Alternation makes the adversary weaker in two-player games

Stratis Skoulakis, Aarhus University

NeurIPS 2023, Spotlight Paper

joint work with

V. Cevher, EPFL

A. Cutkosky, BU

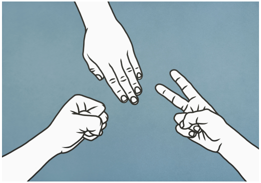
G. Piliouras, DeepMind

A. Kavis, UT Austin

L. Viano, EPFL

Normal Form Games

Rock-Paper-Scissors



Payoff matrix of Alice

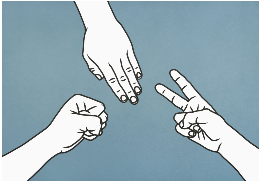
Alice/Bob	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

Payoff matrix of Bob

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

Normal Form Games

Rock-Paper-Scissors



Alice plays Rock

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

Alice gets -1

Bob plays Scissors

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

Bob gets 1

Normal Form Games

Normal-Form Games

Matrices A and B

Payoff matrix A of Alice

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{12}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{n2}	$\cdots$	A_{nm}

Payoff matrix B of Bob

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{12}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{n2}	$\cdots$	B_{nm}

Normal Form Games

Normal-Form Games

Matrices A and B

Alice plays action i

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{1j}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{nj}	$\cdots$	A_{nm}

Alice gets A_{ij}

Bob plays action j

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{1j}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{nj}	$\cdots$	B_{nm}

Bob gets B_{ij}

Normal Form Games

Normal-Form Games - Mixed Strategies

Matrices A and B

Alice plays a prob. distr.

$x = (x_1, \ldots, x_n) \in \Delta_n$

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{12}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{n2}	$\cdots$	A_{nm}

Bob plays a prob. distr.

$y = (y_1, \ldots, y_m) \in \Delta_m$

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{12}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{n2}	$\cdots$	B_{nm}

Normal Form Games

Normal-Form Games - Mixed Strategies

Matrices A and B

Alice plays a prob. distr.

$x = (x_1, \ldots, x_n) \in \Delta_n$

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{12}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{n2}	$\cdots$	A_{nm}

Alice's expected cost

$x^\top A y$

Bob plays a prob. distr.

$y = (y_1, \ldots, y_m) \in \Delta_m$

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{12}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{n2}	$\cdots$	B_{nm}

Bob's expected cost

$x^\top B y$

Normal Form Games

Normal-Form Games - Mixed Strategies

Matrices A and B

Alice plays a prob. distr.

$x = (x_1, \ldots, x_n) \in \Delta_n$

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{12}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{n2}	$\cdots$	A_{nm}

$$x^\top A y$$

Bob plays a prob. distr.

$y = (y_1, \ldots, y_m) \in \Delta_m$

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{12}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{n2}	$\cdots$	B_{nm}

$$x^\top B y$$

Normal-Form Games over Time

- Alice and Bob play the normal-form game (A, B) over T rounds.
  - *Simultaneous play*: Agents simultaneously select their strategies at each round (very well-studied).
  - *Alternating play*: Agents alternatingly update their strategies (this work).

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- Alice selects x_3 / Bob selects y_3

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- Alice selects x_3 / Bob selects $y_3 \longrightarrow$ Alice gets $x_3^\top A y_3$
- $\ldots$

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\longrightarrow$ Alice gets $x_2^\top A y_2$
- Alice selects x_3 / Bob selects y_3 $\longrightarrow$ Alice gets $x_3^\top A y_3$
- $\ldots$

How shoud Alice select her actions over time?

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- Alice selects x_3 / Bob selects $y_3 \longrightarrow$ Alice gets $x_3^\top A y_3$
- $\ldots$

How shoud Alice select her actions over time? *No Regret algorithms*

- Regret Matching [Blackwell '65]
- Hedge [Freund et al. '97]
- Online Gradient Descent [Zinkevic '03]
- Follow the Regularized Leader [Abernethy et al. '10]
- ....

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- ...

