Online Learning in Games

Prof. Volkan Cevher
volkan.cevher@epfl.ch

Lecture 2: Game theory and online learning dynamics (Part I)

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

EE-735 (Spring 2024)

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

Normal form games (Von Neumann [14])

A normal form game features the following ingredients:

- A finite set of, say N, agents.
- Each agent $i \in [N]$ admits
  - A finite set S_i of actions/pure strategies.
  - A function $u_i : S_1 \times \ldots \times S_N \mapsto \mathbb{R}$ where $u_i(s_1, \ldots, s_N)$ denotes agent's i utility if each agent j plays $s_j \in S_j$.

Remarks:

- A normal form games requires $N \times (|S_1| \times \ldots \times |S_N|)$ values to be described.
- By taking $N \to \infty$, we obtain the mean-field games regime [7].
- Each action s_i can be represented as an one-hot vector encoding of the space S_i.

Notation

- A vector $s := (s_1, \ldots, s_N) \in S_1 \times \ldots S_N$ is called a pure strategy profile.
- The vector $s_{-i} := (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_N)$ denotes the pure strategy profile excluding the agent i.

Mixed strategies (Von Neumann [14])

- A probability distribution $x_i \in \Delta(S_i)$ over S_i is called a mixed strategy of the agent $i \in [N]$.
- A collection $x := (x_1, \ldots, x_N) \in \Delta(S_1) \times \ldots \times \Delta(S_N)$ is called a mixed strategy profile.
- A collection $x_{-i} := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N)$ is called the mixed strategy profile, excluding the agent i.

Expected utility (Von Neumann [14])

Given a mixed strategy profile $x := (x_1, \ldots, x_N) \in \Delta(S_1) \times \ldots \times \Delta(S_N)$

- The (expected) utility of the agent $i \in [N]$ equals the following

$$u_i(x_i, x_{-i}) := \mathbb{E}_{(s_1,\ldots,s_N)\sim x}\left[u_i(s_i, s_{-i})\right].$$

- Equivalently, the expected utility of the agent i equals the following

$$u_i(x_i, x_{-i}) := \langle x_i, v_i(x_{-i})\rangle ,$$

  where $[v_i(x_{-i})]_\alpha = \mathbb{E}_{(s_1,\ldots,s_N)\sim x}\left[u_i(\alpha, s_{-i})\right]$ for all $\alpha \in S_i$.

Remark: ◦ The agent-i's utility $u_i(x, x_{-i})$ is a linear function with respect to its mixed strategy $x_i \in \Delta(S_i)$.

Bimatrix normal form games

Bimatrix games

Also known as bilinear games, bimatrix games features the following ingredients:

- The game has only two agents, where $N = 2$.
- $A \in \mathbb{R}^{n \times m}$ encodes the utilities of agent x and $B \in \mathbb{R}^{m \times n}$ encodes the utilities of agent y.
- Given the mixed strategy profile $(x, y) \in \Delta_n \times \Delta_m$, we can express the individual utilities as follows

$$\underbrace{u_x(x, y) := x^\top A y}_{x\text{'s utility}} \quad \text{and} \quad \underbrace{u_y(x, y) := y^\top B x.}_{y\text{'s utility}}$$

Battle of the sexes (Luce et al. 1957)

A Bob and Alice want to decide how to spend Saturday. Their utility profiles are given as follows:

Bob / Alice	Football	Ballet
Football	10,7	0,0
Ballet	0,0	7,10

Remark: ◦ If Bob and Alice go to Ballet together their payoffs are respectively $(7, 10)$.

Nash equilibrium

Definition (Nash [9])

A pure strategy profile $s^\star := (s_1^\star, \ldots, s_n^\star) \in S_1 \times \ldots \times S_N$ is a Pure Nash Equilibrium (PNE) if and only if

$$u_i(s_i^\star, s_{-i}^\star) \geq u_i(s_i, s_{-i}^\star) \quad \text{for all } s_i \in S_i.$$

Remark: ◦ Pure Nash Equilibrium does not always exists.

