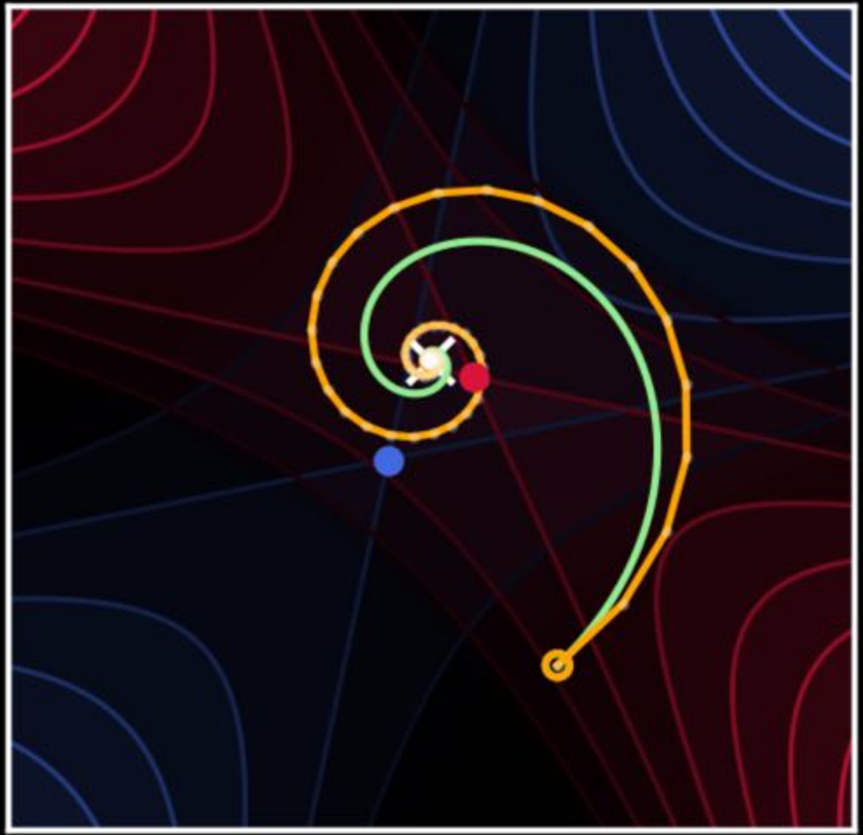
Learning Dynamics of Non-cooperative Agents in Dynamic Environments



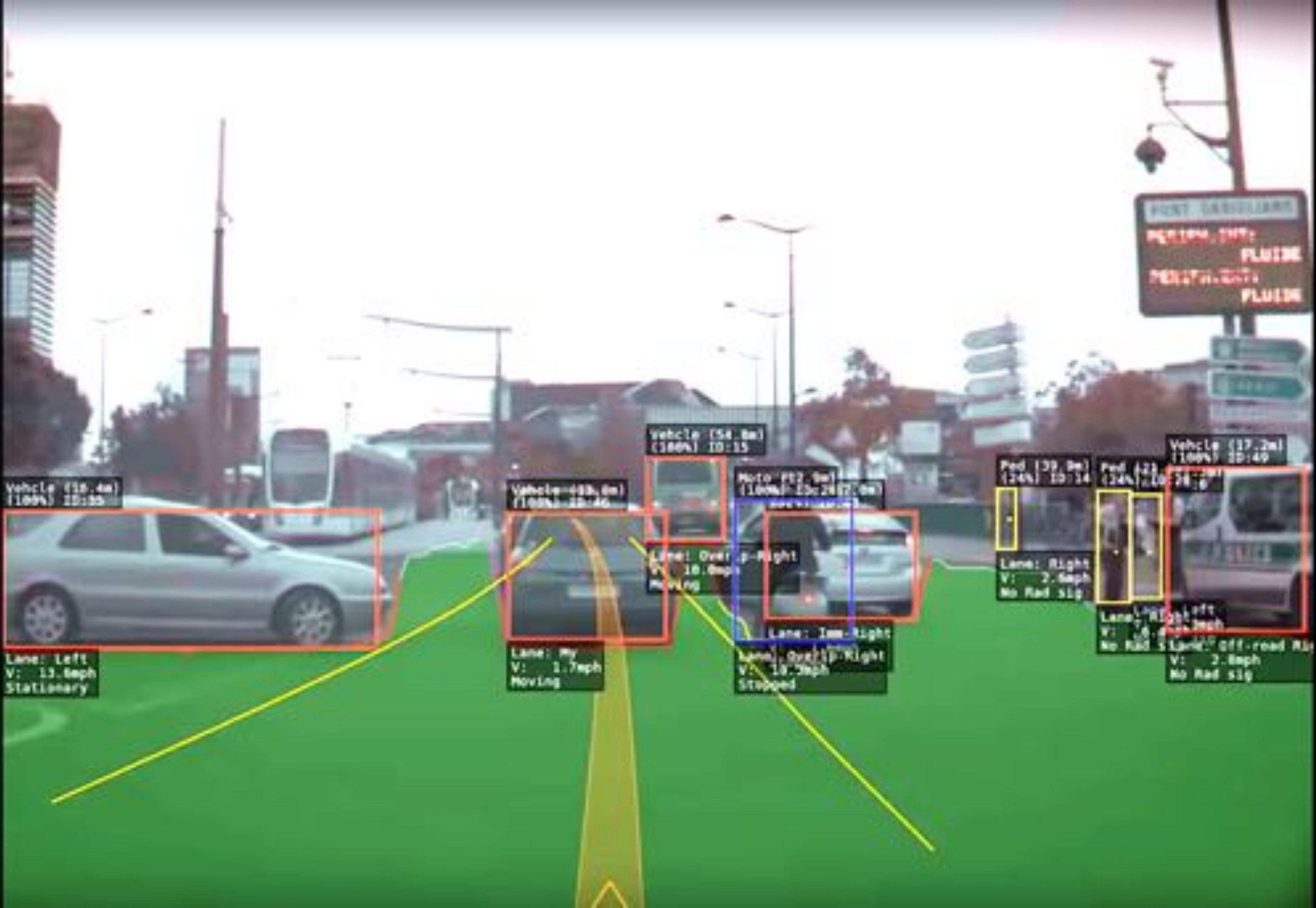
Benjamin J. Chasnov

Electrical and Computer Engineering
University of Washington, Seattle WA

Qualifying Exam, May 2019

Advisors: *Dr. Samuel Burden, Dr. Lillian Ratliff*

Committee: *Dr. Maryam Fazel (chair), Dr. Behçet Açikmeşe, Dr. Kevin Jamieson*


Lane: Left
V: 13.6mph
Stationary
Lane: My
V: 1.7mph
Moving
Vehcle (54.8m)
Lane: Imm-Right
Stopped
Lane: Right
V: 2.6mph
No Rad sig
Vehcle (17.2m)
(100%) ID:49
Lane: Off-road Rig
V: 2.8mph
No Rad sig
FLUIDE
FLUIDE

