}}_{t+1} = \mathbf{h}_{t+1} - \sum_{s=t-D+1}^{t} \mathbf{g}_s$.

Lemma 2 (DOOMD is SOOMD with a bad hint)**.** *DOOMD is SOOMD with* $\tilde{\mathbf{g}}_{t+1} = \tilde{\mathbf{g}}_t + \mathbf{g}_{t-D} - \mathbf{g}_t + \mathbf{h}_{t+1} - \mathbf{h}_t = \mathbf{h}_{t+1} - \sum_{s=t-D+1}^{t} \mathbf{g}_s$.

In App. B, we demonstrate that, as an immediate consequence of our delay-as-optimism perspective, we can provide new regret bounds for ODFTRL and DOOMD. The form of these delayed regret bounds reveals the heightened value of optimism in the presence of delay: in addition to providing an effective guess of a subgradient $\mathbf{g}_t$, an optimistic hint can approximate the missing delayed feedback ($\sum_{s=t-D}^{t-1} \mathbf{g}_s$) and thereby significantly reduce the penalty of delay. If, on the other hand, the hints are a poor proxy for the missing loss subgradients, we still only pay the minimax optimal $\sqrt{D+1}$ penalty for delayed feedback. It remains to choose the regularization parameter λ to achieve this minimax optimal regret rate.

3. Tuning-free Learning with Optimism and Delay

As an application of our ODFTRL and DOOMD analysis, we introduce and analyze delayed and optimistic versions of two popular tuning-free online learning algorithms: regret matching (RM) (Blackwell, 1956; Hart & Mas-Colell, 2000) and regret matching+ (RM+) (Tammelin et al., 2015). RM was developed to find correlated equilibria in two-player games and is commonly used to minimize regret over the simplex. RM+ is a modification of RM designed to accelerate convergence and used to solve the game of Heads-up Limit Texas Hold'em poker (Bowling et al., 2015).

Our generalizations, *delayed optimistic regret matching* (DORM)

$$\mathbf{w}_{t+1} = \tilde{\mathbf{w}}_{t+1}/\langle \mathbf{1}, \tilde{\mathbf{w}}_{t+1}\rangle \quad \text{for} \qquad \text{(DORM)}$$
$$\tilde{\mathbf{w}}_{t+1} \triangleq \max(\mathbf{0}, (\mathbf{r}_{1:t-D} + \mathbf{h}_{t+1})/\lambda)^{q-1} \quad \text{and}$$
$$\mathbf{r}_{t-D} \triangleq \mathbf{1}\langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D}\rangle - \mathbf{g}_{t-D}$$

and *delayed optimistic regret matching+* (DORM+)

$$\mathbf{w}_{t+1} = \tilde{\mathbf{w}}_{t+1}/\langle \mathbf{1}, \tilde{\mathbf{w}}_{t+1}\rangle \quad \text{for} \qquad \text{(DORM+)}$$
$$\tilde{\mathbf{w}}_{t+1} \triangleq \max\left(\mathbf{0}, \tilde{\mathbf{w}}_t^{p-1} + (\mathbf{r}_{t-D} + \mathbf{h}_{t+1} - \mathbf{h}_t)/\lambda\right)^{q-1},$$
$$\mathbf{r}_{t-D} \triangleq \mathbf{1}\langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D}\rangle - \mathbf{g}_{t-D}, \quad \mathbf{h}_0 \triangleq \mathbf{0}, \quad \tilde{\mathbf{w}}_0 \triangleq \mathbf{0},$$

allow for delay D, regularization parameter λ, optimistic hints $\mathbf{h}_t$, and a parameter $q \geq 2$ and its conjugate exponent $p = q/(q-1)$. We refer to $\mathbf{r}_t$ as the *instantaneous regret* of each expert with respect to the play $\mathbf{w}_t$ and the linearized loss vector $\mathbf{g}_t$, and note that DORM and DORM+ recover the standard RM and RM+ algorithms when $D = 0$, $\lambda = 1$, $q = 2$, and $\mathbf{h}_t = \mathbf{0}$, $\forall t$.

