No-Regret Learning in Finite Games

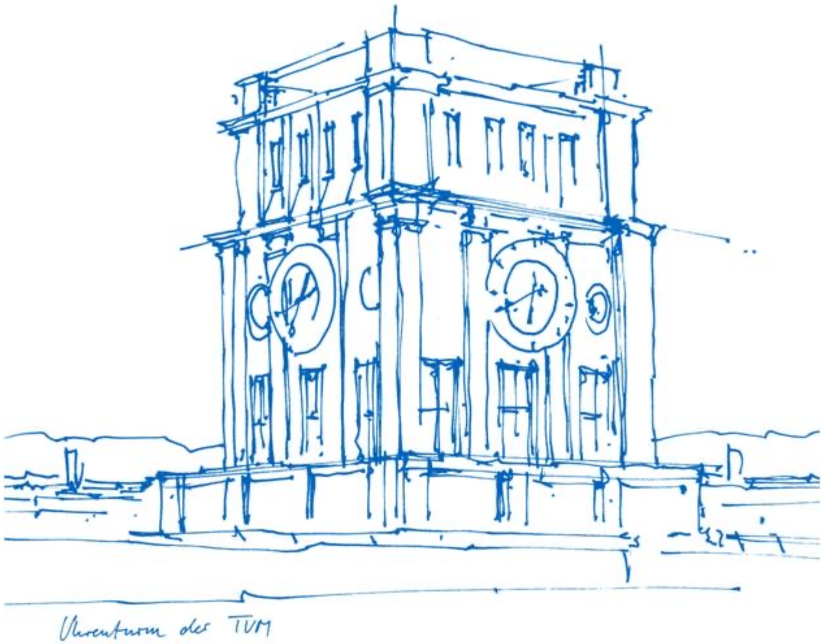
Henning Heyen

Technical University of Munich

Department of Informatics

Decision Sciences & Systems (DSS)

Munich, 11 March 2022



Outline

Motivation

The Problem:

- Nash Existence vs. PPAD completeness
- **Can no-regret dynamics learn Nash equilibria (NE)?**

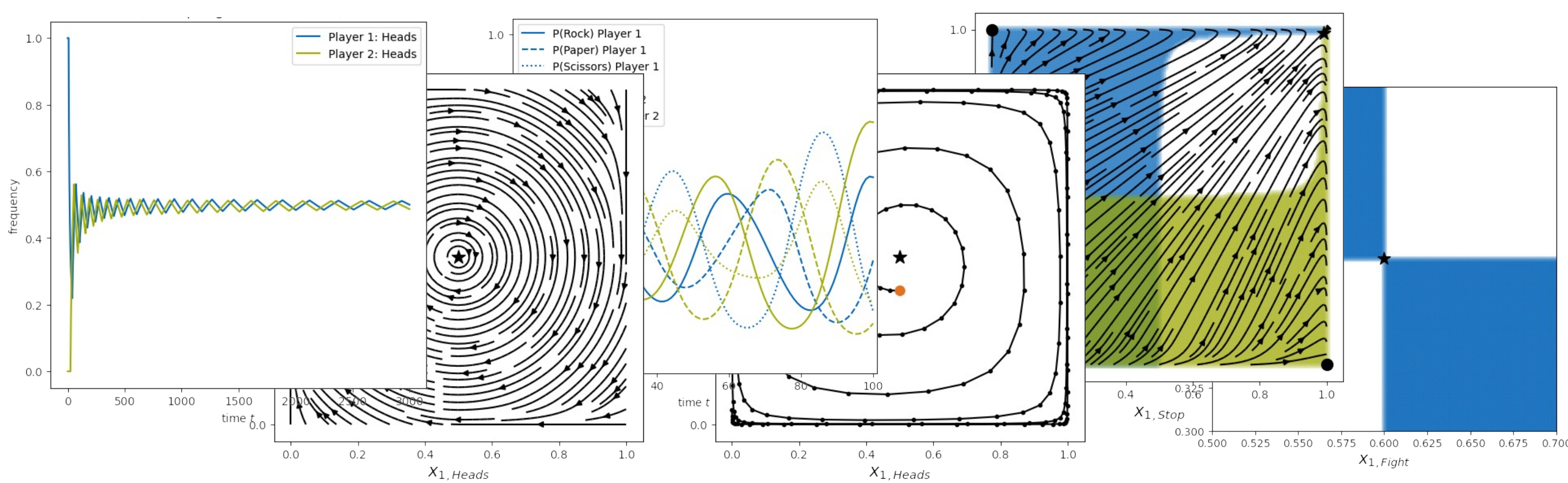
- We know they converge to the games set of coarse correlated equilibria
- However we observe Nash convergence for some games

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Motivation

Contribution:

- Overview on sufficient conditions for which no-regret learning converges to Nash Equilibria
- Visualization of no-regret algorithms in simple two player games



Outline

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

The Basic Model

Online Convex Optimization

input: convex action set $\mathcal{X}$, sequence of convex loss functions $l_t : \mathcal{X} \to \mathbb{R}$
for $t = 1, 2, \ldots, T$ **do**
  select action $x_t \in \mathcal{X}$
  incur loss $l_t(x_t)$
  update $x_t \leftarrow x_{t+1}$
end

Regret

$$reg(T) = \underbrace{\sum_{t=1}^{T} l_t(x_t)}_{\text{Loss of algorithm}} - \underbrace{\min_{x \in \mathcal{X}} \sum_{t=1}^{T} l_t(x)}_{\text{Loss of best fixed action}}$$

No Regret

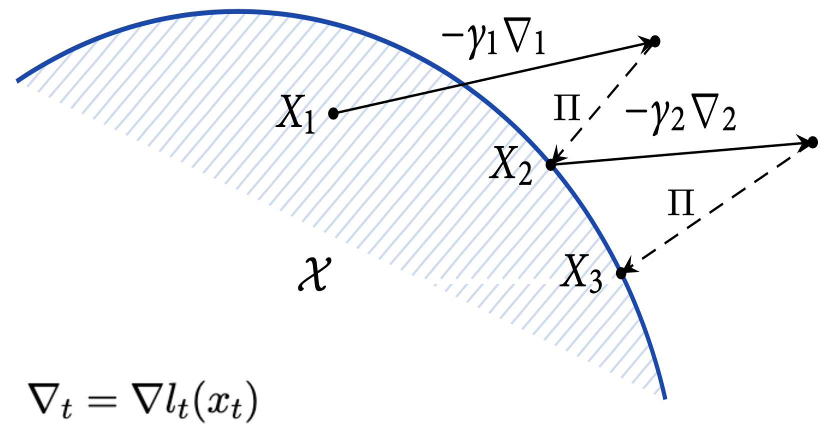
$$reg(T) = o(T)$$

→ Achievable?

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Projected Online Gradient Descent (POGD)

input: step size sequence $\gamma_t > 0$
for $t = 1, 2, \ldots, T$ **do**
  incur loss $l_t(x_t)$
  receive feedback $v_t \leftarrow -\nabla l_t(x_t)$
  update $x_{t+1} = \Pi(x_t + \gamma_t v_t)$
end



$$\Pi(x) = \underset{x' \in \mathcal{X}}{\operatorname{argmin}} \|x' - x\|_2^2$$

↳ Euclidean Projection

Assume l_t **is:**
- L-Lipschitz

Set: $\gamma \sim \frac{1}{L\sqrt{T}}$

→ $reg(T) = \mathcal{O}(\sqrt{T})$

→ **POGD** is a **no regret** algorithm [4]

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Online Mirror Descent

Let $x \leftarrow x_t$, $y \leftarrow \gamma_t v_t$ and $x^+ \leftarrow x_{t+1}$

We can rewrite **POGD** as

$$x^+ = \Pi(x+y) = \operatorname*{argmin}_{x' \in \mathcal{X}} \left\{ \|x+y-x'\|_2^2 \right\}$$

$$= \operatorname*{argmin}_{x' \in \mathcal{X}} \left\{ \|x-x'\|_2^2 + \|y\|_2^2 + 2\langle y, x-x' \rangle \right\}$$

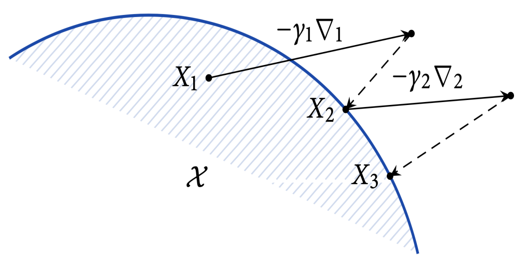
$$= \operatorname*{argmin}_{x' \in \mathcal{X}} \left\{ \langle y, x-x' \rangle + D(x', x) \right\}$$

where $D(x',x) = \frac{1}{2}\|x'-x\|_2^2 = \frac{1}{2}\|x'\|_2^2 - \frac{1}{2}\|x\|_2^2 - \langle x, x'-x \rangle$

POGD

$$x_{t+1} = \Pi(x_t + \gamma_t v_t)$$

$$\Pi(x) = \operatorname*{argmin}_{x' \in \mathcal{X}} \|x'-x\|_2^2$$



Bregman Divergence $\quad D(x',x) = h(x') - h(x) - \langle \nabla h(x), x'-x \rangle$

Online Mirror Descent

Bregman Divergence $D(x', x) = h(x') - h(x) - \langle \nabla h(x), x' - x \rangle$

Online Mirror Descent (OMD)

Generic **mirror map** $Q : \mathcal{X} \times \mathcal{Y} \to \mathcal{X}$

$$Q_x(y) = \underset{x' \in \mathcal{X}}{\operatorname{argmin}} \left\{ \langle y, x - x' \rangle + D(x', x) \right\}$$

Update Step

$$x_{t+1} = Q_{x_t}(\gamma_t v_t)$$

Assume l_t **is:**

- L-Lipschitz

Set: $\gamma \sim \frac{1}{L\sqrt{T}}$ → $reg(T) = \mathcal{O}(\sqrt{T})$

Assume h **is:**

- K-strongly convex
- Continuous

→ **OMD** is a **no regret** algorithm [4]

Online Mirror Descent (OMD) allows to use other norms than the Euclidean norm

Online Mirror Descent

OMD

$$\boxed{\begin{aligned} x_{t+1} &= Q_{x_t}(\gamma_t v_t) \\ Q_x(y) &= \operatorname*{argmin}_{x' \in \mathcal{X}} \left\{ \langle y, x - x' \rangle + D(x', x) \right\} \end{aligned}}$$

Euclidean regularization: $$\boxed{h(x) = \frac{1}{2}\|x\|_2^2}$$

Mirror map: $$Q_x(y) = \operatorname*{argmin}_{x' \in \mathcal{X}} \left\{ \langle y, x - x' \rangle + \frac{1}{2}\|x' - x\|_2^2 \right\} = \Pi(x + y)$$

OMD = POGD

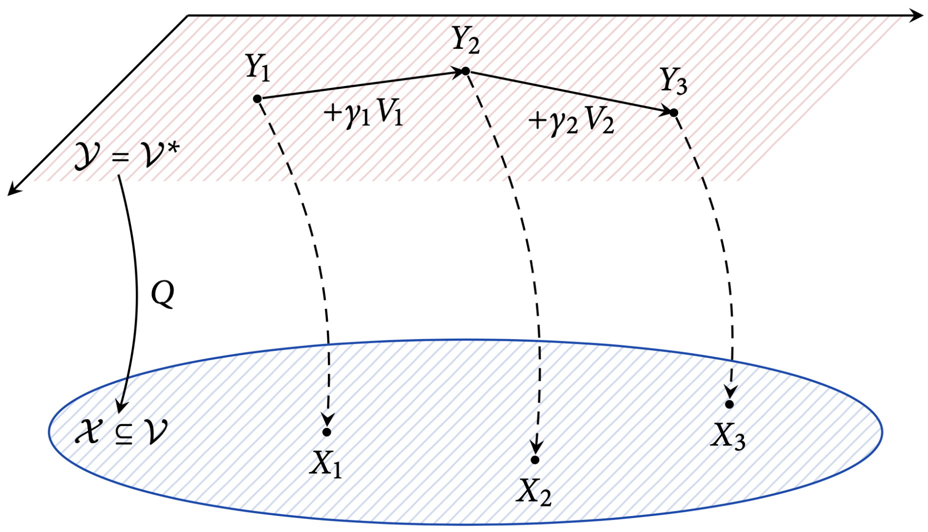
entropic regularization: $$\boxed{h(x) = \sum_{a \in \mathcal{A}} x_a \log(x_a)\ ,} \qquad \mathcal{X} = \Delta(\mathcal{A})$$

Mirror map: $$Q_x(y) = \frac{(x_a \exp(y_a))_{a \in \mathcal{A}}}{\sum_{a \in \mathcal{A}} x_a \exp(y_a)}$$ **entropic gradient descent (EGD)**

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Derived No-Regret Algorithms

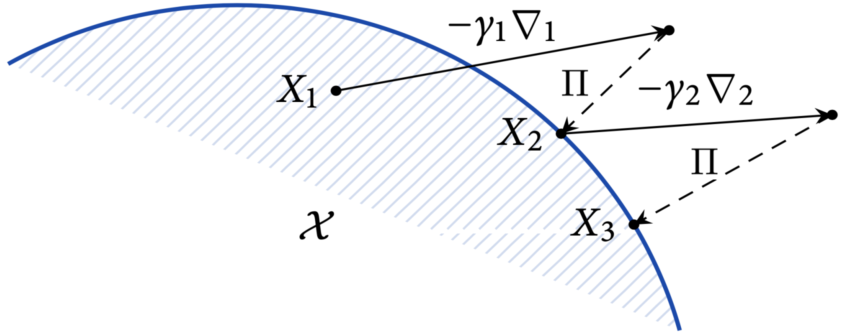
EGD



$$x_{t+1} = Q_{x_t}(\gamma_t v_t)$$

$$Q_x(y) = \frac{(x_a \exp(y_a))_{a \in \mathcal{A}}}{\sum_{a \in \mathcal{A}} x_a \exp(y_a)}$$

POGD



$$x_{t+1} = \Pi(x_t + \gamma_t v_t)$$

$$\Pi(x) = \underset{x' \in \mathcal{X}}{\operatorname{argmin}} \|x' - x\|_2^2$$

Outline

1. Motivation
2. No-Regret Learning
3. **Finite Games**
4. Literature Review
5. Simulations
6. Q&A

Finite Games

	Rock	*Paper*	*Scissors*
Rock	$0,0$	$-1,1$	$1,-1$
Paper	$1,-1$	$0,0$	$-1,1$
Scissors	$-1,1$	$1,-1$	$0,0$

Finite Game $\Gamma \equiv (\mathcal{N}, (\mathcal{A}_i)_{i\in\mathcal{N}}, (u_i)_{i\in\mathcal{N}})$

- Players $\mathcal{N} = \{1, \dots, N\}$
- Pure strategies $a_i \in \mathcal{A}_i$ and action space of pure strategies $\mathcal{A} = \prod_i \mathcal{A}_i$
- Utility function $u_i : \mathcal{A} \to \mathbb{R}$

Mixed Extension

- Mixed Strategies $x = (x_1, \dots, x_N) = (x_i, x_{-i}) \in \mathcal{X} = \prod_i \mathcal{X}_i$

$$\mathcal{X}_i = \Delta(\mathcal{A}_i) = \{x_i : \textstyle\sum_{a_i\in\mathcal{A}_i} x_{i,a_i} = 1 \land x_i \geq 0\}$$ → **convex**

$$u_i(x) = \sum_{a_1\in\mathcal{A}_1} \cdots \sum_{a_N\in\mathcal{A}_N} x_{1,a_1} \dots x_{N,a_N} u_i(a_1, \dots, a_N)$$ → **concave** (linear)

→ **concave game**

Descent vs. Ascent

	Optimization	**Game Theory**
Objective	Minimize loss	Maximize utility
Regret	$reg(T) = \sum_{t=1}^{T} l_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} l_t(x)$	$reg_i(T) = \max_{x_i \in \Delta} \sum_{t=1}^{T} u_i(x_i, x_{-i}^t) - \sum_{t=1}^{T} u_i(x^t)$
Algorithms	• Projected online Gradient Descent (**POGD**) • Entropic Gradient Descent (**EGD**)	• Projected online Gradient Ascent (**POGA**) • Entropic Gradient Ascent (**EGA**)

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Nash Equilibrium

Mixed Nash equilibria (MNE): $u_i(x_i^*, x_{-i}^*) \leq u_i(x_i, x_{-i}^*) \qquad \forall x_i \in \mathcal{X}_i, i \in \mathcal{N}$

	Rock	*Paper*	*Scissors*
Rock	$0,0$	$-1,1$	$1,-1$
Paper	$1,-1$	$0,0$	$-1,1$
Scissors	$-1,1$	$1,-1$	$0,0$

$$x_i^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \quad \forall i \in \mathcal{N}$$

Pure Nash equilibria (PNE): $u_i(a_i^*, a_{-i}^*) \leq u_i(a_i, a_{-i}^*) \qquad \forall a_i \in \mathcal{A}_i, i \in \mathcal{N}$

Strict PNE: $u_i(a_i^*, a_{-i}^*) < u_i(a_i, a_{-i}^*) \qquad \forall a_i \in \mathcal{A}_i \setminus \{a_i^*\}, i \in \mathcal{N}$

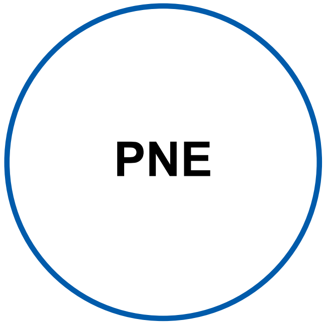
Example:

	H	T
H	$2,3$	$1,2$
T	$1,2$	$2,2$

$x^* = (H,H)$	*strict* PNE
$x^* = (T,T)$	*weak* PNE

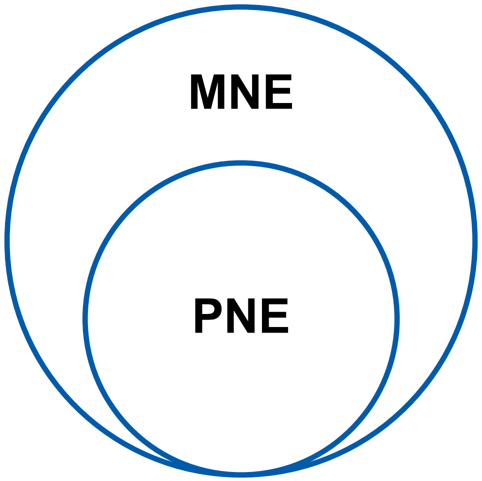
$x_1^*(0,0) \quad x_2^*(1/2,1/2) \quad MNE$