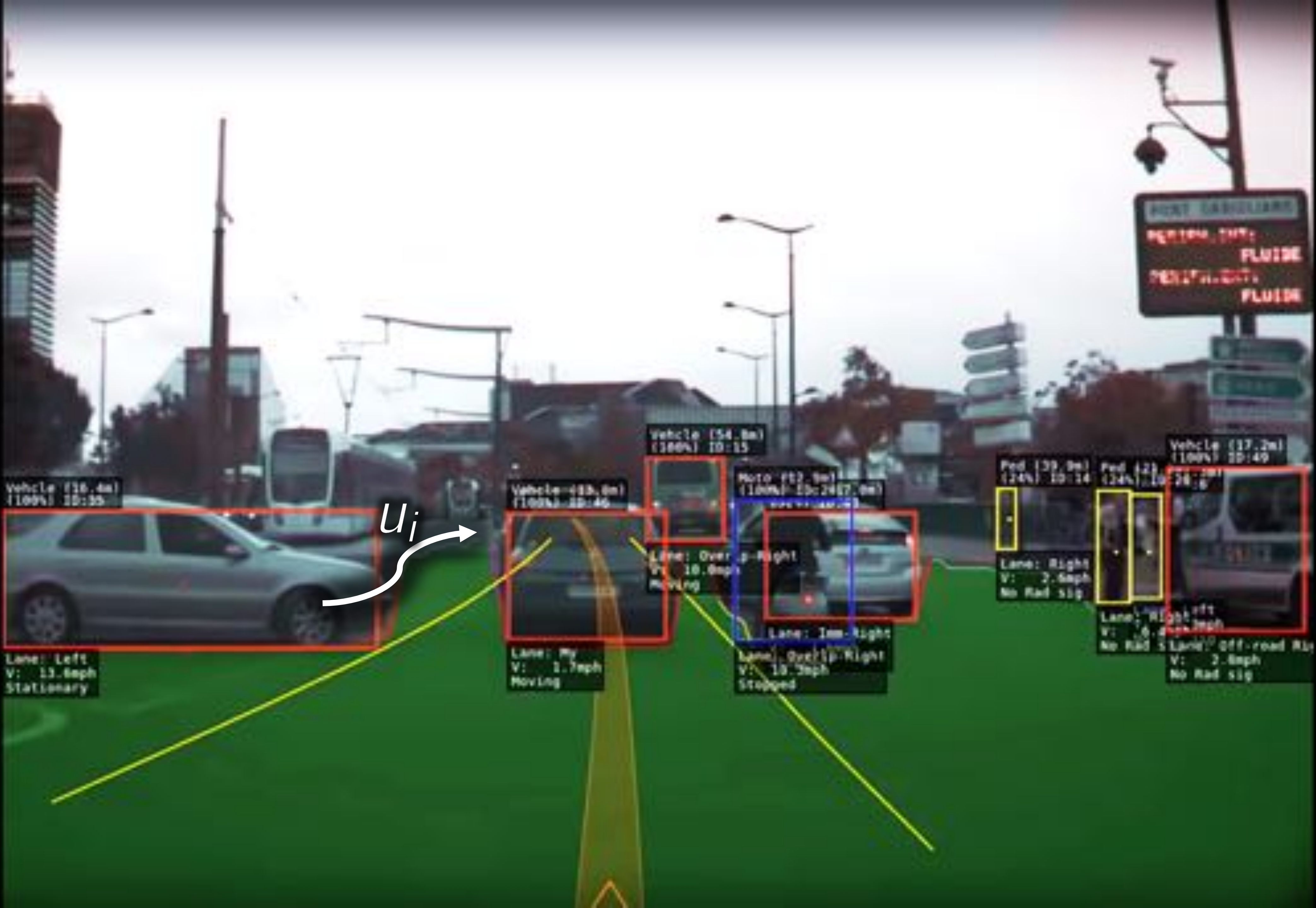
n agents

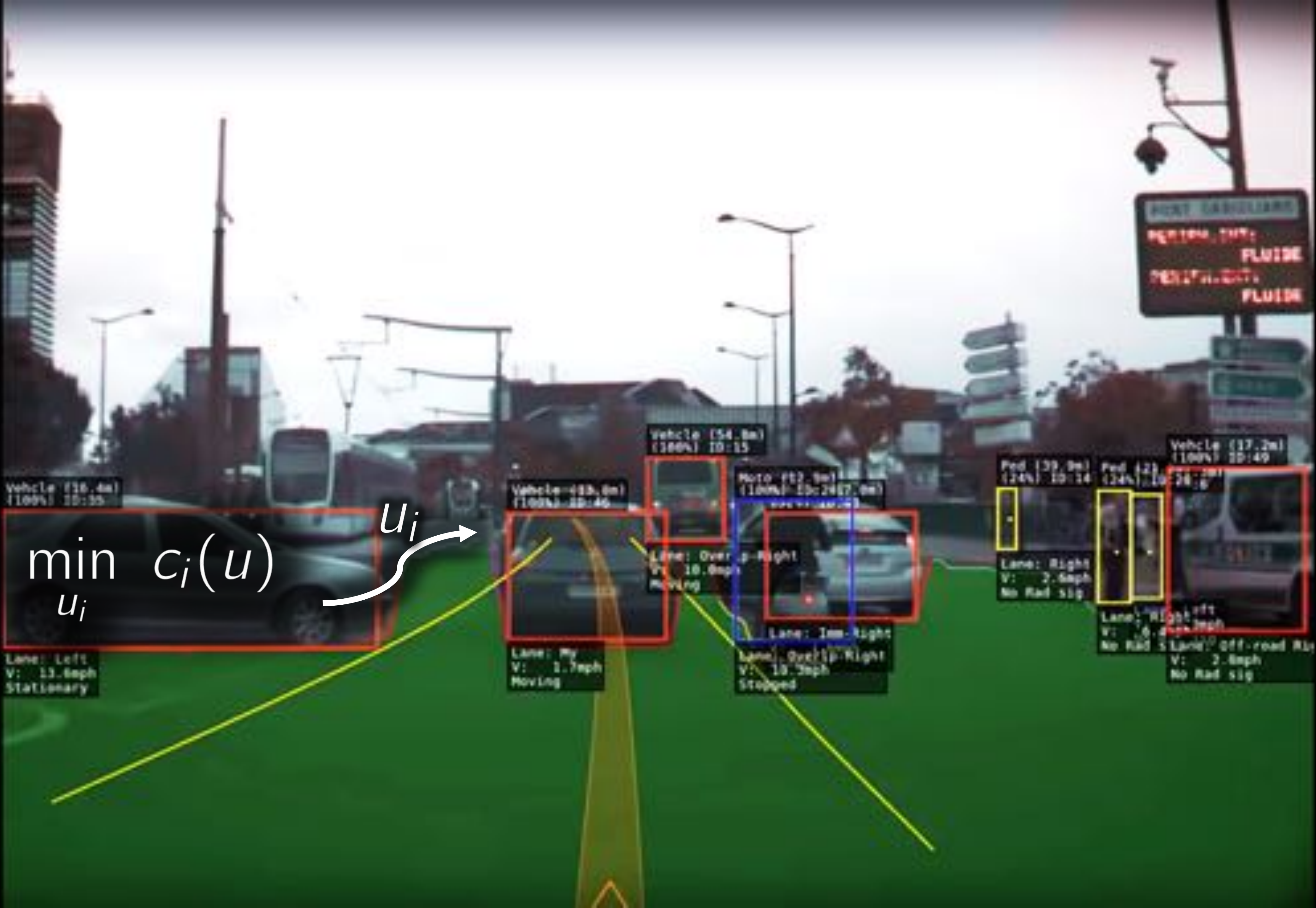

n agents
$\underset{u_i}{\text{minimize}}\ \text{cost}_i(u)$

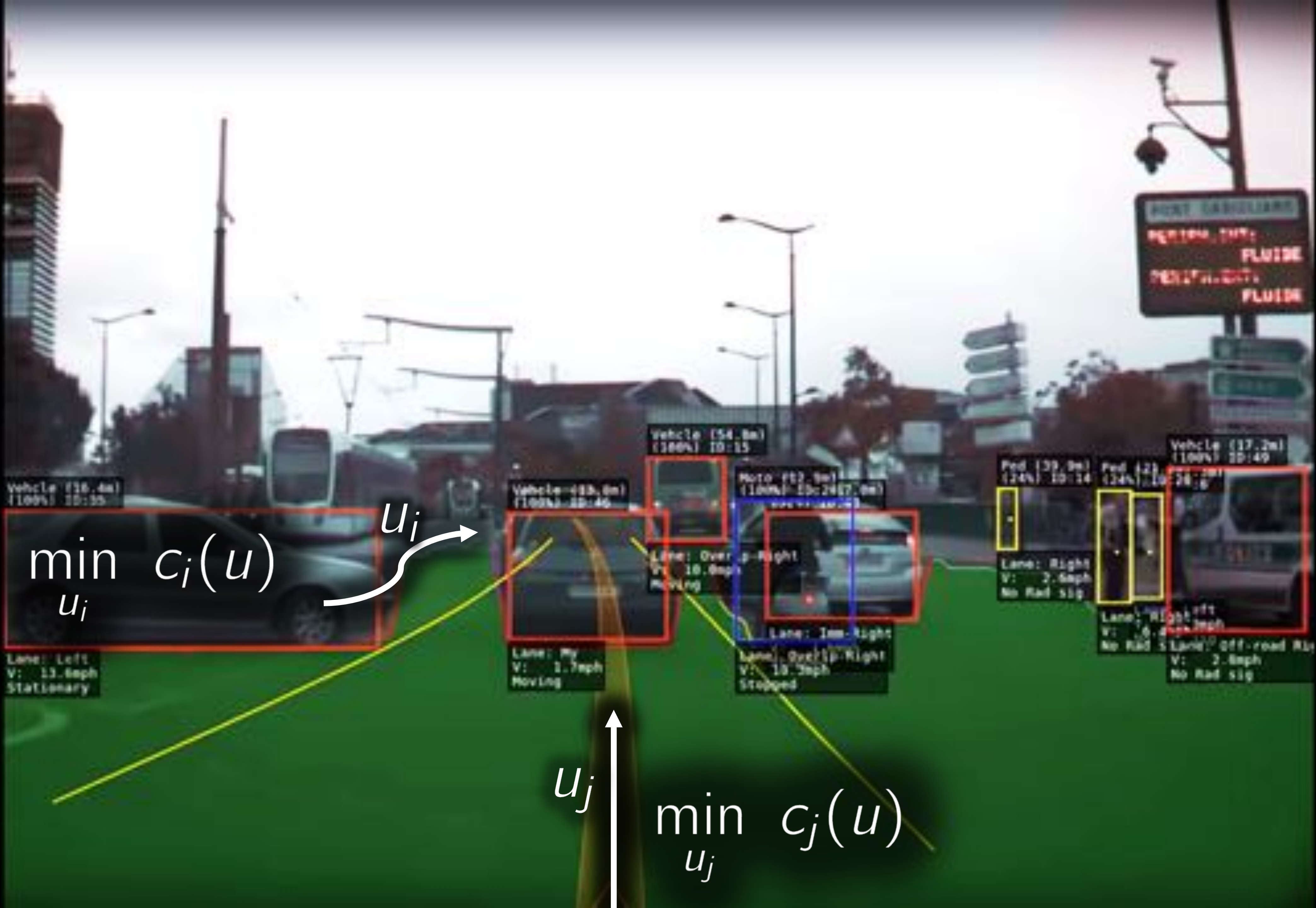
n agents

$\underset{u_i}{\text{minimize}}\ \text{cost}_i(u)$

$u = (u_1, \ldots, u_n)$

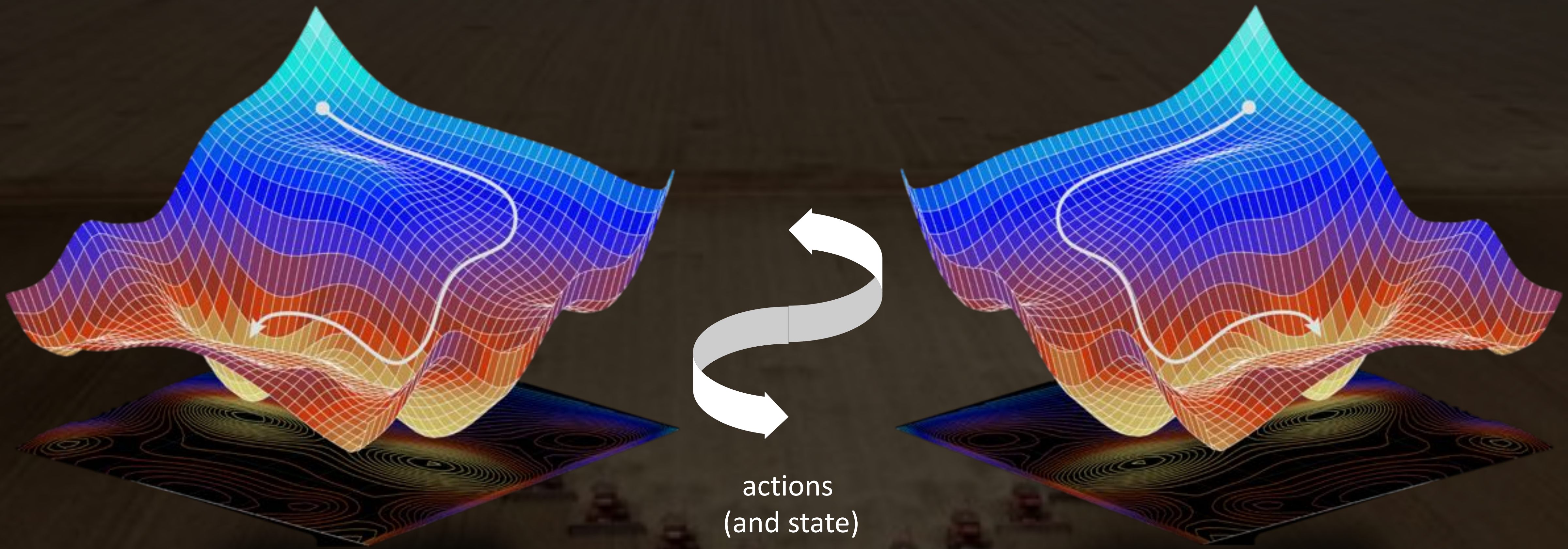

u_i
Lane: Left
V: 13.6mph
Stationary
Lane: My
V: 1.7mph
Moving
No Rad sig
FLUIDE
FLUIDE


$\min_{u_i} \; c_i(u)$
u_i
FLUIDE
FLUIDE

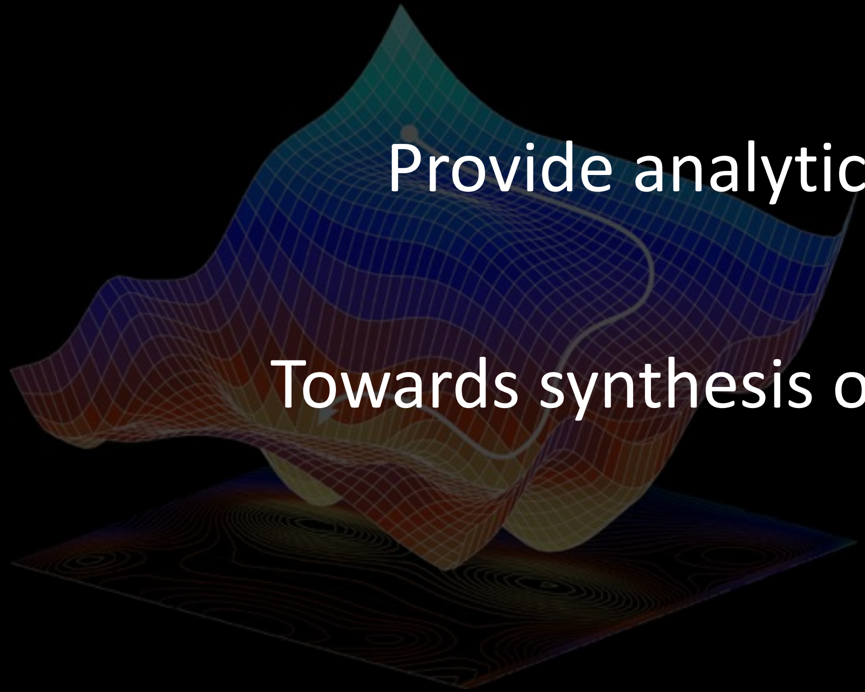


u_i
$\min_{u_i} \; c_i(u)$
u_j
$\min_{u_j} \; c_j(u)$
FLUIDE
FLUIDE


$\underset{u=(u_1,\ldots,u_n)}{\text{minimize}} \ \text{cost}(u)$
$\underset{u_i}{\text{minimize}} \ \text{cost}_i(u)$

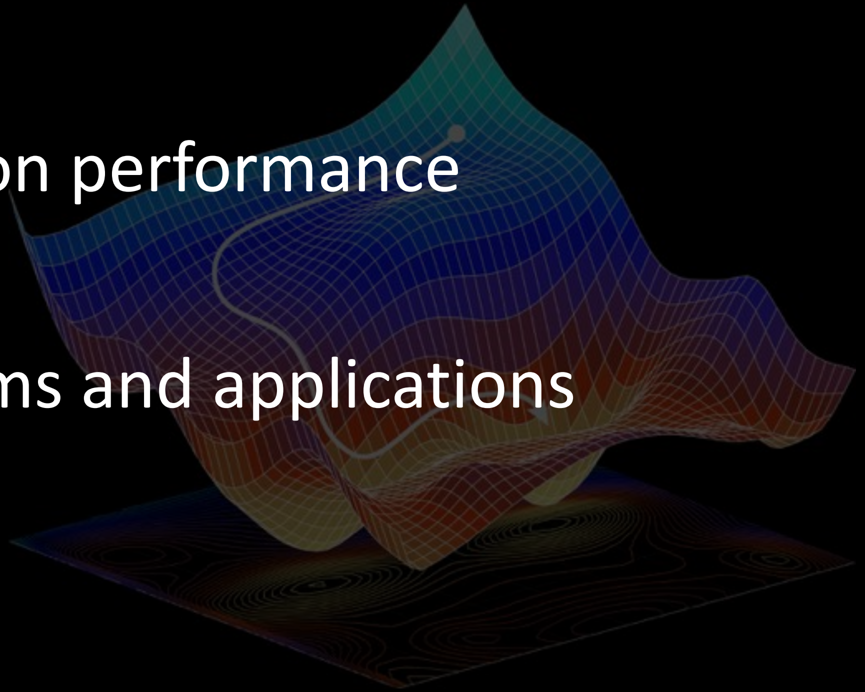


$$\min_{u_i} \; c_i(u) \qquad \qquad \min_{u_j} \; c_j(u)$$



Provide analytical guarantees on performance

Towards synthesis of new algorithms and applications



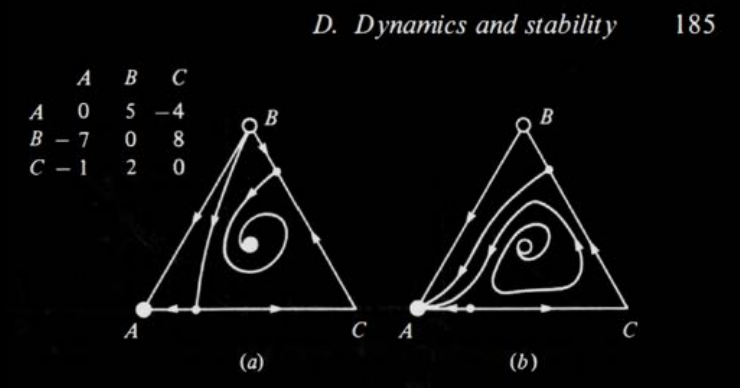
Overview

- **Background and related literature:**
  - Non-cooperative games
  - Learning agents
- **Our contributions:**
  - *Learning dynamics in games*
    - Theorem on non-asymptotic convergence rate for gradient learning
    - Obstacles to convergence (non-Nash attractors, unstable Nash, limit cycles)
  - *Application to dynamic games*
    - LQ games (open loop, feedback, stochastic)
    - Towards realistic real world applications
  - *Publications and presentations*

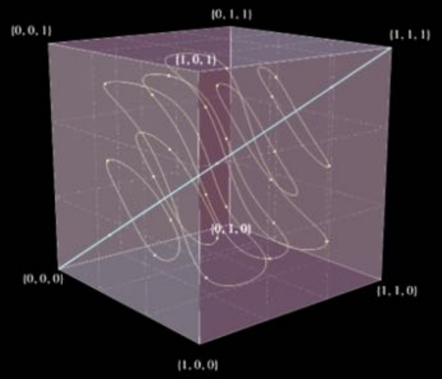
$t = s = 2$

	1	2	row minima
1	1	−1	−1
2	−1	1	−1
column maxima	1	1	

“Theory of Games and Economic behavior”
Von Neumann & Morgenstern
1943



Evolutionary games and stable strategies
Smith, Hofbauer, Sigmund …
1980s



Dynamical systems perspective of learning in games
Benaım, Hirsch, …
Papadimitriou, Mertikopoulos …
2000s – now

1928
Von Neumann

1951
Nash
“Non-cooperative games”

1990s
Fudenberg & Levine
“The Theory of Learning in Games”

Now
Ratliff, Burden, Mazumdar, Chasnov …
Convergence analysis of gradient-based methods

Existence of Equilibrium Points

A proof of this existence theorem based on Kakutani's generalized fixed point theorem was published in Proc. Nat. Acad. Sci. U. S. A., 36, pp. 48–49. The proof given here is a considerable improvement over that earlier version and is based directly on the Brouwer theorem. We proceed by constructing a continuous transformation T of the space of n-tuples such that the fixed points of T are the equilibrium points of the game.