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- $\ldots$

Hedge Algorithm [Freund and Schapire '97] Godel Prize '03

$$x_{t+1}(i) = \frac{x_t(i) \cdot e^{-\gamma [A y_t]_i}}{\sum_{j=1}^n x_t(j) \cdot e^{-\gamma [A y_t]_j}}$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\longrightarrow$ Alice gets $x_2^\top A y_2$
- ...

Hedge Algorithm [Freund and Schapire '97] Godel Prize '03

$$x_{t+1}(i) = \frac{x_t(i) \cdot e^{-\gamma [A y_t]_i}}{\sum_{j=1}^n x_t(j) \cdot e^{-\gamma [A y_t]_j}}$$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies $y_1, \ldots, y_T$, the regret of Alice $\mathcal{R}(T)$

$$\mathcal{R}(T) := \underbrace{\sum_{t=1}^T x_t^\top A y_t}_{\text{cost of Alice}} - \underbrace{\min_{i \in [n]} \sum_{t=1}^T [A y_t]_i}_{\text{cost of best action}} \leq \tilde{\mathcal{O}}\left(\sqrt{T}\right)$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- Alice selects x_3 / Bob selects $y_3 \longrightarrow$ Alice gets $x_3^\top A y_3$
- $\ldots$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies,

$$\underbrace{\frac{1}{T}\sum_{t=1}^{T} x_t^\top A y_t}_{\text{time-average cost}} \leq \underbrace{\frac{1}{T}\min_{x \in \Delta_n}\sum_{t=1}^{T} x^\top A y_t}_{\text{best fixed action}} + \tilde{\mathcal{O}}\left(\frac{\sqrt{T}}{T}\right) \to 0$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- $\ldots$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies $y_1, \ldots, y_T$, the regret of Alice $\mathcal{R}(T)$

$$\mathcal{R}(T) \quad := \quad \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} \quad - \quad \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \quad \leq \quad \tilde{\mathcal{O}}\left(\sqrt{T}\right)$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- $\dots$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies $y_1, \dots, y_T$, the regret of Alice $\mathcal{R}(T)$

$$\mathcal{R}(T) \quad := \quad \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} \quad - \quad \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \quad \leq \quad \tilde{\mathcal{O}}\left(\sqrt{T}\right)$$

Can Alice do better?

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- $\ldots$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies $y_1, \ldots, y_T$, the regret of Alice $\mathcal{R}(T)$

$$\mathcal{R}(T) \quad := \quad \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} \quad - \quad \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \quad \leq \quad \tilde{\mathcal{O}}\left(\sqrt{T}\right)$$

Can Alice do better? No!

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects $y_1 \longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects $y_2 \longrightarrow$ Alice gets $x_2^\top A y_2$
- $\ldots$

Theorem (folkore)

Bob ***can always select*** *$y_1, \ldots, y_T$, the regret of Alice*

$$\mathcal{R}(T) \quad := \quad \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} \quad - \quad \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \quad \geq \quad \Omega(\sqrt{T})$$

$\mathcal{O}(\sqrt{T})$ regret is the best Alice can get in **simultaneous play**!

Regret Lower Bound on Simultaneous play

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Regret Lower Bound on Simultaneous play

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
- Bob is indifferent so it selects $y_t = (1, 0)$ or $(0, 1)$ with prob. $1/2$.

Regret Lower Bound on Simultaneous play

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
- Bob is indifferent so it selects $y_t = (1,0)$ or $(0,1)$ with prob. $1/2$.
- No matter what $x_t \in \Delta_n$ Alice selects,

$$\mathbb{E}[x_t^\top A y_t] = 0 \text{ and thus } \sum_{t=1}^{T} \mathbb{E}[x_t^\top A y_t] = 0$$

Regret Lower Bound on Simultaneous play

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
- Bob is indifferent so it selects $y_t = (1, 0)$ or $(0, 1)$ with prob. $1/2$.
- No matter what $x_t \in \Delta_n$ Alice selects,

$$\mathbb{E}[x_t^\top A y_t] = 0 \text{ and thus } \sum_{t=1}^{T} \mathbb{E}[x_t^\top A y_t] = 0$$

- However $\min_{i \in \{1,2\}} [A y_t]_i = -\Theta(\sqrt{T})$ since $[A y_t]_i = 1$ or -1 with probability $1/2$.