Solution Concepts Hierarchy



Pure Nash Equilibria

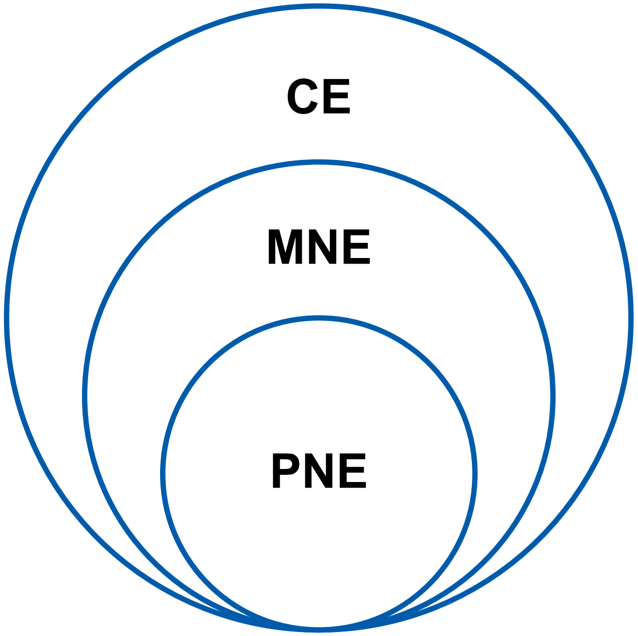
Solution Concepts Hierarchy



Mixed Nash Equilibria

Pure Nash Equilibria

Solution Concepts Hierarchy

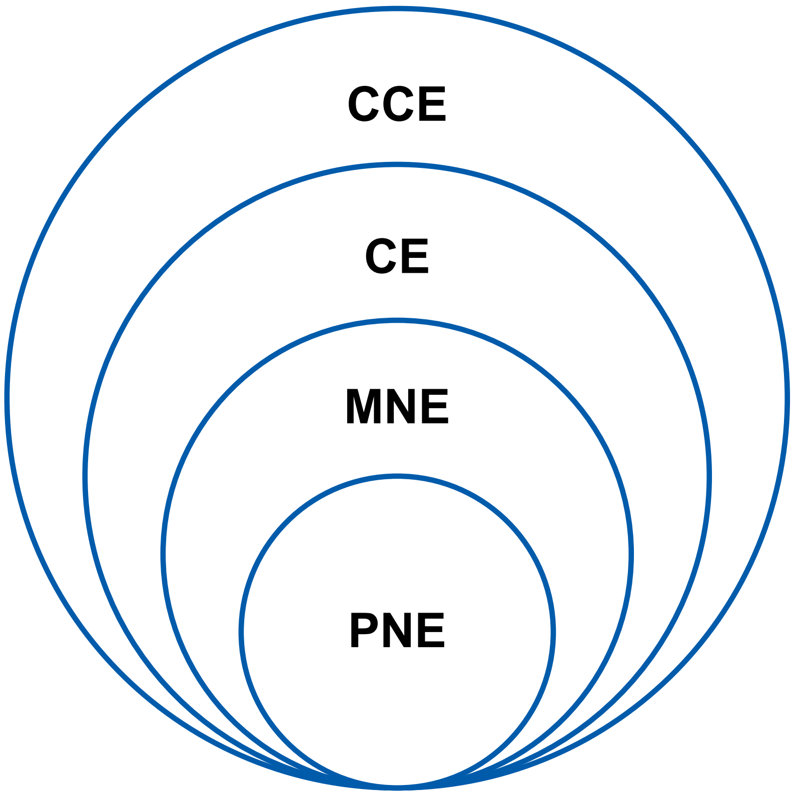


Correlated Equilibria

Mixed Nash Equilibria

Pure Nash Equilibria

Solution Concepts Hierarchy



Coarse Correlated Equilibria

May contain strictly dominated strategies [4]

Correlated Equilibria

Mixed Nash Equilibria

Pure Nash Equilibria

The empirical frequency of play in no-regret algorithms converges to the games set of **CCE** [3]

Outline

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Literature Review

Convergence:
1. Time-average (frequency of play)
2. Last iterate (trajectories)

Under **Follow the Regularized Leader…**

- *The empirical frequency of play converges to Nash in two player zero sum games with interior equilibria* [6]

$x^* \in \mathcal{X}$ variational stable (**stable**) if there exists a neighbourhood U of x^* such that

$$\langle v(x), x - x^* \rangle \leq 0 \qquad \forall x \in U$$

When $U = \mathcal{X}$ then x^* is variational globally stable (**globally stable**)

- *An interior equilibrium can not be asymptotically stable* [3]
- *Stable states are equivalent to strict Nash equilibria* [4]
- *If x^* is globally stable then it's the games unique Nash equilibrium* [4]
- *Last-iterate convergence globally to globally stable equilibria* [4]
- *If x^* is stable then it is locally attracting with high probability* [4]

Outline

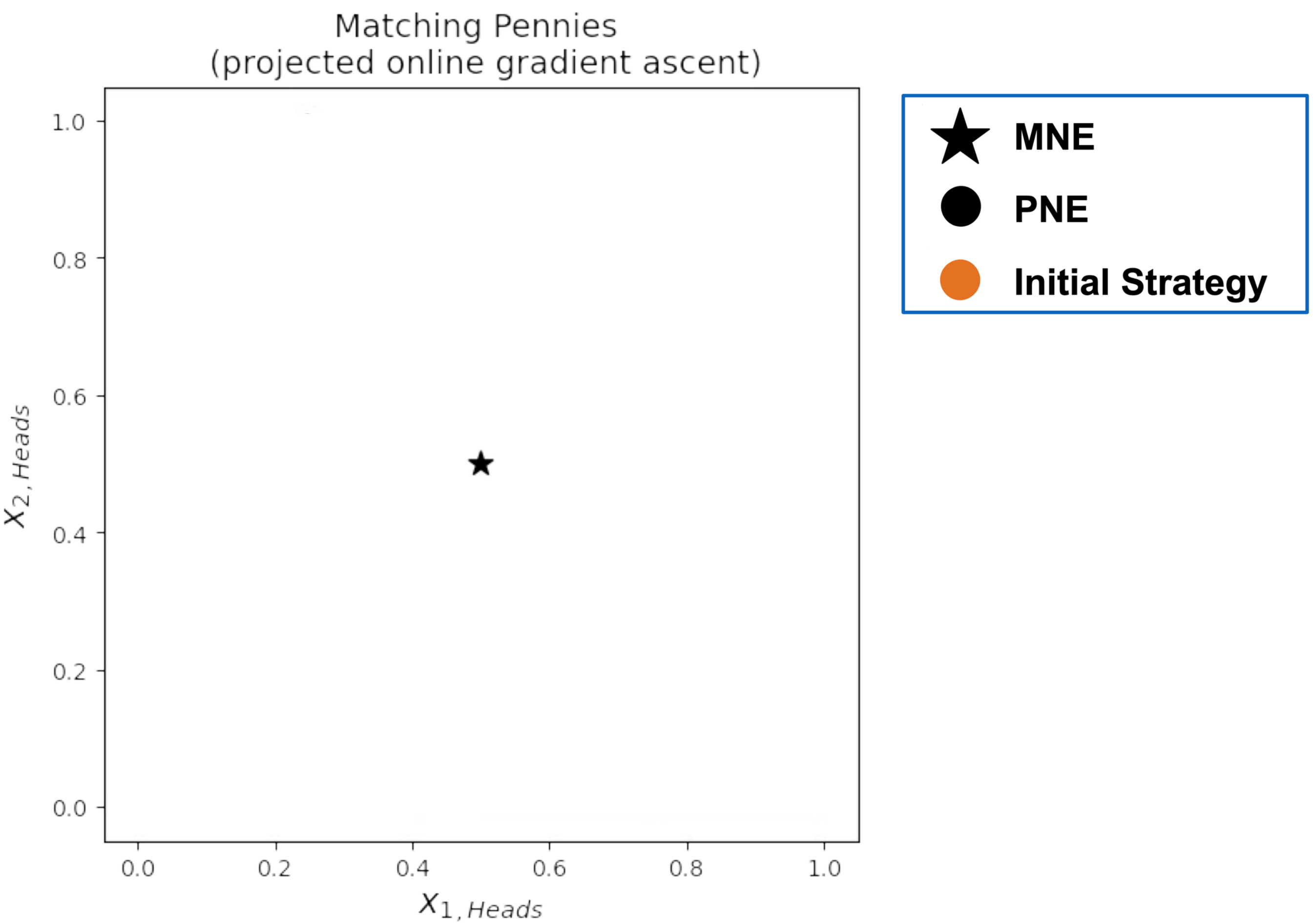
1. Motivation
2. No-Regret Learning
3. Finite Games
4. Literature Review
5. **Simulations**
6. Q&A

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Simulations

	Heads	*Tails*
Heads	$1, -1$	$-1, 1$
Tails	$-1, 1$	$1, -1$

$$x_i^* = (1/2, 1/2) \qquad \forall i \in \mathcal{N}$$

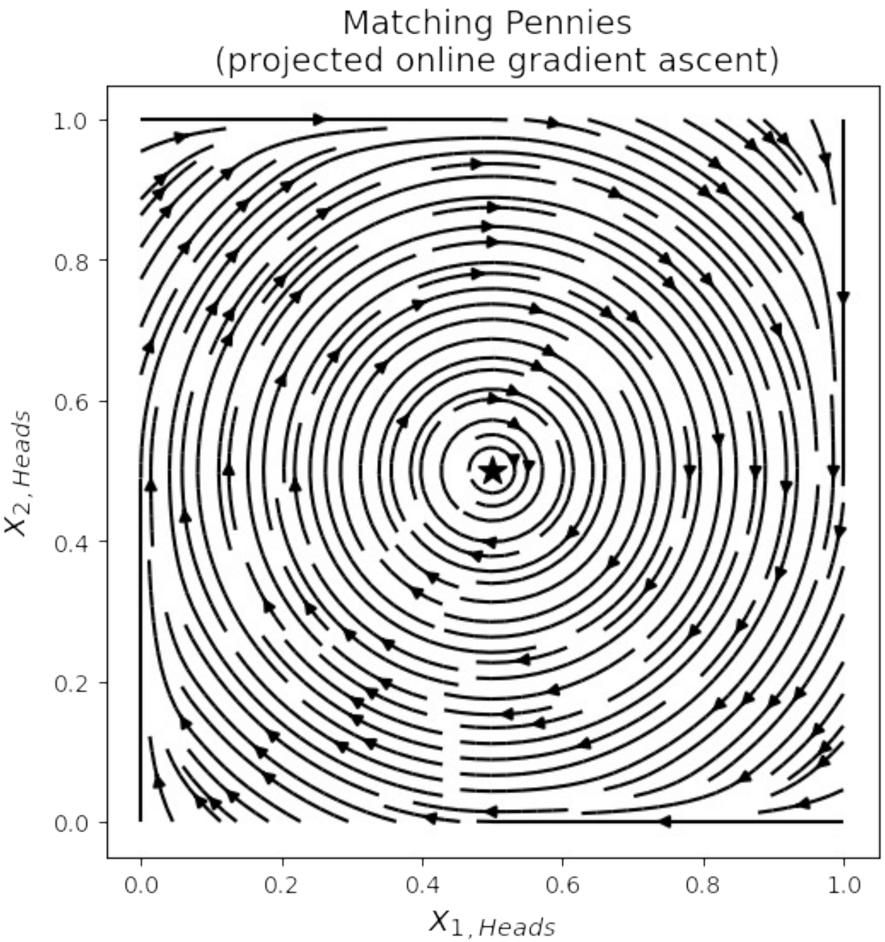


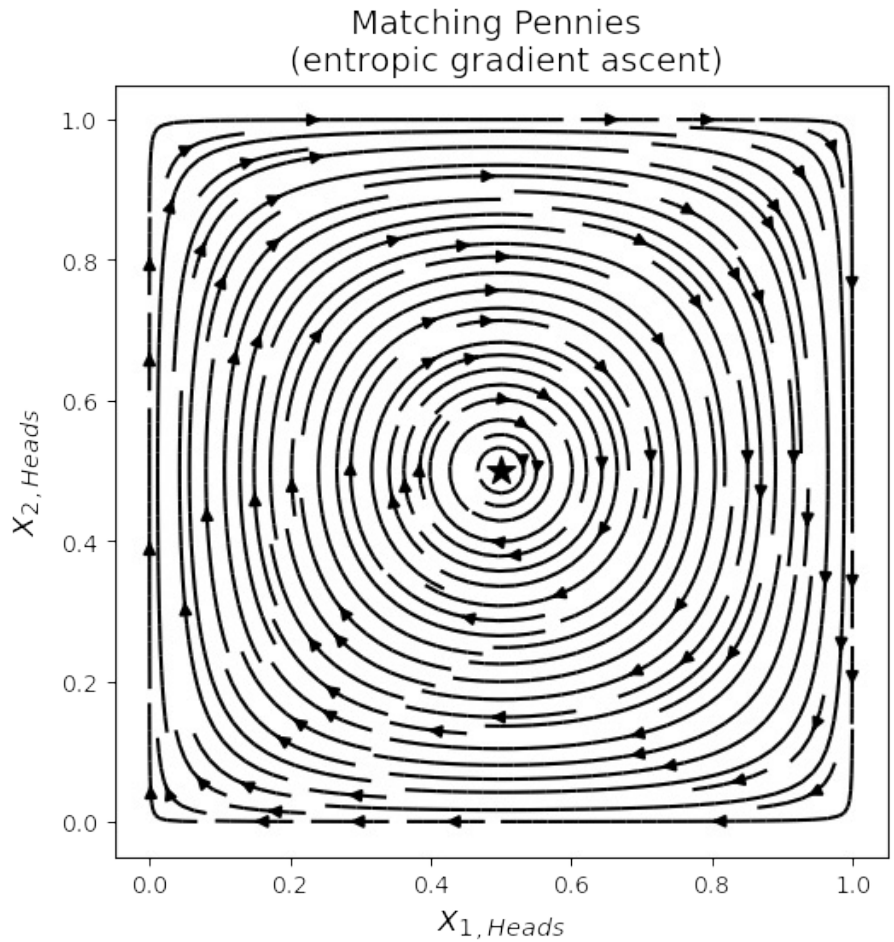
Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Matching Pennies