THEOREM 1. *Every finite game has an equilibrium point.*

PROOF. Let $\mathfrak{s}$ be an n-tuple of mixed strategies, $p_i(\mathfrak{s})$ the corresponding pay-off to player i, and $p_{i\alpha}(\mathfrak{s})$ the pay-off to player i if he changes to his α^{th} pure strategy $\pi_{i\alpha}$ and the others continue to use their respective mixed strategies from $\mathfrak{s}$. We now define a set of continuous functions of $\mathfrak{s}$ by

$$\varphi_{i\alpha}(\mathfrak{s}) = \max(0, p_{i\alpha}(\mathfrak{s}) - p_i(\mathfrak{s}))$$

and for each component s_i of $\mathfrak{s}$ we define a modification s_i' by

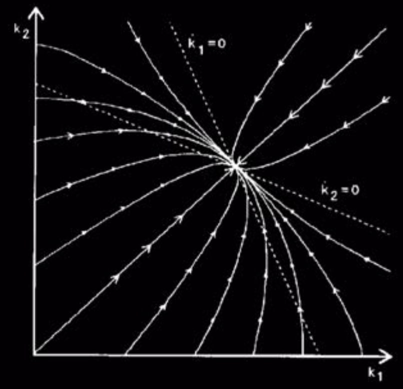


Figure 13.5

Definition: continuous game (2 players)

A 2-player continuous game consists of a joint action/strategy/choice-variable

$$u = (u_1, u_2) \in U_1 \times U_2 = U$$

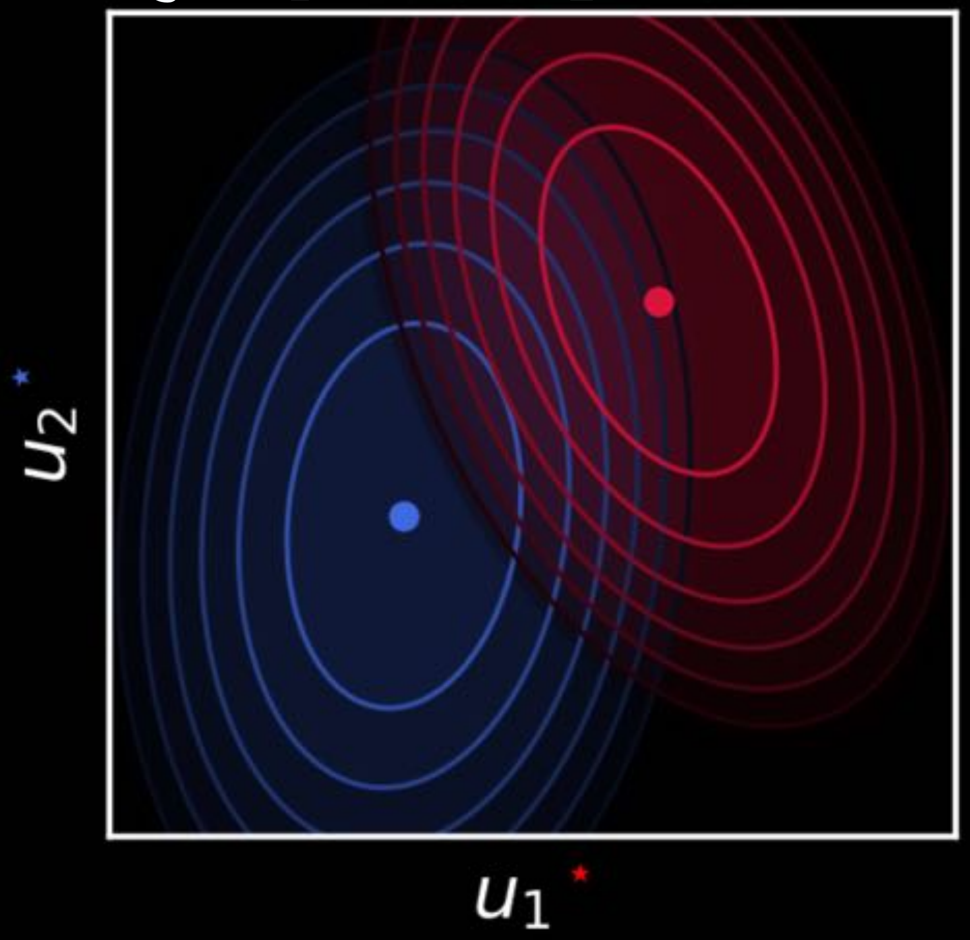
with agent 1's cost function

$$c_1(u) : U \to \mathbb{R}$$

and agent 2's cost function

$$c_2(u) : U \to \mathbb{R}$$

e.g. $U_1 = \mathbb{R},\ U_2 = \mathbb{R}$



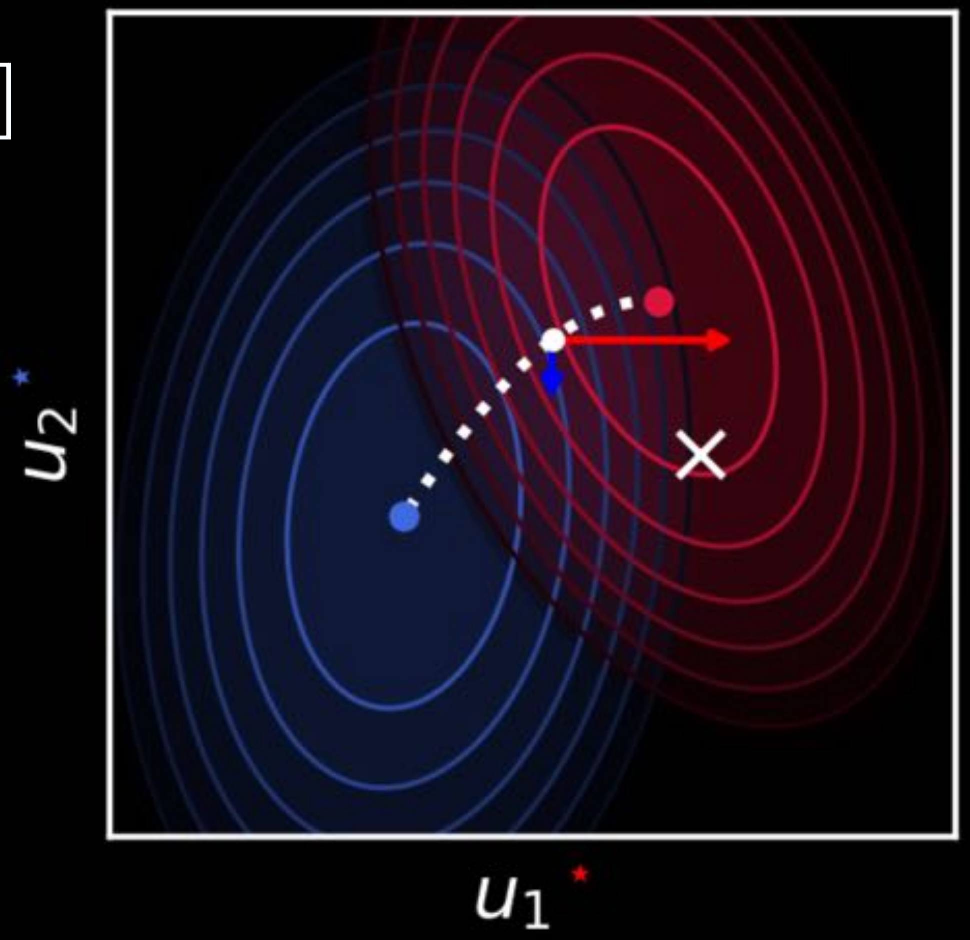
Two different perspectives

Cooperative

$$\min_{u} \; \theta c_1(u) + (1-\theta)c_2(u), \quad \theta \in [0,1]$$

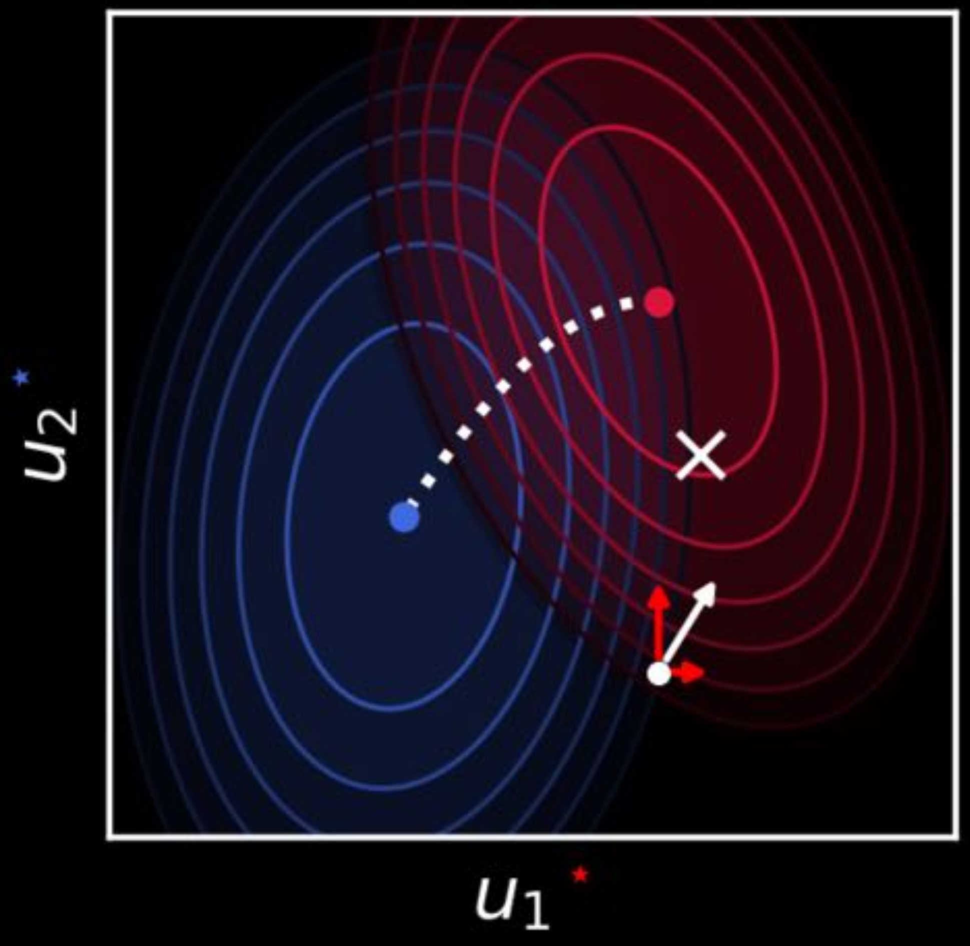
Non-cooperative

$$\min_{u_1} \; c_1(u) \text{ and } \min_{u_2} \; c_2(u)$$



Gradient dynamics

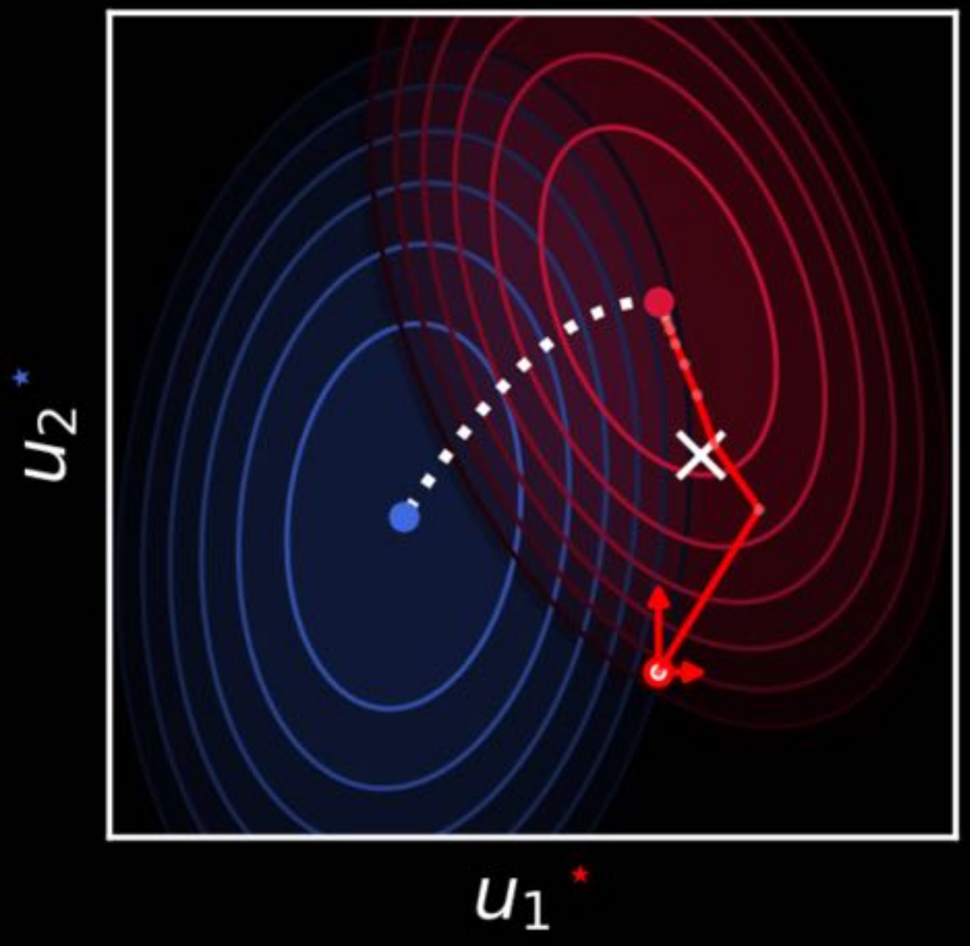
$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_1(u) \end{bmatrix}$$



