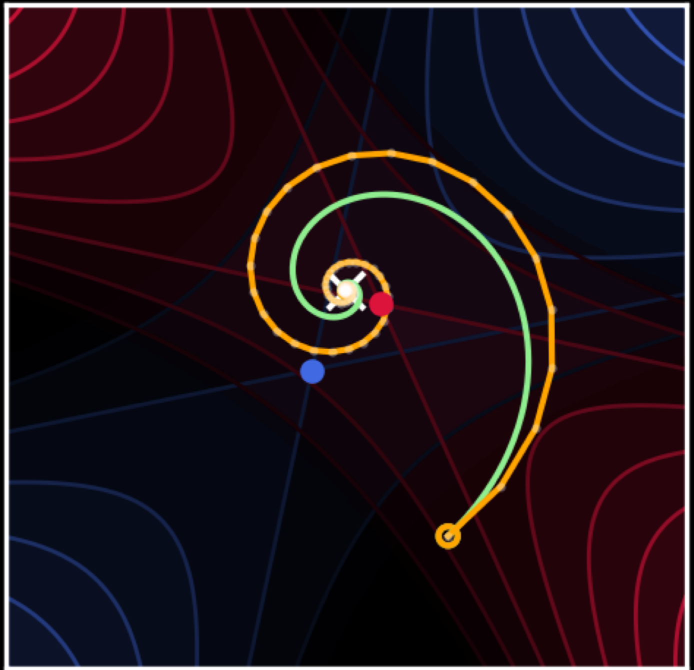
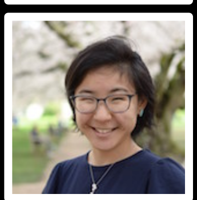
Sensorimotor game dynamics in coupled human-machine tasks




