Online Learning in Games

Prof. Volkan Cevher
volkan.cevher@epfl.ch

Lecture 3: Game theory and online learning dynamics (Part II)

Laboratory for Information and Inference Systems (LIONS)
École Polytechnique Fédérale de Lausanne (EPFL)

EE-735 (Spring 2024)

lions@epfl swisscom E4S Enterprise for Society EPFL

Acknowledgements

These slides would not have been possible without the help of

Kimon Antonakopoulos, Thomas Pethick and Stratis Skoulakis

License Information for Online Learning in Games Slides

Outline of this lecture

Online learning in potential games

◦ Potential games capture cooperative settings at which agents share common interests.

Definition (Potential games Monderer et al. [1])

An N-player game is a *potential game* if and only if a function $\Phi : \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \mapsto \mathbb{R}$ such that for each agent $i \in [N]$,

$$\Phi(x_i, x_{-i}) - \Phi(x_i', x_{-i}) = u_i(x_i, x_{-i}) - u_i(x_i', x_{-i}).$$

Matching Rock-Paper-Scissors

Bob/Alice	Rock	Paper	Scissors
Rock	1 / 1	0 / 0	0 / 0
Paper	0 / 0	1 / 1	0 / 0
Scissors	0 / 0	0 / 0	1 / 1

Potential Function $\rightarrow \Phi(x_1, x_2) := x_1^\top A x_2$

Remarks:

◦ Potential function does not depend on i.

◦ All finite congestion games are potential games Rosenthal [8]; Example 2.3.

◦ Potential games also capture opinion formation games Kleinberg et al. [3].

Congestion games (definition)

Remark: ○ Congestion games traditionally refer to *cost* instead of *utility*.

Definition

A congestion game is an N-player game where we have

- A directed graph $G(V, E)$.
- Each player $i \in [N]$ admits a source/target pair $(s_i, t_i) \in V \times V$.
- A strategy space of agent $i \in [N]$, $\mathcal{P}_i :=$ {the set of all paths from $s_i \in V$ to $t_i \in V$}.
- The *load* $\ell_e(p_1, \ldots, p_N)$ of edge $e \in E$ under the strategy profile $p = (p_1, \ldots, p_N)$ equals

$$\ell_e(p_1, \ldots, p_N) := \underbrace{|i \in [N] : \; e \in P_i|}_{\#\text{agents selecting the edge}} .$$

- The cost of agent $i \in [N]$ under strategy profile $p = (p_1, \ldots, p_N)$ equals

$$\mathrm{Cost}_i(p_i, p_{-i}) := \sum_{e \in p_i} c_e\left(\ell_e(p)\right),$$

  where $c_e(\ell) \geq 0$ is the cost of traversing edge $e \in E$ with load ℓ ($c_e(\ell)$ is non-decreasing with ℓ).

Congestion games (properties)

Questions

- Do Congestion games admit Mixed NE? Yes (by Nash Theorem).
- Do Congestion games admit Pure NE? Yes.

Theorem (Rosenthal [8])

Let the potential function $\Phi(p_1, \ldots, p_N) := \sum_{e \in E} \sum_{i=1}^{\ell_e(p)} c_e(i)$ *then it holds that*

$$\Phi(p_i, p_{-i}) - \Phi(p'_i, p_{-i}) = \mathrm{Cost}_i(p_i, p_{-i}) - \mathrm{Cost}_i(p'_i, p_{-i}).$$

Corollary

Congestion games admit at least one pure Nash Equilibrium.

Proof.

