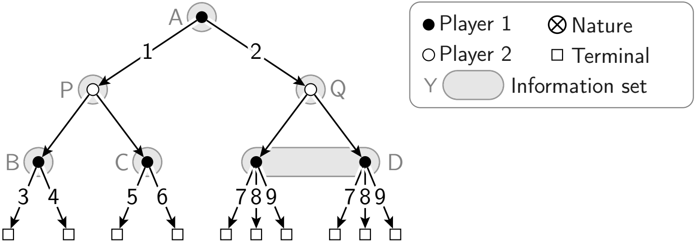
Polynomial-Time Linear-Swap Regret Minimization in Imperfect-Information Sequential Games

Gabriele Farina, **Charilaos (Charis) Pipis**
MIT
{gfarina, chpipis}@mit.edu

Extensive-form games (EFGs)



Hindsight Rationality and Learning in Games

Given a set Φ of strategy transformations $\phi : \mathcal{X} \to \mathcal{X}$, a Φ-regret minimizer is an *online learning* algorithm that minimizes

$$\Phi\text{-Reg}^{(T)} := \max_{\phi \in \Phi} \sum_{t=1}^{T} u^{(t)}(\phi(\boldsymbol{x}^{(t)})) - \sum_{t=1}^{T} u^{(t)}(\boldsymbol{x}^{(t)})$$

Hindsight Rationality and Learning in Games

Given a set Φ of strategy transformations $\phi : \mathcal{X} \to \mathcal{X}$, a Φ-regret minimizer is an *online learning* algorithm that minimizes

$$\Phi\text{-Reg}^{(T)} := \max_{\phi \in \Phi} \sum_{t=1}^{T} u^{(t)}(\phi(\boldsymbol{x}^{(t)})) - \sum_{t=1}^{T} u^{(t)}(\boldsymbol{x}^{(t)})$$

The size of the set Φ is a natural notion of rationality.

Hindsight Rationality and Learning in Games

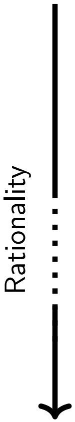
Given a set Φ of strategy transformations $\phi : \mathcal{X} \to \mathcal{X}$, a Φ-regret minimizer is an *online learning* algorithm that minimizes

$$\Phi\text{-Reg}^{(T)} := \max_{\phi \in \Phi} \sum_{t=1}^{T} u^{(t)}(\phi(\boldsymbol{x}^{(t)})) - \sum_{t=1}^{T} u^{(t)}(\boldsymbol{x}^{(t)})$$

The size of the set Φ is a natural notion of rationality.

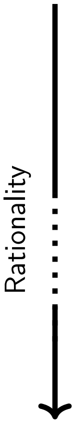
Celebrated Result: If all players of the game are Φ-regret minimizers, then the empirical frequency of play converges to the set of Φ-equilibria.

A spectrum of notable regret minimizers



A spectrum of notable regret minimizers

external regret: Φ contains all constant transformations $\to$ CCE



A spectrum of notable regret minimizers

Rationality

external regret: Φ contains all constant transformations $\rightarrow$ CCE

trigger regret: Φ contains all trigger deviations $\rightarrow$ EFCE

A spectrum of notable regret minimizers

Rationality

external regret: Φ contains all constant transformations $\rightarrow$ CCE

trigger regret: Φ contains all trigger deviations $\rightarrow$ EFCE

swap regret: Φ contains all possible transformations $\rightarrow$ CE

A spectrum of notable regret minimizers

Rationality

external regret: Φ contains all constant transformations $\rightarrow$ CCE

trigger regret: Φ contains all trigger deviations $\rightarrow$ EFCE

swap regret: Φ contains all possible transformations $\rightarrow$ CE
(efficient algorithm not known in EFGs)

A spectrum of notable regret minimizers

Rationality ↓

external regret: Φ contains all constant transformations $\to$ CCE

trigger regret: Φ contains all trigger deviations $\to$ EFCE

swap regret: Φ contains all possible transformations $\to$ CE
(efficient algorithm not known in EFGs)

Big Challenge: Largest *tractable* Φ in extensive-form games?