Regret Lower Bound on Simultaneous play

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
- Bob is indifferent so it selects $y_t = (1,0)$ or $(0,1)$ with prob. $1/2$.
- No matter what $x_t \in \Delta_n$ Alice selects,

$$\mathbb{E}[x_t^\top A y_t] = 0 \text{ and thus } \sum_{t=1}^{T} \mathbb{E}[x_t^\top A y_t] = 0$$

- However $\min_{i \in \{1,2\}} [Ay_t]_i = -\Theta(\sqrt{T})$ since $[Ay_t]_i = 1$ or -1 with probability $1/2$.
- No matter how Alice her sequence $x_1, \ldots, x_T$:

$$\mathbb{E}\left[\sum_{t=1}^{T} x_t^\top A y_t - \min_{i \in [n]} \sum_{t=1}^{T} [Ay_t]_i\right] = \Theta(\sqrt{T})$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\longrightarrow$ Alice gets $x_2^\top A y_2$
- ...

Theorem (folkore)

*Bob **can always select** $y_1, \ldots, y_T$, the regret of Alice*

$$\mathcal{R}(T) \quad := \quad \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} \quad - \quad \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \quad \geq \quad \Omega(\sqrt{T})$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\longrightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\longrightarrow$ Alice gets $x_2^\top A y_2$
- $\dots$

Theorem (folkore)

Bob ***can always select*** $y_1, \dots, y_T$*, the regret of Alice*

$$\mathcal{R}(T) \quad := \quad \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\textit{cost of Alice}} \quad - \quad \underbrace{\min_{i\in[n]} \sum_{t=1}^{T} [A y_t]_i}_{\textit{cost of best action}} \quad \geq \quad \Omega(\sqrt{T})$$

What if Alice and Bob play in alternating turns?

Solving Heads'up Poker

Online Learning + Simultaneous Play

Polaris [Bowling et al. AAMAS '09] → Decent Performance

Solving Heads'up Poker

Online Learning + Simultaneous Play

Polaris [Bowling et al. AAMAS '09] → Decent Performance

Online Learning + Alternating Play

Cepheus [Oskari et al. IJCAI '15], Libratus [Brown et al. IJCAI '17] → **Beat Human Experts!**

Solving Heads'up Poker

Online Learning + Simultaneous Play

Polaris [Bowling et al. AAMAS '09] → Decent Performance

Online Learning + Alternating Play

Cepheus [Oskari et al. IJCAI '15], Libratus [Brown et al. IJCAI '17] → **Beat Human Experts!**

Regret Guarantees → Faster Training!

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects x_1

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects y_2

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects x_3

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$
- $\ldots$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$
- ...

Can Alice have regret $\mathcal{R}(T)$ better than $\mathcal{O}(\sqrt{T})$?

$$\mathcal{R}(T) := \underbrace{\sum_{k=0}^{T/2} x_{2k+1}^\top A(y_{2k} + y_{2k+2})}_{\text{Alice's cost}} - \underbrace{\min_{x \in \Delta_n} \sum_{k=0}^{T/2} x^\top A(y_{2k} + y_{2k+2})}_{\text{best fixed action}}$$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$
- ...

Can Alice have regret $\mathcal{R}(T)$ better than $\mathcal{O}(\sqrt{T})$?

$$\mathcal{R}(T) := \underbrace{\sum_{k=0}^{T/2} x_{2k+1}^\top A(y_{2k} + y_{2k+2})}_{\text{Alice's cost}} - \underbrace{\min_{x \in \Delta_n} \sum_{k=0}^{T/2} x^\top A(y_{2k} + y_{2k+2})}_{\text{best fixed action}}$$

Previous $\Omega(\sqrt{T})$ lower bound fails!

Regret Lower Bound Fails!!!

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Regret Lower Bound Fails!!!