Let $p^\star := \arg\min_p \Phi(p)$. For any agent $i \in [n]$,

$$\mathrm{Cost}_i(p_i^\star, p_{-i}^\star) - \mathrm{Cost}_i(p_i, p_{-i}^\star) = \Phi(p_i^\star, p_{-i}^\star) - \Phi(p_i, p_{-i}^\star) \leq 0 \text{ for all paths } p_i \in \mathcal{P}_i$$

□

Congestion games (potential function)

Theorem (Rosenthal [8])

Let the potential function $\Phi(p_1, \ldots, p_N) := \sum_{e \in E} \sum_{i=1}^{\ell_e(p)} c_e(i)$ *then it holds that*

$$\Phi(p_i, p_{-i}) - \Phi(p_i', p_{-i}) = \mathrm{Cost}_i(p_i, p_{-i}) - \mathrm{Cost}_i(p_i', p_{-i}).$$

Proof.

$$\begin{aligned}
\Phi(p_i, p_{-i}) - \Phi(p_i', p_{-i}) &= \sum_{e \in E} \sum_{i=1}^{\ell_e(p_i, p_{-i})} c_e(i) - \sum_{e \in E} \sum_{i=1}^{\ell_e(p_i', p_{-i})} c_e(i) \\
&= \sum_{e \in p_i / p_i'} \left(\sum_{i=1}^{\ell_e(p_i, p_{-i})} c_e(i) - \sum_{i=1}^{\ell_e(p_i, p_{-i})-1} c_e(i) \right) \\
&+ \sum_{e \in p_i' / p_i} \left(\sum_{i=1}^{\ell_e(p_i', p_{-i})-1} c_e(i) - \sum_{i=1}^{\ell_e(p_i', p_{-i})} c_e(i) \right).
\end{aligned}$$

Congestion games (potential function)

Proof.

$$
\begin{aligned}
\Phi(p_i, p_{-i}) - \Phi(p_i', p_{-i}) &= \sum_{e \in p_i / p_i'} c_e(\ell(p_i, p_{-i})) - \sum_{e \in p_i' / p_i} c_e(\ell(p_i', p_{-i})) \\
&= \sum_{e \in p_i} c_e(\ell(p_i, p_{-i})) - \sum_{e \in p_i} c_e(\ell(p_i', p_{-i})) \\
&= \mathrm{Cost}_i(p_i, p_{-i}) - \mathrm{Cost}_i(p_i', p_{-i})
\end{aligned}
$$

□

Online learning in potential games

◦ Potential games capture cooperative settings at which agents share common interests.

Definition (Potential games Monderer et al. [1])

An N-player game is a *potential game* if and only if a function $\Phi : \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \mapsto \mathbb{R}$ such that for each agent $i \in [N]$,

$$\Phi(x_i, x_{-i}) - \Phi(x_i', x_{-i}) = u_i(x_i, x_{-i}) - u_i(x_i', x_{-i}).$$

Matching Rock-Paper-Scissors

Bob/Alice	Rock	Paper	Scissors
Rock	1 / 1	0 / 0	0 / 0
Paper	0 / 0	1 / 1	0 / 0
Scissors	0 / 0	0 / 0	1 / 1

Potential Function $\rightarrow \Phi(x_1, x_2) := x_1^\top A x_2$

Question

Does no-regret dynamics converge to NE in potential games?

Online gradient ascent in potential games

Convergence of the online gradient ascent

At each round $t \geq 1$,

1. Each agent i selects mixed strategy $x_i^t \in \mathcal{X}_i$.
2. Each agent i experiences utility $u_i(x_i^t, x_{-i}^t)$.
3. Each agent i updates its mixed strategy x_i^{t+1},

$$x_i^{t+1} := \Pi_{\mathcal{X}_i}\left(x_i^t + \frac{1}{L}\nabla_{x_i} u_i(x_i^t, x_{-i}^t)\right)$$

where $L > 0$ is the smoothness constant of $\Phi(x)$.

Theorem (Best-Iterate Convergence)

Let a potential game at which all agents update their mixed strategies according to Online Gradient Ascent (with $1/L$ step-size). In case $u_i(x_i, x_{-i})$ is concave with respect to x_i, then there exists a round $t^\star \in \{1, T\}$ such that

$$\max_{x_i \in \Delta_i} u_i(x_i, x_{-i}^{t^\star}) - u_i(x_i^{t^\star}, x_{-i}^{t^\star}) \leq \mathcal{O}\left(|\mathcal{X}_i| \sqrt{\frac{L\Delta_0}{T}}\right)$$

where $\Delta_0 := \max_x \Phi(x) - \Phi(x^0)$ and $|\mathcal{X}_i|$ denotes the diameter of $\mathcal{X}_i$.