Benjamin Chasnov

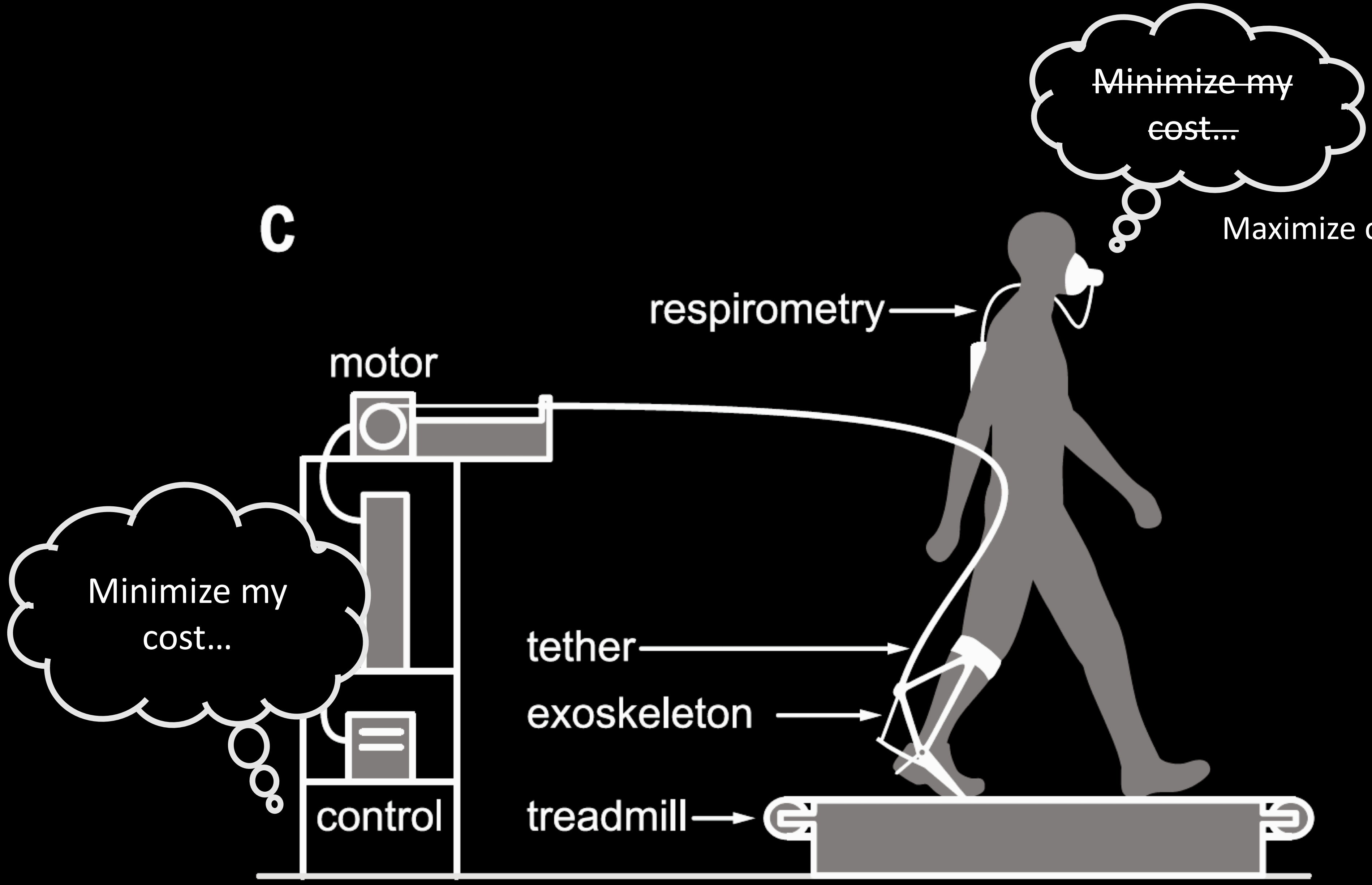

Momona Yamagami


Lillian J. Ratliff


Samuel A. Burden

AMP Lab
http://depts.washington.edu/amplify

Electrical and Computer Engineering
University of Washington, Seattle WA

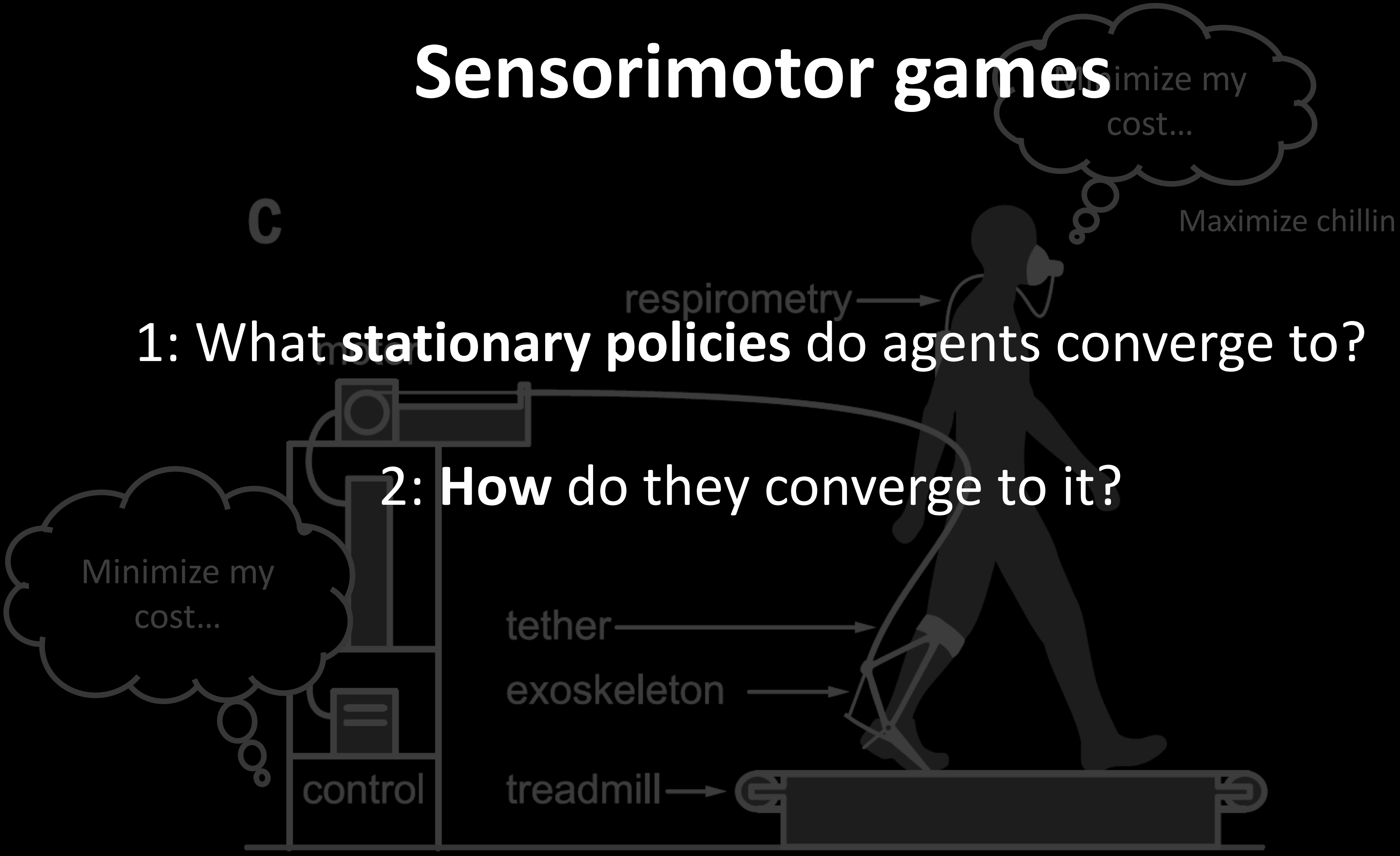


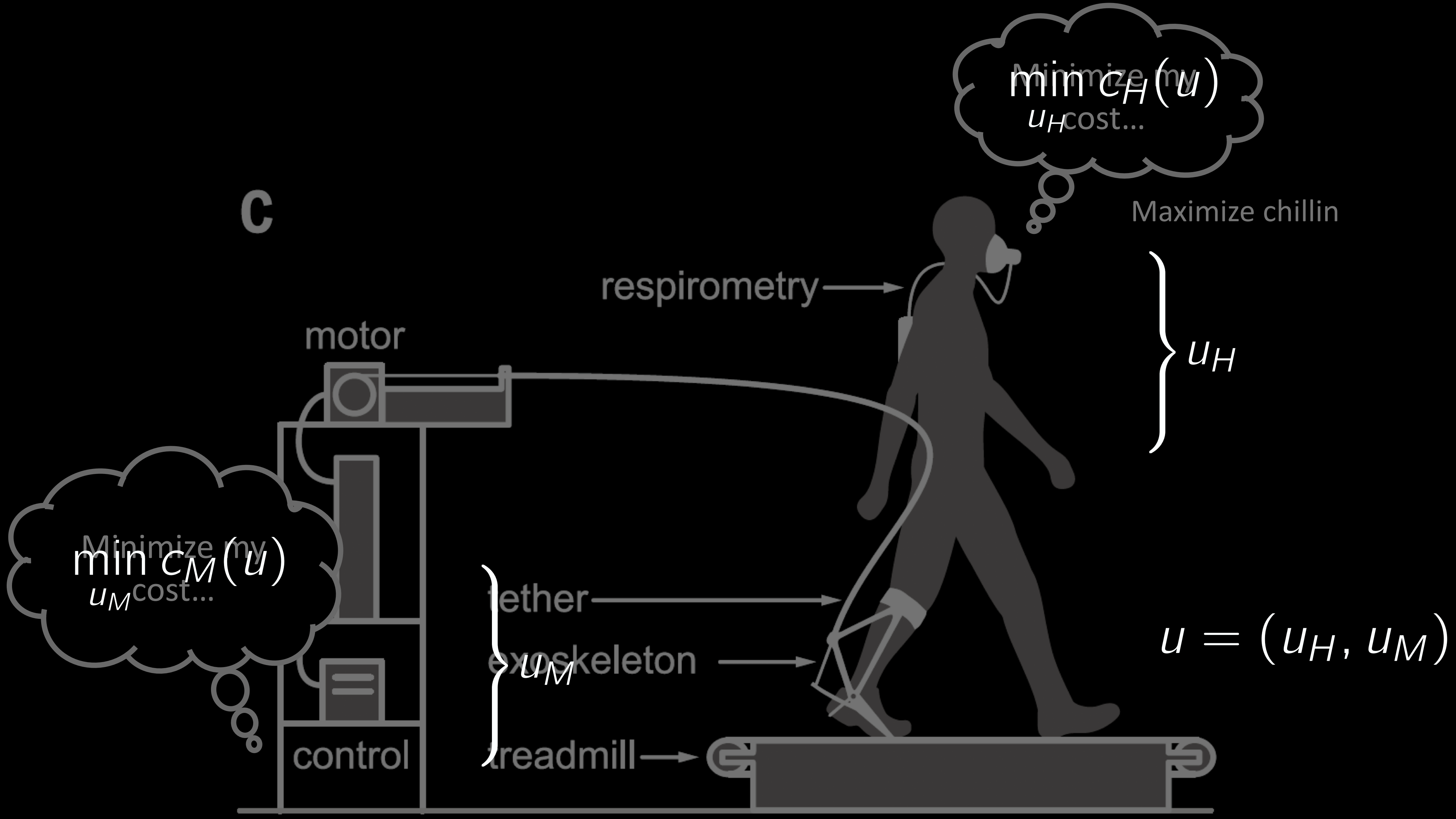
Zhang, J., Fiers, P., Witte, K. A., Jackson, R. W., Poggensee, K. L., Atkeson, C. G., & Collins, S. H. (2017). Human-in-the-loop optimization of exoskeleton assistance during walking. *Science*, *356*(6344), 1280-1284.

Sensorimotor games

1: What **stationary policies** do agents converge to?

2: **How** do they converge to it?





Zhang, J., Fiers, P., Witte, K. A., Jackson, R. W., Poggensee, K. L., Atkeson, C. G., & Collins, S. H. (2017). Human-in-the-loop optimization of exoskeleton assistance during walking. *Science*, *356*(6344), 1280-1284.

Learning to make decisions by optimization

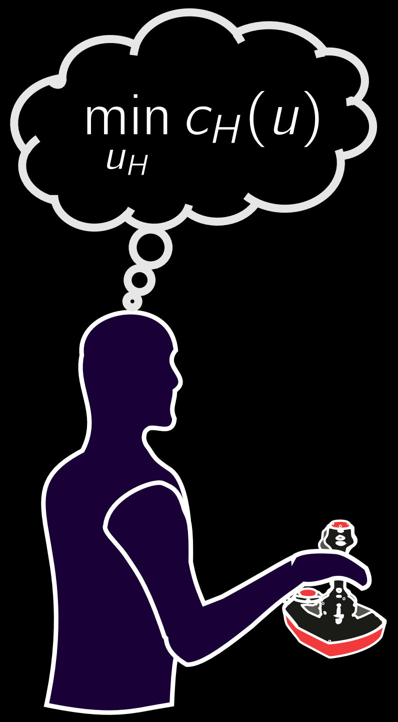
A “rational” human minimizes its cost $\boldsymbol{c_{uH}(u)}$ by descending its steepest gradient, $\boldsymbol{D_{uH}c_H(u)}$.