While these updates may look unfamiliar, we show in App. E that they are special cases of the ODFTRL and DOOMD algorithms. Specifically, we connect DORM to ODFTRL and DORM+ to DOOMD, which enables us to extend previous regret bounds to DORM and DORM+. Additionally, under mild conditions detailed in App. E, we highlight a remarkable property:

> The normalized DORM and DORM+ iterates $\mathbf{w}_t$ are *independent* of the choice of regularization parameter λ.

This result, shown in App. E, implies that DORM and DORM+ are *automatically* optimally tuned with respect to λ, even when run with a default value of $\lambda = 1$.

4. Self-tuned Learning with Optimism and Delay

We now analyze an adaptive version of ODFTRL with time-varying regularization $\lambda_t \psi$ and develop strategies for automatically tuning λ_t in the presence of optimism and delay. Our objective is to achieve the minimax optimal regret rate *and* to find a setting of λ_t that performs well in pratical applications. As noted by Erven et al. (2011); de Rooij et al. (2014); Orabona (2019), the effectiveness of an adaptive regularization setting λ_t that uses an upper bound on regret relies heavily on the tightness of that bound. Our next result introduces analyzes a new tuning strategy inspired by the popular AdaHedge algorithm (Erven et al., 2011) and based on a new tighter bound on ODFTRL regret:

Fix $\alpha > 0$, and consider the *delayed AdaHedge-style* (AdaHedgeD) regularizer sequence within an ODFTRL update:

$$\lambda_{t+1} = \tfrac{1}{\alpha}\textstyle\sum_{s=1}^{t-D} \delta_t \qquad \text{for} \qquad \text{(AdaHedgeD)}$$
$$\delta_t \triangleq \min\left(F_{t+1}(\mathbf{w}_t, \lambda_t) - F_{t+1}(\bar{\mathbf{w}}_t, \lambda_t), \langle \mathbf{g}_t, \mathbf{w}_t - \bar{\mathbf{w}}_t\rangle\right)_+$$
$$\text{with} \quad \bar{\mathbf{w}}_t = \operatorname{argmin}_{\mathbf{w}\in\mathbf{W}} F_{t+1}(\mathbf{w}, \lambda_t) \qquad (1)$$
$$\text{and} \quad F_{t+1}(\mathbf{w}, \lambda_t) \triangleq \lambda_t \psi(\mathbf{w}) + \langle \mathbf{g}_{1:t}, \mathbf{w}\rangle.$$

Remarkably, as we show in App. I, this setting of adaptive regularization yields a minimax optimal $\mathcal{O}(\sqrt{(D+1)T} + D)$ dependence on the delay parameter and nearly matches the regret of the optimal constant λ tuning in hindsight.

5. Experiments

We apply the online learning techniques developed in this paper to the problem of adaptive ensembling for subseasonal forecasting. Our experiments are based on the public subseasonal forecasting codebase of Flaspohler et al. (2021) that uses $d = 6$ physics-based numerical models and machine learning models (CFSv2++, Climatology++, LocalBoosting , MultiLLR , Persistence++, and Salient++) to predict temperature and precipitation 2-6 weeks ahead. In this mid-range climate forecasting task, forecast feedback is delayed; the models make $D = 2$ or 3 forecasts depending on the forecast horizon before receiving feedback. We use delayed, optimistic online learning to play a time-varying convex combination of the d input models, such that $\mathbf{w}_t \in \triangle_{d-1}$. Our objective is to compete with the best input model over a year-long prediction period ($T = 26$ semimonthly dates). The loss function for each forecast date is the geographic root-mean squared error (RMSE) across 514 locations in the Western United States.

We consider four subseasonal prediction tasks – predicting temperature and precipitation at two horizons, weeks 3-4 and weeks 5-6 – and evaluate yearly regret and mean RMSE for each year from 2011-2020. Unless otherwise specified, all online learning algorithms use the previous gradient

Table 1: **Average RMSE of the 2011-2020 semimonthly forecasts**: The average RMSE for online learning algorithms (left) and individual models (right) over a 10-year evaluation period. The top performing model for each task is bolded and shown in green.