Online gradient ascent in potential games

Proof.

By $u_i(x_i, x_{-i}) - u_i(x_i', x_{-i}) = \Phi(x_i, x_{-i}) - \Phi(x_i', x_{-i}) \to \nabla_{x_i} u_i(x_i, x_{-i}) = \nabla_{x_i} \Phi(x_i, x_{-i})$. Thus online gradient ascent admits the following equivalent form

$$x^{t+1} := \Pi_{\mathcal{X}} \left(x^t + \frac{1}{L} \nabla \Phi(x^t) \right) \text{ where } \mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$$

The latter implies that $\left(x^t + 1/L \cdot \nabla \Phi(x^t) - x^{t+1}\right)^\top (x - x^{t+1}) \leq 0$ for all $x \in \mathcal{X}$. Now by the L-smoothness of $\Phi(x)$ we get that

$$\begin{aligned}
\Phi(x^{t+1}) &\geq \Phi(x^t) + \nabla \Phi(x^t)^\top (x^{t+1} - x^t) - \frac{L}{2} \|x^t - x^{t+1}\|^2 \\
&= \Phi(x^t) + \underbrace{L(x^t + \frac{1}{L}\Phi(x^t) - x^{t+1})^\top (x^{t+1} - x^t)}_{\geq 0} + \frac{L}{2} \|x^t - x^{t+1}\|^2 \\
&\geq \Phi(x^t) + \frac{L}{2} \|x^t - x^{t+1}\|^2
\end{aligned}$$

We thus get that

$$\|x^t - x^{t+1}\|^2 \leq 2 \frac{\Phi(x^{t+1}) - \Phi(x^t)}{L}$$

Online gradient ascent in potential games

Proof.

By telescopic summation

$$\sum_{t=1}^{T} \|x^t - x^{t+1}\|^2 \leq 2\frac{\max_{x\in\mathcal{X}} \Phi(x) - \Phi(x^0)}{L} \leq 2\frac{\Delta_0}{L}$$

The latter implies that there exists $t^* \in \{1, T\}$ such that

$$\|x^{t^\star} - x^{t^\star+1}\| \leq \sqrt{2\frac{\Delta_0}{LT}}$$

To simplify notation let $x^* := x^{t^\star}$ and $y^* := x^{t^\star+1}$. As a result, for each agent $i \in [N]$: $\|x_i^\star - y_i^\star\| \leq \sqrt{2\frac{\Delta_0}{LT}}$.

Hence $\|\underbrace{\Pi_{\mathcal{X}_i}\left(x_i^\star + \frac{1}{L}\nabla_{x_i} u_i(x_i^\star, x_{-i}^\star)\right)}_{y_i^\star} - x_i^\star\| \leq \sqrt{2\frac{\Delta_0}{LT}}$ and thus

$$\left(x_i^\star + 1/L \cdot \nabla_{x_i} u_i(x^\star) - y_i^\star\right)^\top (x_i - y_i^\star) \leq 0 \text{ for all } x_i \in \mathcal{X}_i$$

□

Online gradient ascent in potential games

Proof.

Since $u_i(x_i, x_{-i})$ is concave with respect to x_i, for any $x_i \in \mathcal{X}_i$

$$
\begin{aligned}
u_i(x_i, x_{-i}^\star) - u_i(x_i^\star, x_{-i}^\star) &\leq \nabla_{x_i} u_i(x_i^\star, x_{-i}^\star)^\top (x_i - x_i^\star) \\
&= \nabla_{x_i} u_i(x_i^\star, x_{-i}^\star)^\top (x_i - y_i^\star) + \nabla_{x_i} u_i(x_i^\star, x_{-i}^\star)^\top (y_i^\star - x_i^\star) \\
&\leq L(y_i^\star - x_i^\star)^\top (x_i - y_i^\star) + \nabla_{x_i} u_i(x_i^\star, x_{-i}^\star)^\top (y_i^\star - x_i^\star) \\
&\leq L|\mathcal{X}_i| \|y_i^\star - x_i^\star\| + L\|y_i^\star - x_i^\star\|^2 \\
&\leq L|\mathcal{X}_i| \sqrt{2\frac{\Delta_0}{LT}} + 2\frac{\Delta_0}{T}
\end{aligned}
$$

□

Recap on online learning in games

Question

- What if each agent i selects a no-regret algorithm to update its strategy x_i^t?
- Where does the overall strategy vector x_t converges to?