	Heads	*Tails*
Heads	$1, -1$	$-1, 1$
Tails	$-1, 1$	$1, -1$

$$x_i^* = (1/2, 1/2) \qquad \forall i \in \mathcal{N}$$

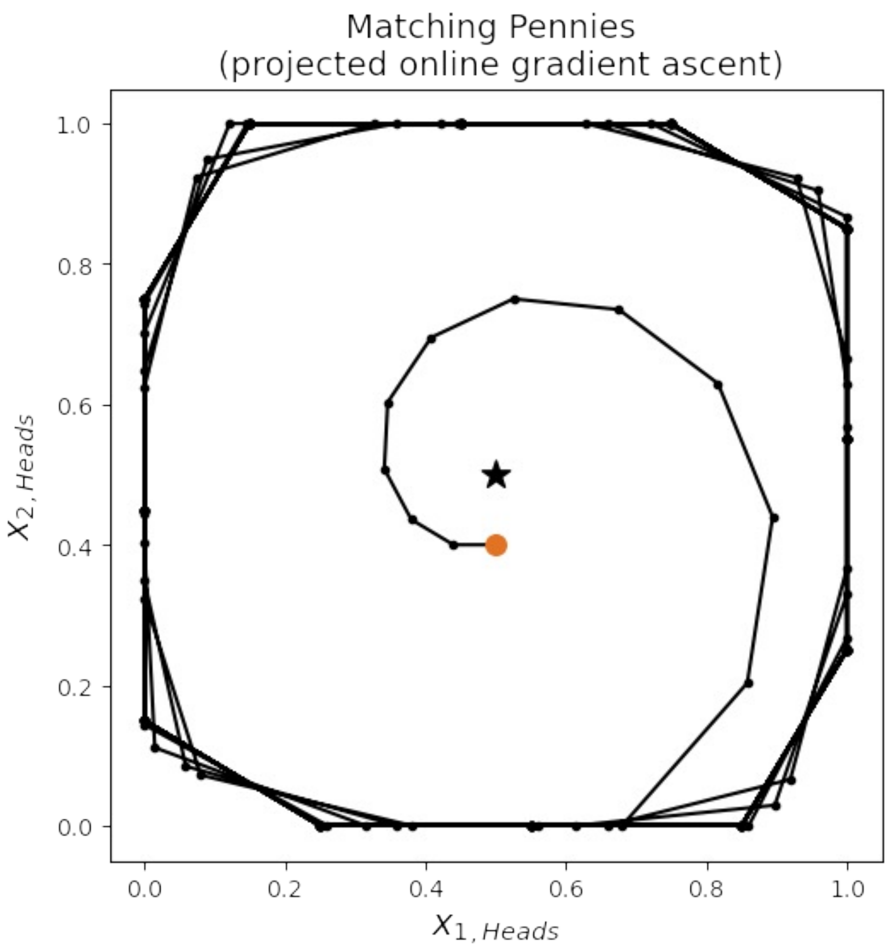


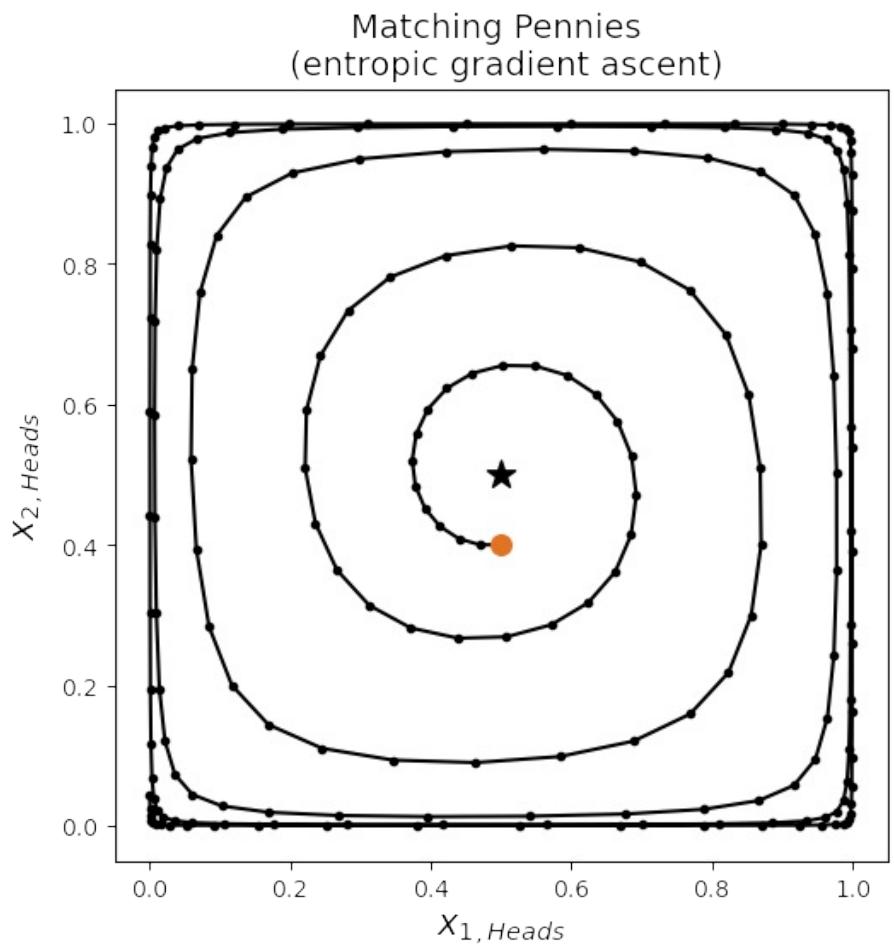


Matching Pennies

	$Heads$	$Tails$
$Heads$	$1, -1$	$-1, 1$
$Tails$	$-1, 1$	$1, -1$

$$x_i^* = (1/2, 1/2) \qquad \forall i \in \mathcal{N}$$



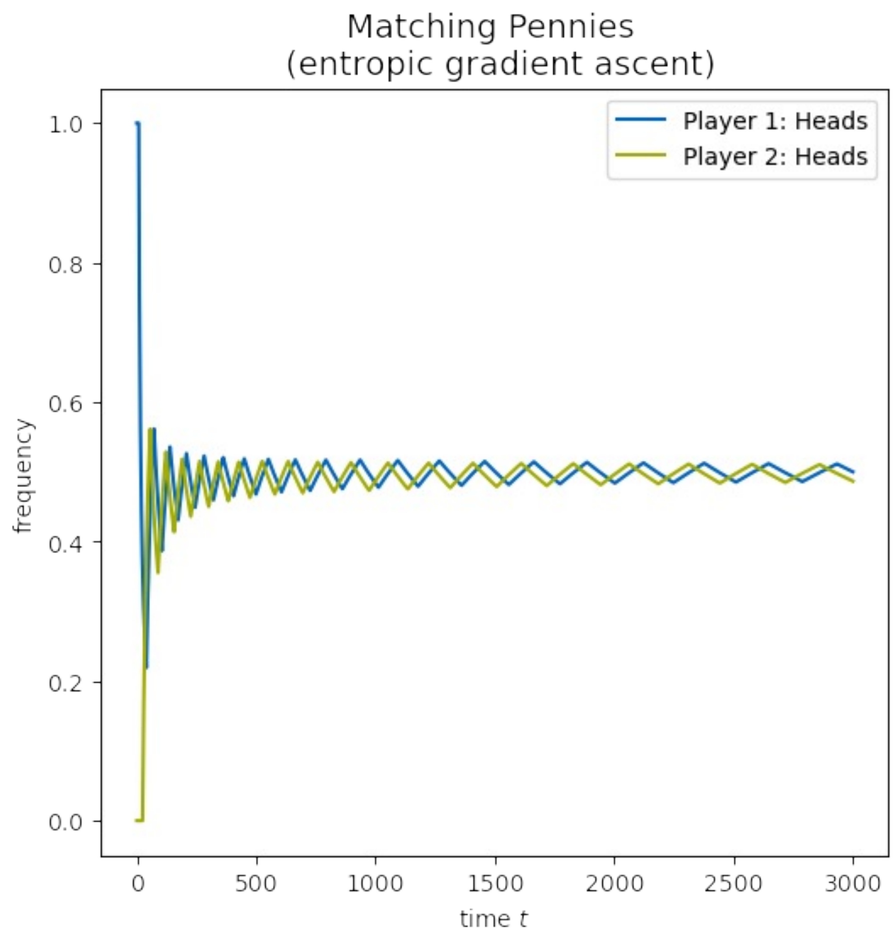


Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Matching Pennies

	$Heads$	$Tails$
$Heads$	$1, -1$	$-1, 1$
$Tails$	$-1, 1$	$1, -1$

$$x_i^* = (1/2, 1/2) \qquad \forall i \in \mathcal{N}$$

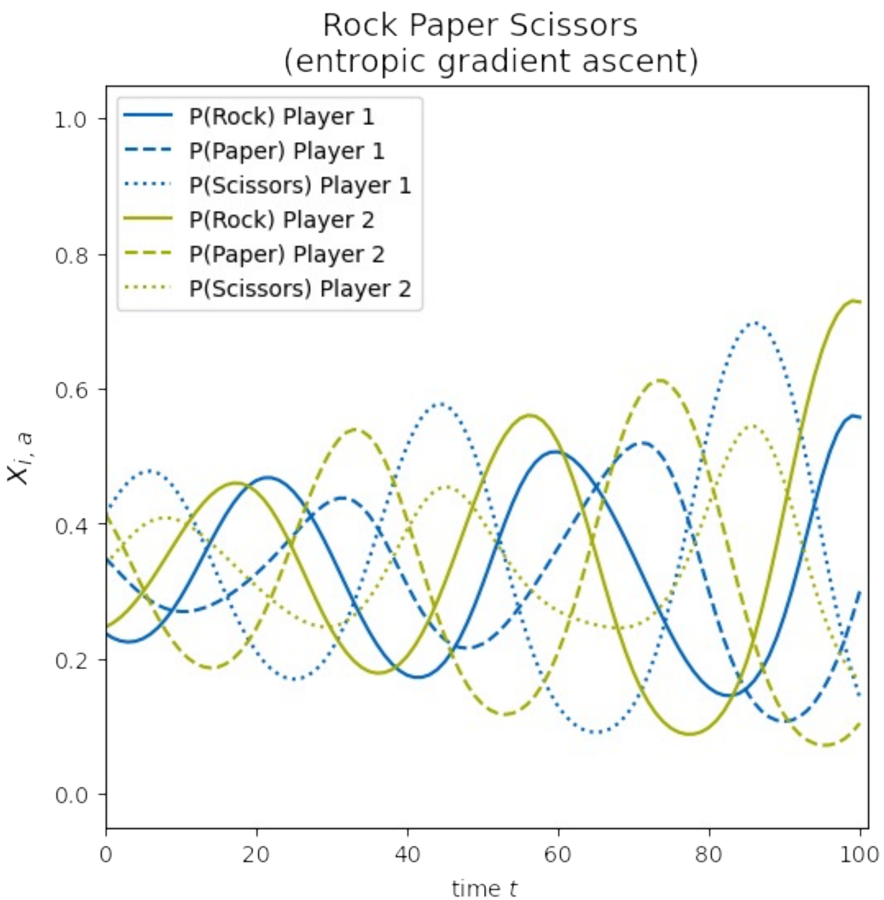


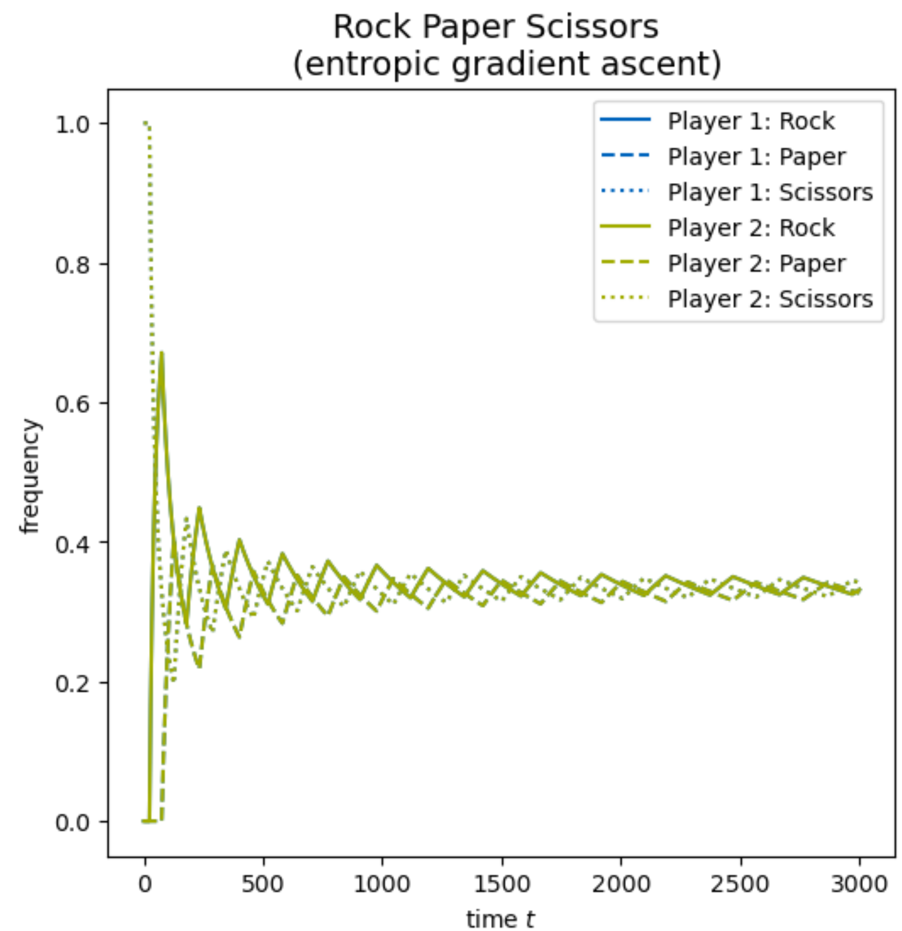
Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Rock Paper Scissors

	Rock	*Paper*	*Scissors*
Rock	$0, 0$	$-1, 1$	$1, -1$
Paper	$1, -1$	$0, 0$	$-1, 1$
Scissors	$-1, 1$	$1, -1$	$0, 0$

$$x_i^* = (1/3, 1/3, 1/3) \qquad \forall i \in \mathcal{N}$$



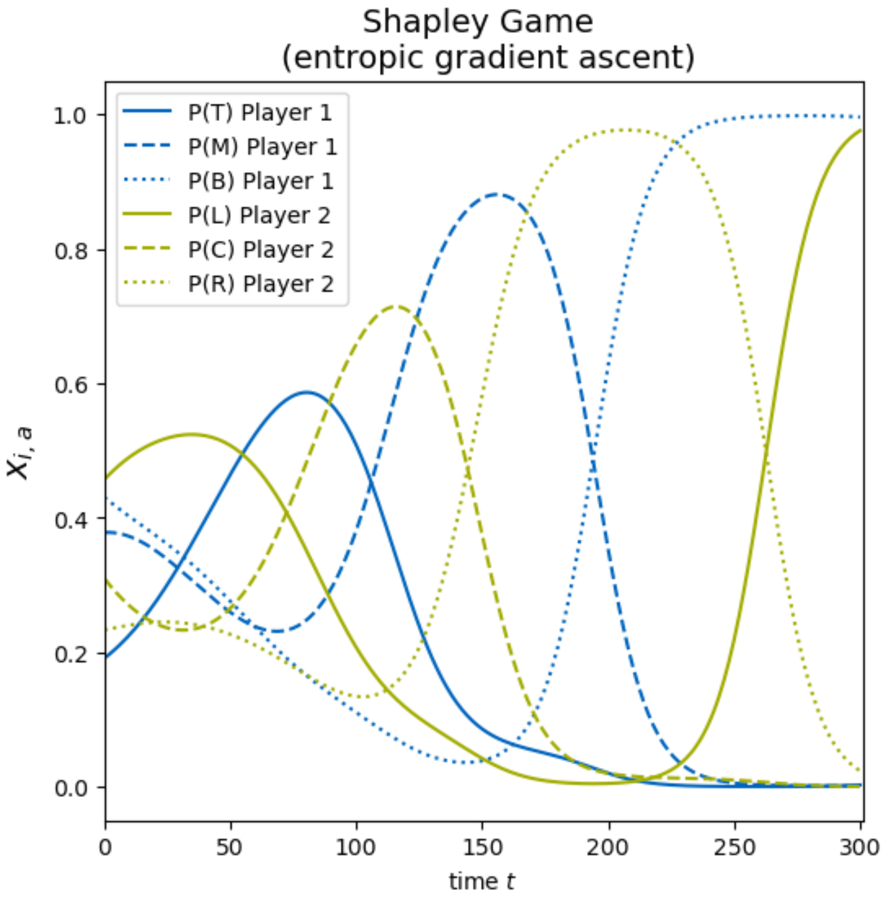


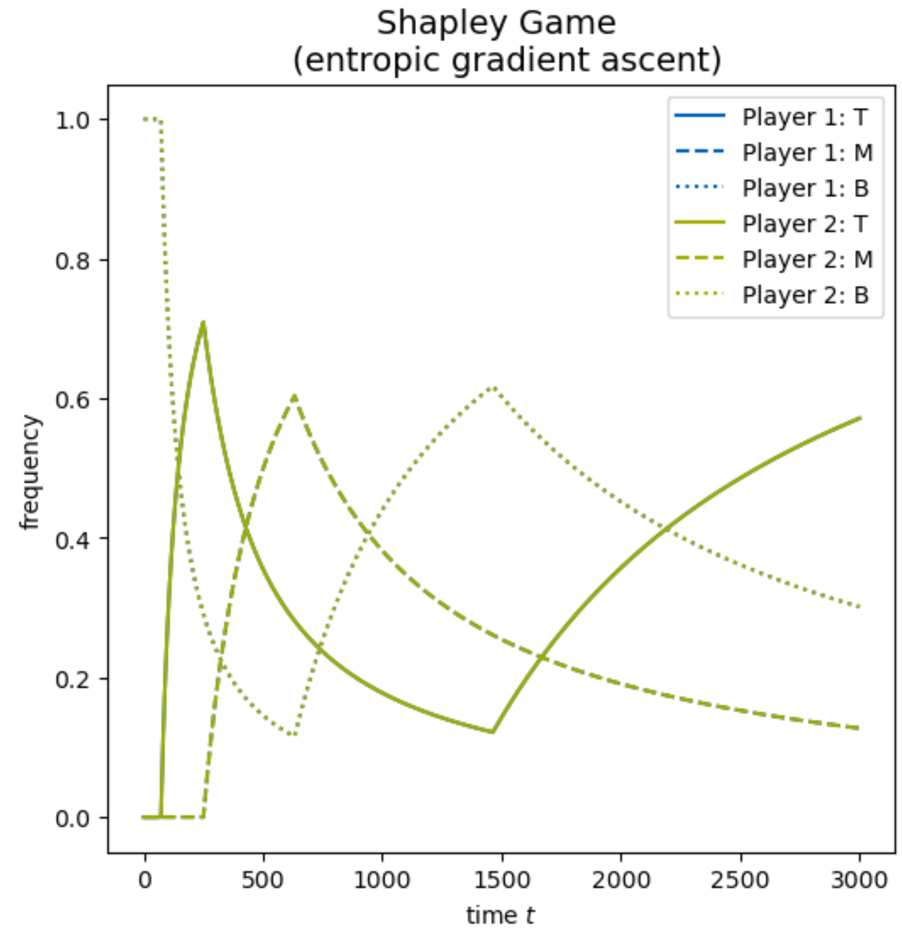
Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Shapley Game