	AdaHedgeD	DORM	DORM+	CFSv2++	Clim.++	LocalBoosting	MultiLLR	Persist.++	Salient++
Precip. 34w	21.837	21.737	**21.675**	21.978	21.986	22.357	22.431	21.973	23.344
Precip. 56w	21.987	21.957	**21.838**	22.004	21.993	22.383	22.570	22.030	23.257
Temp. 34w	2.287	2.259	**2.247**	2.277	2.319	2.394	2.352	2.253	2.508
Temp. 56	2.321	2.318	2.304	**2.278**	2.317	2.440	2.368	2.284	2.569

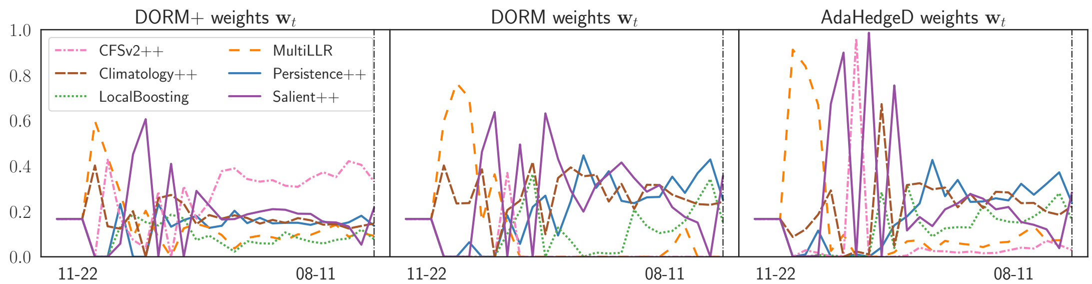

Figure 1: **Impact of regularization:** The plays $\mathbf{w}_t$ of online learning algorithms used to combine the input models for the Temp. 34w task in the 2020 evaluation year. DORM and AdaHedgeD are both FTRL-based algorithms and have similar plays; AdaHedgeD appears to be less regularized. DORM+, on the other hand, is an OMD-based algorithm and was designed to handle applications where the "best" expert model changes frequently. DORM+ is the top learning algorithm for this subseasonal forecasting task, indicating the importance of this adaptivity property.

optimism strategy $\mathbf{h}_t = (D+1)\mathbf{g}_{t-D-1}$. See App. M for full experimental details and App. N for algorithmic details.

5.1. Competing with the best input model

Our primary objective in online learning is to achieve zero average regret, i.e., to perform as well as the best input model in the competitor set $\mathbf{U}$. To evaluate model regret, we run our three delayed online learning algorithms — DORM, DORM+, and AdaHedgeD— on all four subseasonal prediction tasks and measure their average RMSE loss.

The average yearly RMSE for the three online learning algorithms and the six input models is shown in Table 1. The DORM+ algorithm outperforms the best input model for all tasks except Temp. 56w. All online learning algorithms achieve negative regret for both precipitation tasks. Fig. 1 shows an example of the weights played by the three algorithms. Fig. 2 shows the yearly cumulative regret (in terms of the RMSE loss) of the online learning algorithms over the 10-year evaluation period. There are several years (e.g., 2012, 2014, 2020) in which all online learning algorithms achieve negative regret, outperforming the best input forecasting model. The consistently low regret year-to-year of DORM+ makes it a promising candidate for real-world delayed subseasonal forecasting.

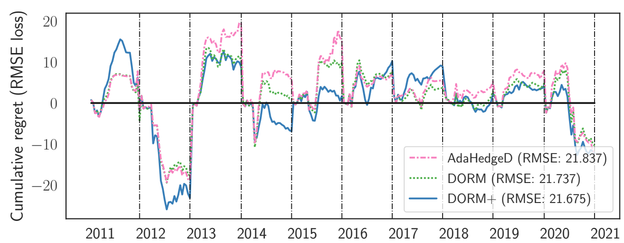

Figure 2: **Overall performance**: Yearly cumulative regret under RMSE loss for the three delayed online learning algorithms presented, over the 10-year evaluation period for the Precip. 34w task. The zero line corresponds to the performance of the best input model in a given year. Negative values indicate that the online learner outperformed the best input model in a given year.

