Multi-Agent Online Optimization with Delays

- **Yu-Guan Hsieh** (Université Grenoble Alpes, LJK)

Keywords: Online learning, multi-agent, delayed feedback, asynchronicity, adaptive methods.

Abstract: Online learning provides a powerful paradigm for modeling sequential decision making tasks, with various applications ranging from recommender systems, network routing, to trend forecasting. In its most basic form, a learner repeatedly takes actions, and the cost of an action, possibly along with some gradient-based feedback, is *immediately* revealed to the learner after that action is taken, *without* any delay. The goal of the learner is to minimize the cumulative cost of the taken actions.

Our work [1] extends the online learning framework to consider *multi-agent* environments with *asynchronous* activation and *delayed* feedback. These delays can be *i*) inherent to the problem, *ii*) due to computational overload, or *iii*) caused by communication latency between the agents. Formally, at each time slot, an agent becomes active, is required to make a prediction x_t, and suffers loss $f_t(x_t)$ where f_t is a unknown convex cost function. Meanwhile, the feedback element $g_t \in \partial f_t(x_t)$ only arrives at the agents after a certain delay that varies across the agents. The performance of a learning algorithm up to time T is measured by the regret

$$\mathbf{Reg}_T(u) = \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(u)$$

for u an arbitrary comparator action. This setup is relevant for wireless sensor networks or any online learning algorithm deployed in a distributed manner. Two instances covered by the setup are illustrated in Figure 1.

Subsequently, we provide *adaptive* algorithms based on Nesterov's dual averaging template that achieve optimal regret bound in this challenging situation. Importantly, our method only exploits local information available to each agent and does not require any synchronization. In particular, there is no need to maintain a global clock for the updates of the agents. These results extend a series of work [2, 3] on adaptive methods for delayed online learning to the multi-agent setup.

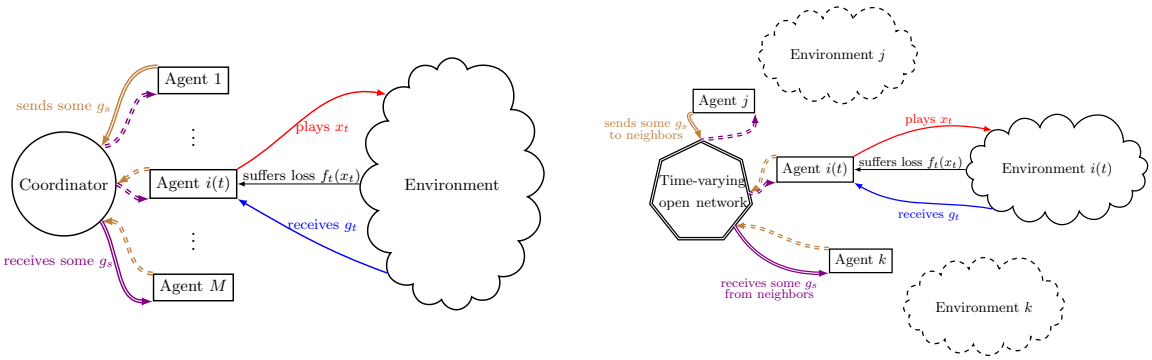

Figure 1: In the considered setup, a network of agents collaborate to minimize the total regret. We do not put any restriction on how the feedback is actually communicated. This can for example be done either through a coordinator-worker structure (left) or a decentralized open network (right).

References:

[1] Yu-Guan Hsieh, Franck Iutzeler, Jérôme Malick, and Panayotis Mertikopoulos. Multi-agent online optimization with delays: Asynchronicity, adaptivity, and optimism. To appear in *The Journal of Machine Learning Research*, 2022.

[2] Pooria Joulani, Andras Gyorgy, and Csaba Szepesvári. Delay-tolerant online convex optimization: Unified analysis and adaptive-gradient algorithms. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[3] Brendan McMahan and Matthew Streeter. Delay-tolerant algorithms for asynchronous distributed online learning. In *NIPS '14: Proceedings of the 28th International Conference on Neural Information Processing Systems*, pages 2915–2923, 2014