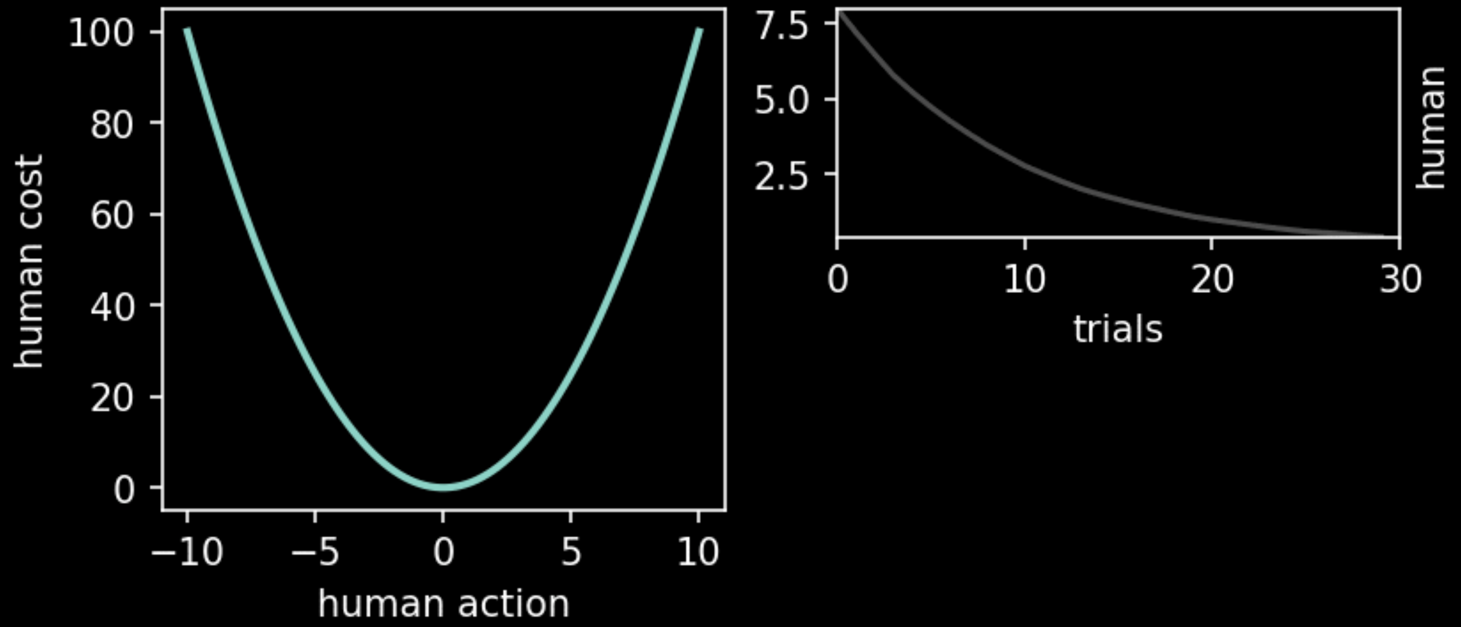
$$u_H^+ = u_H - \gamma D_{u_H} c_H(u)$$

Learning to make decisions by optimization

A "rational" human minimizes its cost $\boldsymbol{c_{uH}(u)}$ by descending its steepest gradient, $\boldsymbol{D_{uH}c_H(u)}$.

$$u_H^+ = u_H - \gamma D_{u_H} c_H(u)$$



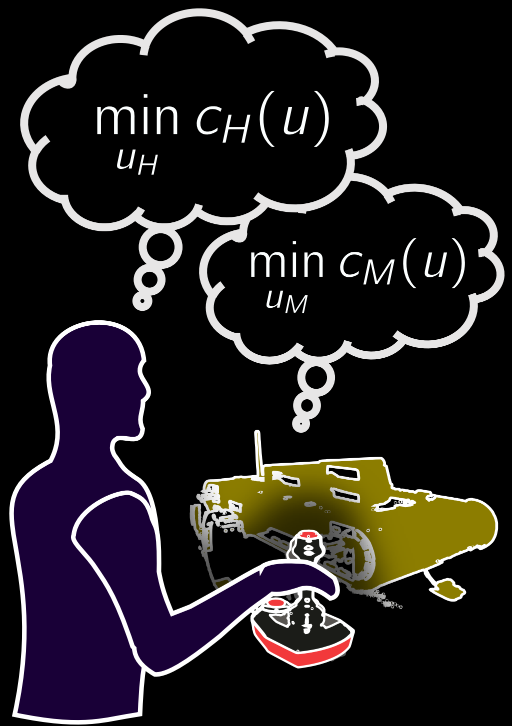


Coupled learning: game dynamics

A group of optimization agents minimize their *own* cost
with respect to their *own* action

$$u_H^+ = u_H - \gamma D_{u_H} c_H(u)$$

$$u = (u_H, u_M)$$

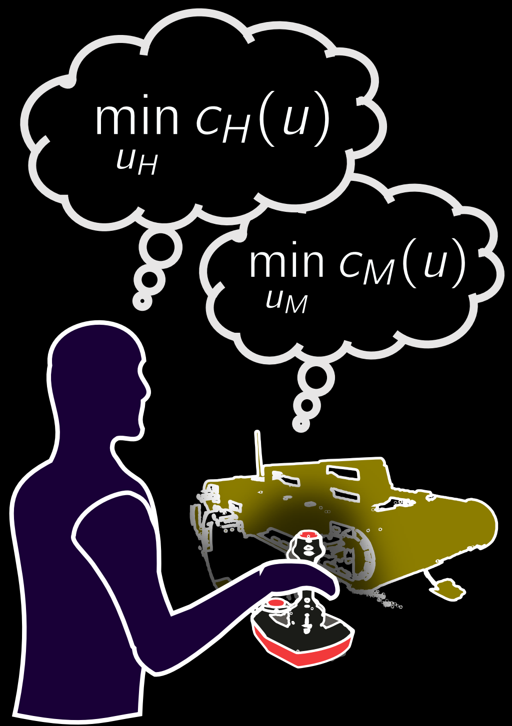


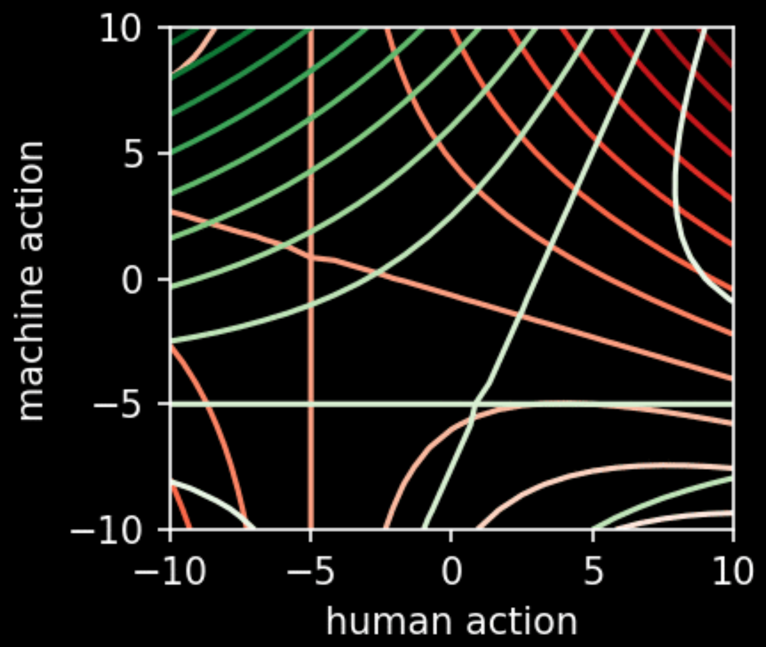
Coupled learning: game dynamics

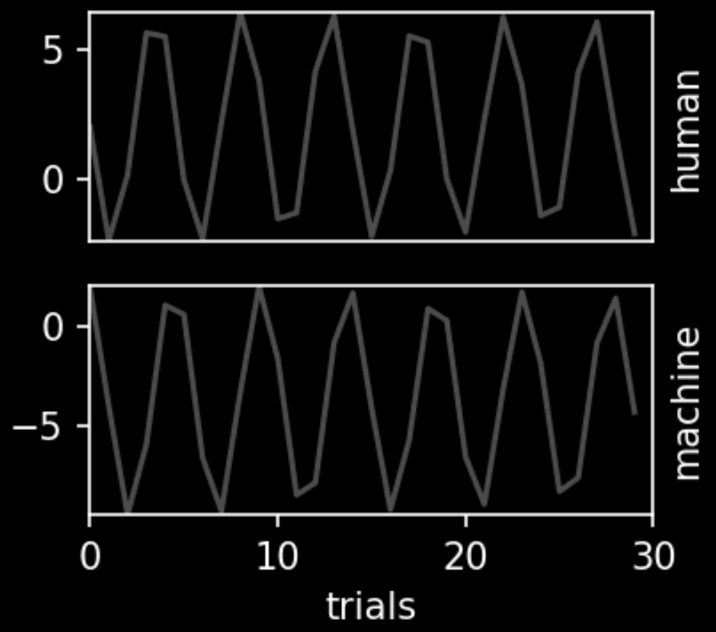
A group of optimization agents minimize their *own* cost
with respect to their *own* action

