Faster No-Regret Learning for Extensive-Form Correlated Equilibria

Ioannis Anagnostides[†], **Gabriele Farina**[†], **Christian Kroer**[¶],
Andrea Celli[‡], **Tuomas Sandholm**[†,⋆,&,♣]

[†]Carnegie Mellon University, [¶]Columbia University, [‡]Bocconi University, [⋆]Strategy Robot,
[&]Optimized Markets, [♣]Strategic Machine

Introduction

Nash and Correlated Equilibria

Nash equilibrium has been central in many recent landmark results in AI revolving around zero-sum game solving. However, it suffers from many drawbacks in multiplayer general-sum games:

- Equilibrium selection: An equilibrium strategy may perform poorly against the "wrong" equilibrium
- Computational Intractability: Nash equilibria are hard to compute in general games [5]

A competing notion of rationality, proposed by Aumann [1], is that of **correlated equilibrium**.

Unlike Nash equilibrium, there are uncoupled no-regret learning algorithms converging to correlated equilibria in general games.

Accelerated No-Regret Learning Dynamics for Correlated Equilibria in Normal-Form Games

There has been a considerable amount of interest in developing faster no-regret learning algorithms in normal-form games that outperform the adversarial $\Theta(T^{-1/2})$ barrier:

- [3, 6]: $\widetilde{O}(T^{-1})$ convergence to Nash equilibria in zero-sum games
- [7]: $O(T^{-3/4})$ convergence to coarse correlated equilibrium
- [2]: $O(T^{-3/4})$ convergence to correlated equilibria
- [4]: $\widetilde{O}(T^{-1})$ convergence to coarse correlated equilibria

Much less is known for extensive-form games!

Open Question

Are there faster no-regret learning dynamics for extensive-form correlated equilibria?

Extensive-Form Games

Extensive-form games substantially generalize normal-form games by allowing both simultaneous and sequential moves, as well as imperfect-information. Most strategic interactions in real-world applications involve imperfect-information.

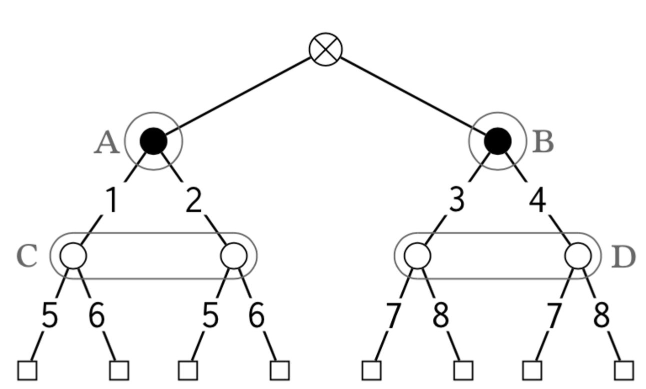


Fig. 1. Example of a two-player EFG.

Theoretical Results

Extensive-Form Correlated Equilibrium

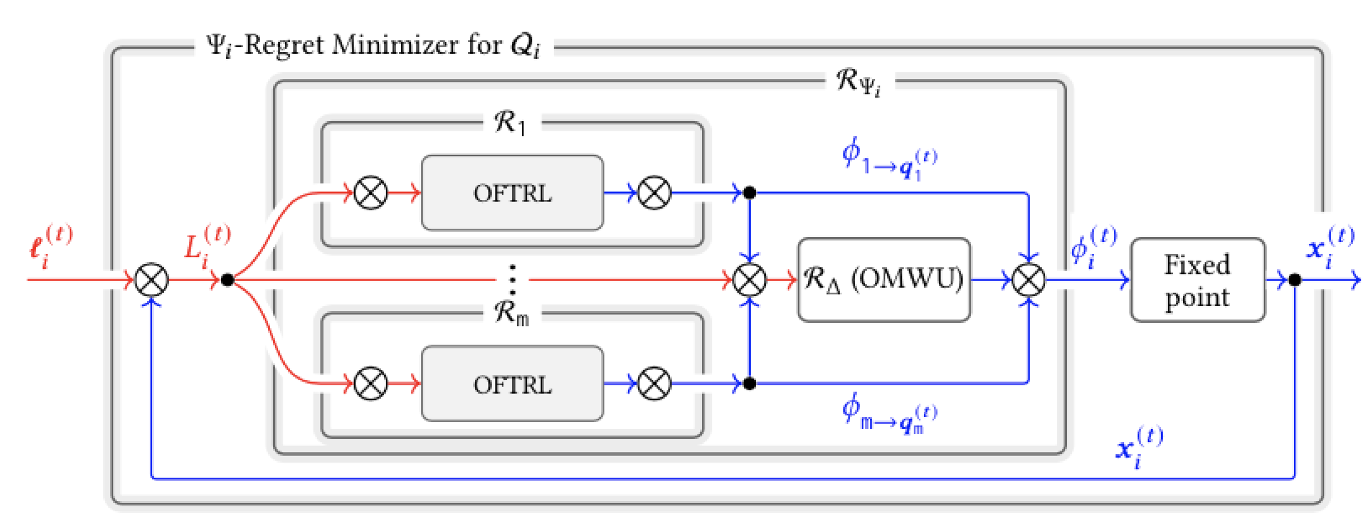
In an **extensive-form correlated equilibrium (EFCE)**, recommended actions are gradually revealed to players as they reach infosets. Thus, the mediator must take into account the evolution of beliefs throughout the game.