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
- Bob is indifferent so it selects $y_t = (1, 0)$ or $(0, 1)$ with prob. $1/2$.
- Alice selects $x_t = (1, 0)$ if $y_{t-1} = (1, 0)$ and $x_t = (0, 1)$ if $y_{t-1} = (0, 1)$. Thus,

$$\sum_{t=1}^{T} \mathbb{E}[x_t^\top A(y_{t-1} + y_t)] = -T$$

Regret Lower Bound Fails!!!

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
- Bob is indifferent so it selects $y_t = (1,0)$ or $(0,1)$ with prob. $1/2$.
- Alice selects $x_t = (1,0)$ if $y_{t-1} = (1,0)$ and $x_t = (0,1)$ if $y_{t-1} = (0,1)$. Thus,

$$\sum_{t=1}^{T} \mathbb{E}[x_t^\top A(y_{t-1} + y_t)] = -T$$

- However $\min_{i \in \{1,2\}} [A(y_t + y_{t-1})]_i = -\Theta(\sqrt{T})$ since $[Ay_t]_i = 1$ or -1 with probability $1/2$.

Regret Lower Bound Fails!!!

- Let $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
- Bob is indifferent so it selects $y_t = (1,0)$ or $(0,1)$ with prob. $1/2$.
- Alice selects $x_t = (1,0)$ if $y_{t-1} = (1,0)$ and $x_t = (0,1)$ if $y_{t-1} = (0,1)$. Thus,

$$\sum_{t=1}^{T} \mathbb{E}[x_t^\top A(y_{t-1} + y_t)] = -T$$

- However $\min_{i\in\{1,2\}}[A(y_t + y_{t-1})]_i = -\Theta(\sqrt{T})$ since $[Ay_t]_i = 1$ or -1 with probability $1/2$.
- No matter how Alice her sequence $x_1, \ldots, x_T$:

$$\mathbb{E}\left[\sum_{t=1}^{T} x_t^\top A(y_t + y_{t-1}) - \min_{i\in[n]} \sum_{t=1}^{T} [A(y_t + y_{t-1})]_i\right] = -T + \Theta(\sqrt{T})$$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$
- $\dots$

Can Alice have regret $\mathcal{R}(T)$ better than $\mathcal{O}(\sqrt{T})$?

$$\mathcal{R}(T) := \underbrace{\sum_{k=0}^{T/2} x_{2k+1}^\top A(y_{2k} + y_{2k+2})}_{\text{Alice's cost}} - \underbrace{\min_{x \in \Delta_n} \sum_{k=0}^{T/2} x^\top A(y_{2k} + y_{2k+2})}_{\text{best fixed action}}$$

Previous $\Omega(\sqrt{T})$ lower bound fails!

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$
- $\ldots$

Theorem (Skoulakis et al., NeurIPS '23 spotlight])

In alternating play*, Alice can always guarantee*

- $\tilde{\mathcal{O}}(T^{1/3})$ *regret for general games.*
- $\mathcal{O}(\log T)$ *for* $n = 2$ *actions (different algorithm).*

Alternating Turns

Previous Results

- If **both** Alice and Bob use *Gradient Descent* in **unconstrained zero-sum games** $\longrightarrow \mathcal{O}(1)$ regret [Bailey et al. COLT 2020]
- If **both** Alice and Bob use *Hedge* in **zero-sum games** $\longrightarrow \mathcal{O}(T^{1/3})$ regret [Wibisono et al. NeurIPS 2022]

Alternating Turns

Previous Results

- If **both** Alice and Bob use *Gradient Descent* in **unconstrained zero-sum games** $\longrightarrow \mathcal{O}(1)$ regret [Bailey et al. COLT 2020]
- If **both** Alice and Bob use *Hedge* in **zero-sum games** $\longrightarrow \mathcal{O}(T^{1/3})$ regret [Wibisono et al. NeurIPS 2022]

Previous results make assumptions on Bob's behavior.

Alternating Turns

Previous Results

- If **both** Alice and Bob use *Gradient Descent* in **unconstrained zero-sum games** $\longrightarrow \mathcal{O}(1)$ regret [Bailey et al. COLT 2020]
- If **both** Alice and Bob use *Hedge* in **zero-sum games** $\longrightarrow \mathcal{O}(T^{1/3})$ regret [Wibisono et al. NeurIPS 2022]

Previous results make assumptions on Bob's behavior.