$$D_j c_i(u) \equiv \frac{\partial c_i(u)}{\partial u_j} \in \mathbb{R}^{d_j}$$

Gradient dynamics

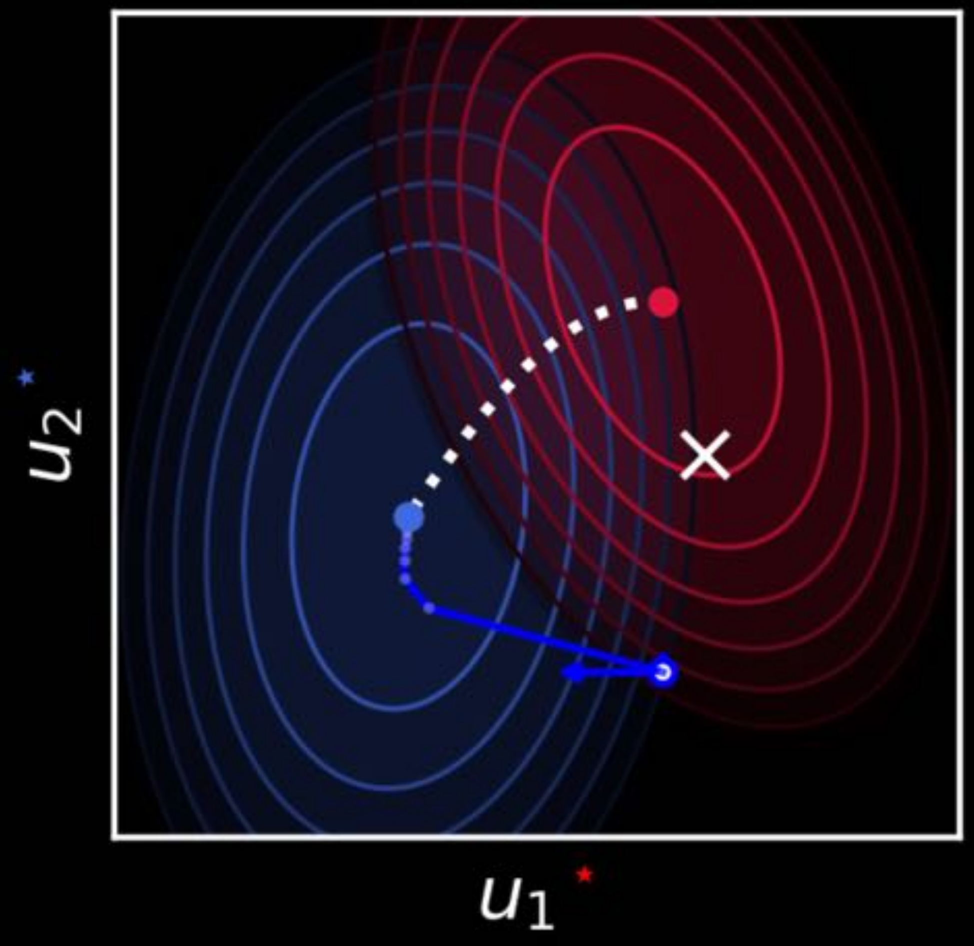
$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_1(u) \end{bmatrix}$$



Gradient dynamics

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_1(u) \end{bmatrix}$$

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_2(u) \\ D_2 c_2(u) \end{bmatrix}$$

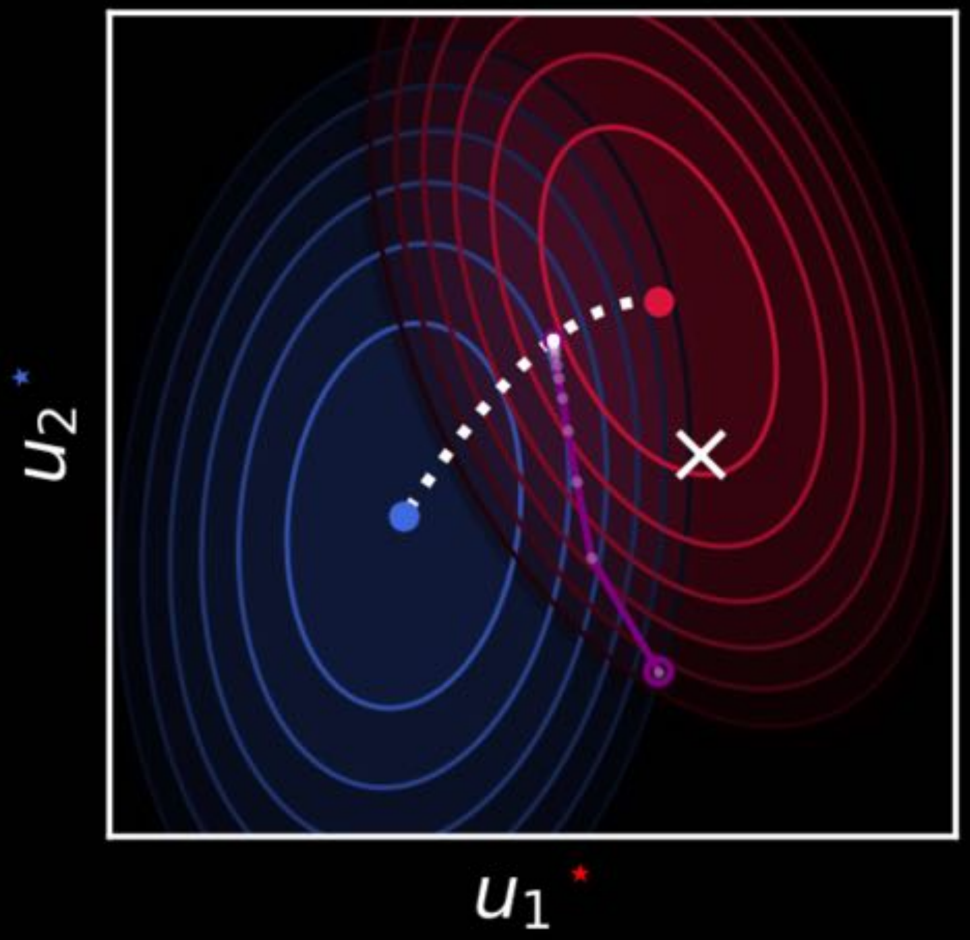


Cooperative dynamics

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_1(u) \end{bmatrix}$$

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_2(u) \\ D_2 c_2(u) \end{bmatrix}$$

$$u^+ = u - \gamma \theta D c_1(u) + (1-\theta) D c_2(u)$$



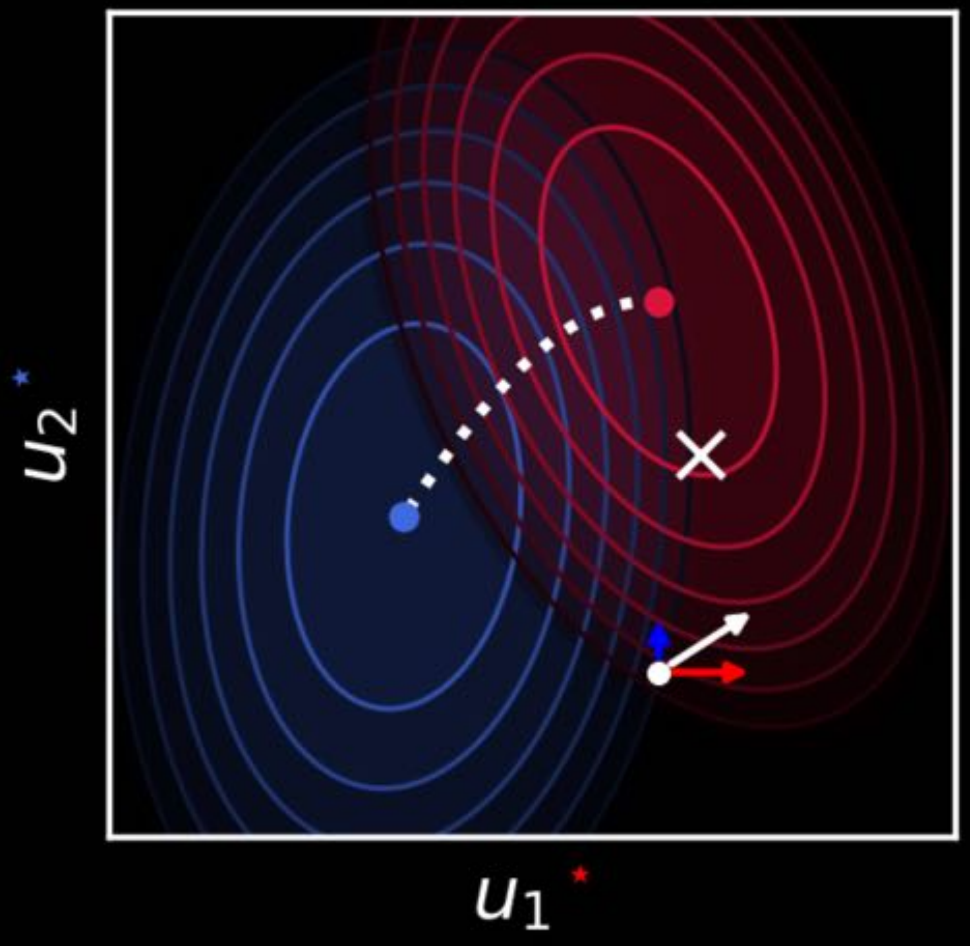
Game vector field

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_1(u) \end{bmatrix}$$

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_2(u) \\ D_2 c_2(u) \end{bmatrix}$$

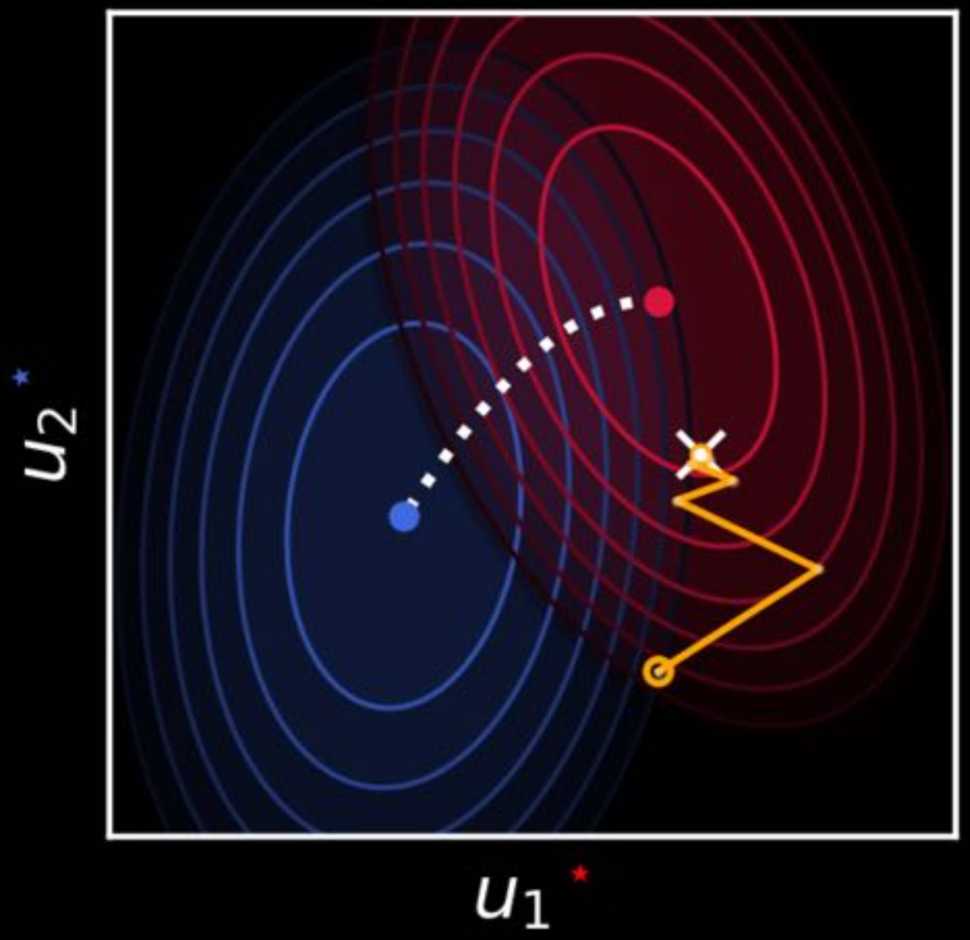
$$u^+ = u - \gamma \theta D c_1(u) + (1-\theta) D c_2(u)$$

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_2(u) \end{bmatrix}$$



Non-cooperative perspective

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_2(u) \end{bmatrix}$$



Definition: differential Nash equilibrium

A strategy is a *Nash equilibrium* if no agent can do better by unilaterally changing its strategy.