6. Conclusion

In this work, we overcame the challenges of delayed feedback and short regret horizons in online learning with optimism. We developed three practical non-replicated, self-tuned and tuning-free algorithms with optimal regret guarantees — DORM, DORM+, and AdaHedgeD. Our "delay as optimism" reduction and refined analysis of optimistic learning produced novel regret bounds for both optimistic and delayed online learning and elucidated the connections between these two problems. Within the subseasonal forecasting domain, we demonstrated that delayed online learning methods can produce state-of-the art forecasting ensembles robustly from year-to-year. Our results highlighted DORM+ as a particularly promising candidate for subseasonal forecasting due to its tuning-free nature and adaptivity when the best input model changes frequently. Through theoretical and experimental validation, we have presented DORM, DORM+, and AdaHedgeD as practical and robust algorithms for delayed time-series forecasting that can be applied in a variety of application domains to improve the quality of sequential decision-making.

References

Agarwal, A. and Duchi, J. C. Distributed delayed stochastic optimization. In Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011.

Blackwell, D. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.

Bowling, M., Burch, N., Johanson, M., and Tammelin, O. Heads-up limit hold'em poker is solved. *Science*, 347(6218):145–149, 2015. ISSN 0036-8075. doi: 10.1126/science.1259433.

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

Chiang, C.-K., Yang, T., Lee, C.-J., Mahdavi, M., Lu, C.-J., Jin, R., and Zhu, S. Online optimization with gradual variations. In Mannor, S., Srebro, N., and Williamson, R. C. (eds.), *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23, pp. 6.1–6.20, Edinburgh, Scotland, 25–27 Jun 2012.

de Rooij, S., van Erven, T., Grünwald, P. D., and Koolen, W. M. Follow the leader if you can, hedge if you must. *Journal of Machine Learning Research*, 15(37):1281–1316, 2014.

Erven, T., Koolen, W. M., Rooij, S., and Grünwald, P. Adaptive hedge. In Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems*, volume 24, pp. 1656–1664. Curran Associates, Inc., 2011.

Flaspohler, G., Orabona, F., Cohen, J., Mouatadid, S., Oprescu, M., Orenstein, P., and Mackey, L. Replication Data for: Online Learning with Optimism and Delay, 2021. URL https://doi.org/10.7910/DVN/IOCFCY.

Hart, S. and Mas-Colell, A. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

Hsieh, Y.-G., Iutzeler, F., Malick, J., and Mertikopoulos, P. Multi-agent online optimization with delays: Asynchronicity, adaptivity, and optimism. *arXiv preprint arXiv:2012.11579*, 2020.

Hwang, J., Orenstein, P., Cohen, J., Pfeiffer, K., and Mackey, L. Improving subseasonal forecasting in the western U.S. with machine learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2325–2335, 2019.

Joulani, P., Gyorgy, A., and Szepesvári, C. Online learning under delayed feedback. In *International Conference on Machine Learning*, pp. 1453–1461, 2013.

Joulani, P., Gyorgy, A., and Szepesvári, C. Delay-tolerant online convex optimization: Unified analysis and adaptive-gradient algorithms. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

Joulani, P., György, A., and Szepesvári, C. A modular analysis of adaptive (non-) convex optimization: Optimism, composite objectives, and variational bounds. In *International Conference on Algorithmic Learning Theory*, pp. 681–720. PMLR, 2017.

Kamalaruban, P. Improved optimistic mirror descent for sparsity and curvature. *arXiv preprint arXiv:1609.02383*, 2016.

Koolen, W., Van Erven, T., and Grunwald, P. Learning the learning rate for prediction with expert advice. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems*, volume 27, pp. 2294–2302. Curran Associates, Inc., 2014.

Korotin, A., V'yugin, V., and Burnaev, E. Adaptive hedging under delayed feedback. *Neurocomputing*, 397:356–368, 2020.

Liu, J. and Wright, S. J. Asynchronous stochastic coordinate descent: Parallelism and convergence properties. *SIAM Journal on Optimization*, 25(1):351–376, 2015.

Liu, J., Wright, S., Ré, C., Bittorf, V., and Sridhar, S. An asynchronous parallel stochastic coordinate descent algorithm. In *International Conference on Machine Learning*, pp. 469–477. PMLR, 2014.

McMahan, B. and Streeter, M. Delay-tolerant algorithms for asynchronous distributed online learning. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems*, volume 27, pp. 2915–2923. Curran Associates, Inc., 2014.

McMahan, H. B. A survey of algorithms and analysis for adaptive online learning. *The Journal of Machine Learning Research*, 18(1): 3117–3166, 2017.