$$u_H^+ = u_H - \gamma D_{u_H} c_H(u)$$
$$u_M^+ = u_M - \gamma D_{u_M} c_M(u)$$

$$u = (u_H, u_M)$$





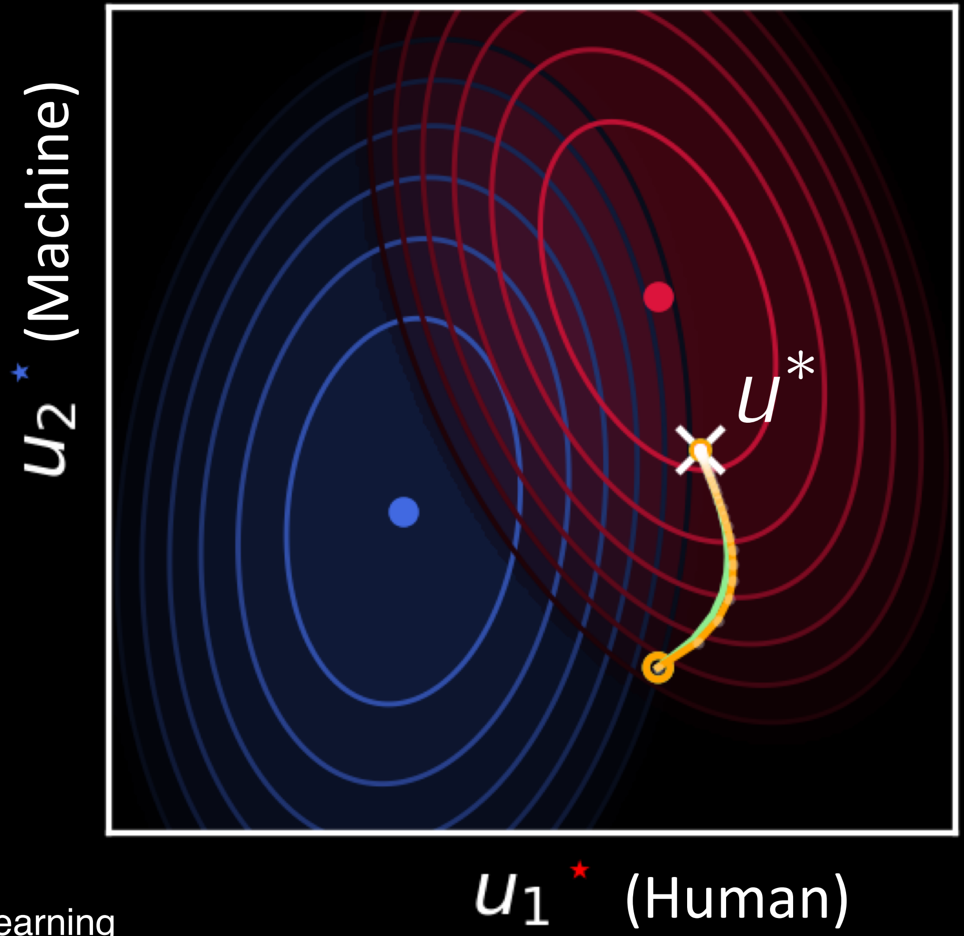


Learning: a dynamical systems perspective

$$u_H^+ = u_H - \gamma D_{u_H} c_H(u)$$
$$u_M^+ = u_M - \gamma D_{u_M} c_M(u)$$

$$\updownarrow$$

$$\dot{u} = -\omega(u)$$
$$\approx -\underbrace{J(u^*)}(u - u^*)$$

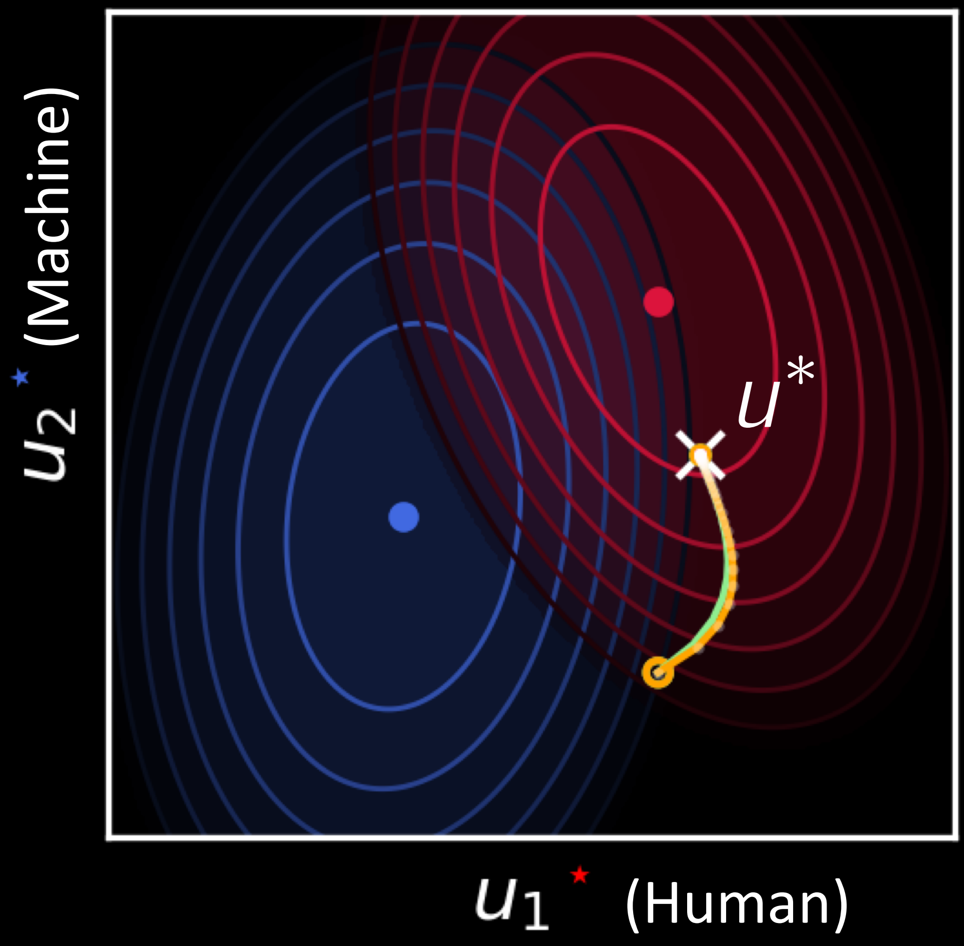


B. Chasnov, L. Ratliff, E. Mazumdar, S. Burden. “Convergence Analysis of Gradient-Based Learning with Non-Uniform Learning Rates in Non-Cooperative Multi-Agent Settings” UAI 2019