A *differential Nash equilibrium* satisfies first order conditions

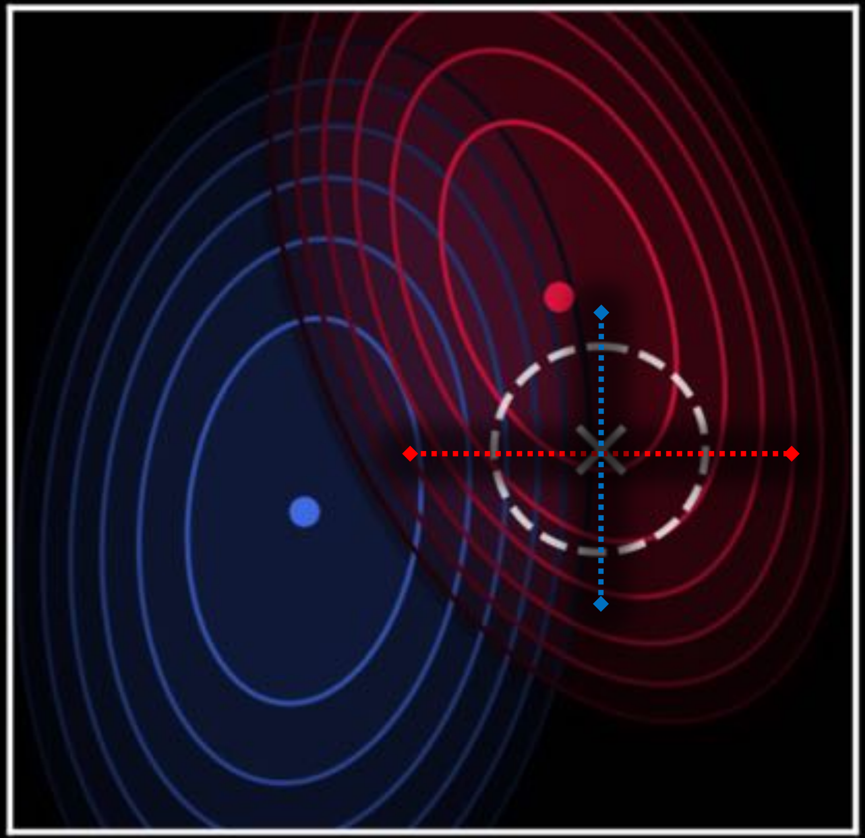
$$D_1 c_1(u^*) = 0, \;\; D_2 c_2(u^*) = 0$$

and second order conditions

$$D_{11} c_1(u^*) > 0, \;\; D_{22} c_2(u^*) > 0$$

for twice continuously-differentiable cost functions.

u_2

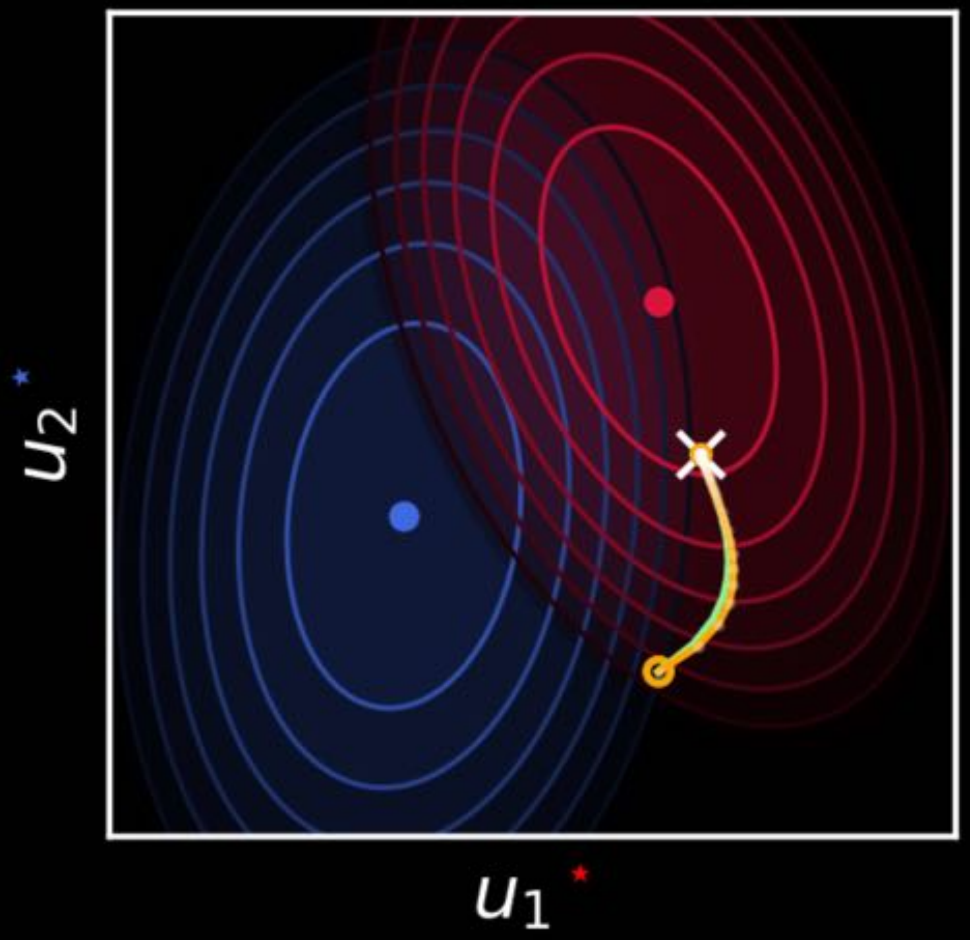


u_1

Ratliff, Lillian J., Samuel A. Burden, and S. Shankar Sastry. "Genericity and structural stability of non-degenerate differential Nash equilibria." *2014 American Control Conference*. IEEE, 2014.

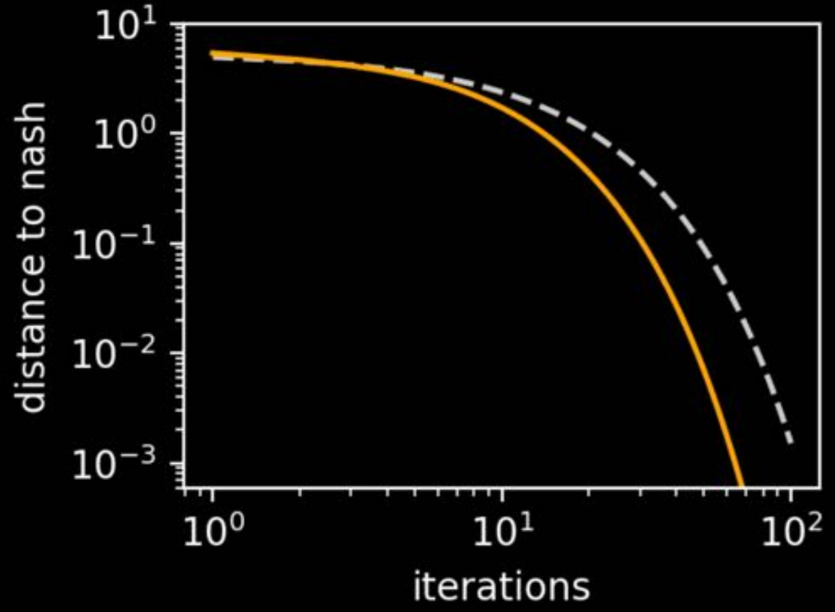
Part I: Learning dynamics in games

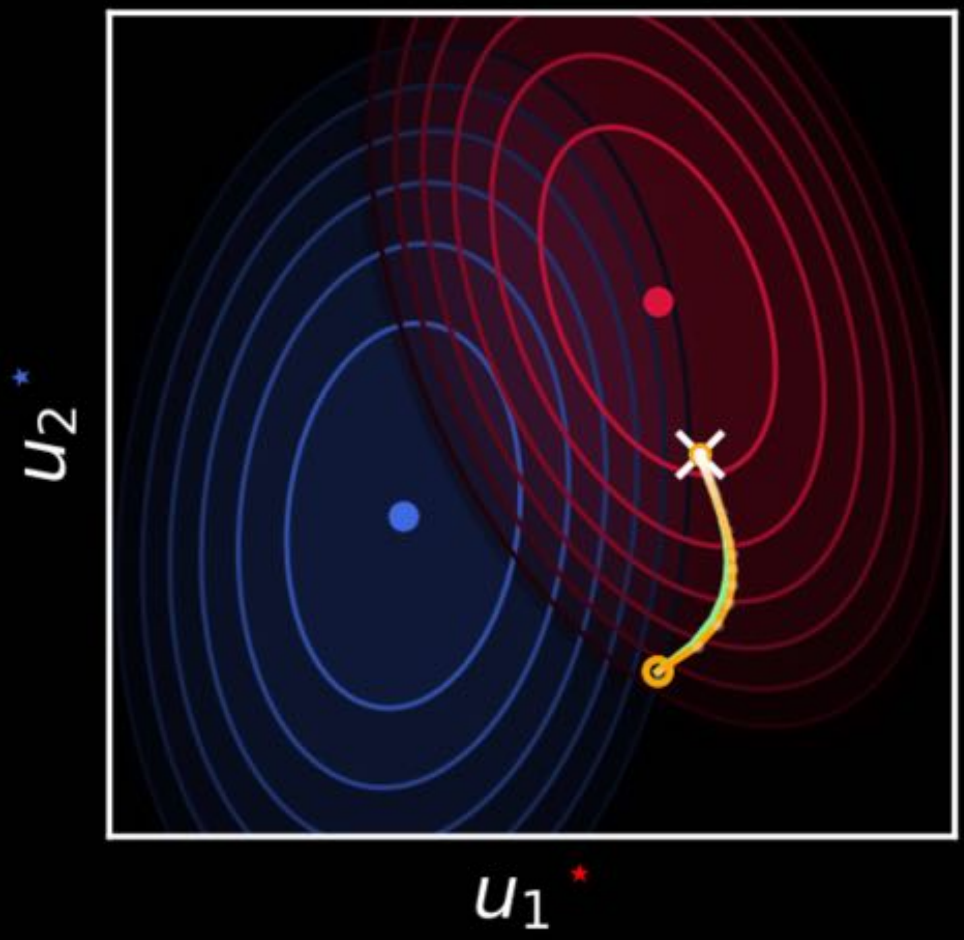
$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_2(u) \end{bmatrix}$$



Goal: Non-asymptotic convergence guarantees

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_2(u) \end{bmatrix}$$





Linearization of learning dynamics

$$u^+ = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_2(u) \end{bmatrix}$$

$$\approx [I - \gamma J(u)](u - u^*)$$

Fixed points of vector field *ω(u)*

$$D_1 c_1(u^*) = 0,\ D_2 c_2(u^*) = 0$$

Jacobian of vector field *ω(u)*

$$J = D\omega = \begin{bmatrix} D_{11} c_1 & D_{12} c_1 \\ D_{21} c_2 & D_{22} c_2 \end{bmatrix}$$

Non-asymptotic convergence guarantee

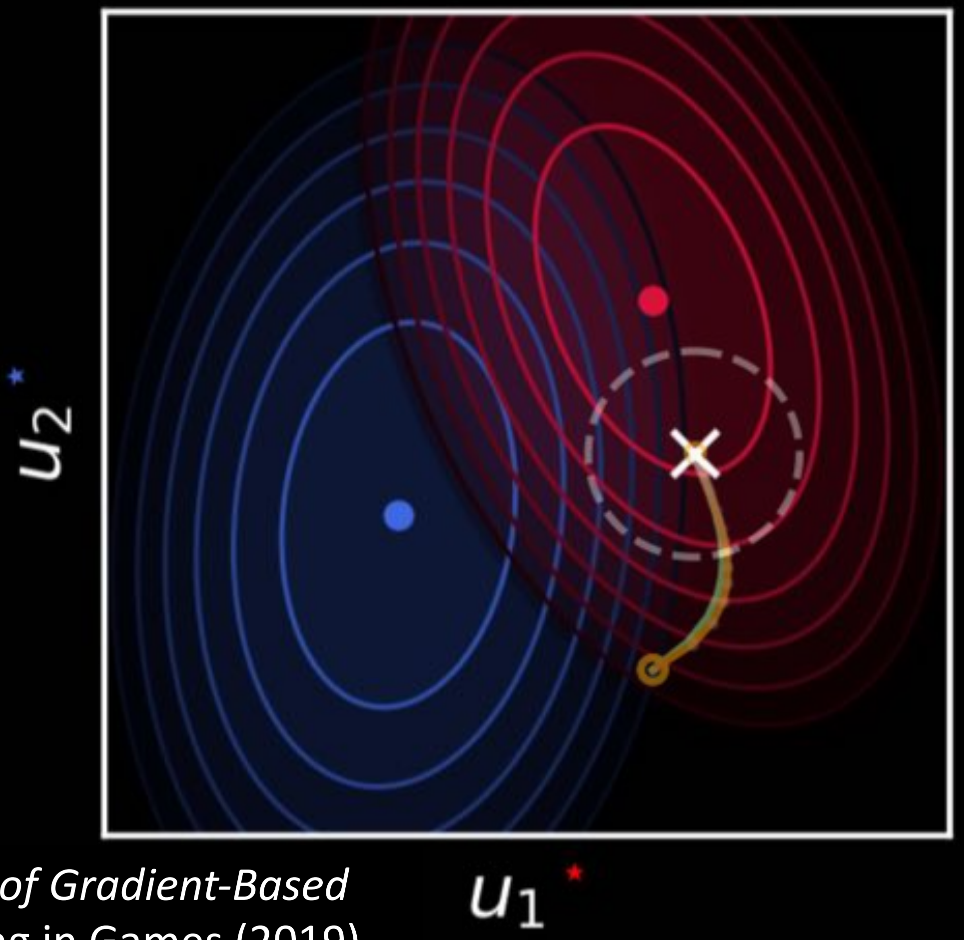
Theorem: If agents have learning rates $\gamma = \alpha/\beta^2$ where singular values α, β are

$$\alpha = \min_{u \in B_r(u^*)} \sigma_{\min}(J(u) + J(u)^T)/2$$

$$\beta = \max_{u \in B_r(u^*)} \sigma_{\max} J(u)$$

and $u^{(0)}$ is initialized in a region of attraction of a local Nash equilibrium, then the iterates $u^{(k)}$ will be bounded by

$$\|u^{(k)} - u^*\| \leq \exp\left(-\sqrt{\frac{\alpha}{2\beta}}k\right)\|u^{(0)} - u^*\|$$