	L	C	R
T	$1,0$	$0,1$	$0,0$
M	$0,0$	$1,0$	$0,1$
B	$0,1$	$0,0$	$1,0$

$$x_i^* = (1/3, 1/3, 1/3) \qquad \forall i \in \mathcal{N}$$



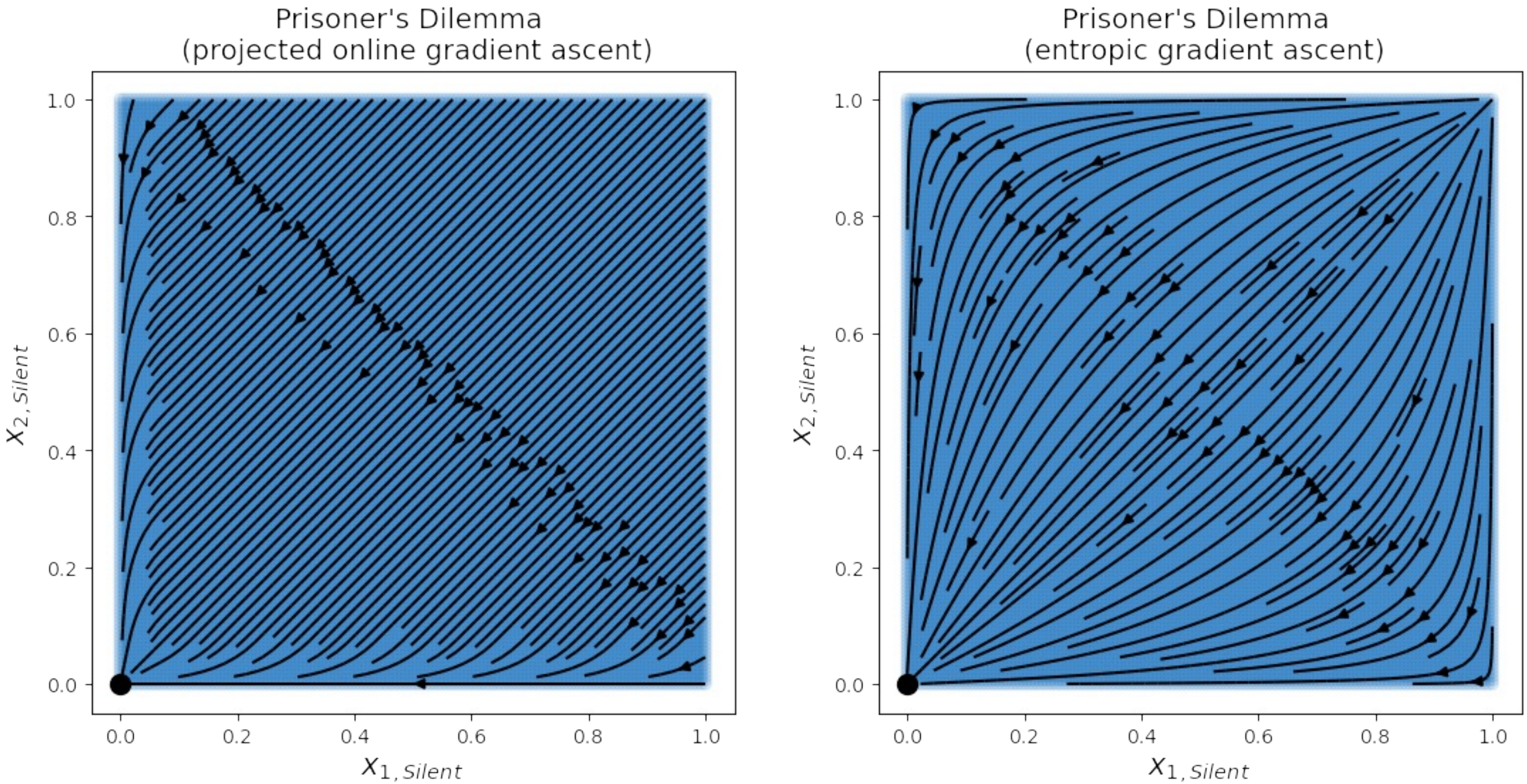


Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Prisoner‘s Dilemma

	$Silent$	$Betray$
$Silent$	$-1,-1$	$-3,0$
$Betray$	$0,-3$	$-2,-2$

$x^* = (Betray, Betray)$ $\quad$ $strict$ PNE

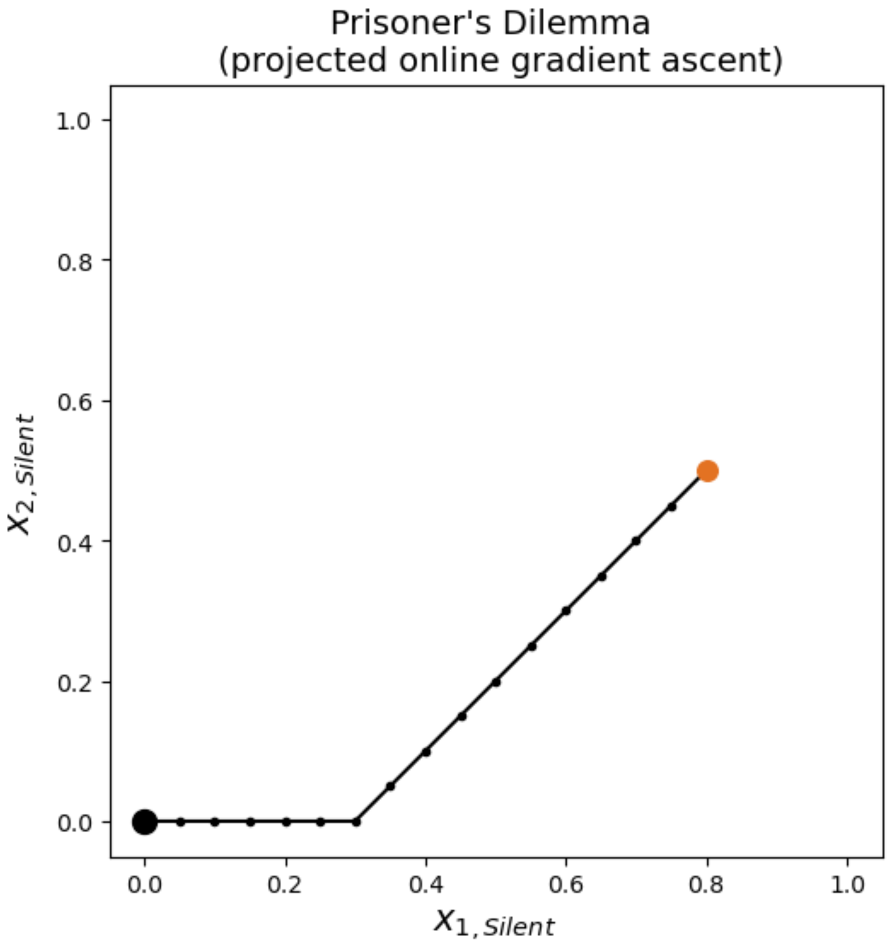


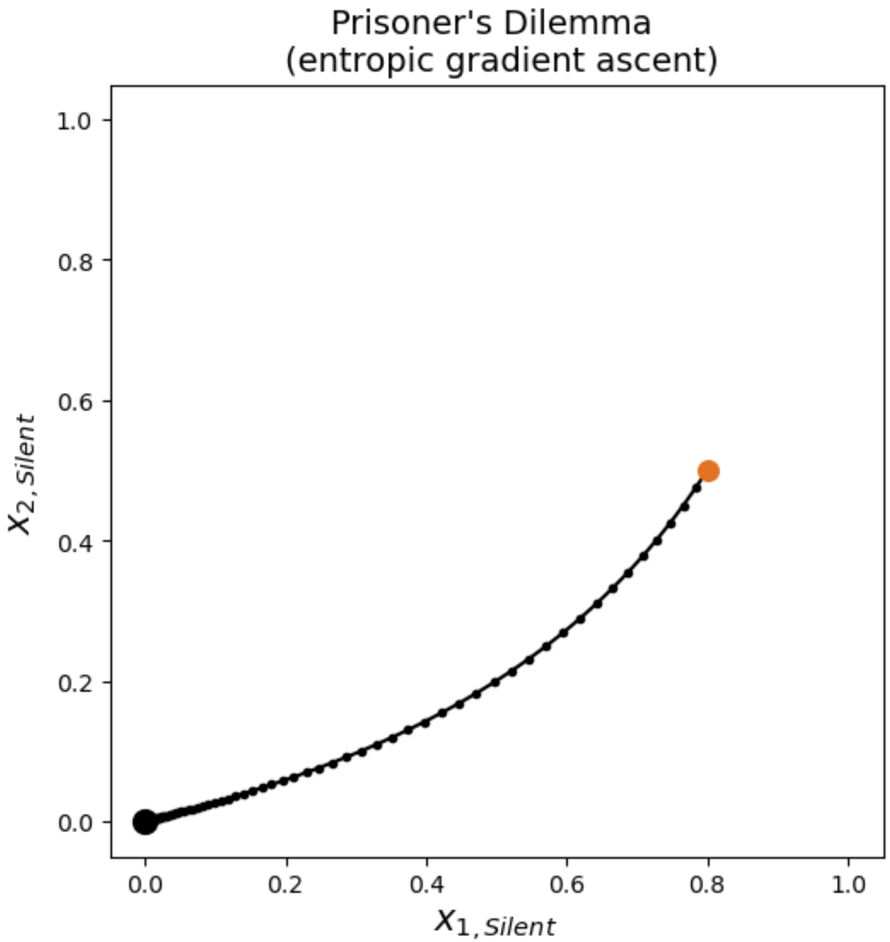
Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

Simulations

	Silent	*Betray*
Silent	$-1,-1$	$-3,0$
Betray	$0,-3$	$-2,-2$

$x^* = (Betray, Betray)$ *strict* PNE





Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

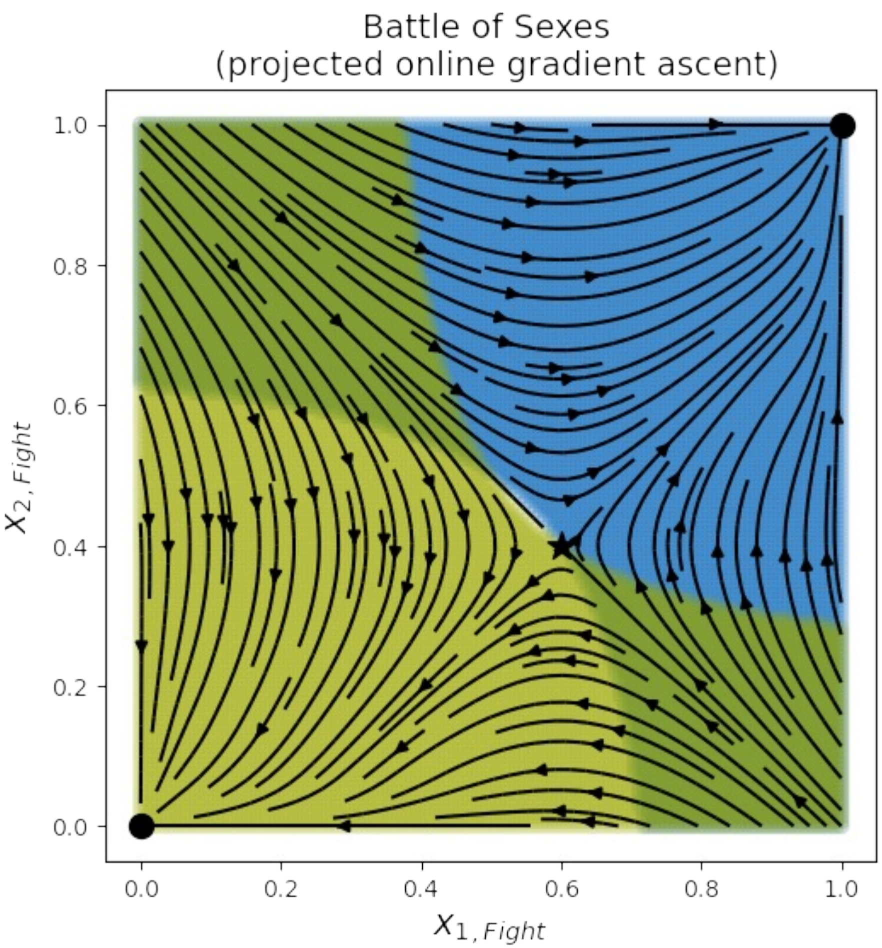
Battle of Sexes

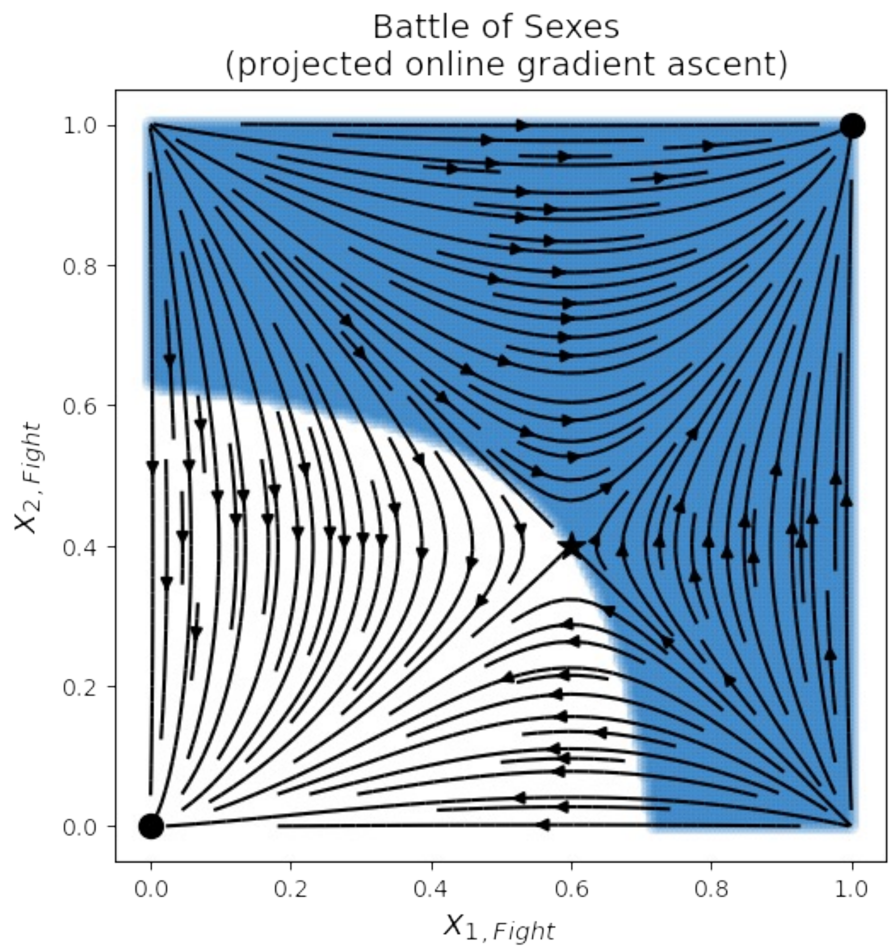
	Fight	*Ballet*
Fight	$3, 2$	$0, 0$
Ballet	$0, 0$	$2, 3$

$x^* = (Fight, Fight)$ *strict* PNE

$x^* = (Ballet, Ballet)$ *strict* PNE

$x_1^* = (3/5, 2/5)$ $x_2^* = (2/5, 3/5)$ MNE





Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

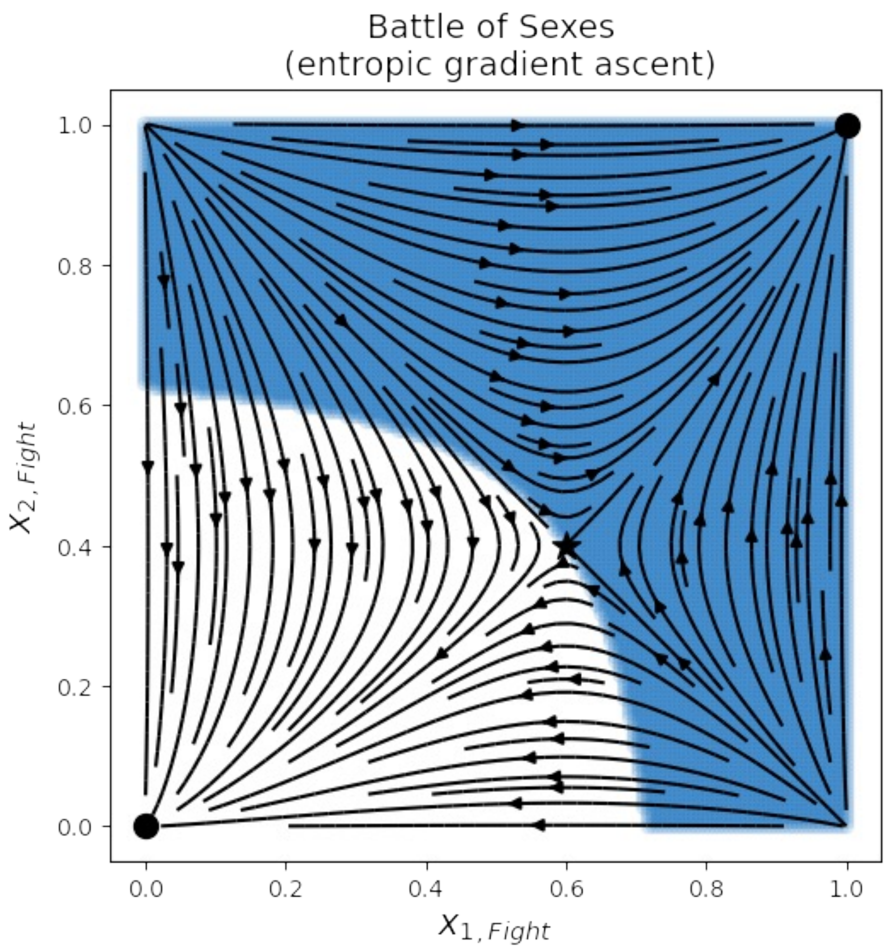
Battle of Sexes

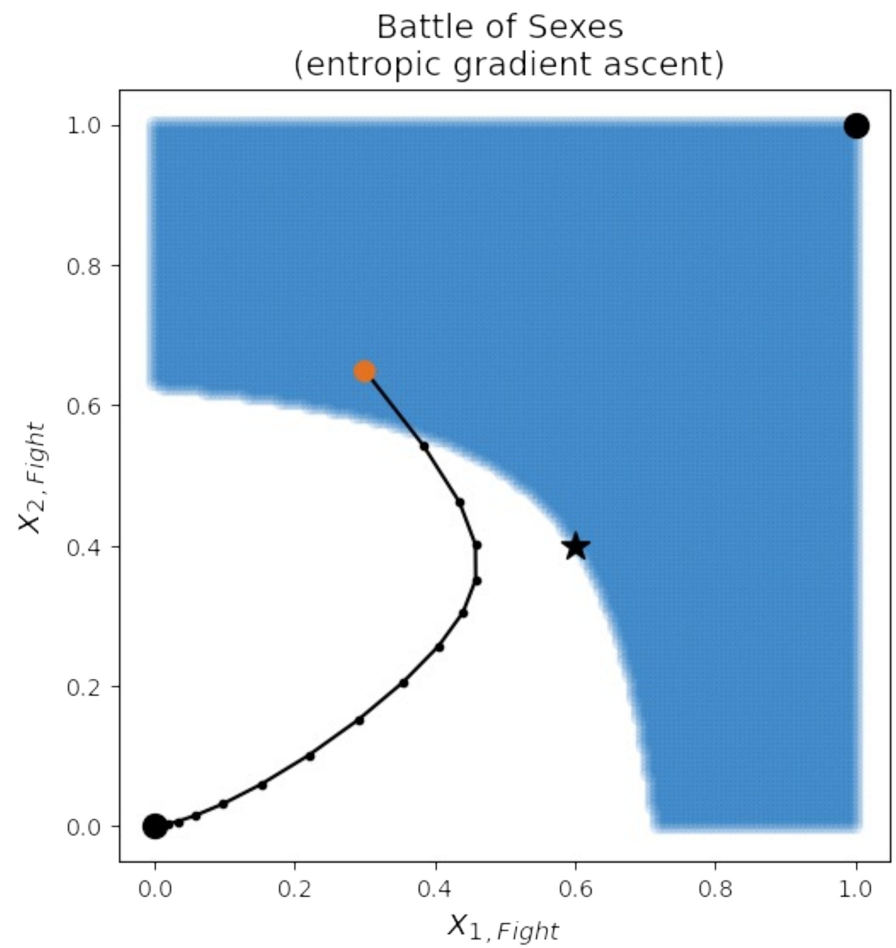
	Fight	*Ballet*
Fight	$3, 2$	$0, 0$
Ballet	$0, 0$	$2, 3$

$x^* = (Fight, Fight)$ *strict* PNE

$x^* = (Ballet, Ballet)$ *strict* PNE

$x_1^* = (3/5, 2/5)$ $x_2^* = (2/5, 3/5)$ MNE





Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

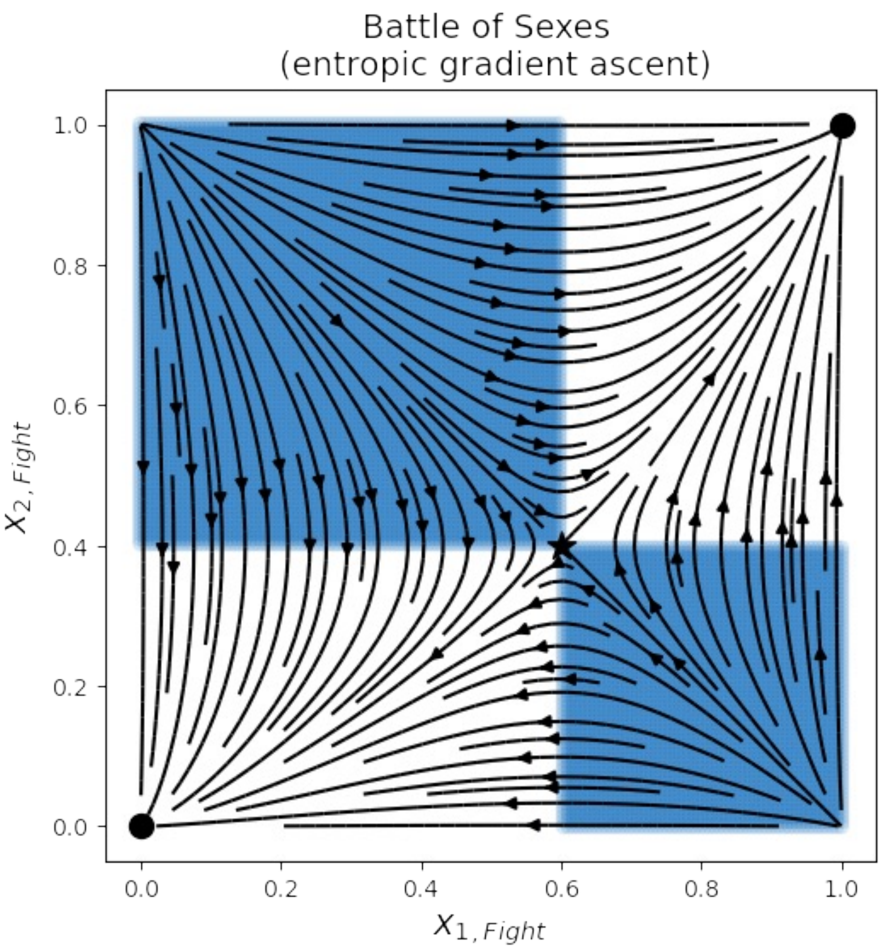
Battle of Sexes

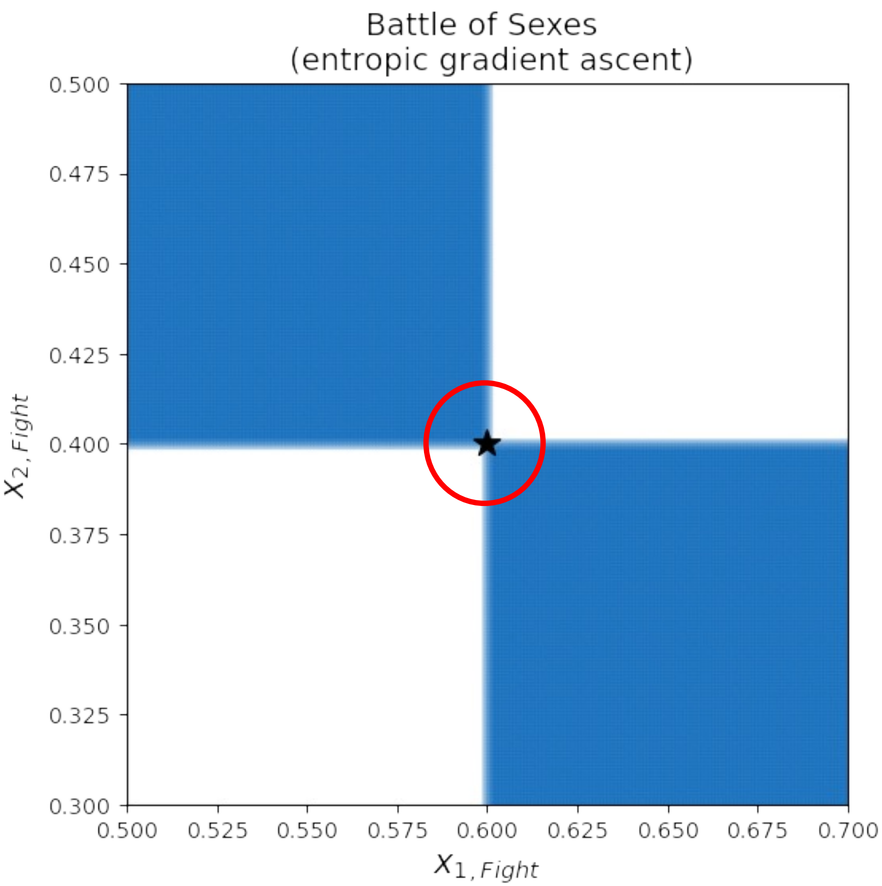
	Fight	*Ballet*
Fight	$3, 2$	$0, 0$
Ballet	$0, 0$	$2, 3$