Learning: a dynamical systems perspective

$$\dot{u} = -\omega(u)$$
$$\approx -\underbrace{J(u^*)}(u - u^*)$$

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



Predictions: learning in games

1. What: stationary policies?

Agents may converge to local maxima

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

Nash

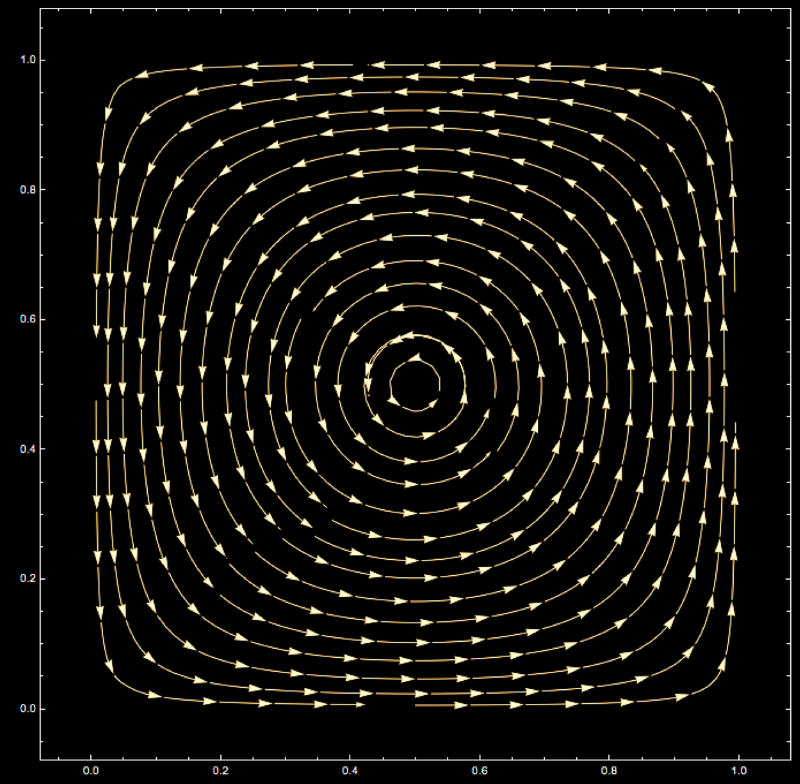
$$J(u^*) = \begin{bmatrix} + & \\ & - \end{bmatrix}$$

Non-Nash

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).

2. How: transient behaviors?

Agent policies may oscillate endlessly

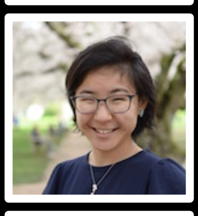


Mertikopoulos, Papadimitriou and Piliouras. "Cycles in adversarial regularized learning." In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2018.

thank you!

bchasnov@uw.edu
http://students.washington.edu/bchasnov/


Benjamin Chasnov


Momona Yamagami


Lillian J. Ratliff


Samuel A. Burden

Electrical and Computer Engineering
University of Washington, Seattle WA

COMPUTATIONAL
NEUROSCIENCE
CENTER

Definition: differential Nash equilibrium

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

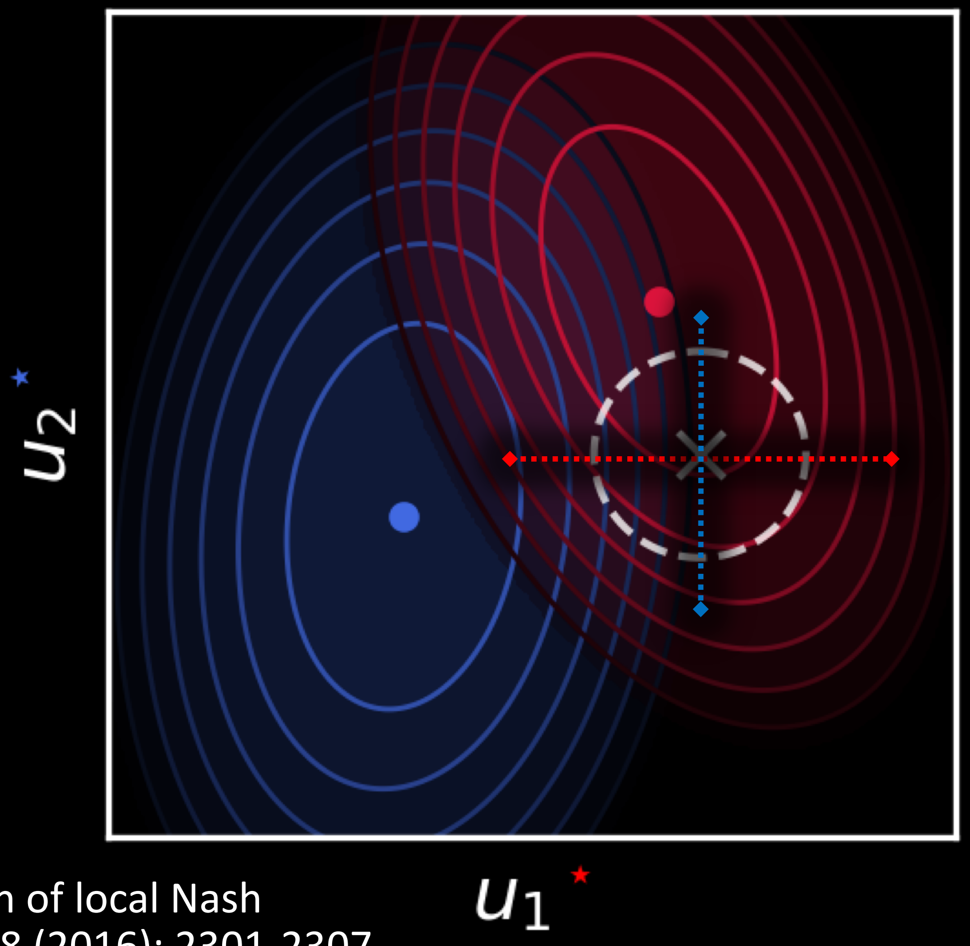
A strategy is a *differential Nash equilibrium* if it satisfies first order conditions

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

and second order conditions

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

for twice continuously-differentiable costs.



Lillian J. Ratliff, Samuel A. Burden, and S. Shankar Sastry. "On the characterization of local Nash equilibria in continuous games." *IEEE Transactions on Automatic Control* 61, no. 8 (2016): 2301-2307.