Definition (Nash [9])

A mixed strategy profile $x^\star := (x_1^\star, \ldots, x_n^\star) \in \Delta(S_1) \times \ldots \times \Delta(S_N)$ is a Mixed Nash Equilibrium (MNE) if and only if

$$u_i(x_i^\star, x_{-i}^\star) \geq u_i(x_i, x_{-i}^\star) \quad \text{for all } x_i \in \Delta(S_i).$$

Remark: ◦ No agent i can increase their utility by deviating to another mixed strategy $x_i' \in \Delta(S_i)$.

Theorem (Nash [9])

Every normal form game admits at least one Nash equilibrium $x^\star$.

Remark: ◦ A normal form game may admit more than one Nash equilibrium.

An example with multiple Nash equilibria

Battle of the sexes (Luce et al. 1957)

A Bob and Alice want to decide how to spend Saturday night. Their utility profiles are given as follows

Bob / Alice	Football	Ballet
Football	10,7	0,0
Ballet	0,0	7,10

Remarks:

- Battle of the sexes admits multiple Nash Equilibrium $x^\star \in \Delta_2 \times \Delta_2$.
  - $x_A^\star = (1, 0)$ and $x_B^\star = (1, 0)$.
  - $x_A^\star = (0, 1)$ and $x_B^\star = (0, 1)$.
  - $x_A^\star = (7/17, 10/17)$ and $x_B^\star = (7/17, 10/17)$.

On the existence of the Nash equilibrium

Theorem (Brouwer 1910)

Let $\Phi : \mathcal{D} \mapsto \mathcal{D}$ be a continuous mapping, where $\mathcal{D}$ is a compact space. Then, there exists a fixed point $x^\star \in \mathcal{D}$ such that $\Phi(x^\star) = x^\star$

Proof.

The proof relies on Brouwer's fixed point theorem. Let $p_i(s_i, x) = u_i(s_i, x_{-i}) - u_i(x_i, x_{-i})$. Given a mixed strategy profile x, consider the strategy profile x' as follows, which we use to define a continuous mapping G:

$$x'_{is_i} \leftarrow \frac{x_{is_i} + \max(p_i(s_i, x), 0)}{1 + \sum_{\bar{s}_i \in S_i} \max(p_i(\bar{s}_i, x), 0)},$$

which *greedily* reinforces actions with higher utilities under x_{-i}. Note that $\sum_{s_i \in S_i} x_{is_i} = 1$, which is the first term in the denominator in order to normalize x'_{is_i} to be in the simplex.

Consider the mapping $G(x) := (G_1(x), \ldots, G_n(x))$, where $G_i(x) = x'_i$. Since G is a continuous mapping in a compact space, we know that $x^\star = G(x^\star)$. Hence, for each agent $i \in N$, it holds that

$$x^\star_{is_i} = \frac{x^\star_{is_i} + \max(p_i(s_i, x^\star), 0)}{1 + \sum_{\bar{s}_i \in S_i} \max(p_i(\bar{s}_i, x^\star), 0)} \quad \text{for each } s_i \in S_i.$$

□

On the existence of the Nash equilibrium

Proof.

In case $\sum_{s_i \in S_i} \max(p_i(s_i, x^*), 0) = 0$ then $u_i(s_i, x^\star_{-i}) \leq u_i(x^\star_i, x^\star_{-i})$ for all $s \in S_i$. Let us assume that

$$\sum_{s_i \in S_i} \max(p_i(s_i, x^\star), 0) > 0.$$

Let $S_i^+ = \{s_i \in S_i : \max(p_i(s_i, x^*), 0) > 0\}$. Then $x^\star_{is_i} = 0$ for all $s \notin S_i^+$ and thus $\sum_{s_i \in S_i^+} x^\star_{is_i} = 1$.
However for all $s_i \in S_i^+$, it holds that