Theorem (Skoulakis et al., NeurIPS '23 spotlight])

In alternating play*, Alice can always guarantee*

- $\tilde{\mathcal{O}}(T^{1/3})$ *regret for general games.*
- $\mathcal{O}(\log T)$ *for* $n = 2$ *actions (different algorithm).*

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$
- ...

	Simultaneous Play	Alternating Play
General Games	$\tilde{\mathcal{O}}(\sqrt{T})$ tight	$\tilde{\mathcal{O}}(T^{1/3})$
2 actions for Alice	$\tilde{\mathcal{O}}(\sqrt{T})$ tight	$\tilde{\mathcal{O}}(\log T)$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$
- ...

Our algorithm (general simplex)

$$x_{2k+1} \leftarrow \mathrm{argmin}_{x \in \Delta_n} \left[3 \underbrace{x^\top A y_{2k}}_{\text{exploits } y_{2k}} + \underbrace{2 \sum_{k'=0}^{k-1} x^\top A y_{2k'}}_{\text{reinforces good past actions}} - \underbrace{\gamma \cdot \sum_{i=1}^{n} \log x_i}_{\text{prevents overfitting}} \right]$$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects $x_1 \longrightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \longrightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \longrightarrow$ Alice gets $x_3^\top A y_2$
- $\ldots$

Our algorithm (general simplex)

$$x_{2k+1} \leftarrow \mathrm{argmin}_{x \in \Delta_n} \left[3 \underbrace{x^\top A y_{2k}}_{\text{exploits } y_{2k}} + \underbrace{2 \sum_{k'=0}^{k-1} x^\top A y_{2k'}}_{\text{reinforces good past actions}} - \underbrace{\gamma \cdot \sum_{i=1}^{n} \log x_i}_{\text{prevents overfitting}} \right]$$

Setting $\gamma = \mathcal{O}(T^{1/3}) \rightarrow \mathcal{O}(T^{1/3})$ alternating regret

Our Algorithm (General Case)

Algorithm 1: Our method

for rounds $t = 1, \ldots, T$ **do**

Alice *selects* sets $c_t \to Ay_t$ and selects the mixed strategy $x_t \in \Delta_2$,

$$x_t := \mathrm{argmin}_{x \in \Delta_n} \left[2\gamma c_{t-1}^\top x + \gamma \sum_{s=1}^{t-1} (c_s + c_{s-1})^\top x \underbrace{- \sum_{i=1}^{2} \log x_i / \gamma}_{\text{log-barrier penalty}} \right]$$

Alice **suffers** cost $(c_t + c_{t-1})^\top x_t$

end for

Analysis

Be the Regularized Leader

Let $y_{t+1} := \min_{x \in \Delta_n} \left[\gamma(c_t + c_{t-1})^\top x + \gamma \sum_{s=1}^{t-1} (c_s + c_{s-1})^\top x - \sum_{i=1}^{2} \log x_i\right]$.
Then,

$$\sum_{t=1}^{T} (c_t + c_{t-1})^\top \cdot y_{t+1} - \min_{x \in \Delta_n} \sum_{t=1}^{T} (c_t + c_{t-1})^\top \cdot x \leq \mathcal{O}\left(\frac{\log T}{\gamma}\right)$$

- **Regret Decomposition**

$$\sum_{t=1}^{T} (c_{t-1} + c_t)^\top (x_t - x^\star)$$
$$= \underbrace{\sum_{t=1}^{T} (c_{t-1} + c_t)^\top (y_{t+1} - x^\star)}_{\leq \mathcal{O}(\log T/\gamma)} + \sum_{t=1}^{T} (c_{t-1} + c_t)^\top (x_t - y_{t+1})$$