Answer

Depends on the class of game!

- Zero-sum Games $\rightarrow$ Nash Equilibrium (*time-average and last-iterate convergence*)
- Potential Games$\rightarrow$ Nash Equilibrium (*best-iterate convergence*)
- General Normal Form Games $\rightarrow$ Coarse Correlated Equilibrium (*time-average convergence*)

Coarse correlated equilibrium (definition)

Coarse Correlated Equilibrium [Aumman (Nobel Prize 2005)]

Let an n-player Normal Form Game with actions sets $\mathcal{S}_1, \ldots, \mathcal{S}_n$. A *Coarse Correlated Equilibrium* σ is a joint probability distribution $\sigma \in \Delta(\mathcal{S}_1 \times \ldots \times \mathcal{S}_n)$ such that

$$\mathrm{E}_{s \sim \sigma}\left[u_i(s_i, s_{-i})\right] \geq \mathrm{E}_{s \sim \sigma}\left[u_i(s_i', s_{-i})\right] \quad \text{for all } s_i' \in \mathcal{S}_i.$$

Explaining CCE

- A *mediator* selects a pure strategy profile $s = (s_1, \ldots, s_n) \sim \sigma$ and suggests each agent i to play s_i.
- If agent i assumes that all other agents follow the *mediator*'s then because of

  $$\mathrm{E}_{s \sim \sigma}\left[u_i(s_i, s_{-i})\right] \geq \mathrm{E}_{s \sim \sigma}\left[u_i(s_i', s_{-i})\right] \quad \text{for all } s_i' \in \mathcal{S}_i$$

  it is agent's i favor to *also follow the mediator's advice*.
- If σ is a CCE then all agents finally follow the mediator's advice.

Example of coarse correlated equilibrium

Some potential game [1]

Bob/Alice	A	B	C
A	1 / 1	-1 / -1	0 / 0
B	-1 / -1	1 / 1	0 / 0
C	0 / 0	0 / 0	-1 / -1

A Coarse correlated equilibrium σ

Bob/Alice	A	B	C
A	1/3	0	0
B	0	1/3	0
C	0	0	1/3

[1] https://www.cis.upenn.edu/~aaroth/courses/slides/agt17/lect08.pdf

Example of Coarse correlated equilibrium

- The mediator suggests $= \begin{cases} A \text{ to Bob and } A \text{ to Alice} & \text{with prob.} 1/3 \\ B \text{ to Bob and } B \text{ to Alice} & \text{with prob.} 1/3 \\ C \text{ to Bob and } C \text{ to Alice} & \text{with prob.} 1/3 \end{cases}$
- If both Bob and Alice follow the mediator's advice then Bob's payoff equals

$$\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 1 - \frac{1}{3} \cdot 1 = \frac{1}{3}$$

- Bob *assumes* that Alice follows the mediator's advice. Under this assumption if Bob selects action B then its expected payoff is

$$\underbrace{(-1) \cdot \Pr_\sigma[\text{Alice selects A}] + (1) \cdot \Pr_\sigma[\text{Alice selects B}] + (0) \cdot \Pr_\sigma[\text{Alice selects C}] = 0}_{\text{Bob's expected utility by selecting action } B}$$

- As a result, by assuming that Alice follows the mediator's advice, then Bob prefers following the mediator's advice ($1/3$ expected payoff) than playing B (0 expected payoff). Similarly for actions A and C (0 expected payoff).
- Respectively for Alice.