Spectrum of the Jacobian

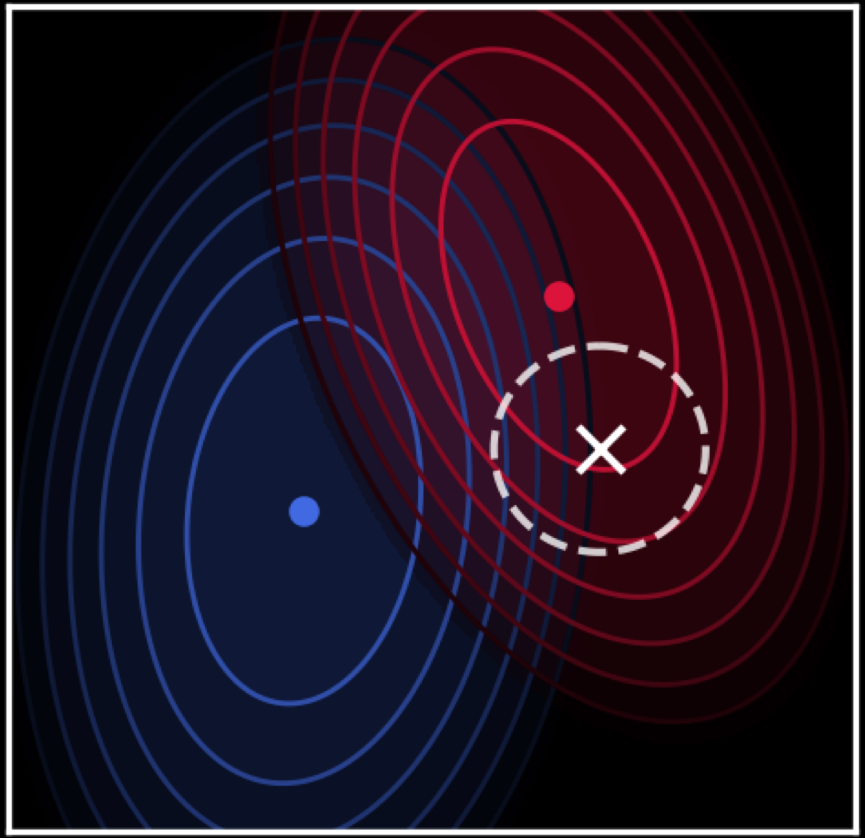
$$\dot{u} = -\omega(u)$$

$$\approx -\underbrace{J(u^*)}_{}(u - u^*)$$

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

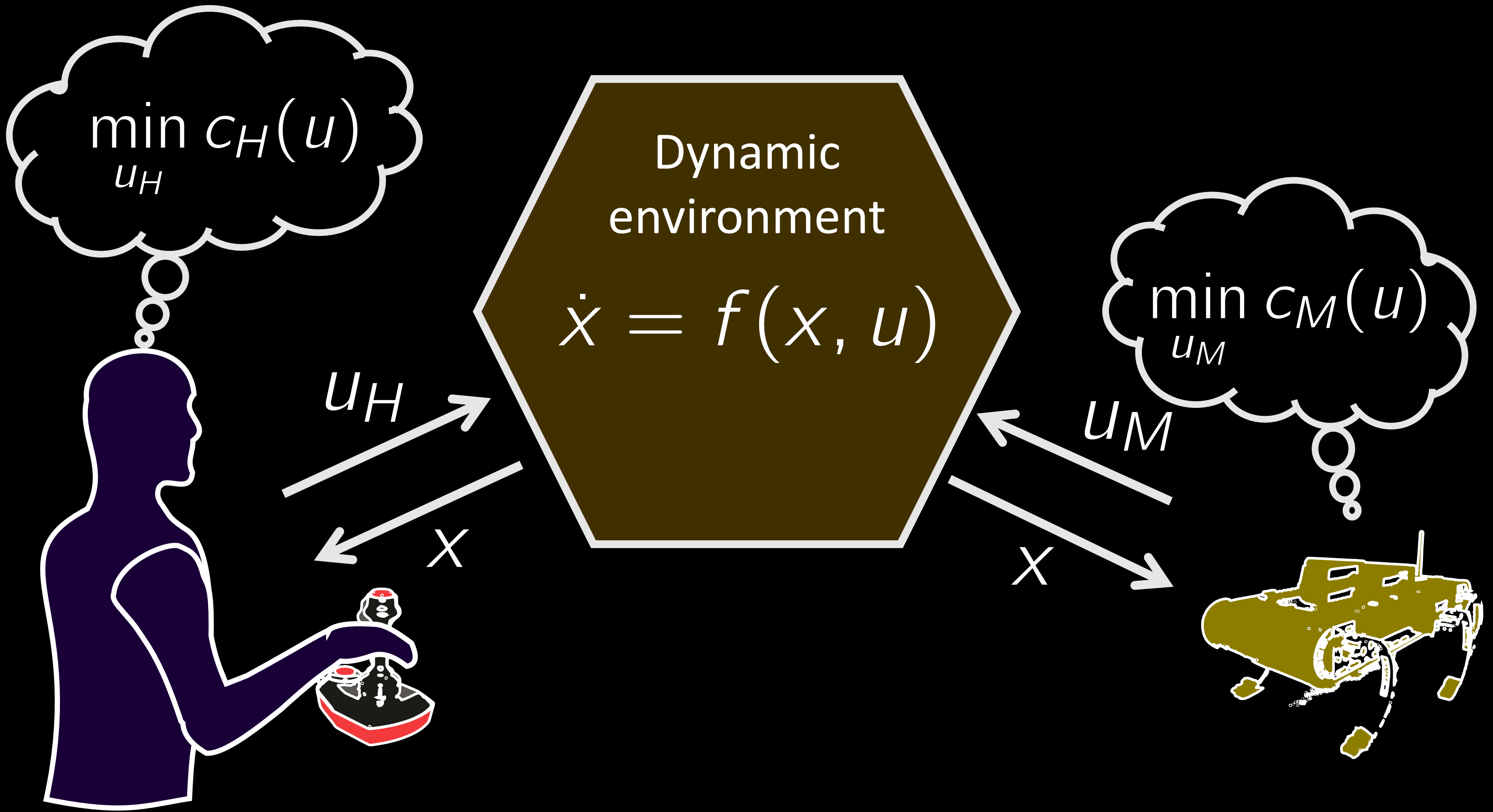
If $\mathrm{blockdiag}_i(J) > 0$ at u^* $\forall i$, then u^* is Nash.

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



A model of decision-making

Optimal control guides autonomous controllers to make decisions in a dynamic environment.