$$\sum_{s \in S_i^+} x^\star_{is_i} \cdot u_i(s_i, x_{-i}) > \sum_{s_i \in S_i^+} x^\star_{is_i} \cdot u_i(x^\star_i, x^\star_{-i}) = u_i(x^\star_i, x^\star_{-i}),$$

which is a contradiction since $\sum_{s_i \in S_i^+} x^\star_{is_i} \cdot u_i(s_i, x^\star_{-i}) = u_i(x^\star_i, x^\star_{-i})$. □

A critique on Nash equilibrium

Pros

- A *steady state* at which no agent wants to deviate.
- It always exists for any normal form game.

Cons

- Most likely there is no polynomial-time algorithm even approximating NE.
  - Computing Nash Equilibrium (even for bi-matrix games) is $\mathbb{PPAD}$-complete Daskalakis et al. [5], Chen et al. [3]
  - Approximating NE up to a small universal constant $\epsilon > 0$ in polynomial-time collapses the Exponential Time Hypothesis for $\mathbb{PPAD}$ Rubinstein [13].
- Nash Equilibrium is not very informative on how to play games.
  - There are multiple NE. Thus, the choice $x_i^\star \in \Delta(S_i)$ needs to be made for any single agent $i \in [N]$.
  - Even if Nash Equilibrium is unique, agent $i \in [N]$ needs the knowledge of the utility functions of the other agents $(u_1(\cdot), \ldots, u_N(\cdot))$ to compute $x_i^\star \in \Delta(S_i)$.

Repeated games and online learning

Repeated game-play in normal form games

- Consider a normal form game with N-agents.
- At each round $t = 1, \ldots, T$:
  1. Each agent $i \in [N]$ (secretly) selects a mixed strategy $x_i^t \in \Delta(S_i)$.
  2. All agents simultaneously reveal the selected strategies.
  3. Each agent $i \in [N]$ experiences utility

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

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

Repeated games and online learning

Repeated game-play in normal form games

- Consider a normal form game with N-agents.
- At each round $t = 1, \dots, T$:
  1. Each agent $i \in [N]$ (secretly) selects a mixed strategy $x_i^t \in \Delta(S_i)$.
  2. All agents simultaneously reveal the selected strategies.
  3. Each agent $i \in [N]$ experiences utility

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

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

Question

How should an agent $i \in [N]$ select $x_i^t \in \Delta(S_i)$ at each round $t \geq 1$ so as to maximize its overall utility?

Answer:

Online learning!

Online learning in games

Simultaneous game-play in normal form games (Agent's i perspective)

1. Agent i selects a strategy $x_i^t \in \Delta(S_i)$.
2. All other agents select a strategy profile x_{-i}^t.
3. Agent i experiences utility $\langle x_i^t, v_i(x_{-i}^t)\rangle$ and receives feedback $v_i(x_{-i}^t)$.

Corollary

If an agent i uses a no-regret algorithm $\mathcal{A}$ to select its mixed strategy $x_i^t \in \Delta(S_i)$, then it holds that

$$\max_{x_i \in \Delta(S_i)} \sum_{t=1}^{T} u_i(x_i, x_{-i}^t) - \sum_{t=1}^{T} u_i(x_i^t, x_{-i}^t) \leq \mathcal{R}_{\mathcal{A}}(T).$$

Remarks:
- ◦ The time-averaged utility of each agent approaches the utility of the best fixed action: $\mathcal{R}_{\mathcal{A}}(T)/T \to 0$.
- ◦ Online Learning is a *rational behavior* for selfish agents trying to maximize their individual utilities.

Online learning in games and convergence to equilibrium

Questions

- What if all agents select a no-regret algorithm to update their mixed strategies?
- 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*)

Remark: ◦ Potential games and Coarse Correlated Equilibrium will be defined up next.