$x^* = (Fight, Fight)$ *strict* PNE

$x^* = (Ballet, Ballet)$ *strict* PNE

$x_1^* = (3/5, 2/5)$ $x_2^* = (2/5, 3/5)$ MNE





Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

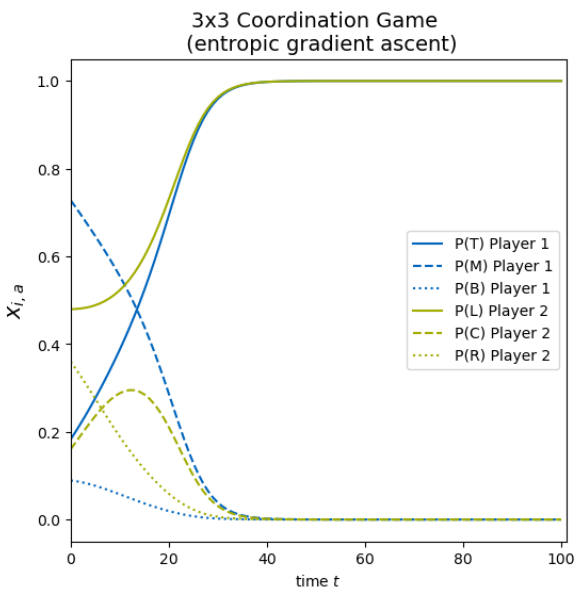
Coordination Game

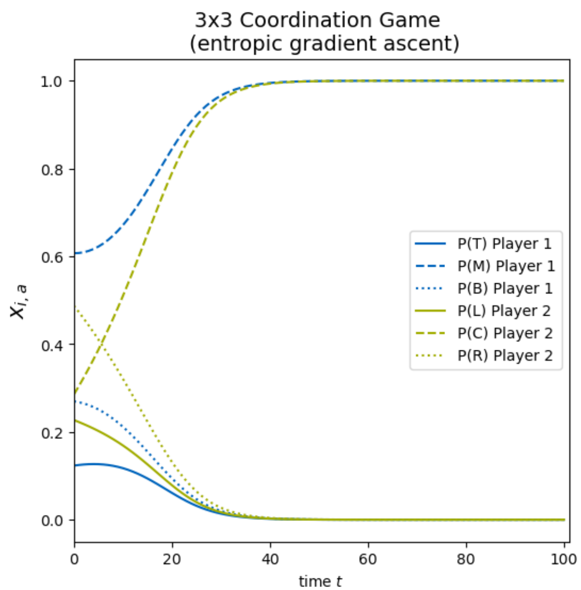
	L	C	R
T	$3, 3$	$0, 0$	$0, 0$
M	$0, 0$	$2, 2$	$0, 0$
B	$0, 0$	$0, 0$	$1, 1$

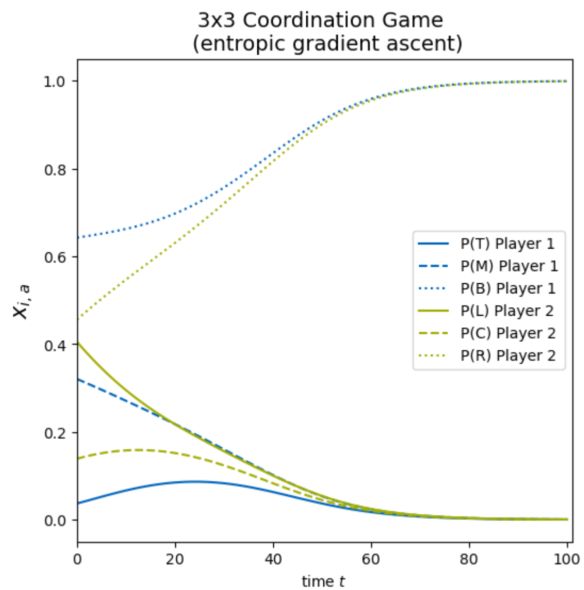
$x^* = (T, L)$ *strict* PNE

$x^* = (M, C)$ *strict* PNE

$x^* = (B, R)$ *strict* PNE







Relative frequencies (10000 initial strategies):

```
{'(T,L)': 0.64, '(M,C)': 0.28, '(B,R)': 0.08}
```

Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

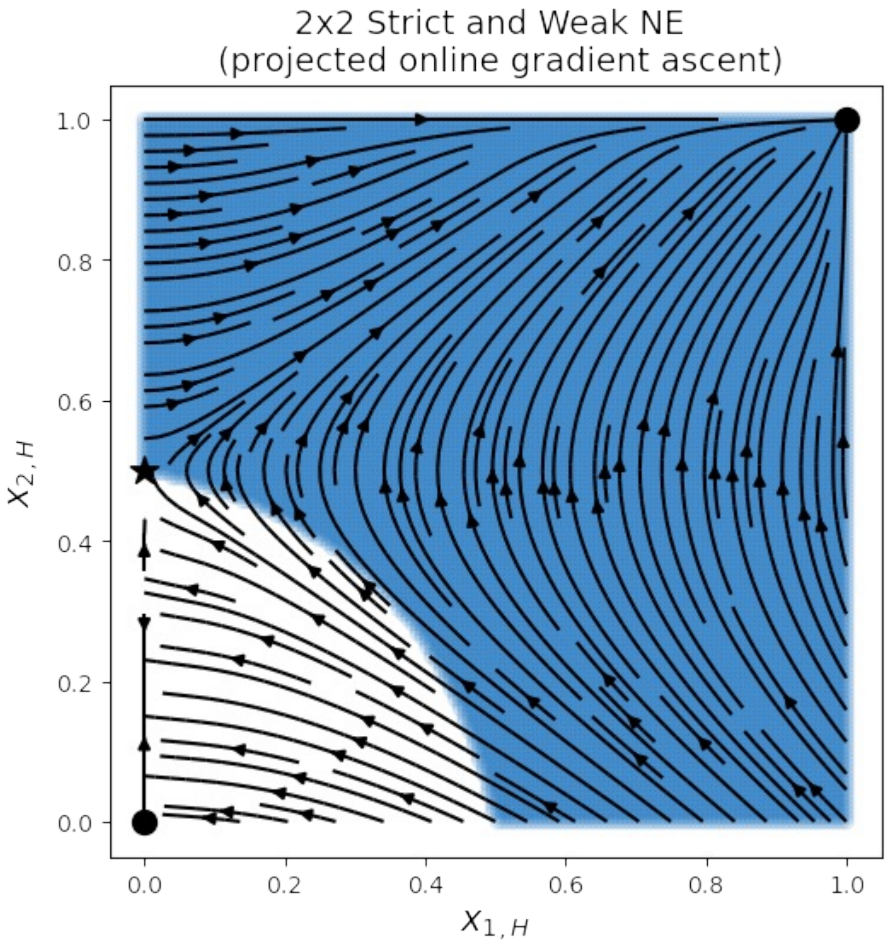
Strict and Weak 2x2

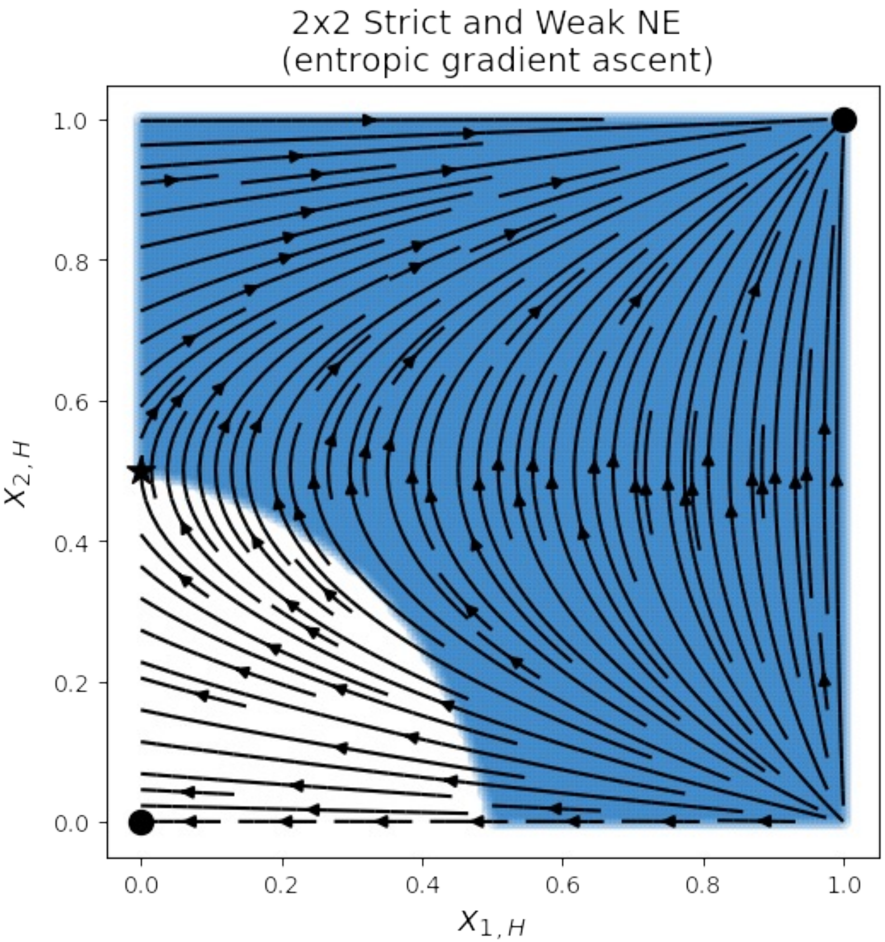
	H	T
H	$2,3$	$1,2$
T	$1,2$	$2,2$

$x^* = (H,H)$ *strict* PNE

$x^* = (T,T)$ *weak* PNE

$x_1^*(0,0)$ $x_2^*(1/2,1/2)$ MNE





Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

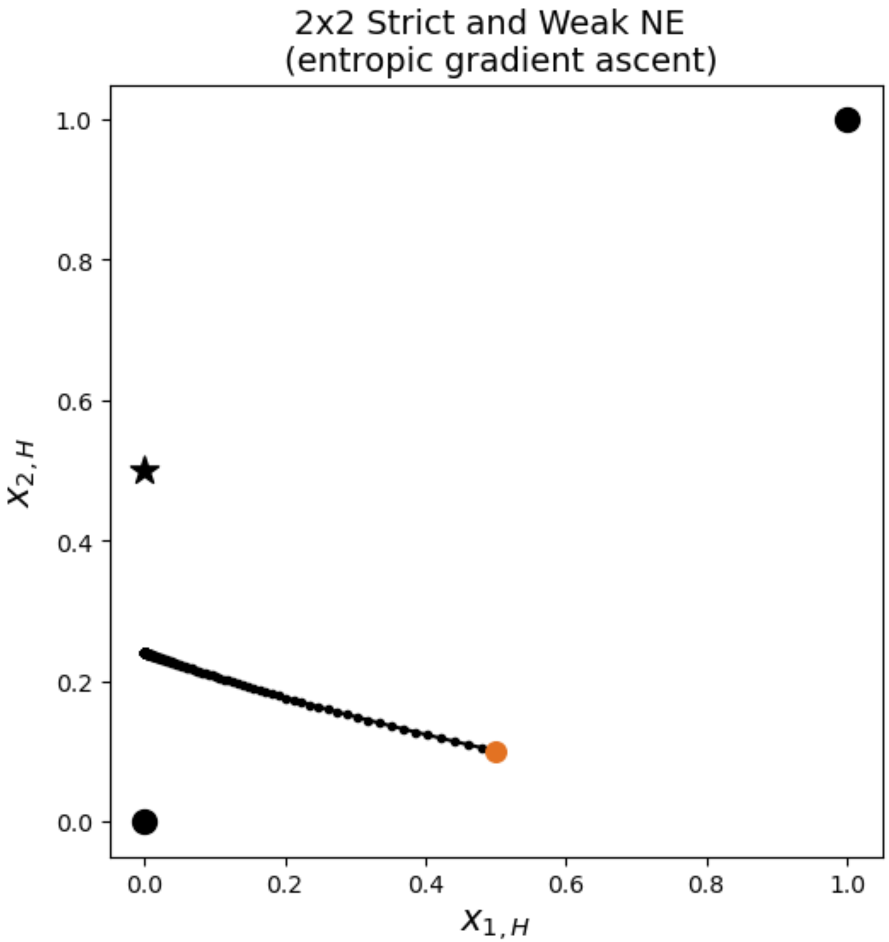
Strict and Weak 2x2

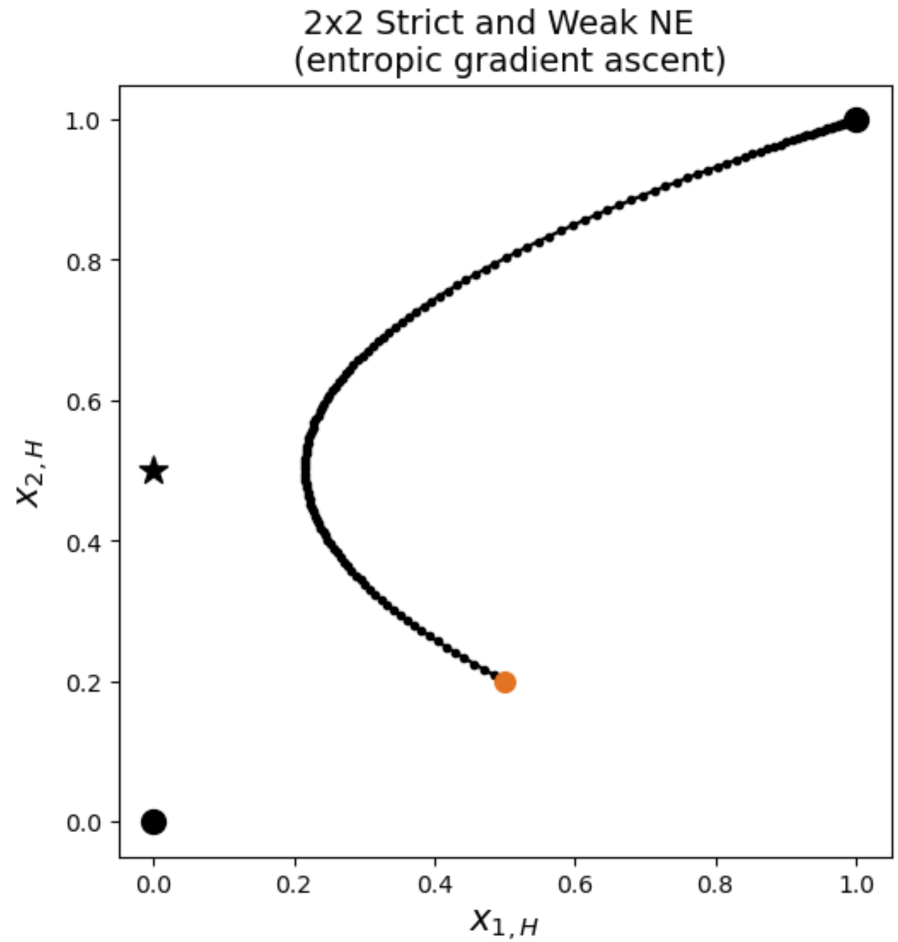
	H	T
H	$2,3$	$1,2$
T	$1,2$	$2,2$

$x^* = (H,H)$ *strict* PNE

$x^* = (T,T)$ *weak* PNE

$x_1^*(0,0)$ $x_2^*(1/2,1/2)$ MNE





Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

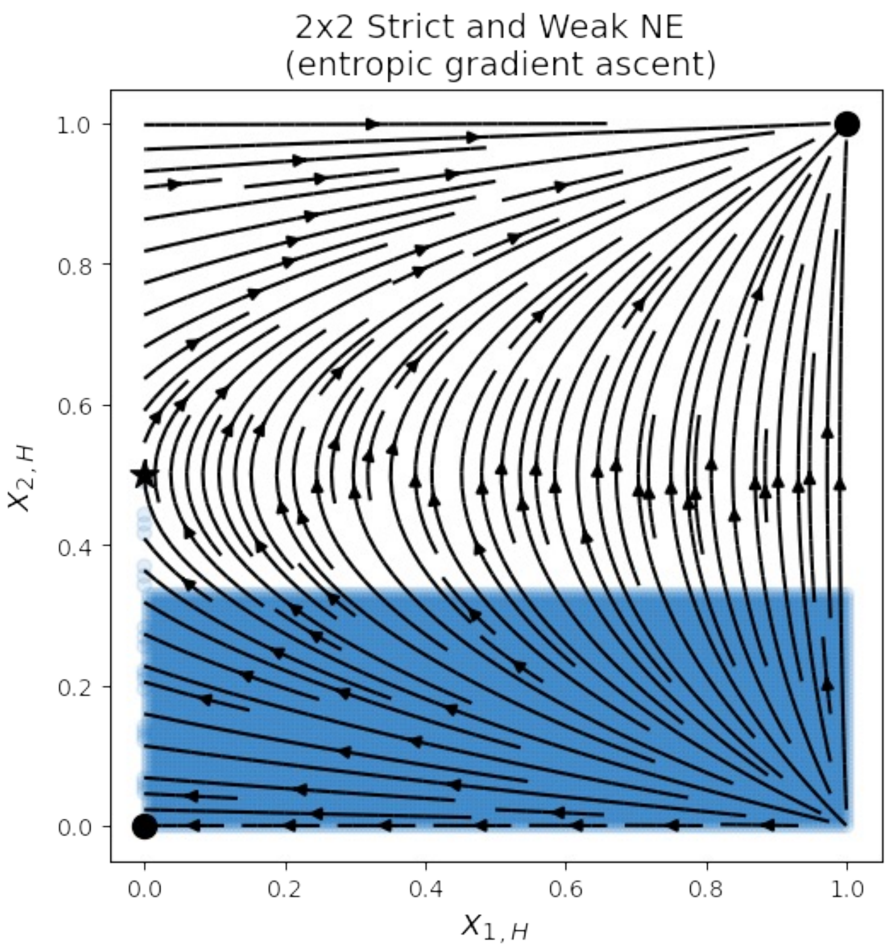
Strict and Weak 2x2

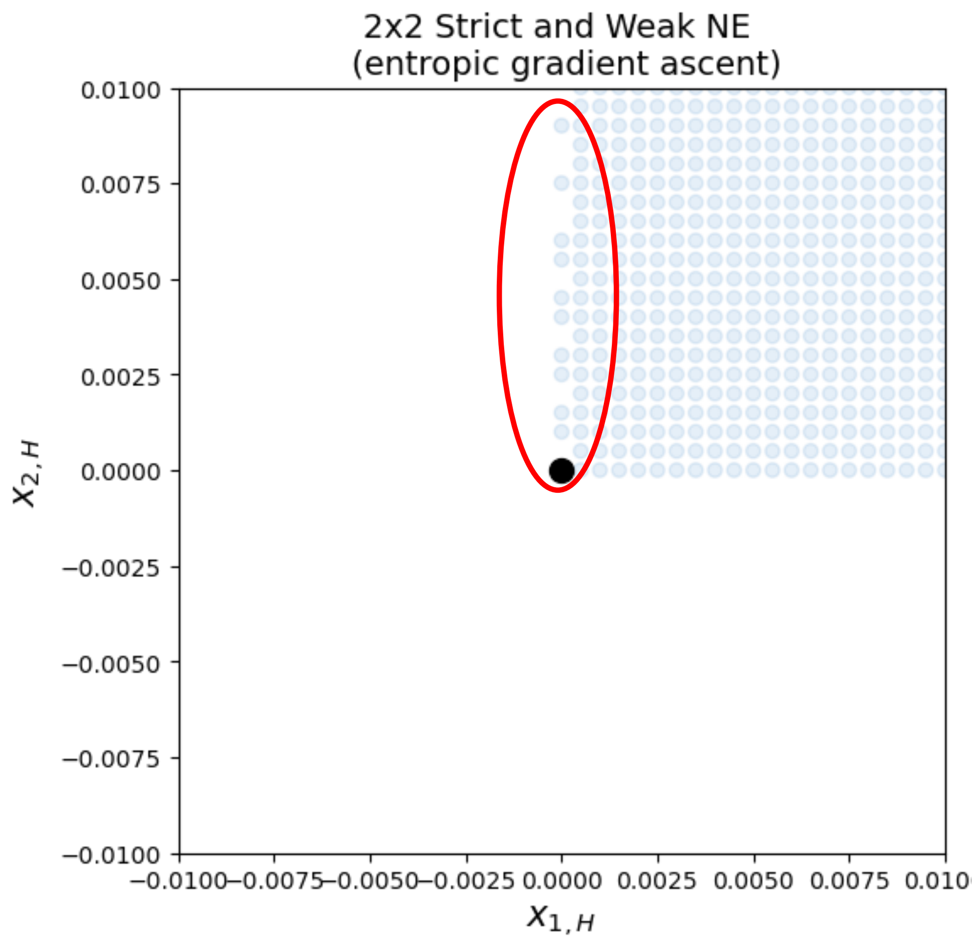
	H	T
H	$2,3$	$1,2$
T	$1,2$	$2,2$

$x^* = (H,H)$ *strict* PNE

$x^* = (T,T)$ *weak* PNE

$x_1^*(0,0)$ $x_2^*(1/2,1/2)$ MNE





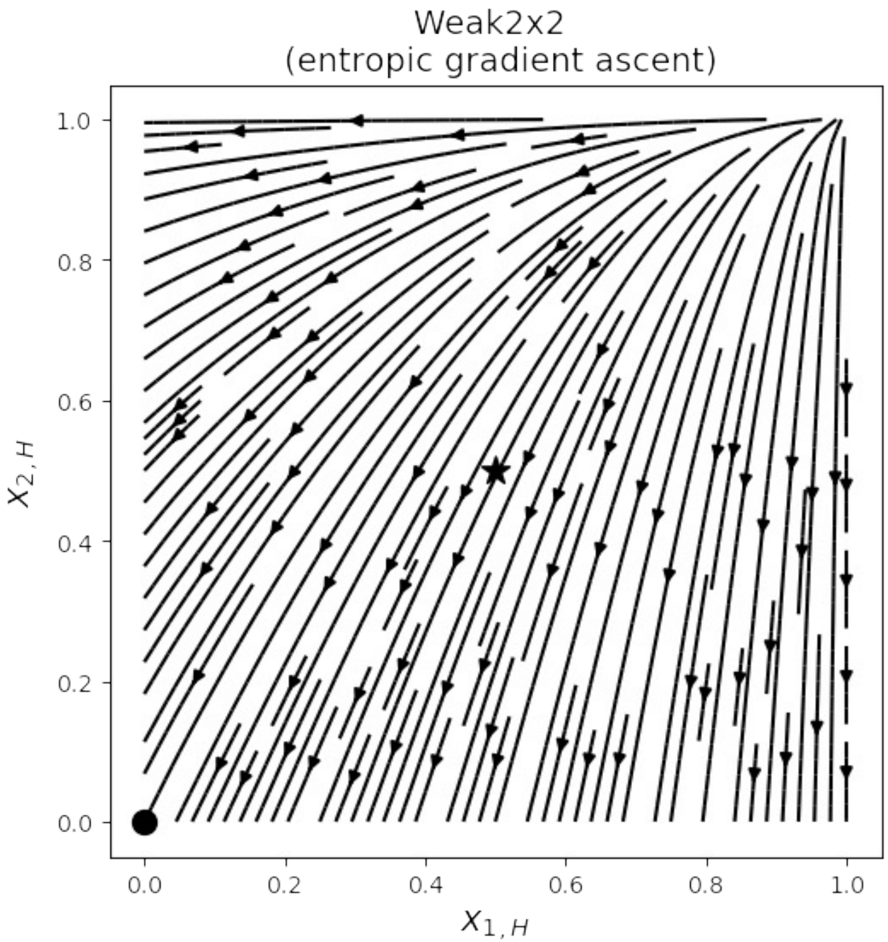
Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

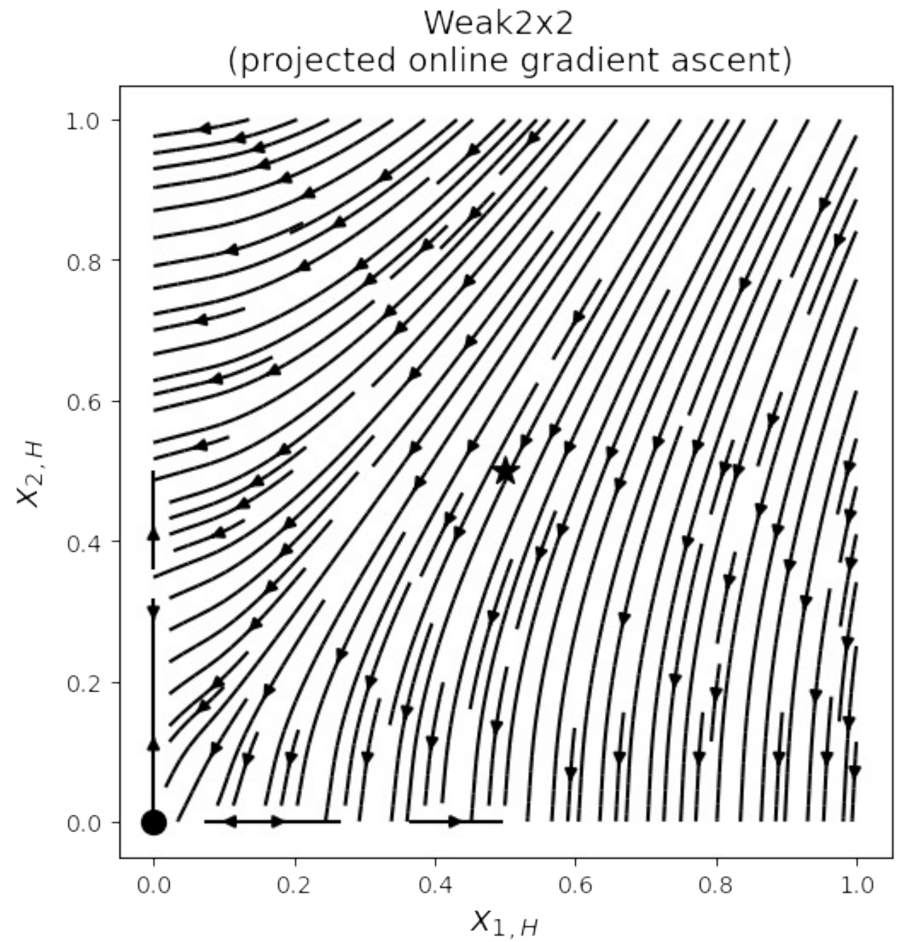
Weak 2x2

	H	T
H	$1, 1$	$1, 2$
T	$0, 2$	$2, 2$

$x^* = (T, T)$ *weak* PNE

$x_1^* = x_2^* = (1/2, 1/2)$ MNE





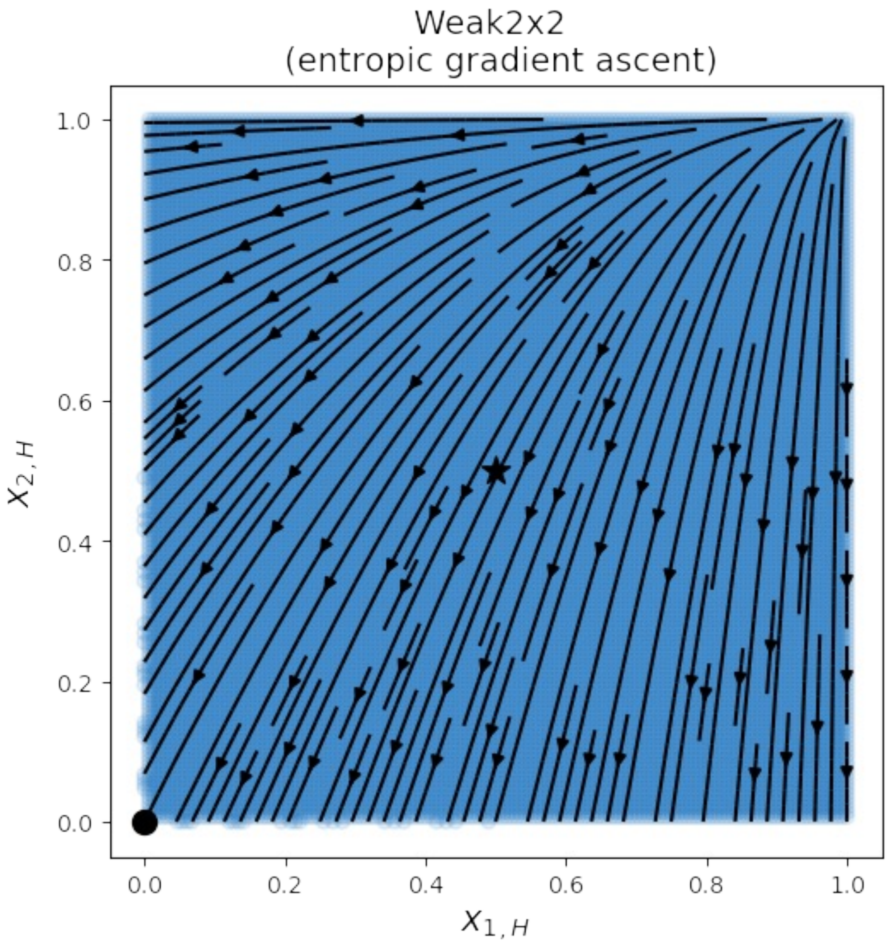
Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

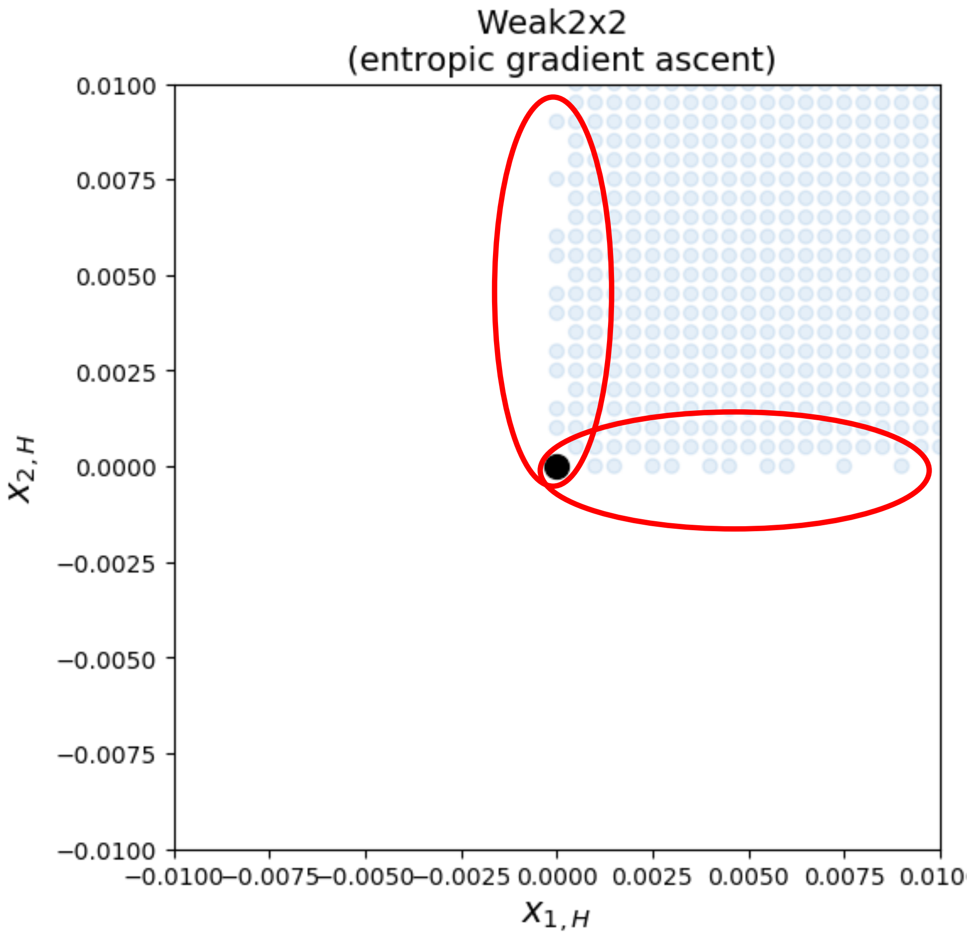
Weak 2x2

	H	T
H	$1, 1$	$1, 2$
T	$0, 2$	$2, 2$

$x^* = (T, T)$ *weak* PNE

$x_1^* = x_2^* = (1/2, 1/2)$ MNE





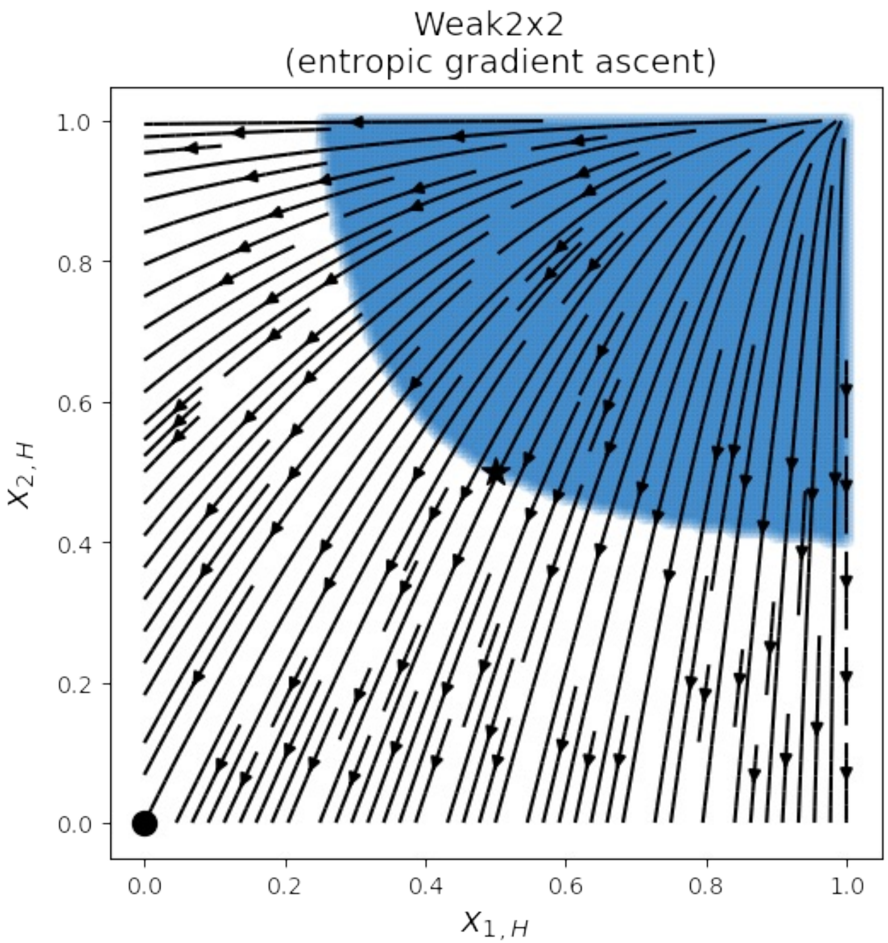
Decision Sciences & Systems (DSS)
Department of Informatics (I18)
Technical University of Munich

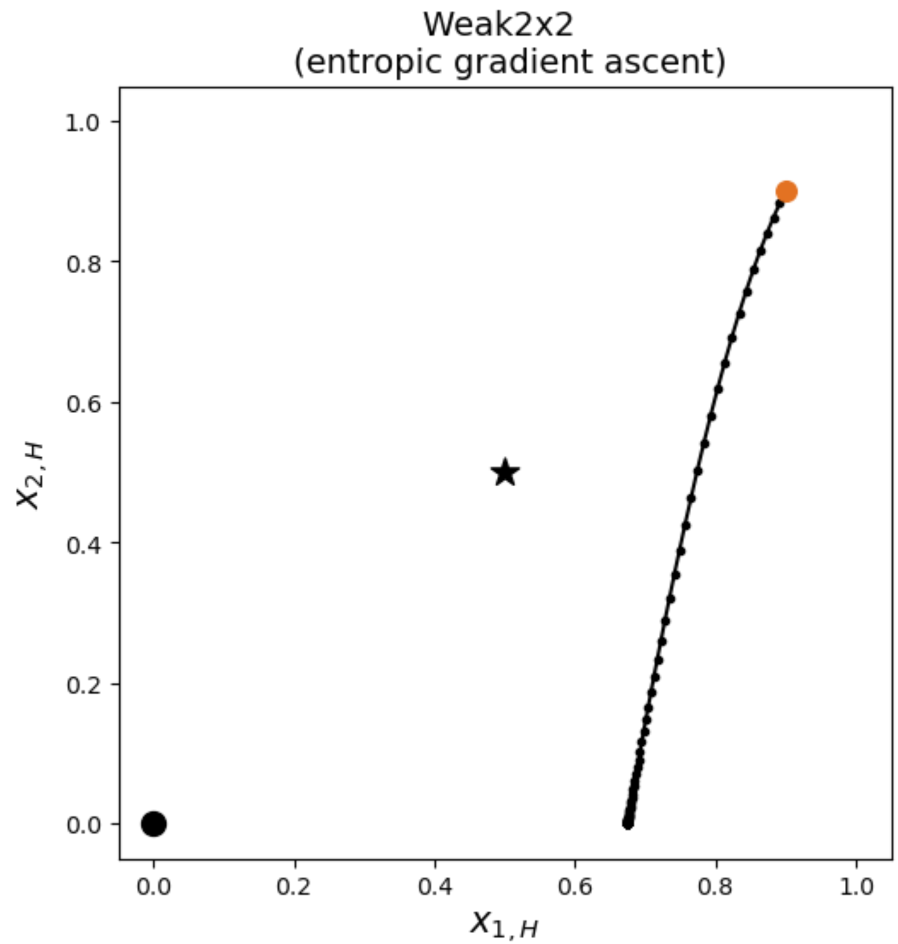
Weak 2x2

	H	T
H	$1, 1$	$1, 2$
T	$0, 2$	$2, 2$

$x^* = (T, T)$ *weak* PNE

$x_1^* = x_2^* = (1/2, 1/2)$ MNE





References

[1] X. Chen and X. Deng. “Settling the complexity of two-player Nash equilibrium.” In: 2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS’06). IEEE. 2006, pp. 261–272.