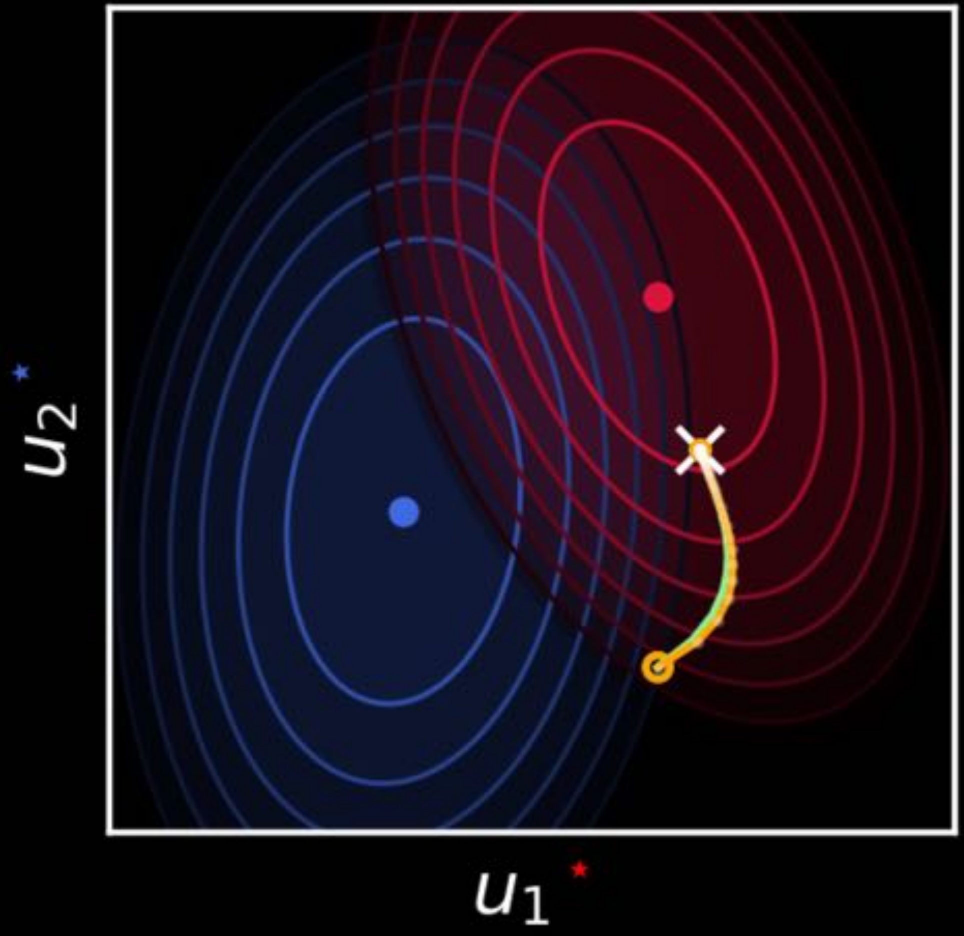
Chasnov, Ratliff, Calderone, Mazumdar, Burden, *"Finite-Time Convergence of Gradient-Based Learning in Continuous Games."* AAAI Workshop on Reinforcement Learning in Games (2019).

Dynamical systems perspective

$$u^{+} = u - \gamma \begin{bmatrix} D_1 c_1(u) \\ D_2 c_2(u) \end{bmatrix}$$

$\updownarrow$ (with appropriate γ)

$$\dot{u} = -\omega(u)$$



Mazumdar, Eric, and Lillian J. Ratliff. *"On the convergence of gradient-based learning in continuous games." arXiv:1804.05464* (2018).

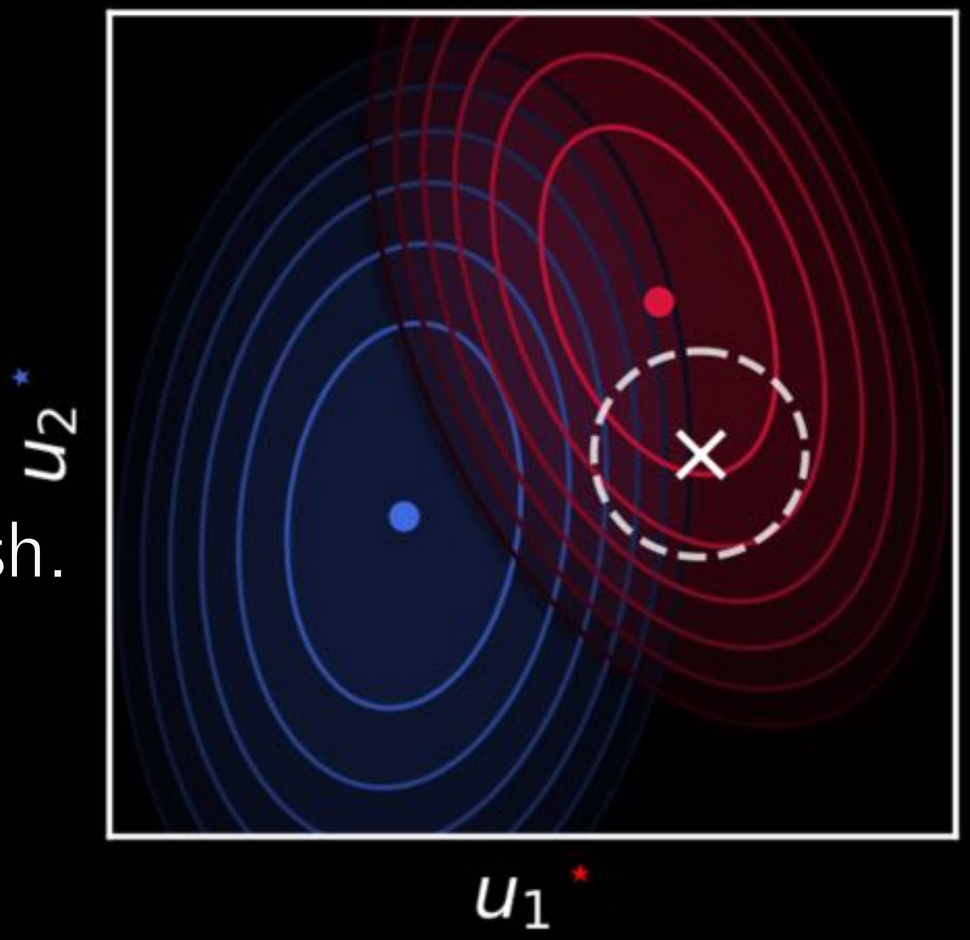
Spectrum of the Jacobian

$$\begin{aligned}\dot{u} &= -\omega(u)\\ &= -\underbrace{J(u)}_{}\,u\end{aligned}$$

If $\text{spec}(J) \subset \mathbb{C}_+^\circ$ at u^*, then u^* is stable.

If $\text{blockdiag}_i(J) > 0$ at u^* $\forall i$, then u^* is Nash.

$$J = D\omega = \begin{bmatrix} D_{11}c_1 & D_{12}c_1 \\ D_{21}c_2 & D_{22}c_2 \end{bmatrix}$$



Issue 1: not all stable equilibria are Nash

$$\mathrm{spec}(J) \subset \mathbb{C}_+^\circ$$

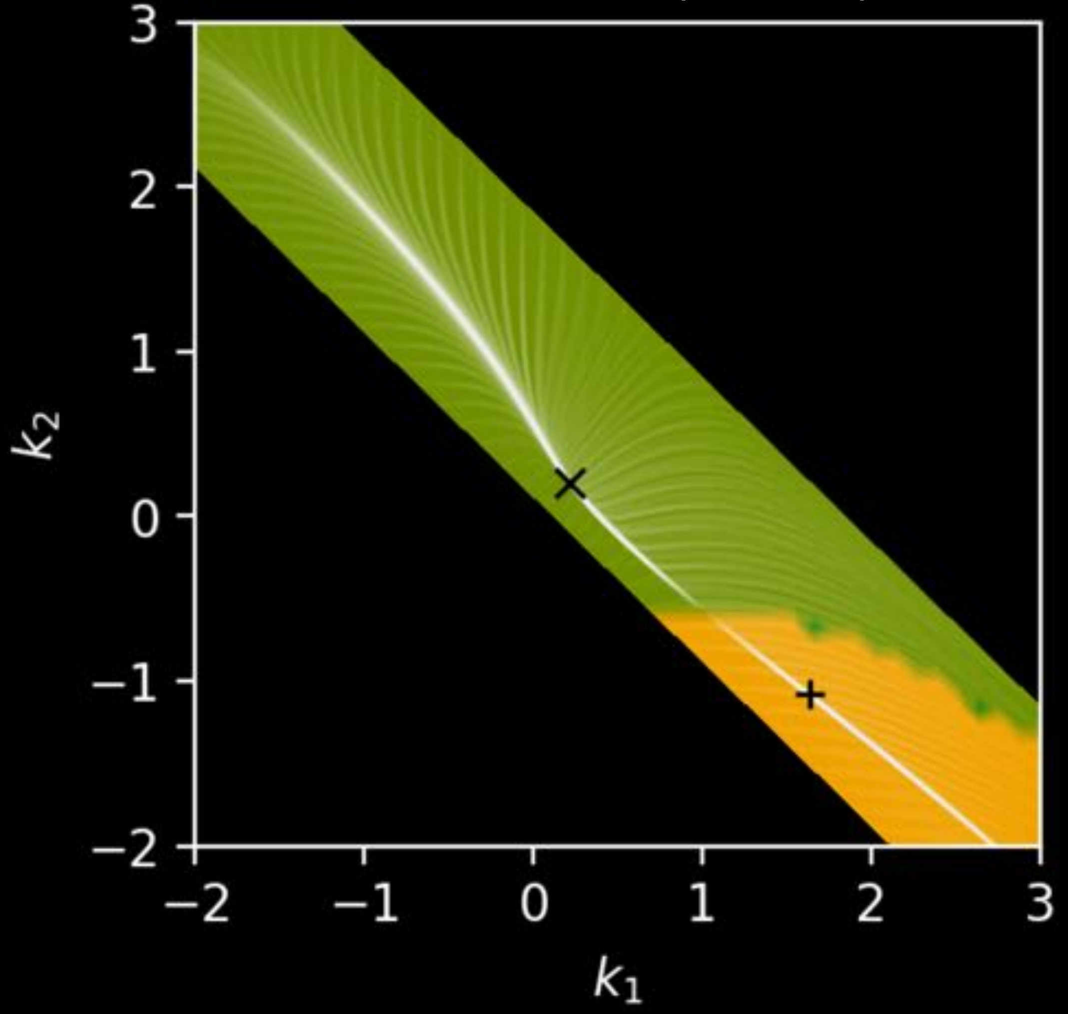
$$J(u^*) = \begin{bmatrix} + & \\ & + \end{bmatrix}$$

Nash

$$J(u^*) = \begin{bmatrix} + & \\ & - \end{bmatrix}$$

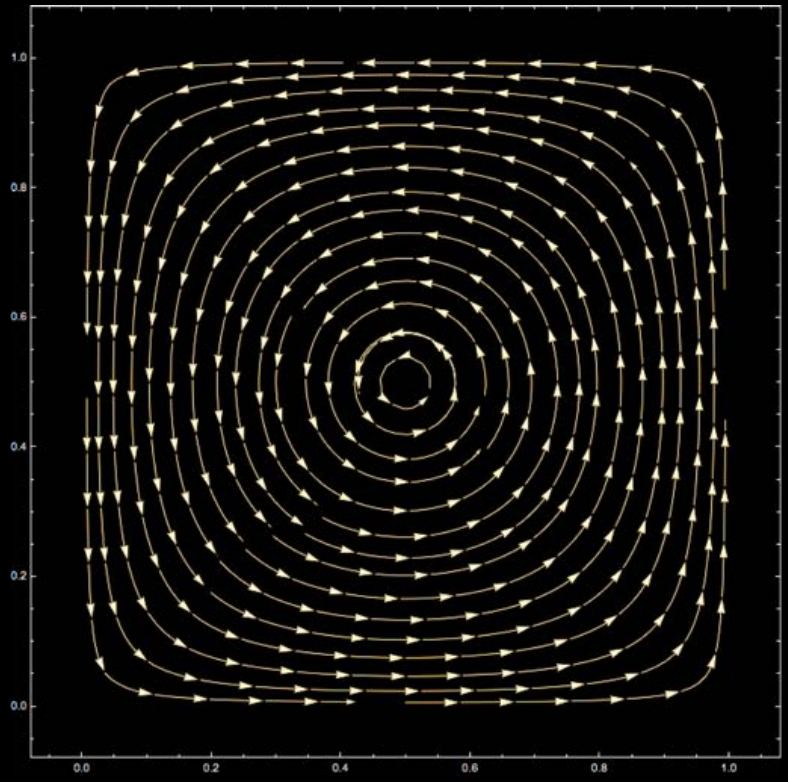
Non-Nash

e.g. $U_1 \times U_2 = (k_1, k_2)$

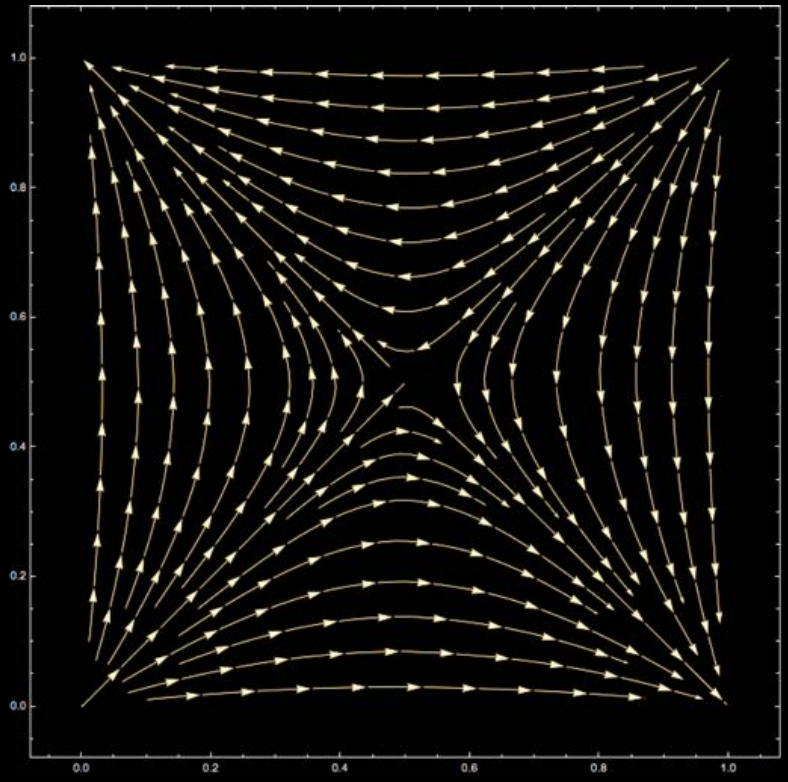


Mazumdar, Eric V., Michael I. Jordan, and S. Shankar Sastry. "On finding local nash equilibria (and only local nash equilibria) in zero-sum games." *arXiv preprint arXiv:1901.00838* (2019).