Online learning in zero-sum games

- Zero-sum games capture extremely antagonistic settings at which $u_1(s_1, s_2) = -u_2(s_2, s_1)$.
- The sum of utilities always equals to 0.

Zero-sum games [Von Neumann [14]]

The zero-sum games are characterized by the following:

- We can consider them as a bimatrix normal form game (A, B) where $B = -A^\top$.
- In a zero-sum game, the x-agent tries to minimize $x^\top Ay$ and the y-agent tries to maximize $x^\top Ay$.

Rock-paper-scissors

In this classical game, Bob and Alice plays rock ($>$ scissors and $<$ paper), paper ($>$ rock and $<$ scissors), and scissors ($>$ paper and $<$ rock) with the following payoffs:

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

There is a generalization of this game called rock-paper-scissors-lizard-spock.

Online learning in zero-sum games

- Zero-sum games capture extremely antagonistic settings at which $u_1(s_1, s_2) = -u_2(s_2, s_1)$.
- The sum of utilities always equals to 0.

Zero-sum games [Von Neumann [14]]

The zero-sum games are characterized by the following:

- We can consider them as a bimatrix normal form game (A, B) where $B = -A^\top$.
- In a zero-sum game, the x-agent tries to minimize $x^\top Ay$ and the y-agent tries to maximize $x^\top Ay$.

Nash equilibrium for zero-sum games

A mixed strategy profile $(x^\star, y^\star)$ is a Nash equilibrium in zero-sum bimatrix games if the following holds

$$\max_{y \in \Delta_m} x^{\star\top} Ay \leq x^{\star\top} Ay^\star \leq \min_{x \in \Delta n} x^\top Ay^\star.$$

Remark: ◦ The Nash equilibrium in zero-sum games is also known as minimax equilibrium.

No regret dynamics in zero-sum games

Theorem (Folkore)

If the x-agent uses an online learning algorithm $\mathcal{A}$ and the y-agent an online learning algorithm $\mathcal{B}$ to respectively update their strategies, then the following holds

$$\max_{y\in\Delta_m} \hat{x}^\top Ay - \frac{\mathcal{R}_\mathcal{A}(T)+\mathcal{R}_\mathcal{B}(T)}{T} \leq \hat{x}^\top A\hat{y} \leq \min_{x\in\Delta_n} x^\top A\hat{y} + \frac{\mathcal{R}_\mathcal{A}(T)+\mathcal{R}_\mathcal{B}(T)}{T},$$

where the time-averaged vector is defined as follows $(\hat{x},\hat{y}) := \sum_{t=1}^T (x^t, y^t)/T$.

Remarks:

- The above results holds no matter the selected algorithms $\mathcal{A}$ and $\mathcal{B}$.
- One can compute an ϵ-approximate Nash Equilibrium of a Zero-Sum game in $\mathcal{O}\left(nm\frac{\log n+\log m}{\epsilon^2}\right)$ time complexity by running the Hedge algorithm.
  - The regret of Hedge equals $\mathcal{R}_{\text{Hedge}}(T) := \mathcal{O}\left(\sqrt{T\log n}\right)$.
  - Computing the vectors Ay_t and $x_t^\top A$ requires $\mathcal{O}(nm)$ time-complexity.
  - See further Lecture 1.

No regret dynamics in zero-sum games (proof)

Proof.

- By the no-regret property of x-agent, it holds that

$$\sum_{t=1}^{T} x^{t\top} A y^t - \min_{x \in \Delta_n} \sum_{t=1}^{T} x^\top A y^t \leq \mathcal{R}_{\mathcal{A}}(T).$$

- By the no-regret property of y-agent, it holds that

$$\max_{y \in \Delta_m} \sum_{t=1}^{T} x_t^\top A y - \sum_{t=1}^{T} x^{t\top} A y^t \leq \mathcal{R}_{\mathcal{B}}(T).$$