Control of dynamical systems

Dynamic environment

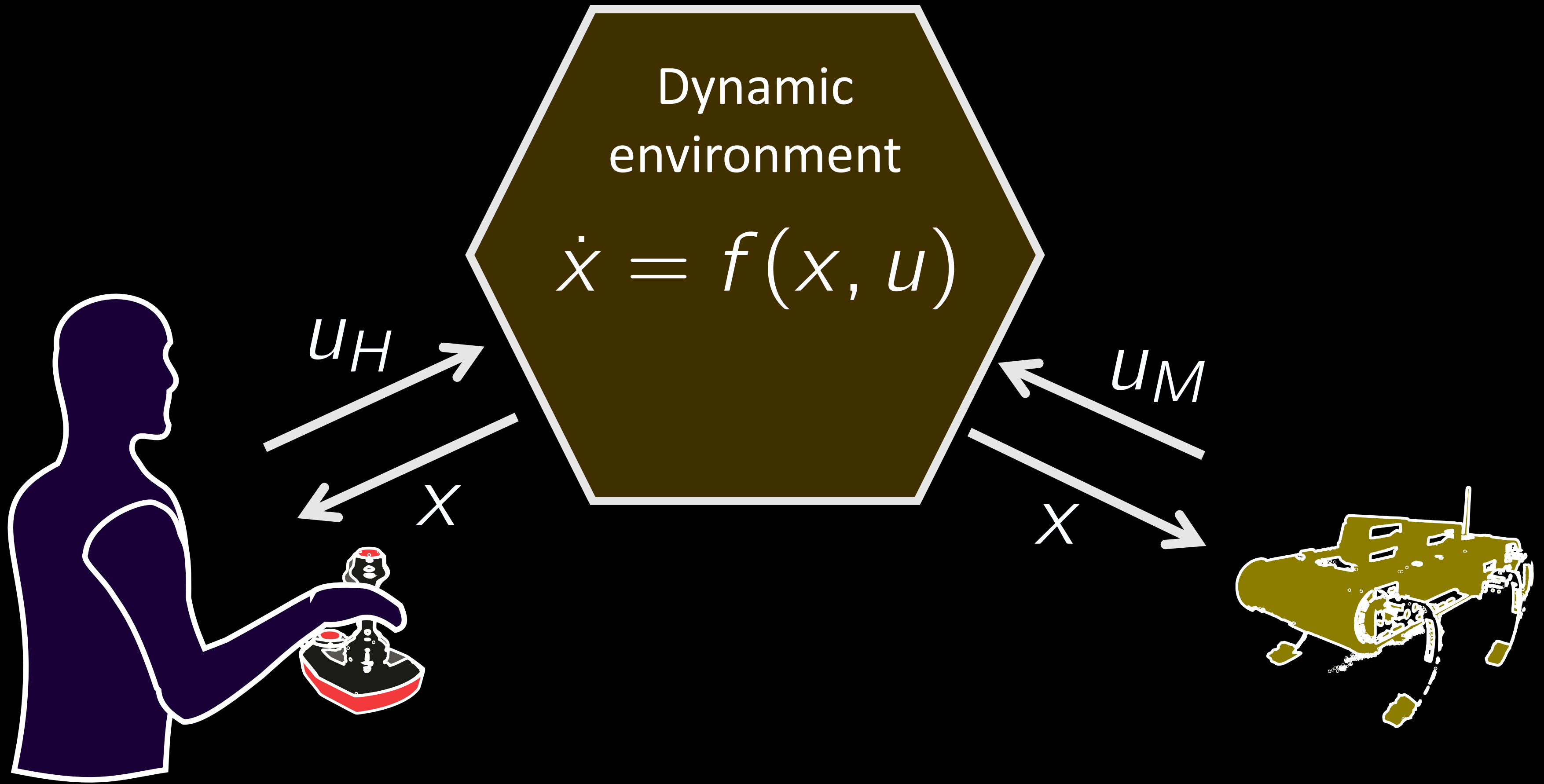
$$\dot{x} = f(x, u)$$

$$u = (u_H, u_M)$$

CITE SPIE PAPER

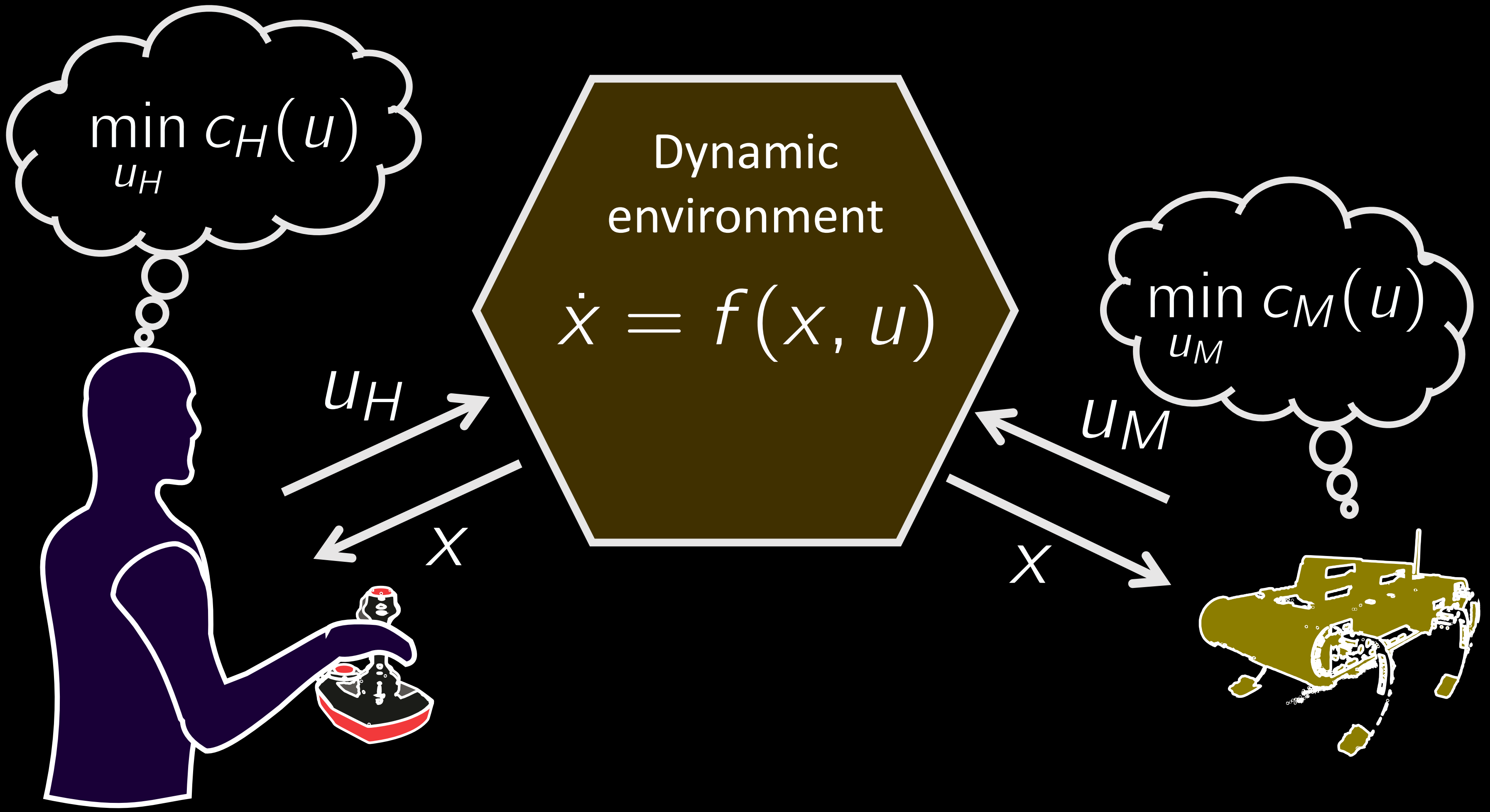
Sensorimotor games

Human and machine jointly control a dynamical system by designing control (policies) $\boldsymbol{u_H}$, $\boldsymbol{u_M}$ in feedback with observing $\boldsymbol{x}$.



A model of decision-making

Optimal control guides autonomous controllers to make decisions in a dynamic environment.

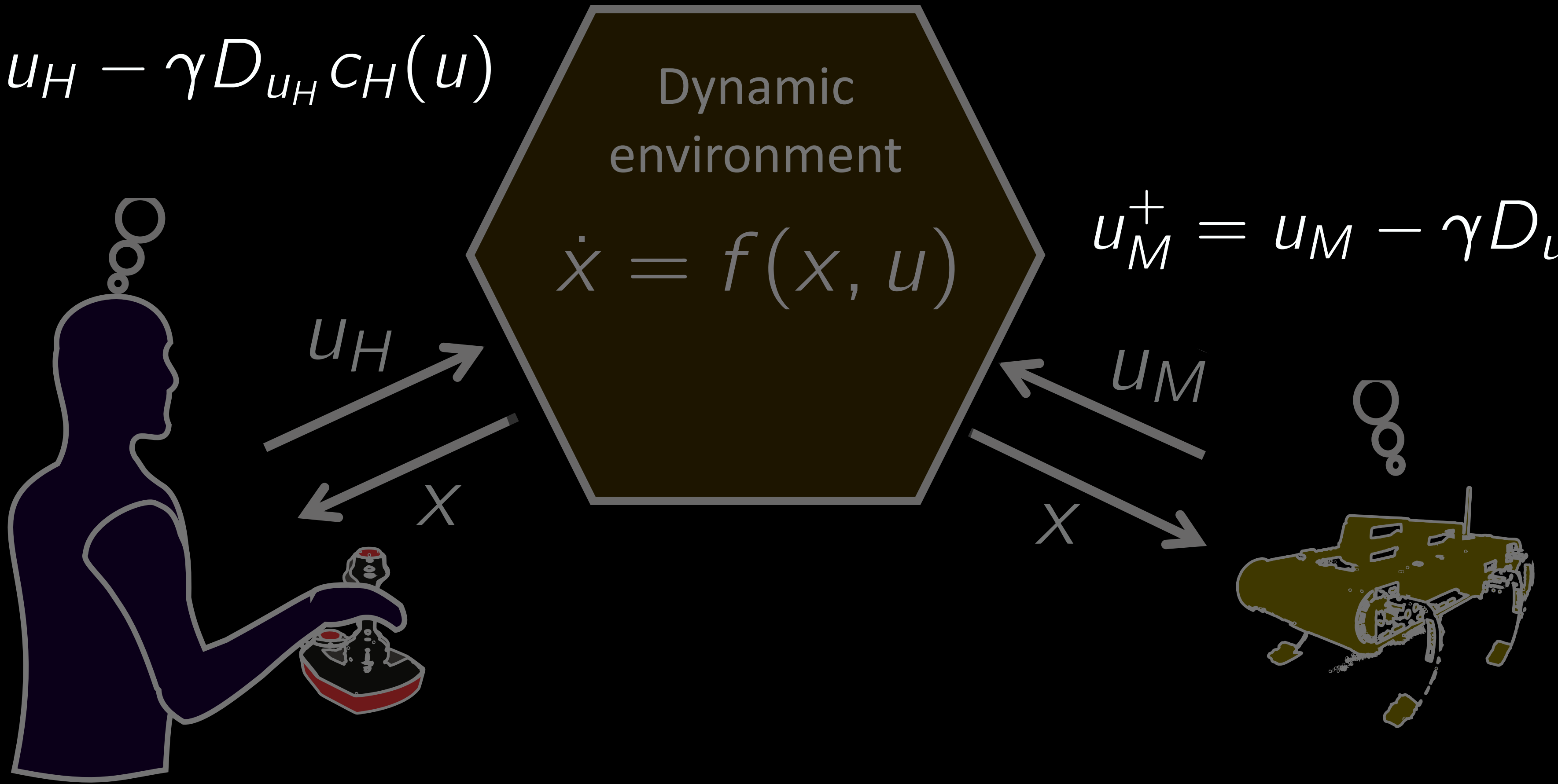


Learning to make decisions by optimization

A first order method for seeking local minima of each agents’ cost.