Issue 2: not all Nash equilibria are attractors



Zero-sum game



Partnership game

Hofbauer and Sigmund. *Evolutionary games and population dynamics*. Cambridge university press, 1998.

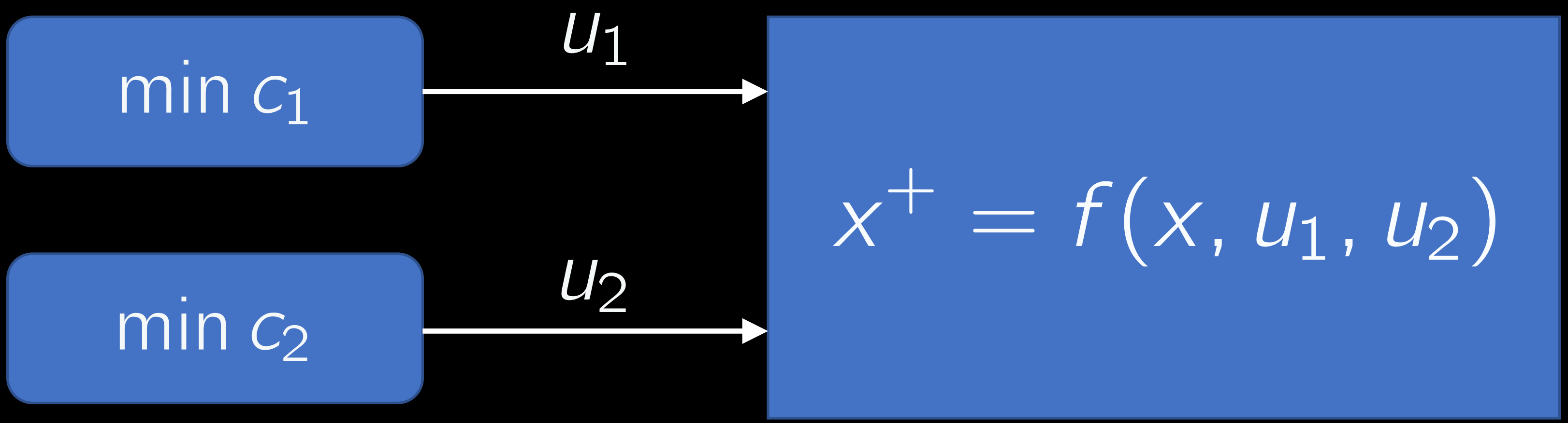
Mertikopoulos, Papadimitriou and Piliouras. "Cycles in adversarial regularized learning." In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2018.

Part II: Towards application in dynamic games

$$x^+ = f(x, u_1, u_2)$$

$$\min_{u_1} \; c_1(x, u), \; \min_{u_2} \; c_2(x, u)$$

Open loop dynamic games

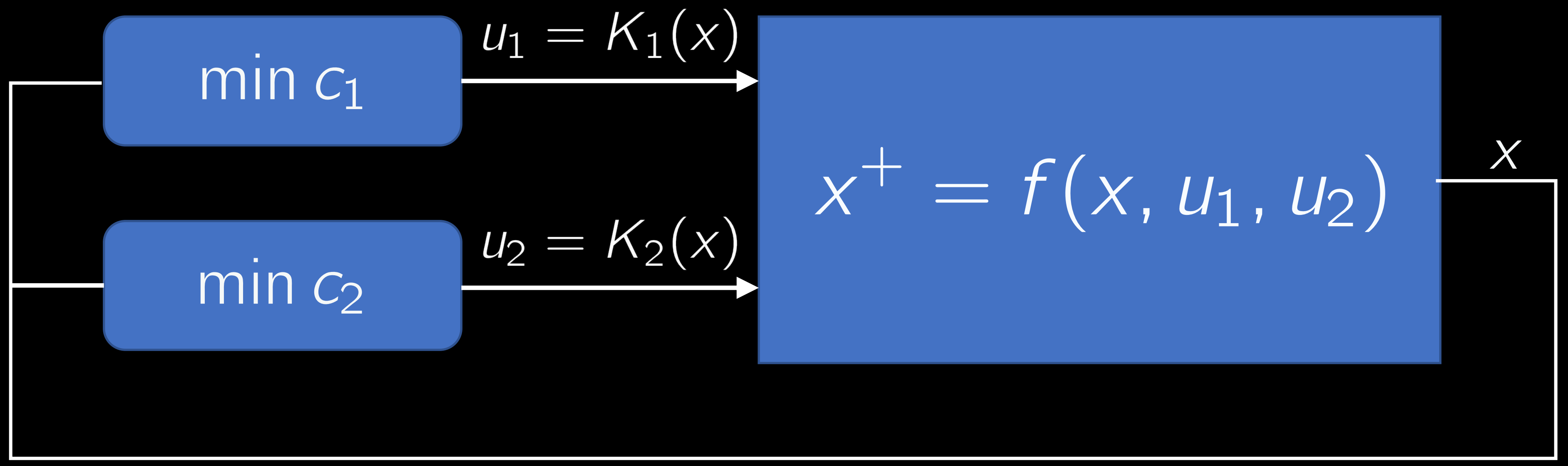


$$\frac{\partial}{\partial u_1} c_1(x_0, u)$$

$$\frac{\partial}{\partial u_2} c_2(x_0, u)$$

Basar, Tamer, and Geert Jan Olsder. “*Dynamic noncooperative game theory*.” Vol. 23. Siam, 1999.

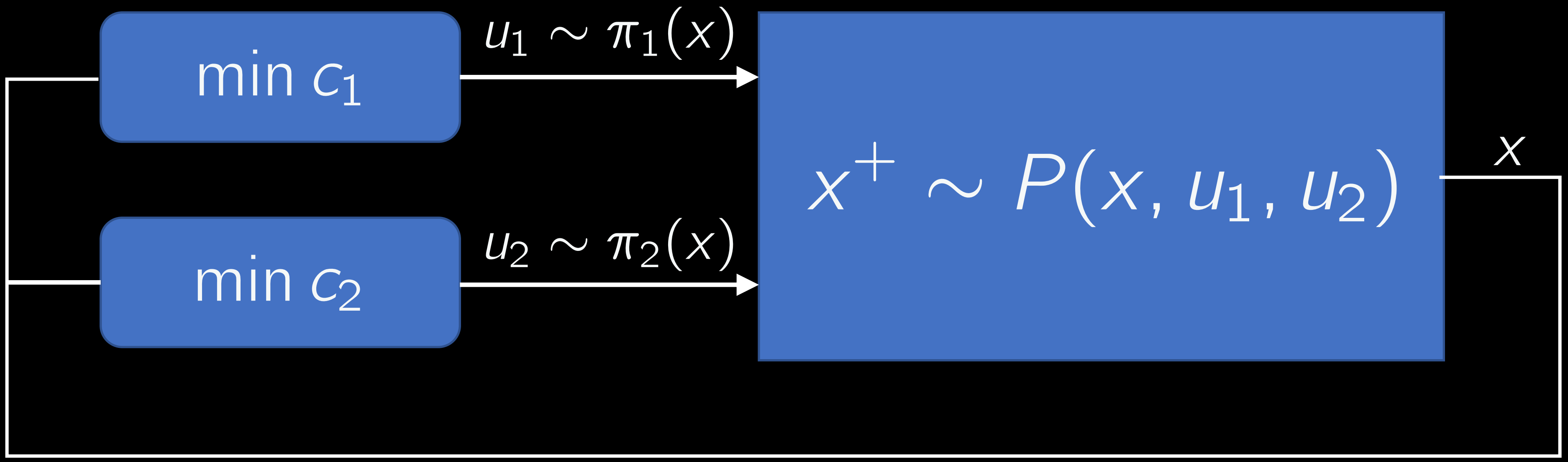
Closed loop dynamic games



$$\frac{\partial}{\partial K_i} c_i(x, K)$$

Basar, Tamer, and Geert Jan Olsder. "*Dynamic noncooperative game theory*." Vol. 23. Siam, 1999.

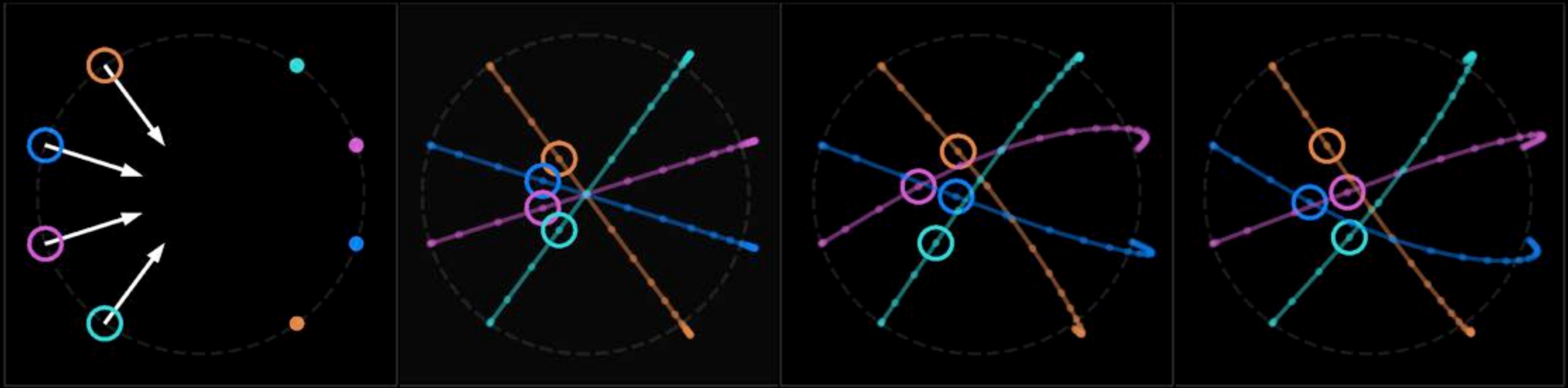
Stochastic games



$$\widehat{\frac{\partial}{\partial \theta_i} c_i(\theta)}$$

Shapley, Lloyd S. "Stochastic games." *Proceedings of the national academy of sciences* (1953).

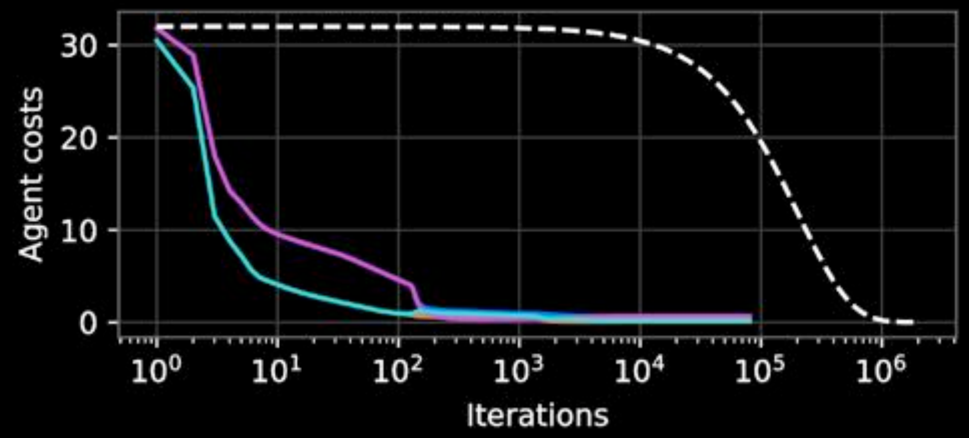
Open loop dynamic game



Initialization

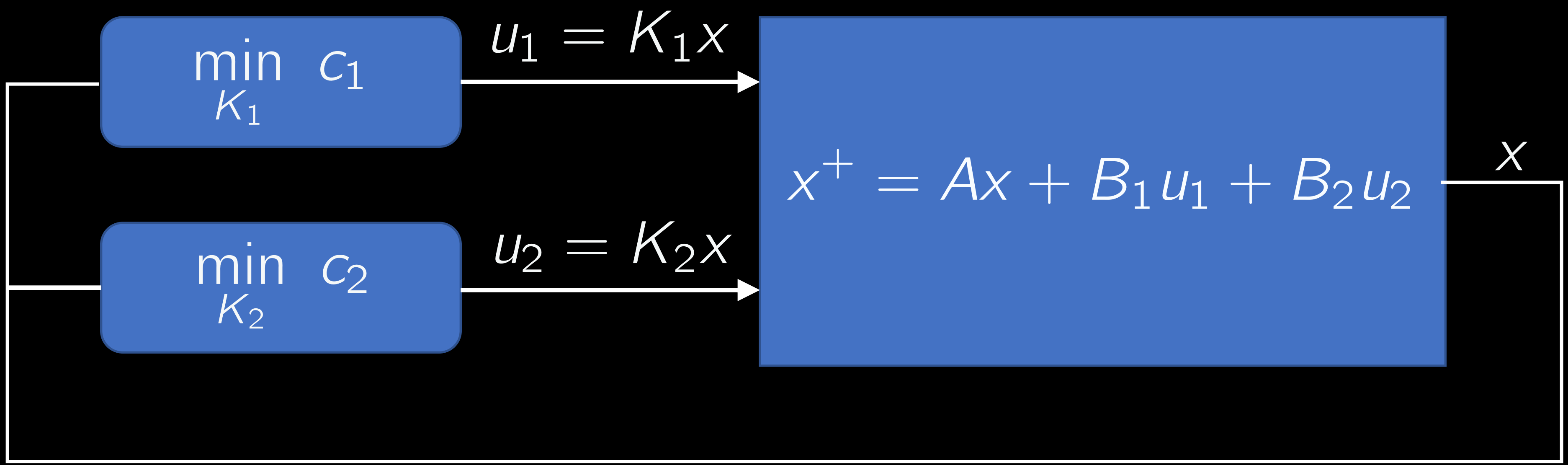
Nash equilibrium

Nash equilibrium



Chasnov, Ratliff, Calderone, Mazumdar, Burden, *"Finite-Time Convergence of Gradient-Based Learning in Continuous Games."* AAAI Workshop on Reinforcement Learning in Games (2019).