Coarse correlated equilibrium vs Nash equilibrium

CCE vs NE

- If $x^\star$ is a Nash Equilibrium then $x^\star$ is a Coarse Correlated Equilibrium ($\mathrm{NE} \subseteq \mathrm{CCE}$).
- Nash Equilibrium is a *product* probability distribution $x^\star \in \Delta(A_1) \times \ldots \times \Delta(A_n)$ while CCE is a *joint* probability distribution $\sigma \in \Delta(A_1 \times \ldots \times A_n)$.
- A CCE σ is not necessarily a NE (previous example).

Online Learning in Normal-Form Games

Definition

Given a mixed strategy profile $x := (x_1, \ldots, x_n)$, $\mu(x)$ denotes the product probability distribution over pure strategy profiles $s = (s_1, \ldots, s_n)$ induced by x, $\mu_s(x) := \Pi_{s_i \in s} x_{is_i} = \Pr_x[s]$.

Theorem (Folkore)

If each agent $i \in [N]$ uses a no-regret algorithm $\mathcal{A}_i$ with regret $\mathcal{R}_i(T)$ to update its strategy $x_i^t \in \Delta_i$. Then the joint time-average distribution $\hat{\sigma} := \sum_{t=1}^{T} \mu(x^t)/T$ is a $\max_{i \in [N]} \mathcal{R}_i(T)/T$-approximate Coarse Correlated Equilibrium.

Online Learning in Normal-Form Games

Proof.

By the no-regret property, for each agent $i \in [N]$

$$\frac{1}{T}\sum_{t=1}^{T} u_i(x_i^t, x_{-i}^t) \geq \max_{s_i' \in S_i} \frac{1}{T}\sum_{t=1}^{T} u_i(s_i', x_{-i}^t) - \frac{\mathcal{R}_i(T)}{T}.$$

Equivalently,

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{s\in\mathcal{S}} \Pr_{(x_i^t, x_{-i}^t)}[(s_i, s_{-i})]\, u_i(s_i, s_{-i}) \geq \max_{s_i' \in S_i} \frac{1}{T}\sum_{t=1}^{T}\sum_{s\in\mathcal{S}} \Pr_{x_{-i}^t}[s_{-i}]\, u_i(s_i', s_{-i}) - \frac{\mathcal{R}_i(T)}{T}$$

$$\sum_{s\in\mathcal{S}} \frac{1}{T}\sum_{t=1}^{T} \Pr_{(x_i^t, x_{-i}^t)}[(s_i, s_{-i})]\, u_i(s_i, s_{-i}) \geq \max_{s_i' \in S_i} \sum_{s\in\mathcal{S}} \frac{1}{T}\sum_{t=1}^{T} \Pr_{x_{-i}^t}[s_{-i}]\, u_i(s_i', s_{-i}) - \frac{\mathcal{R}_i(T)}{T}$$

Since $\hat{\sigma} := \sum_{t=1}^{T} \mu(x^t)/T$, by definition we get that $\hat{\sigma}(s) := \frac{1}{T}\sum_{t=1}^{T} \Pr_{x^t}[s]$ and thus

$$\mathbb{E}_{s\sim\hat{\sigma}}\left[u_i(s_i, s_{-i})\right] \geq \max_{s_i' \in S_i} \mathbb{E}_{s\sim\hat{\sigma}}\left[u_i(s_i', s_{-i})\right] - \frac{\mathcal{R}_i(T)}{T}$$

Faster no-regret dynamics for CCE

So far

If all agents use a no-regret algorithm with $\sqrt{T}$ then the time-average *joint strategy profile* converges with rate $\mathcal{O}(1/\sqrt{T})$ to CCE.

- Can we do better?