[2] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. “The complexity of computing a Nash equilibrium.” In: SIAM Journal on Computing 39.1 (2009), pp. 195–259.

[3] L. Flokas, E.-V. Vlatakis-Gkaragkounis, T. Lianeas, P. Mertikopoulos, and G. Piliouras. “No-regret learning and mixed Nash equilibria: They do not mix.” In: arXiv preprint arXiv:2010.09514 (2020).

[4] P. Mertikopoulos. “Online optimization and learning in games: Theory and applications.” PhD thesis. Grenoble 1 UGA-Université Grenoble Alpes, 2019.

[5] J. Nash. “Non-cooperative games.” In: Annals of mathematics (1951), pp. 286–295.

[6] P. Mertikopoulos and W. H. Sandholm. “Learning in games via reinforcement and regularization.” In: Mathematics of Operations Research 41.4 (2016), pp. 1297–1324.

Q&A

Thank you ☺

Backup Slides

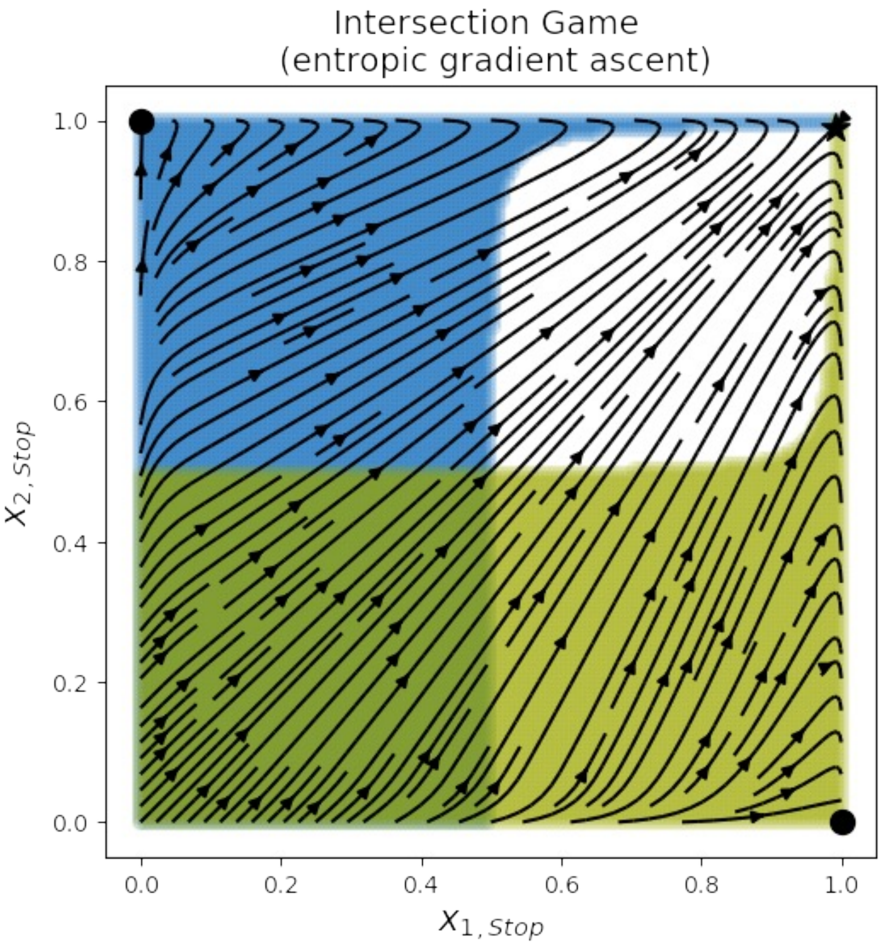
Intersection Game

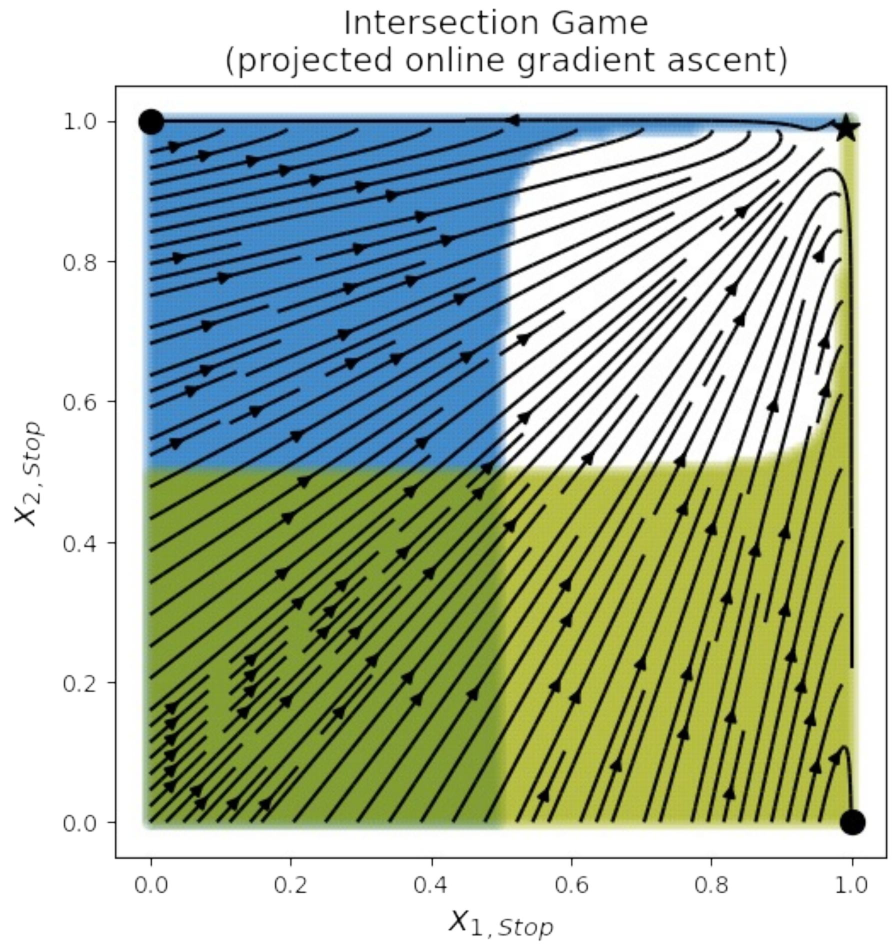
	$Stop$	Go
$Stop$	$0, 0$	$0, 1$
Go	$1, 0$	$-100, -100$

$x^* = (Go, Go)$ *strict* PNE

$x^* = (Stop, Stop)$ *strict* PNE

$x_1^* = x_2^* = (100/101, 1/101)$ MNE





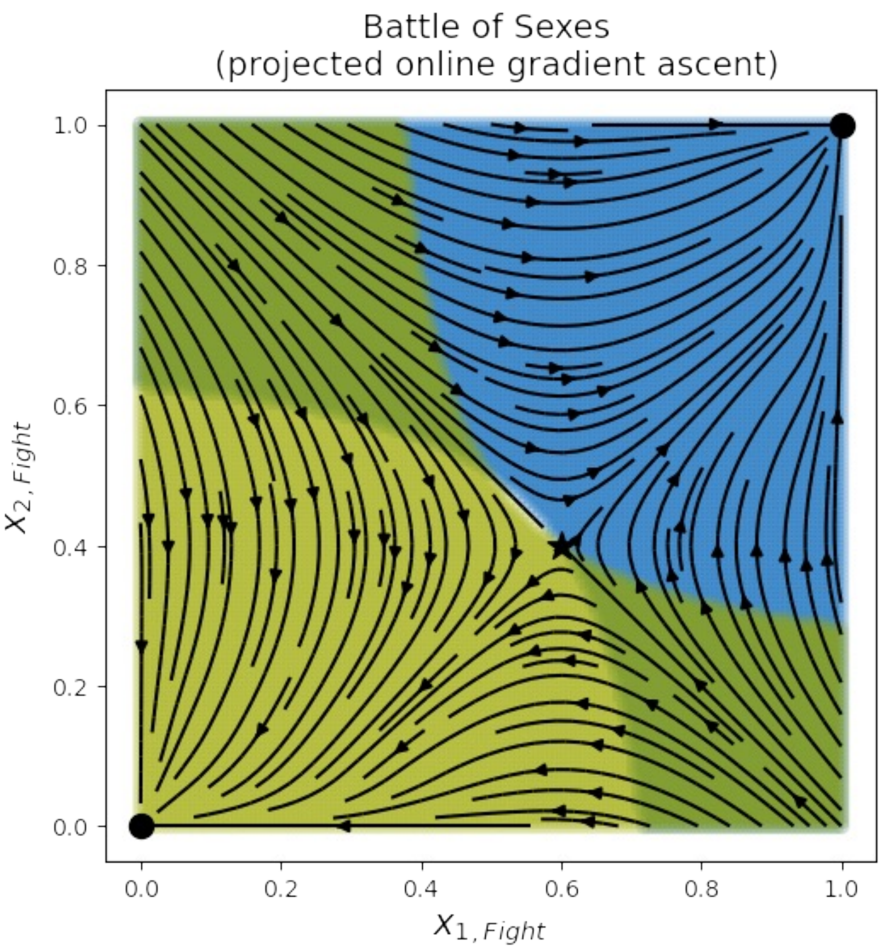
Battle of Sexes

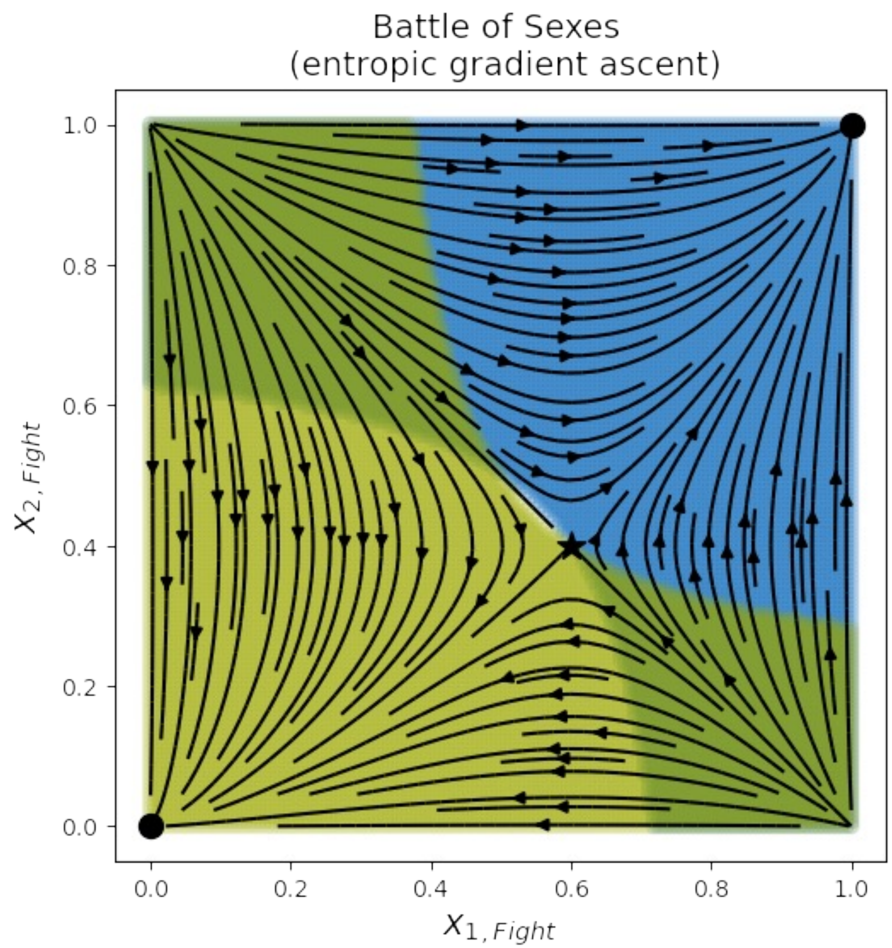
	Fight	*Ballet*
Fight	$3, 2$	$0, 0$
Ballet	$0, 0$	$2, 3$

$x^* = (Fight, Fight)$ *strict* PNE

$x^* = (Betray, Betray)$ *strict* PNE

$x_1^* = (2/3, 1/3)$ $x_2^* = (1/3, 2/3)$ MNE





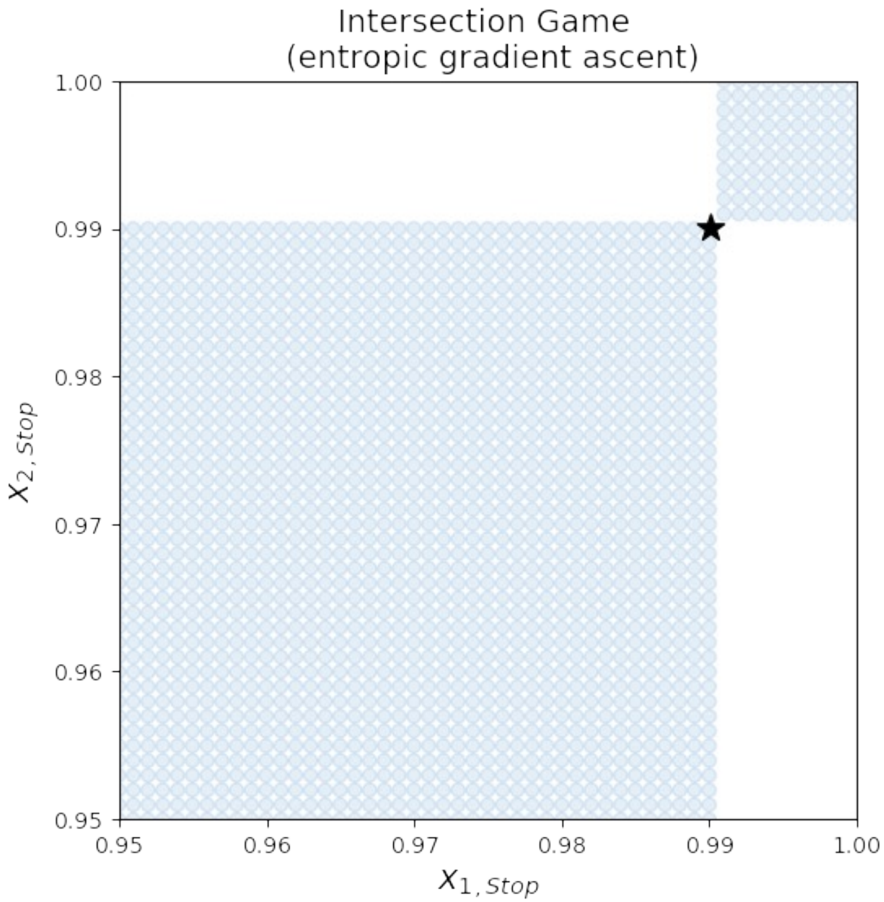
Intersection Game

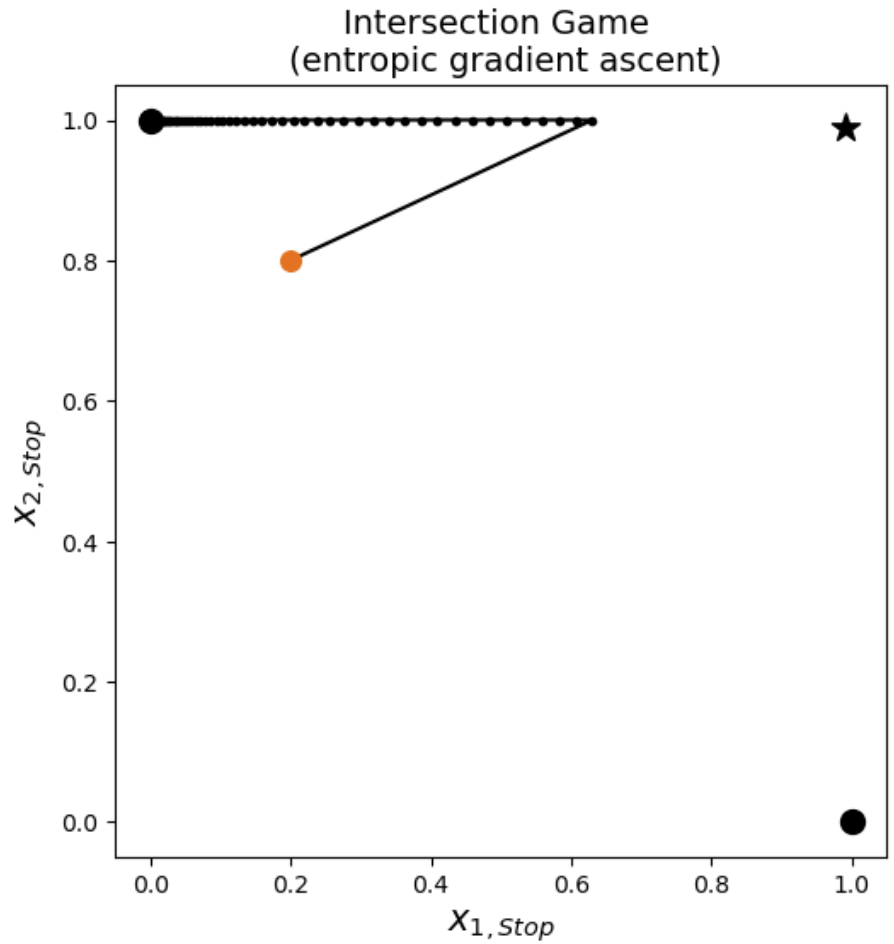
	Stop	*Go*
Stop	$0, 0$	$0, 1$
Go	$1, 0$	$-100, -100$