Linear Quadratic games

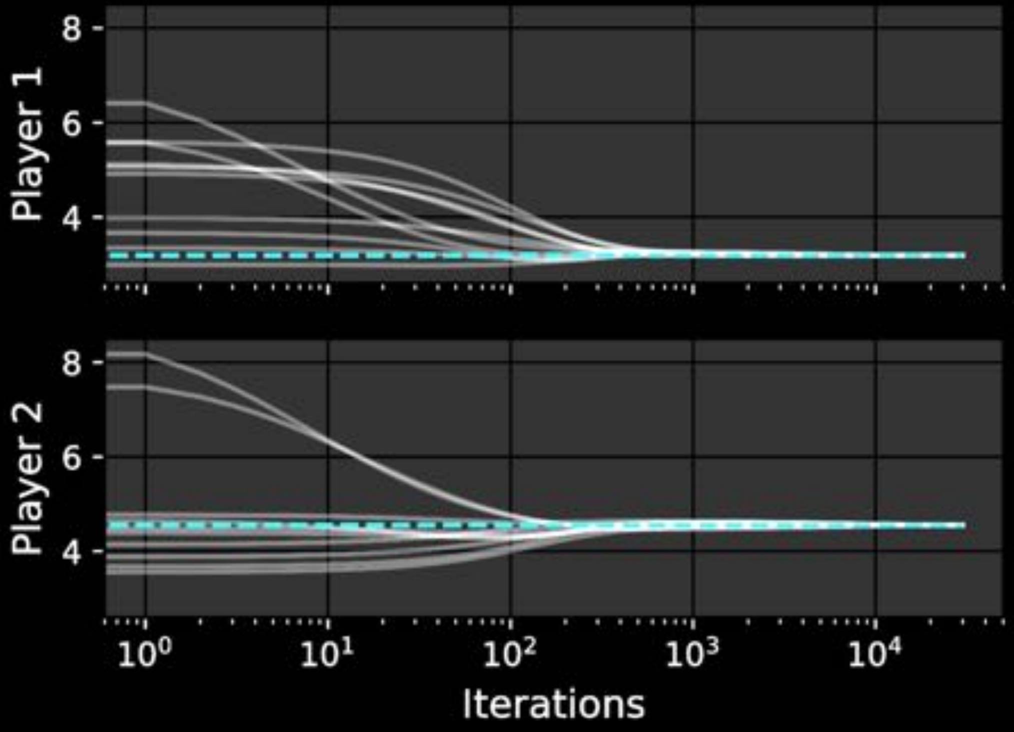


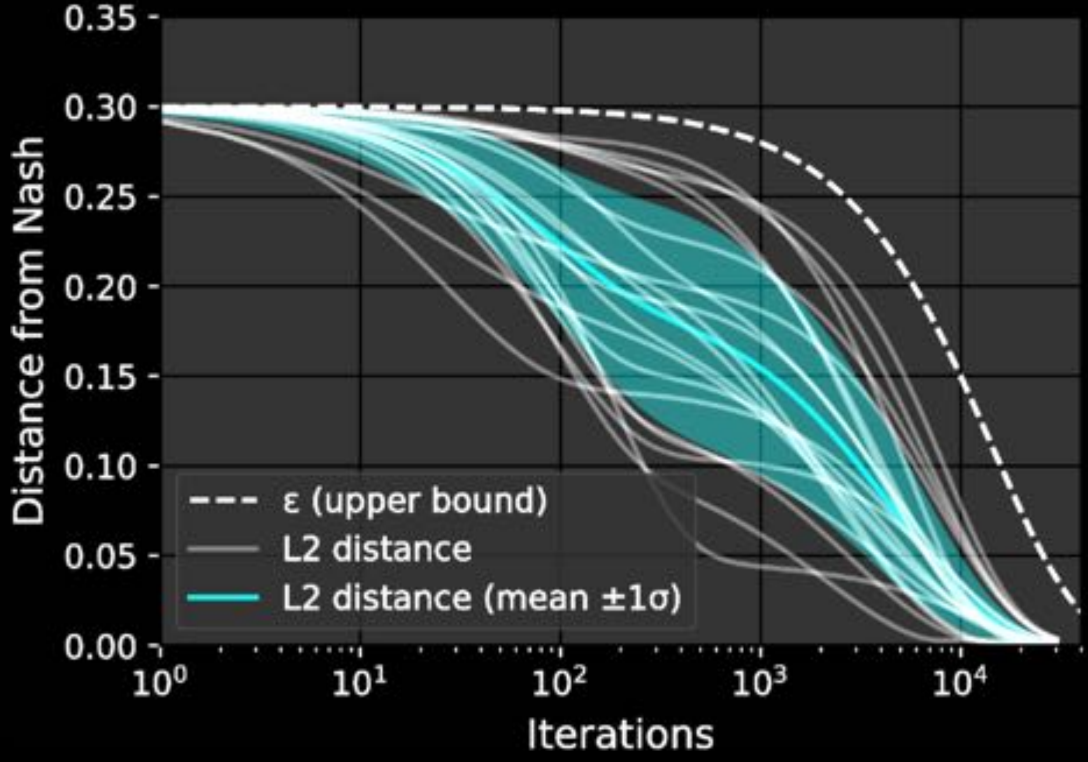
$$c_1(x_0, K_1, K_2) = \sum_{t=0}^{\infty} x^T Q_1 x + u_1^T R_{11} u_1 + u_2^T R_{12} u_2$$

$$c_2(x_0, K_1, K_2) = \sum_{t=0}^{\infty} x^T Q_2 x + u_1^T R_{21} u_1 + u_2^T R_{22} u_2$$

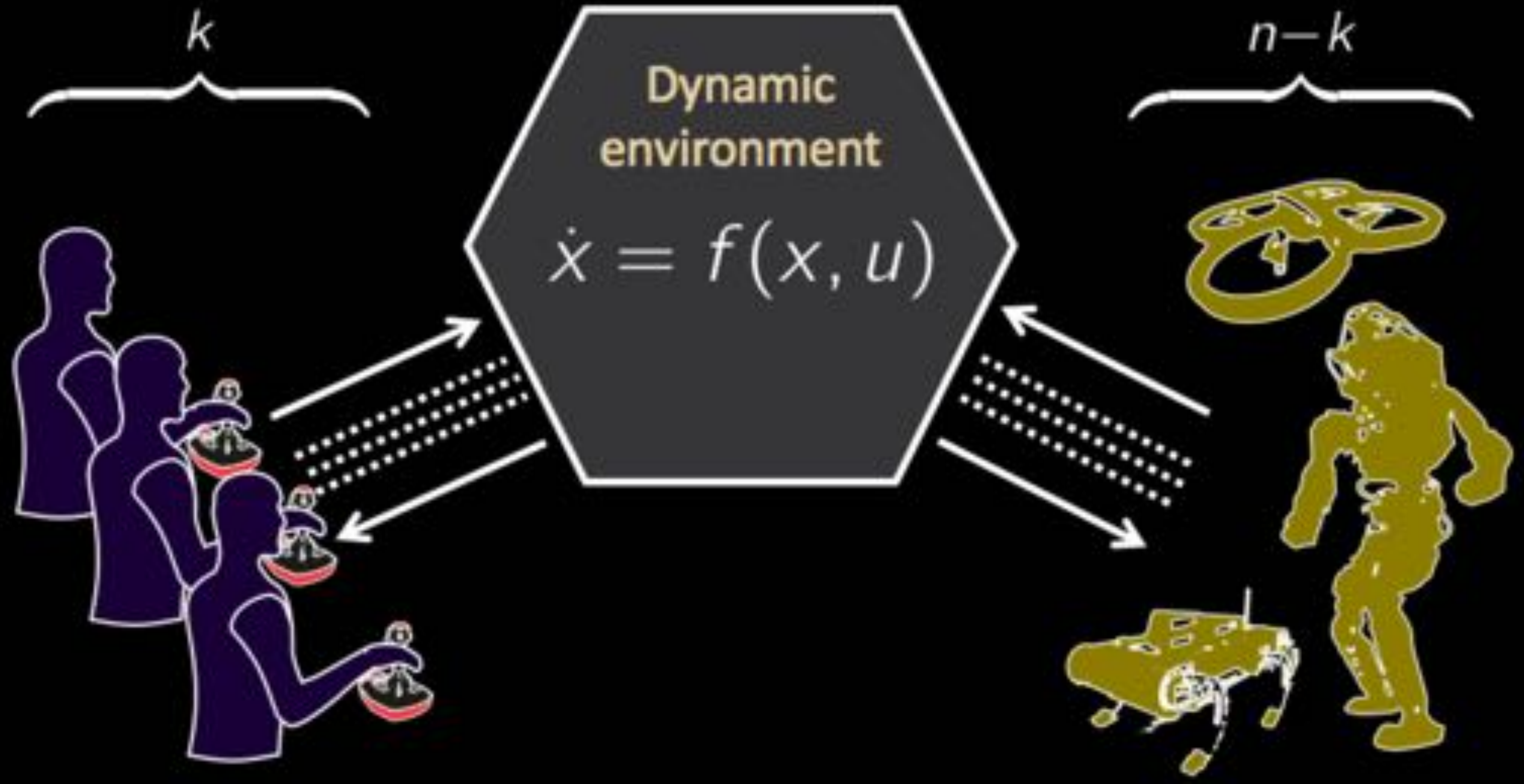
Linear Quadratic game: convergence of gradient method

$$K_1^+ = K_1 - \gamma \nabla_{K_1} c_1(x_0, K_1, K_2)$$
$$K_2^+ = K_2 - \gamma \nabla_{K_2} c_2(x_0, K_1, K_2)$$





Chasnov, Ratliff, Calderone, Mazumdar, Burden, *"Finite-Time Convergence of Gradient-Based Learning in Continuous Games."* AAAI Workshop on Reinforcement Learning in Games (2019).



Chasnov, Yamagami, Parsa, Ratliff, and Burden. *“Experiments with sensorimotor games in dynamic human/machine interaction.”* In SPIE Defense + Security. International Society for Optics and Photonics, 2019.

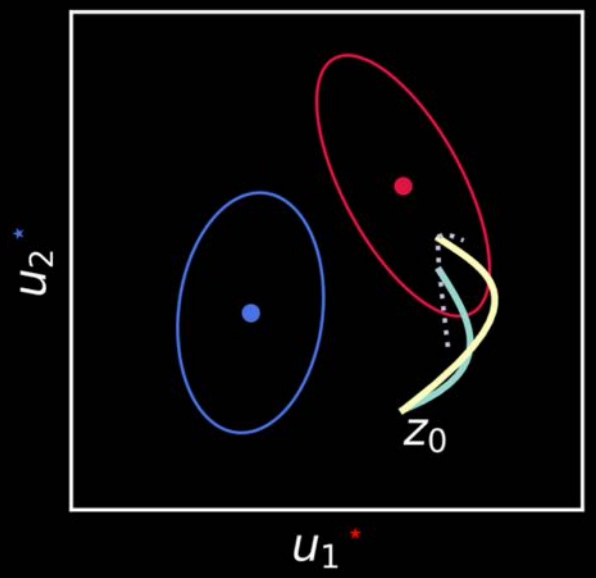
Papers and presentations

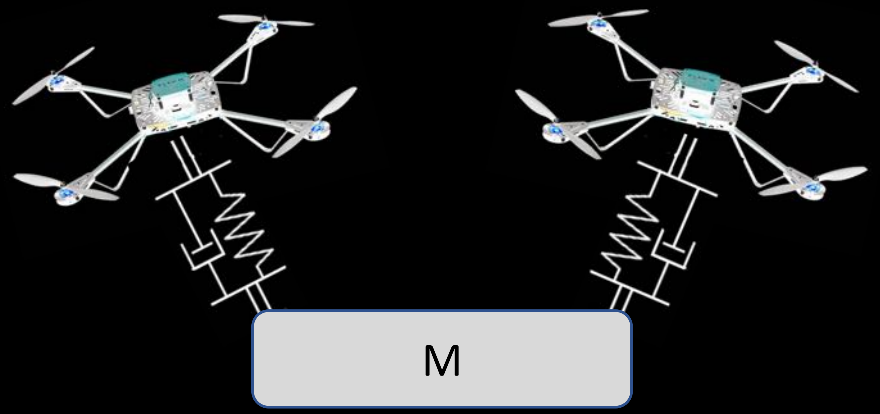
- Papers
  - Chasnov, Ratliff, Calderone, Mazumdar, Burden, *"Finite-Time Convergence of Gradient-Based Learning in Continuous Games."* AAAI Workshop on Reinforcement Learning in Games, 2019.
  - Chasnov, Yamagami, Parsa, Ratliff, Burden. *“Experiments with sensorimotor games in dynamic human/machine interaction.”* In SPIE Defense + Security. International Society for Optics and Photonics, 2019.
  - Chasnov, Ratliff, Mazumdar, and Burden. *“Convergence analysis of gradient-based learning in continuous games.”* The Conference on Uncertainty in Artificial Intelligence (in submission), 2019.
- Presentations
  - Neural Computation and Engineering Connection, Seattle WA, Jan 2019
  - AAAI Workshop on RL in Games, Honolulu HI, February 2019
  - SPIE Defense + Security, Baltimore MD, March 2019
  - Dynamic Walking, Canmore Alberta, June 2019

Future work

- Constrained action space
  - projected descent
- Strategic learning for faster convergence
  - recursive model of agents' learning

- Real world robotic systems
  - dynamically coupled quadcopters
- Human/machine games
  - teleoperation via optimization





Thank you