Recent stream of works

- Optimistic Hedge admit rate $\mathcal{O}(T^{-3/4})$ Syrgannis et al. [9]
- Optimistic Hedge admit rate $\mathcal{O}(\log^4 T \cdot T^{-1})$ Daskalakis et al. [4]
- Optimistic Mirror Descent rate $\mathcal{O}(T^{-1})$ (for variationally stable games) Hsieh et al. [6]
- Clairvoyant MWU admit rate $\mathcal{O}(\log T^{-1} T^{-1})$ Piliouras et al. [7]
- Optimstic FTLR with log-barrier admit rate $\mathcal{O}(\log T \cdot T^{-1})$ Anagnostidis et al. [2]
- Optimstic Hedge admit rate $\mathcal{O}(\log T \cdot T^{-1})$ Mansour et al. [5]

Remark:

- Establishing no-regret dynamics with $\mathcal{O}(T^{-1})$ rate is still open!

Optimistic FTRL with euclidean regularization

Simultaneous game-play in normal form games

- Consider a Normal Form Game with N agents and utilities in $[-1, 1]$.
- At each day $t = 1, \ldots, T$:
  1. Each agent $i \in [N]$ (secretly) selects a mixed strategy $x_i^t \in \Delta(S_i)$ as follows

$$x_i^t \leftarrow \underbrace{\max_{x_i \in \Delta_i} \left[\gamma v_i(x_{-i}^{t-1})^\top x + \gamma \sum_{s=1}^{t-1} v_i(x_{-i}^s)^\top x - \frac{\|x\|^2}{2} \right]}_{\text{Optimistic FTRL}}$$

  2. All agents simultaneously reveal the selected strategies.
  3. Each agent $i \in [N]$ experiences utility

$$u_i(x_i^t, x_{-i}^t) := \mathbb{E}_{s \sim x^t} [u_i(s_i, s_{-i})] = \left\langle x_i^t, v_i(x_{-i}^t) \right\rangle$$

  and receives $v_i(x_{-i}^t)$ (expected utility vector of agent i) as feedback.

Optimistic FTRL with euclidean reguralization

Theorem

If all agents of an normal-form game with $[-1,1]$ payoffs, use Optimistic FTRL with $\gamma := T^{-1/3}$ then the resulting time-average joint strategy profile converges to CCE *with rate* $\mathcal{O}\left(\max_{i\in[N]} |S_i| N \cdot T^{-2/3}\right)$.

Proof.

Let $y_i^t \leftarrow \max_{x_i\in\Delta_i}\left[\gamma v_i(x_{-i}^t)^\top x + \gamma \sum_{s=1}^{t-1} v_i(x_{-i}^s)^\top x - \frac{\|x\|^2}{2}\right]$ (Be the Regularized Leader). Then by the *Be the Leader lemma* (see Lecture 1),

$$\max_{x_i\in\Delta_i} \sum_{t=1}^{T} v_i(x_{-i}^t)^\top x_i - \sum_{t=1}^{T} v_i(x_{-i}^t)^\top y_i^t \leq \mathcal{O}\left(\frac{1}{\gamma}\right)$$

Notice that

$$x_i^t \leftarrow \max_{x_i\in\Delta_i}\left[\gamma v_i(x_{-i}^{t-1})^\top x + \gamma \sum_{s=1}^{t-1} v_i(x_{-i}^s)^\top x - \frac{\|x\|^2}{2}\right]$$

and thus by the strong concavity of the $R(x)$,

$$\|x_i^t - y_i^t\| \leq \gamma \|v_i(x_{-i}^t) - v_i(x_{-i}^{t-1})\|$$

Optimistic FTRL with euclidean reguralization

Proof.

As a result,

$$\begin{aligned}\max_{x_i \in \Delta_i} \sum_{t=1}^{T} v_i(x_{-i}^t)^\top x_i - \sum_{t=1}^{T} v_i(x_{-i}^t)^\top x_i^t &= \max_{x_i \in \Delta_i} \sum_{t=1}^{T} v_i(x_{-i}^t)^\top x_i - \sum_{t=1}^{T} v_i(x_{-i}^t)^\top y_i^t \\ &+ \sum_{t=1}^{T} v_i(x_{-i}^t)^\top (x_i^t - y_i^t) \\ &\leq \mathcal{O}\left(\frac{1}{\gamma}\right) + \gamma \sqrt{|S_i|} \sum_{t=1}^{T} \|v_i(x_{-i}^t) - v_i(x_{-i}^t)\|\end{aligned}$$

Syrgannis et al. [9] showed that $\|v_i(x_{-i}^t) - v_i(x_{-i}^{t-1})\| \leq \sum_{j \neq i} \|x_j^t - x_j^{t-1}\|$. □

Optimistic FTRL with euclidean reguralization

Proof.