A spectrum of notable regret minimizers

Rationality

external regret: Φ contains all constant transformations $\rightarrow$ CCE

trigger regret: Φ contains all trigger deviations $\rightarrow$ EFCE

linear-swap regret (this work)

swap regret: Φ contains all possible transformations $\rightarrow$ CE
(efficient algorithm not known in EFGs)

Big Challenge: Largest *tractable* Φ in extensive-form games?

This work: Φ is the set of all linear transformations in EFGs

Let $\mathcal{Q}$ be the set of all *reduced* sequence-form strategies. Then

$$\Phi := \{\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} : \boldsymbol{A} \in \mathbb{R}^{d \times d}, \text{ with } \boldsymbol{A}\boldsymbol{x} \in \mathcal{Q} \quad \forall\, \boldsymbol{x} \in \mathcal{Q}\}.$$

Resulting notion of regret is *"linear-swap regret"*.

This work: Φ is the set of all linear transformations in EFGs

Let $\mathcal{Q}$ be the set of all *reduced* sequence-form strategies. Then

$$\Phi := \{\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} : \boldsymbol{A} \in \mathbb{R}^{d \times d}, \text{ with } \boldsymbol{A}\boldsymbol{x} \in \mathcal{Q} \quad \forall\, \boldsymbol{x} \in \mathcal{Q}\}.$$

Resulting notion of regret is *"linear-swap regret"*.

Why care?

- in EFGs it contains the trigger deviations used for EFCE

This work: Φ is the set of all linear transformations in EFGs

Let $\mathcal{Q}$ be the set of all *reduced* sequence-form strategies. Then

$$\Phi := \{\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} : \boldsymbol{A} \in \mathbb{R}^{d \times d}, \text{ with } \boldsymbol{A}\boldsymbol{x} \in \mathcal{Q} \quad \forall\, \boldsymbol{x} \in \mathcal{Q}\}.$$

Resulting notion of regret is *"linear-swap regret"*.

Why care?

- in EFGs it contains the trigger deviations used for EFCE
- in NFGs it is equal to the swap deviations and gives CE
- in Bayesian games with a learner vs optimizer, it is necessary to have at least a linear-swap regret minimizer [1]

---

[1] Mansour, Mohri, Schneider, Sivan (2022)

Algorithm for linear-swap regret minimization

Framework [1] to construct a Φ-regret minimizer:

Φ-regret minimizer on action set $\mathcal{X}$.	$\Longrightarrow$	External regret minimizer on action set Φ.

[1] Gordon, Greenwald, Marks (2008)

Algorithm for linear-swap regret minimization

Framework [1] to construct a Φ-regret minimizer:

Φ-regret minimizer on action set $\mathcal{X}$.	$\Longrightarrow$	External regret minimizer on action set Φ.

But the framework alone does *not* give us *efficient* algorithms.

[1] Gordon, Greenwald, Marks (2008)

Algorithm for linear-swap regret minimization

Framework [1] to construct a Φ-regret minimizer:

Φ-regret minimizer on action set $\mathcal{X}$. $\Longrightarrow$ External regret minimizer on action set Φ.

But the framework alone does *not* give us *efficient* algorithms.

Crucial step: We characterize the set of linear-swap transformations using poly num. of constraints.

[1] Gordon, Greenwald, Marks (2008)

Algorithm for linear-swap regret minimization

Framework [1] to construct a Φ-regret minimizer:

Φ-regret minimizer on action set $\mathcal{X}$.	$\Longrightarrow$	External regret minimizer on action set Φ.

But the framework alone does *not* give us *efficient* algorithms.

Crucial step: We characterize the set of linear-swap transformations using poly num. of constraints.

- the characterization depends on the structure of the game tree
- allows us to directly apply online convex optimization methods

[1] Gordon, Greenwald, Marks (2008)

Linear-Deviation Correlated Equilibrium (LCE)

Strictly between normal-form and extensive-form correlated equilibrium:

$$\text{CE} \subset \textbf{LCE} \subset \text{EFCE}$$

And, NP-hard to maximize Social Welfare, by a reduction from SAT.

Conclusion

- We construct a **polynomial-time** algorithm for **linear-swap regret** minimization in **EFGs**.
  - A stronger notion of sequential hindsight rationality that can be efficiently computed.
- These dynamics induce the **linear-deviation correlated equilibrium**.
  - It lies *strictly* between CE and EFCE.
  - Equilibrium selection is NP-hard.