Analysis

Be the Regularized Leader

Let $y_{t+1} := \min_{x \in \Delta_n} \left[\gamma(c_t + c_{t-1})^\top x + \gamma \sum_{s=1}^{t-1}(c_s + c_{s-1})^\top x - \sum_{i=1}^{2} \log x_i\right]$.
Then,

$$\sum_{t=1}^{T}(c_t + c_{t-1})^\top \cdot y_{t+1} - \min_{x \in \Delta_n} \sum_{t=1}^{T}(c_t + c_{t-1})^\top \cdot x \le \mathcal{O}\left(\frac{\log T}{\gamma}\right)$$

- **Regret Decomposition**

$$\sum_{t=1}^{T}(c_{t-1} + c_t)^\top (x_t - x^\star)$$

$$= \underbrace{\sum_{t=1}^{T}(c_{t-1} + c_t)^\top (y_{t+1} - x^\star)}_{\le \mathcal{O}(\log T/\gamma)} + \underbrace{\sum_{t=1}^{T}(c_{t-1} + c_t)^\top (x_t - y_{t+1})}_{\text{upper bound this!}}$$

Analysis

In order to control $\sum_{t=1}^{T}(c_{t-1}+c_t)^\top(x_t-y_{t+1})$, we notice that

- $\underbrace{y_{t+1} := \min_{x\in\Delta_2}\left[\gamma(c_t+c_{t-1})^\top x + \gamma\sum_{s=1}^{t-1}(c_s+c_{s-1})^\top x - \sum_{i=1}^{2}\log x_i\right]}_{\text{Be the Leader}}$
- $\underbrace{x_t := \min_{x\in\Delta_2}\left[2\gamma c_{t-1}^\top x + \gamma\sum_{s=1}^{t-1}(c_s+c_{s-1})^\top x - \sum_{i=1}^{2}\log x_i\right]}_{\text{our Algorithm}}$

Lemma (Skoulakis et al., 2023)

Let $x_t, y_{t+1} \in \Delta_n$ defined as above. Then,

$$x_t - y_{t+1} = \gamma \cdot A(x_t, y_{t+1})(c_t - c_{t-1})$$

where $A(x,y)$ is an $n\times n$ matrix with

$$||A(x,y) - A(x',y')||_2 \leq \mathcal{O}(1)\cdot||(x,y)-(x',y')||_2.$$

Analysis

- **Regret Decomposition**

$$
\begin{aligned}
&\sum_{t=1}^{T}(c_{t-1}+c_t)^\top(x_t-x^\star)\\
&\quad=\underbrace{\sum_{t=1}^{T}(c_{t-1}+c_t)^\top(y_{t+1}-x^\star)}_{\leq\mathcal{O}(\log T/\gamma)}+\underbrace{\sum_{t=1}^{T}(c_{t-1}+c_t)^\top(x_t-y_{t+1})}_{\text{upper bound this!}}\\
&\quad=\mathcal{O}(\log T/\gamma)+\gamma\sum_{t=1}^{T}(c_{t-1}+c_t)^\top A(x_t,y_{t+1})(c_t-c_{t-1})\\
&\quad=\mathcal{O}(\log T/\gamma)+\gamma\sum_{t=1}^{T}c_t^\top\left(A(x_t,y_{t+1})-A(x_{t-1},y_t)\right)c_t\\
&\quad=\mathcal{O}(\log T/\gamma)+\gamma\sum_{t=1}^{T}||c_t||_2^2\cdot||(x_t,y_{t+1})-(x_{t-1},y_t)||_2\\
&\quad=\mathcal{O}(\log T/\gamma)+\mathcal{O}(\gamma^2T)=\mathcal{O}(T^{1/3})\text{ for }\gamma=T^{-1/3}.
\end{aligned}
$$

2 action simplex

- If Alice admits 2 actions $\rightarrow \mathcal{O}(\log T)$ in alternating play!