Since each agent j uses the OFTRL algorithm with step-size γ then $\|x_i^t - x_i^{t-1}\| \leq \mathcal{O}(\gamma\sqrt{|\mathcal{S}_i|})$ and thus

$$\|v_i(x_{-i}^t) - v_i(x_{-i}^{t-1})\| \leq \sum_{j\neq i} \|x_j^t - x_j^{t-1}\| \leq \mathcal{O}\left(\gamma N \sqrt{\max_{i\in[N]} |\mathcal{S}_i|}\right).$$

As a result, we get that

$$\begin{aligned}
\max_{x_i\in\Delta_i} \sum_{t=1}^T v_i(x_{-i}^t)^\top x_i - \sum_{t=1}^T v_i(x_{-i}^t)^\top x_i^t \quad &\leq \quad \mathcal{O}\left(\frac{1}{\gamma}\right) + \gamma\sqrt{|\mathcal{S}_i|}\sum_{t=1}^T \|v_i(x_{-i}^t) - v_i(x_{-i}^t)\| \\
&\leq \quad \mathcal{O}\left(\frac{1}{\gamma}\right) + \gamma^2 N \max_{i\in[N]} |\mathcal{S}_i| \cdot T \\
&\leq \quad \mathcal{O}\left(N \max_{i\in[N]} |\mathcal{S}_i| \cdot T^{1/3}\right)
\end{aligned}$$

□

References I

[1] Potential games.
Games and Economic Behavior, 14(1):124–143, 1996.
(Cited on pages 5 and 10.)

[2] Ioannis Anagnostides, Gabriele Farina, Christian Kroer, Chung-Wei Lee, Haipeng Luo, and Tuomas Sandholm.
Uncoupled learning dynamics with o(log T) swap regret in multiplayer games.
CoRR, abs/2204.11417, 2022.
(Cited on page 22.)

[3] David Bindel, Jon M. Kleinberg, and Sigal Oren.
How bad is forming your own opinion?
In Rafail Ostrovsky, editor, *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 57–66. IEEE Computer Society, 2011.
(Cited on page 5.)

References II

[4] Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich.
Near-optimal no-regret learning in general games.
In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan, editors, *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pages 27604–27616, 2021.
(Cited on page 22.)

[5] Liad Erez, Tal Lancewicki, Uri Sherman, Tomer Koren, and Yishay Mansour.
Regret minimization and convergence to equilibria in general-sum markov games, 2022.
(Cited on page 22.)

[6] Yu-Guan Hsieh, Kimon Antonakopoulos, and Panayotis Mertikopoulos.
Adaptive learning in continuous games: Optimal regret bounds and convergence to nash equilibrium.
In Mikhail Belkin and Samory Kpotufe, editors, *Conference on Learning Theory, COLT 2021, 15-19 August 2021, Boulder, Colorado, USA*, volume 134 of *Proceedings of Machine Learning Research*, pages 2388–2422. PMLR, 2021.
(Cited on page 22.)

References III

[7] Georgios Piliouras, Ryann Sim, and Stratis Skoulakis.
Optimal no-regret learning in general games: Bounded regret with unbounded step-sizes via clairvoyant MWU.
CoRR, abs/2111.14737, 2021.
(Cited on page 22.)

[8] Robert W Rosenthal.
A class of games possessing pure-strategy nash equilibria.
International Journal of Game Theory, 2:65–67, 1973.
(Cited on pages 5, 7, and 8.)

[9] Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E. Schapire.
Fast convergence of regularized learning in games.
In Corinna Cortes, Neil D. Lawrence, Daniel D. Lee, Masashi Sugiyama, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, pages 2989–2997, 2015.
(Cited on pages 22 and 25.)