McQuade, S. and Monteleoni, C. Global climate model tracking using geospatial neighborhoods. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 26, 2012.

Mesterharm, C. On-line learning with delayed label feedback. In *International Conference on Algorithmic Learning Theory*, pp. 399–413. Springer, 2005.

Mohri, M. and Yang, S. Accelerating online convex optimization via adaptive prediction. In *Artificial Intelligence and Statistics*, pp. 848–856. PMLR, 2016.

Monteleoni, C. and Jaakkola, T. Online learning of non-stationary sequences. In Thrun, S., Saul, L., and Schölkopf, B. (eds.), *Advances in Neural Information Processing Systems*, volume 16, pp. 1093–1100. MIT Press, 2004.

Monteleoni, C., Schmidt, G. A., Saroha, S., and Asplund, E. Tracking climate models. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 4(4):372–392, 2011.

Nesterov, Y. Efficiency of coordinate descent methods on huge-scale optimization problems. *SIAM Journal on Optimization*, 22(2): 341–362, 2012.

Nowak, K., Beardsley, J., Brekke, L. D., Ferguson, I., and Raff, D. Subseasonal prediction for water management: Reclamation forecast rodeo I and II. In *100th American Meteorological Society Annual Meeting*. AMS, 2020.

Orabona, F. A modern introduction to online learning. *ArXiv*, abs/1912.13213, 2019.

Orabona, F. and Pál, D. Scale-free algorithms for online linear optimization. In *International Conference on Algorithmic Learning Theory*, pp. 287–301. Springer, 2015.

Orabona, F. and Pál, D. Optimal non-asymptotic lower bound on the minimax regret of learning with expert advice. *arXiv preprint arXiv:1511.02176*, 2015.

Quanrud, K. and Khashabi, D. Online learning with adversarial delays. In Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 28, pp. 1270–1278, 2015.

Rakhlin, A. and Sridharan, K. Online learning with predictable sequences. In Shalev-Shwartz, S. and Steinwart, I. (eds.), *Proceedings of the 26th Annual Conference on Learning Theory*, pp. 993–1019. PMLR, 2013a.

Rakhlin, S. and Sridharan, K. Optimization, learning, and games with predictable sequences. In Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems*, pp. 3066–3074. Curran Associates, Inc., 2013b.

Recht, B., Re, C., Wright, S., and Niu, F. Hogwild!: A lock-free approach to parallelizing stochastic gradient descent. In Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems*, volume 24, pp. 693–701. Curran Associates, Inc., 2011.

Rockafellar, R. T. *Convex analysis*, volume 36. Princeton university press, 1970.

Shalev-Shwartz, S. *Online learning: Theory, algorithms, and applications*. PhD thesis, The Hebrew University, 2007.

Shalev-Shwartz, S. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.

Sra, S., Yu, A. W., Li, M., and Smola, A. AdaDelay: Delay adaptive distributed stochastic optimization. In Gretton, A. and Robert, C. C. (eds.), *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51, pp. 957–965. PMLR, 2016.

Steinhardt, J. and Liang, P. Adaptivity and optimism: An improved exponentiated gradient algorithm. In *International Conference on Machine Learning*, pp. 1593–1601, 2014.

Syrgkanis, V., Agarwal, A., Luo, H., and Schapire, R. E. Fast convergence of regularized learning in games. In Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.

Tammelin, O., Burch, N., Johanson, M., and Bowling, M. Solving heads-up limit texas hold'em. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

Weinberger, M. J. and Ordentlich, E. On delayed prediction of individual sequences. *IEEE Transactions on Information Theory*, 48(7): 1959–1976, 2002.

White, C. J., Carlsen, H., Robertson, A. W., Klein, R. J., Lazo, J. K., Kumar, A., Vitart, F., Coughlan de Perez, E., Ray, A. J., Murray, V., et al. Potential applications of subseasonal-to-seasonal (s2s) predictions. *Meteorological applications*, 24(3):315–325, 2017.

Zinkevich, M., Johanson, M., Bowling, M. H., and Piccione, C. Regret minimization in games with incomplete information. In Platt, J., Koller, D., Singer, Y., and Roweis, S. (eds.), *Advances in Neural Information Processing Systems*, volume 20. Curran Associates, Inc., 2007.