2 action simplex

- If Alice admits 2 actions $\rightarrow \mathcal{O}(\log T)$ in alternating play!
- At round $t = 2k + 1$

2 action simplex

- If Alice admits 2 actions $\rightarrow \mathcal{O}(\log T)$ in alternating play!
- At round $t = 2k + 1$
  - Greedy Best Response $\rightarrow w_k = \arg\min_{x \in \Delta_n} \left[x^\top A y_{2k}\right]$

2 action simplex

- If Alice admits 2 actions $\rightarrow \mathcal{O}(\log T)$ in alternating play!
- At round $t = 2k + 1$
  - Greedy Best Response $\rightarrow w_k = \arg\min_{x \in \Delta_n} \left[x^\top A y_{2k} \right]$
  - Follow the Regularized Leader $\rightarrow$

$$z_k = \arg\min_{x \in \Delta_n} \left[\sum_{k=1}^{t/2} x^\top A y_{2k} + \gamma ||x||^2 \right]$$

2 action simplex

- If Alice admits 2 actions $\rightarrow \mathcal{O}(\log T)$ in alternating play!
- At round $t = 2k+1$
  - Greedy Best Response $\rightarrow w_k = \arg\min_{x \in \Delta_n} \left[x^\top A y_{2k}\right]$
  - Follow the Regularized Leader $\rightarrow$

$$z_k = \arg\min_{x \in \Delta_n} \left[\sum_{k=1}^{t/2} x^\top A y_{2k} + \gamma ||x||^2\right]$$

  - If FTRL (z_k) admits $\mathcal{O}(\sqrt{T})$ regret $\rightarrow$ Greedy BR (w_k) admits $\mathcal{O}(1)$ regret

2 action simplex

- If Alice admits 2 actions $\to \mathcal{O}(\log T)$ in alternating play!
- At round $t = 2k+1$
  - Greedy Best Response $\to w_k = \arg\min_{x \in \Delta_n} \left[x^\top A y_{2k}\right]$
  - Follow the Regularized Leader $\to$

$$z_k = \arg\min_{x \in \Delta_n} \left[\sum_{k=1}^{t/2} x^\top A y_{2k} + \gamma ||x||^2\right]$$

  - If FTRL (z_k) admits $\mathcal{O}(\sqrt{T})$ regret $\to$ Greedy BR (w_k) admits $\mathcal{O}(1)$ regret
  - If Greedy BR (w_k) admits $\mathcal{O}(\sqrt{T})$ regret $\to$ FTRL (z_k) admits $\mathcal{O}(1)$ regret

2 action simplex

- If Alice admits 2 actions $\rightarrow \mathcal{O}(\log T)$ in alternating play!
- At round $t = 2k+1$
  - Greedy Best Response $\rightarrow w_k = \arg\min_{x \in \Delta_n} \left[x^\top A y_{2k}\right]$
  - Follow the Regularized Leader $\rightarrow$

$$z_k = \arg\min_{x \in \Delta_n} \left[\sum_{k=1}^{t/2} x^\top A y_{2k} + \gamma ||x||^2\right]$$

  - If FTRL (z_k) admits $\mathcal{O}(\sqrt{T})$ regret $\rightarrow$ Greedy BR (w_k) admits $\mathcal{O}(1)$ regret
  - If Greedy BR (w_k) admits $\mathcal{O}(\sqrt{T})$ regret $\rightarrow$ FTRL (z_k) admits $\mathcal{O}(1)$ regret
- Alice plays x_{2k+1} in the convex hull of z_k and w_k.

Take-Away Message

Algorithms beyond $\mathcal{O}(\sqrt{T})$ regret lower bounds of simultaneous play.

	Simultaneous Play	Alternating Play
General Games	$\tilde{\mathcal{O}}(\sqrt{T})$ tight	$\tilde{\mathcal{O}}(T^{1/3})$
2 actions for Alice	$\tilde{\mathcal{O}}(\sqrt{T})$ tight	$\tilde{\mathcal{O}}(\log T)$

Take-Away Message

Algorithms beyond $\mathcal{O}(\sqrt{T})$ regret lower bounds of simultaneous play.

	Simultaneous Play	Alternating Play
General Games	$\tilde{\mathcal{O}}(\sqrt{T})$ tight	$\tilde{\mathcal{O}}(T^{1/3})$
2 actions for Alice	$\tilde{\mathcal{O}}(\sqrt{T})$ tight	$\tilde{\mathcal{O}}(\log T)$

Thank you!!