Main Result

Theorem 1. *On any perfect-recall general-sum multiplayer extensive-form game, there exist uncoupled no-regret learning dynamics which lead to a correlated distribution of play that is an $O(T^{-3/4})$-approximate . Here the $O(\cdot)$ notation suppresses game-specific parameters polynomial in the size of the game.*

This substantially improves over the prior best rate of $O(T^{-1/2})$.

Our Construction



Main Ingredients

We develop a general template for performing accelerated Phi-regret minimization.

Phi-regret is a powerful framework for **hindsight rationality**. To employ our template we establish the following components:

1. Predictive regret minimizer for the set of trigger deviation functions
2. Stability analysis of the fixed points

Experiments

We support our theory with experiments on benchmark games:

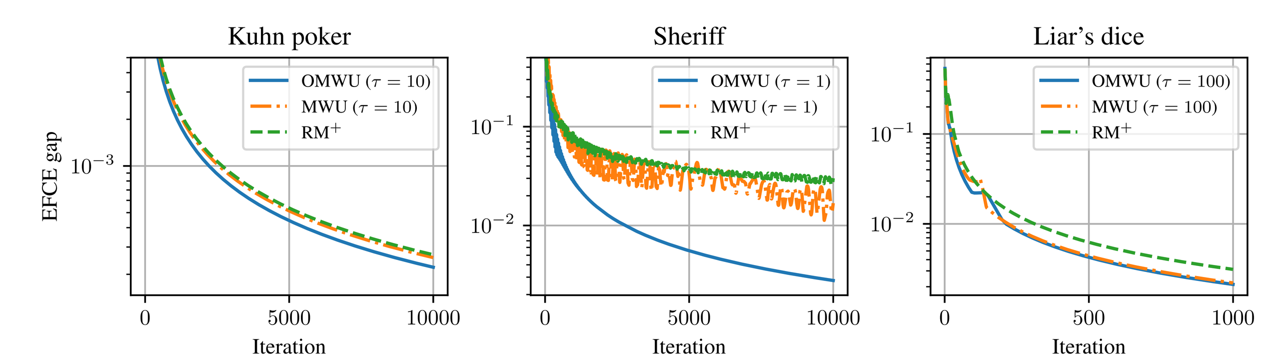
- 3-player Kuhn poker
- 2-player Sheriff
- 2-player Liar's dice

We investigate the performance of dynamics with a CFR decomposition under 3 different local regret minmizers:

(i) multiplicative weights (MW)

(ii) optimistic multiplicative weights (OMW)

(iii) regret matching$^+$



The y-axis illustrates the EFCE-gap. Surprisingly, we observe that OMW substantially outperforms regret matching$^+$ on Sheriff, a game specifically introduced for its interesting correlated equilibria.

References

[1] Robert Aumann. "Subjectivity and Correlation in Randomized Strategies". In: *Journa Mathematical Economics* 1 (1974), pp. 67–96.

[2] Xi Chen and Binghui Peng. "Hedging in games: Faster convergence of external and s regrets". In: 2020.

[3] Constantinos Daskalakis, Alan Deckelbaum, and Anthony Kim. "Near-optimal no-regre gorithms for zero-sum games". In: 92 (2015), pp. 327–348.

[4] Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. "Near-Optimal No-Re Learning in General Games". In: *CoRR* abs/2108.06924 (2021).

[5] Constantinos Daskalakis, Paul Goldberg, and Christos Papadimitriou. "The Complexit Computing a Nash Equilibrium". In: 2006.

[6] Alexander Rakhlin and Karthik Sridharan. "Online Learning with Predictable Sequences" *Conference on Learning Theory*. 2013, pp. 993–1019.

[7] Vasilis Syrgkanis et al. "Fast convergence of regularized learning in games". In: *Advance Neural Information Processing Systems*. 2015, pp. 2989–2997.