By summing the above two inequalities and dividing by T, we have

$$\max_{y \in \Delta_m} \hat{x}^\top A y - \min_{x \in \Delta_n} x^\top A \hat{y} \leq \frac{\mathcal{R}_{\mathcal{A}}(T) + \mathcal{R}_{\mathcal{B}}(T)}{T}.$$

As a result, we obtain the approximate Nash characterization: $\hat{x}^\top A \hat{y} \leq \min_{x \in \Delta_n} x^\top A \hat{y} + \frac{\mathcal{R}_{\mathcal{A}}(T) + \mathcal{R}_{\mathcal{B}}(T)}{T}$. □

Going beyond $\mathcal{O}\left(\sqrt{T}\right)$-regret in zero-sum games

- So far we have seen
  - No-regret algorithms with $\mathcal{O}(\sqrt{T})$-regret.
  - The above regret bound is tight (see Lecture 1).
  - Once both agents of a zero-sum game use no-regret algorithms, we obtain $\mathcal{O}(1/\sqrt{T})$ convergence to NE.
- Daskalakis et al. [4] show the first case where an online learning algorithm
  - guarantees $\mathcal{O}(\sqrt{T})$ in the adversarial case
  - attains $\tilde{\mathcal{O}}(1/T)$ convergence to NE once adopted by both players of the zero-sum game

Remark: ◦ The online learning algorithm [4] is based on the excessive gap technique Nesterov [10].

Optimistic hedge

Remarks:

- Rakhlin et al. [12] simplified the algorithm of Daskalakis et al. [4].
- They propose the so-called *Optimistic Hedge* algorithm.
- *Optimistic methods* were first introduced by Popov [11].

Optimistic Hedge (Rakhlin et al. [12])

At each round $t \geq 1$,

- The learner plays a mixed strategy $x^t \in \Delta_n$.
- The adversary selects a utility vector $u^t \in [-1, 1]^n$.
- The learner observes the utility $u_t^\top x_t$ and updates their mixed strategy $x^{t+1} \in \Delta_n$ as follows

$$x_j^{t+1} := \frac{x_j^t \cdot e^{2\gamma u_j^t - \gamma u_j^{t-1}}}{\sum_{k=1}^n x_k^t \cdot e^{2\gamma u_k^t - \gamma u_k^{t-1}}}$$

Remark:

- More offline optimistic algorithms will be covered in the later lectures.

Optimistic Hedge

Theorem (Rakhlin et al. [12])

If both agents of a zero-sum game use Optimistic Hedge with constant step-size $\gamma \in (0,1)$ *then*

$$\min_{x \in \Delta_n} x^\top A y - \tilde{\mathcal{O}}\left(\frac{1}{T}\right) \leq \hat{x}^\top A \hat{y} \leq \max_{y \in \Delta_m} \hat{x}^\top A y + \tilde{\mathcal{O}}\left(\frac{1}{T}\right)$$

where $(\hat{x}, \hat{y}) := \sum_{t=1}^{T} (x^t, y^t)/T$.

Remark:

- Optimistic Hedge computes an ϵ-approximate NE of a zero-sum game in $\tilde{\mathcal{O}}(nm/\epsilon)$.
- This is $\mathcal{O}(1/\epsilon)$ faster than the Hedge algorithm.

Last-iterate convergence

Optimistic hedge vs hedge in zero-sum games

- The time-average of the trajectory produced by Hedge converges with rate $\tilde{\mathcal{O}}\left(1/\sqrt{T}\right)$ to NE.
- The time-average of the trajectory produced by Optimistic Hedge converges with rate $\tilde{\mathcal{O}}\left(1/T\right)$ to NE Rakhlin et al. [12].
- The time-average of the trajectory produced by Robust Optimistic Hedge converges with rate $\mathcal{O}\left(1/T\right)$ to NE Kangarshahi et al. [8].

Question: ◦ What about the last-iterate convergence?