$x^* = (Go, Go)$ *strict* PNE

$x^* = (Stop, Stop)$ *strict* PNE

$x_1^* = x_2^* = (100/101, 1/101)$ MNE





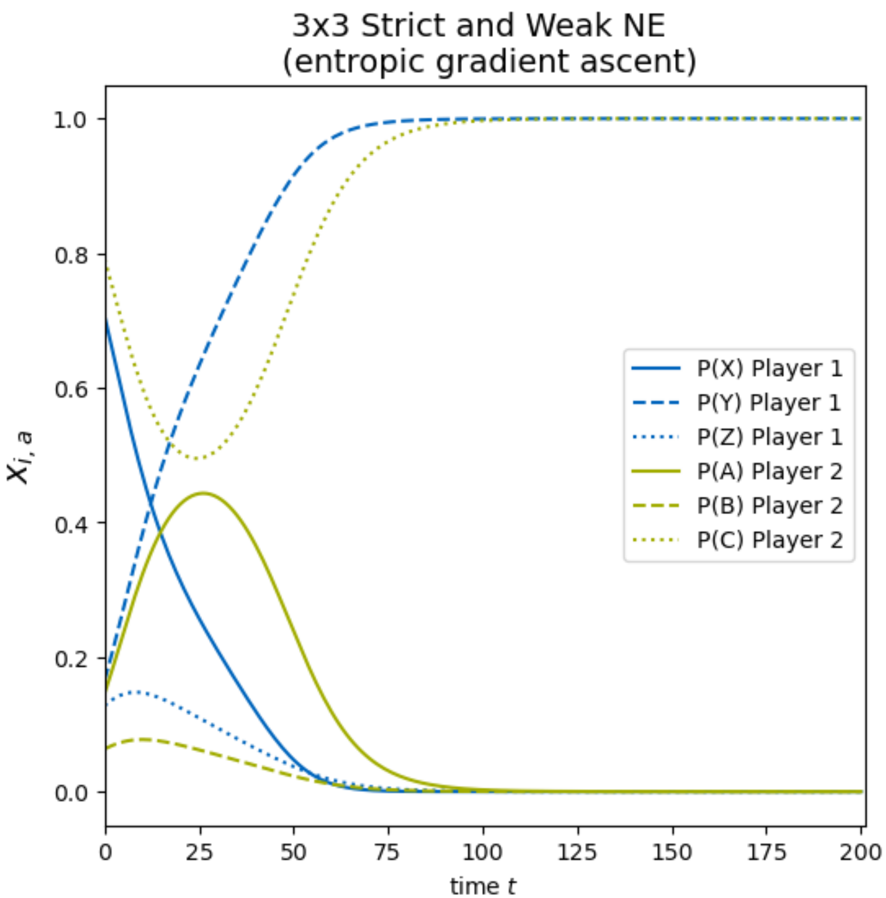
Strict and Weak 3x3

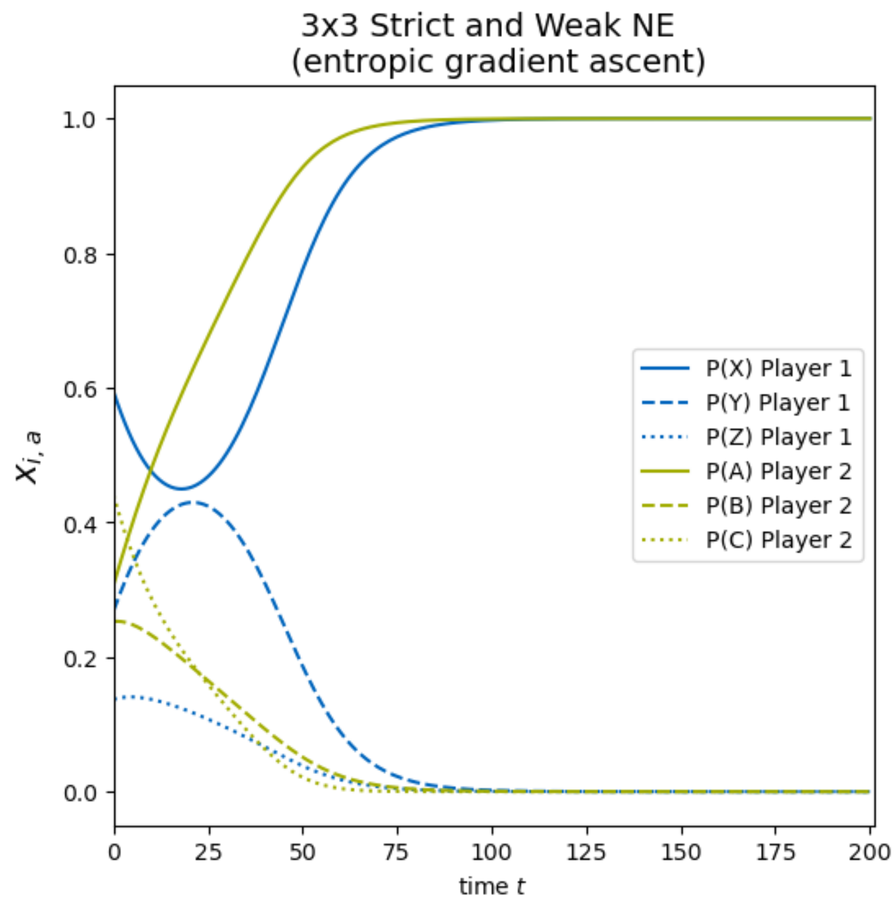
	A	B	C
X	$2,3$	$1,2$	$1,1$
Y	$1,1$	$2,1$	$3,2$
Z	$1,2$	$2,2$	$2,1$

$x^* = (X, A)$ *strict* PNE

$x^* = (Y, C)$ *strict* PNE

$x^* = (Z, B)$ *weak* PNE





Leader Following Policies

Follow the leader

$$x_{t+1} = \underset{x \in \mathcal{X}}{\operatorname{argmin}} \sum_{i=1}^{t} l_i(x)$$

Leader Following Policies

Follow the regularized leader (FTRL)

$$x_{t+1} = \operatorname*{argmin}_{x \in \mathcal{X}} \left\{ \sum_{i=1}^{t} l_i(x) + \frac{1}{\gamma} h(x) \right\}$$

$h : \mathcal{X} \to \mathbb{R}$: regularization function

$\gamma > 0$: step size

Assume l_t **is:**

- L-Lipschitz

Assume h **is:**

- K-strongly convex
- Continuous

Set: $\gamma \sim \dfrac{1}{L\sqrt{T}}$

→ $reg(T) = \mathcal{O}(\sqrt{T})$ → **FTRL** is a **no regret** algorithm [4]

Nash Existence vs. PPAD Completeness

Nash existence: Every finite game with mixed extension yield a Nash equilibrium [5]

PPAD completeness: Finding a Nash equilibrium is **PPAD-complete** [1,2]

Backup Slides

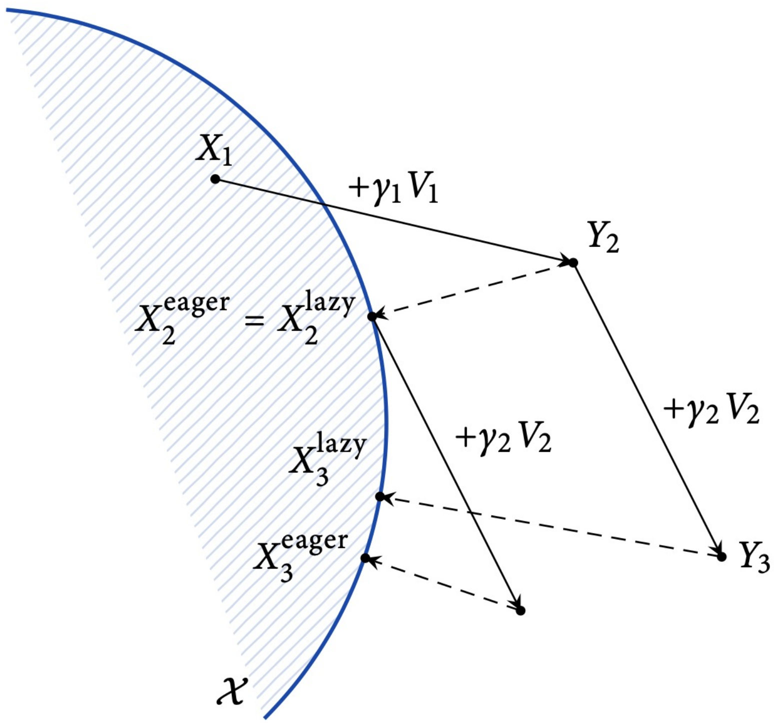
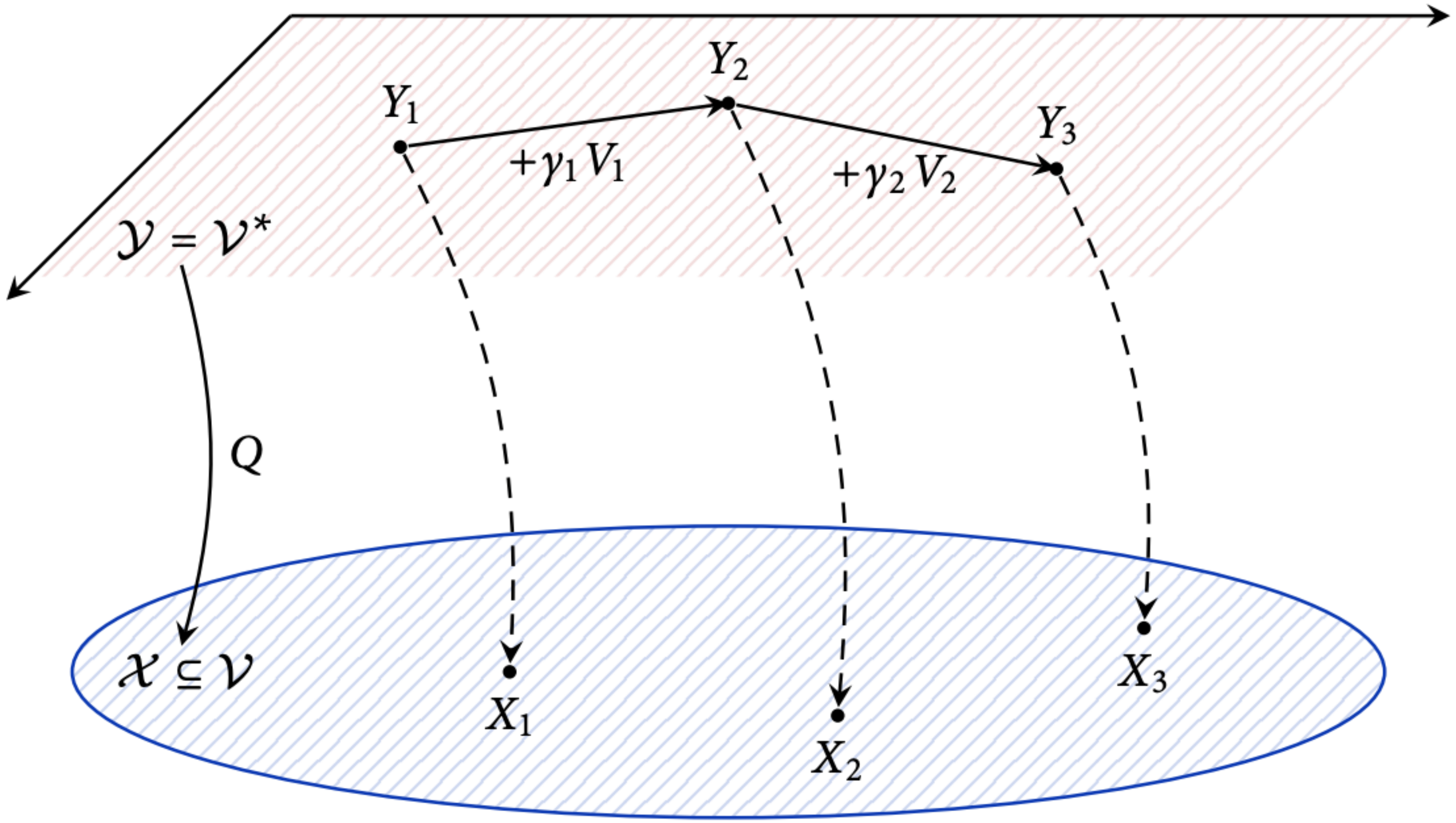


Figure 2.4: Lazy vs. eager gradient descent.

Online Mirror Descent



An Example

	Rock	Paper	Scissors
Rock	$0,0$	$-1,1$	$1,-1$
Paper	$1,-1$	$0,0$	$-1,1$
Scissors	$-1,1$	$1,-1$	$0,0$

defines $u_i : \mathcal{A} \to \mathbb{R}$

Finite Game with Mixed Extensions:

- Players $\mathcal{N} = \{1, 2\}$ called row and column player
- Pure strategies $\mathcal{A}_1 = \mathcal{A}_2 = \{Rock, Paper, Scissor\} \equiv \{R, P, S\}$
- Mixed strategies $x_1 = (1/2, 1/2, 0)$ $x_2 = (1, 0, 0)$ $A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$
- Utility function $u_1(x) = {x_1}^T A x_2$ $u_2(x) = {x_1}^T B x_2$ $B = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$

  $u_1(x) = 1/2$ $u_2(x) = -1/2$
- Gradients $v_1(x) = \nabla_{x_1} u_1(x) = A x_2$ $v_2(x) = \nabla_{x_2} u_2(x) = x_1^T B$

Correlated Equilibria

	$Stop$	Go
$Stop$	$0, 0$	$0, 1$
Go	$1, 0$	$-100, -100$

Correlated equilibria (CE): $\mathbb{E}_{a\sim\mathcal{D}}[u_i(a)] \geq \mathbb{E}_{a\sim\mathcal{D}}[u_i(a_i^*, a_{-i})|a_i]$

Coarse Correlated equilibria (CCE): $\mathbb{E}_{a\sim\mathcal{D}}[u_i(a)] \geq \mathbb{E}_{a\sim\mathcal{D}}[u_i(a_i^*, a_{-i})]$

MNE	$Stop$	Go
$Stop$	98%	$< 1\%$
Go	$< 1\%$	$\approx 0.01\%$

$$p = 100/101$$

CE	$Stop$	Go
$Stop$	0%	50%
Go	50%	0%

$$PNE \subset MNE \subset CE \subset CCE$$

Trajectories vs. time-averages:

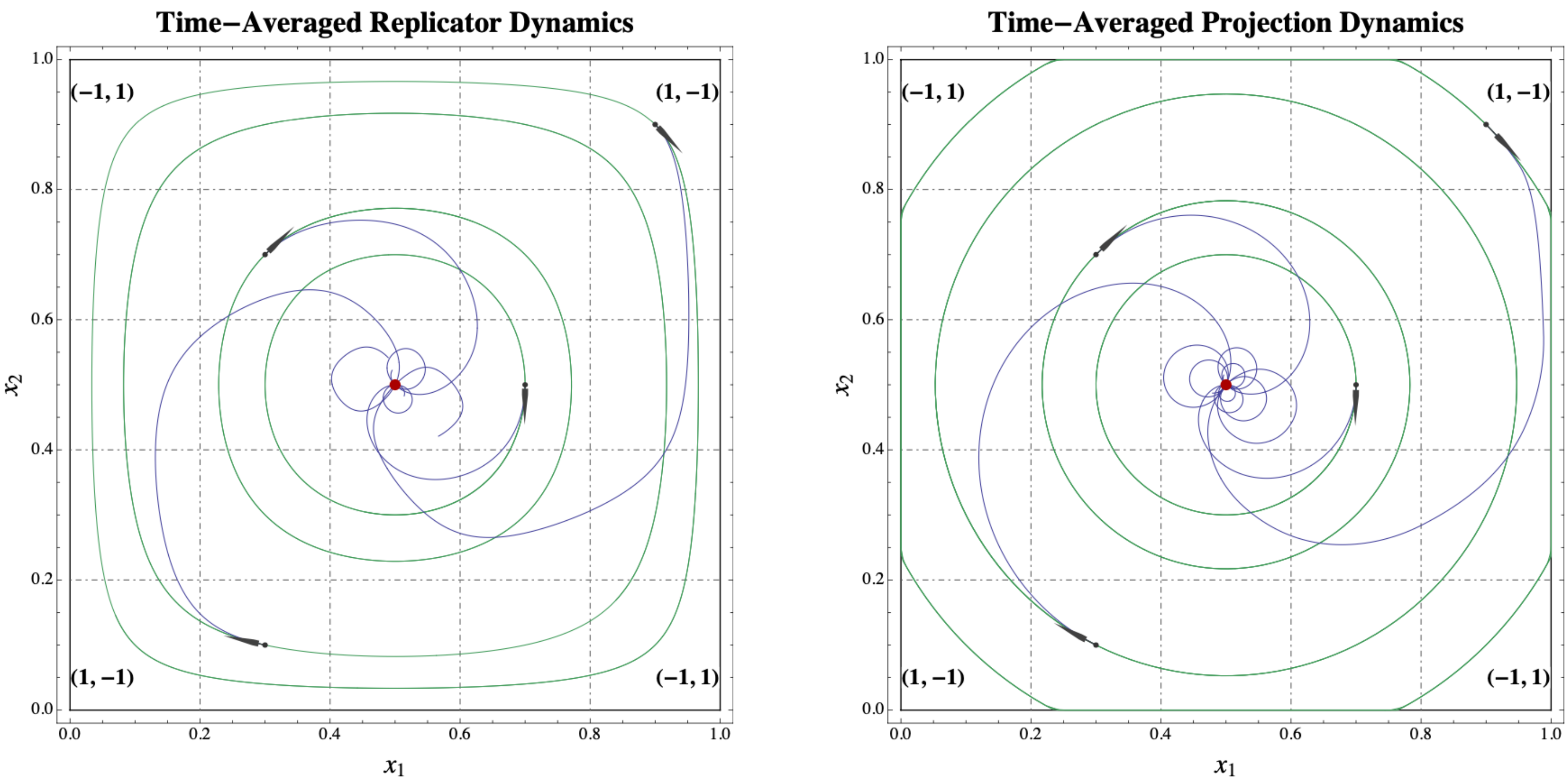
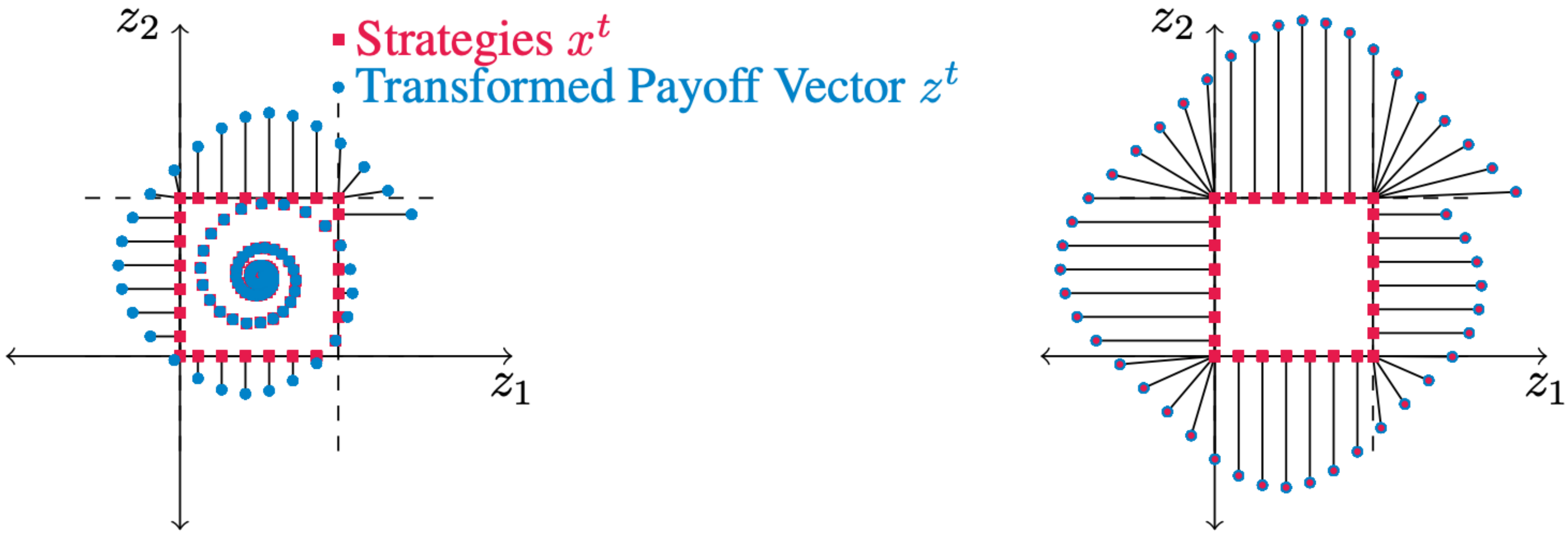


Figure: Behavior of time-averages under replicator (left) and projection dynamics (right)

Backup Slides



(a) Iterations 1-95

(b) Iterations 95-140