Theorem (Bailey et al. [1])

In Zero-sum games Hedge dynamics does not converge in the last-iterate sense.

Last-iterate convergence of optimistic methods

- Asymptotic last-Iterate convergence of Optimistic Hedge Daskalakis et al. [6].
- $\mathcal{O}(1/\sqrt{T})$ last-iterate convergence of Optimistic GDA Cai et al. [2].

Online Learning in Zero-Sum games

Recap

- If both agents use no-regret algorithms then $\mathcal{O}(1/\sqrt{T})$ time-average convergence.
- If both agents use Optimistic Hedge then $\mathcal{O}(1/T)$ time-average convergence and *asymptotic* last-iterate convergence (no rates are known).
- If both agents use Optimistic GDA then $\mathcal{O}(1/\sqrt{T})$ last-iterate convergence.

References I

[1] James P. Bailey and Georgios Piliouras.
Multiplicative weights update in zero-sum games.
In Éva Tardos, Edith Elkind, and Rakesh Vohra, editors, *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018*, pages 321–338. ACM, 2018.
(Cited on page 24.)

[2] Yang Cai, Argyris Oikonomou, and Weiqiang Zheng.
Tight last-iterate convergence of the extragradient method for constrained monotone variational inequalities.
CoRR, abs/2204.09228, 2022.
(Cited on page 24.)

[3] Xi Chen and Xiaotie Deng.
Settling the complexity of two-player nash equilibrium.
In *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 21-24 October 2006, Berkeley, California, USA, Proceedings*, pages 261–272. IEEE Computer Society, 2006.
(Cited on page 12.)

References II

[4] Constantinos Daskalakis, Alan Deckelbaum, and Anthony Kim.
Near-optimal no-regret algorithms for zero-sum games.
In Dana Randall, editor, *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 235–254. SIAM, 2011.
(Cited on pages 21 and 22.)

[5] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou.
The complexity of computing a nash equilibrium.
In Jon M. Kleinberg, editor, *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 71–78. ACM, 2006.
(Cited on page 12.)

[6] Constantinos Daskalakis and Ioannis Panageas.
Last-iterate convergence: Zero-sum games and constrained min-max optimization.
In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 27:1–27:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.
(Cited on page 24.)

References III

[7] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions.
Mean Field Games and Applications, pages 205–266.
Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
(Cited on page 5.)

[8] Ehsan Asadi Kangarshahi, Ya-Ping Hsieh, Mehmet Fatih Sahin, and Volkan Cevher.
Let's be honest: An optimal no-regret framework for zero-sum games.
In Jennifer G. Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 2493–2501. PMLR, 2018.
(Cited on page 24.)

[9] J. F. Nash.
Equilibrium points in \$n\$-person games.
Proceedings of the National Academy of Sciences of the United States of America, 36(48-49), 1950.
(Cited on page 8.)

[10] Yurii Nesterov.
Smooth minimization of non-smooth functions.
Mathematical Programming, 103:127–152, 2005.
(Cited on page 21.)

References IV

[11] Leonid D. Popov.
A modification of the Arrow-Hurwicz method for search of saddle points.
Mathematical notes of the Academy of Sciences of the USSR, 28:845–848, 1980.
(Cited on page 22.)

[12] Alexander Rakhlin and Karthik Sridharan.
Optimization, learning, and games with predictable sequences.
In Christopher J. C. Burges, Léon Bottou, Zoubin Ghahramani, and Kilian Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*, pages 3066–3074, 2013.
(Cited on pages 22, 23, and 24.)

[13] Aviad Rubinstein.
Inapproximability of nash equilibrium.
In Rocco A. Servedio and Ronitt Rubinfeld, editors, *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 409–418. ACM, 2015.
(Cited on page 12.)

References V

[14] J. von Neumann and O. Morgenstern.
Theory of games and economic behavior.
Princeton University Press, 1947.
(Cited on pages 5, 6, 17, and 18.)