$$u_H^+ = u_H - \gamma D_{u_H} c_H(u)$$

$$u_M^+ = u_M - \gamma D_{u_M} c_M(u)$$



Prediction 1: not all stable equilibria are Nash

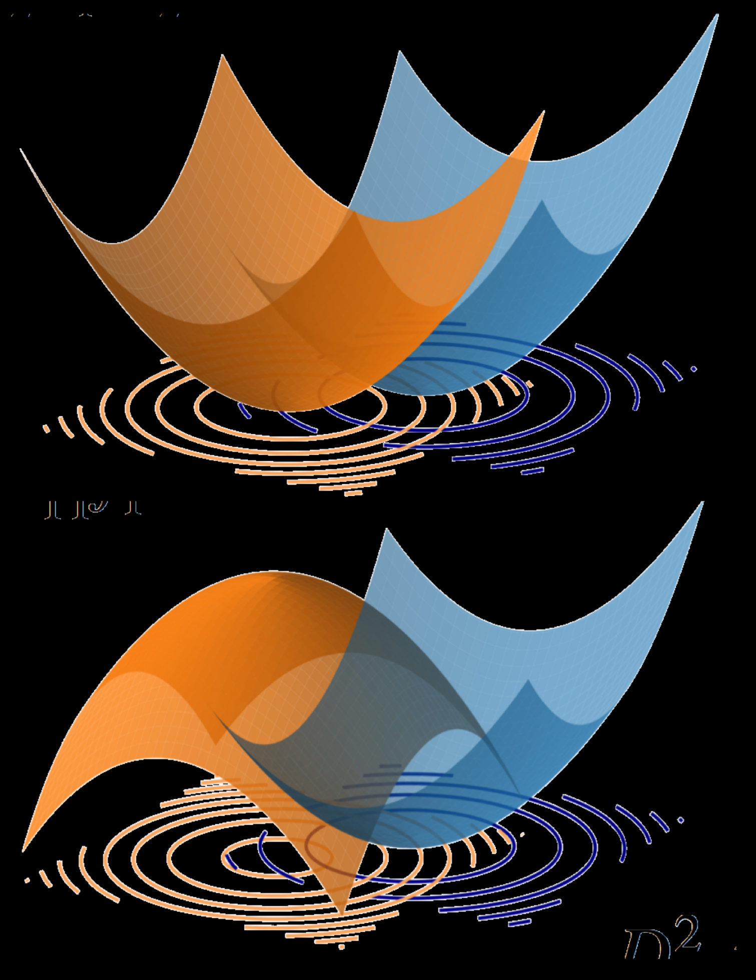
$$\mathrm{spec}(J) \subset \mathbb{C}_+^\circ$$

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

Nash

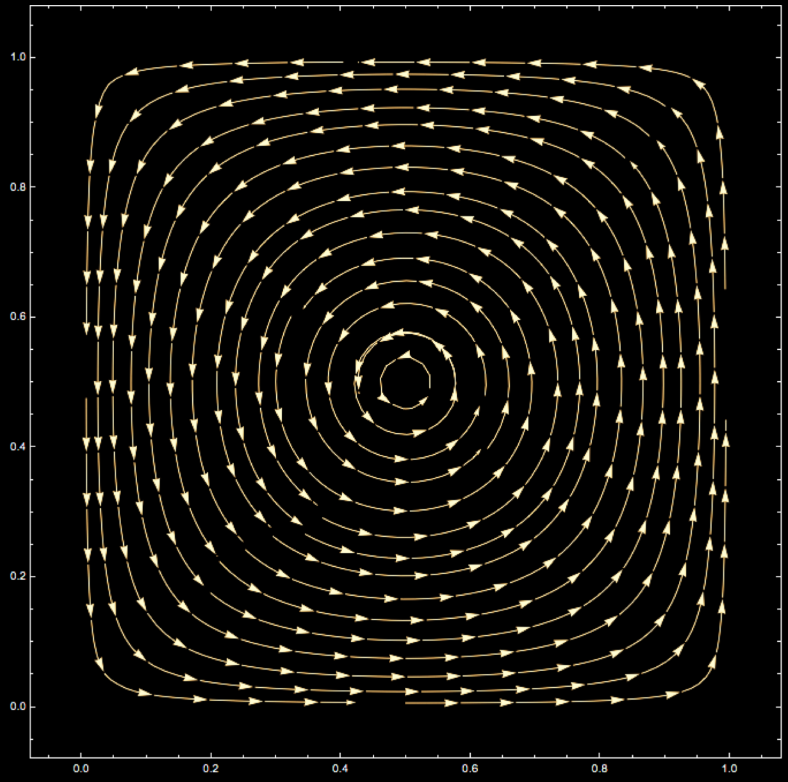
$$J(u^*) = \begin{bmatrix} + & \\ & - \end{bmatrix}$$

Non-Nash

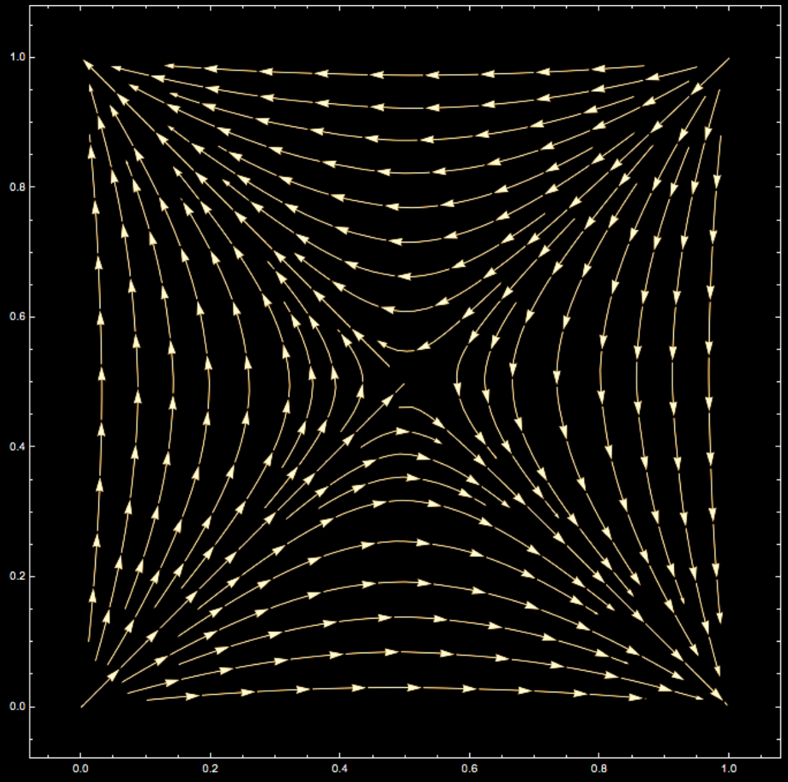


Mazumdar, Eric V., Michael I. Jordan, and S. Shankar Sastry. "On finding local nash equilibria (and only local nash equilibria) in zero-sum games." *arXiv preprint arXiv:1901.00838* (2019).

Prediction 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.

Future work

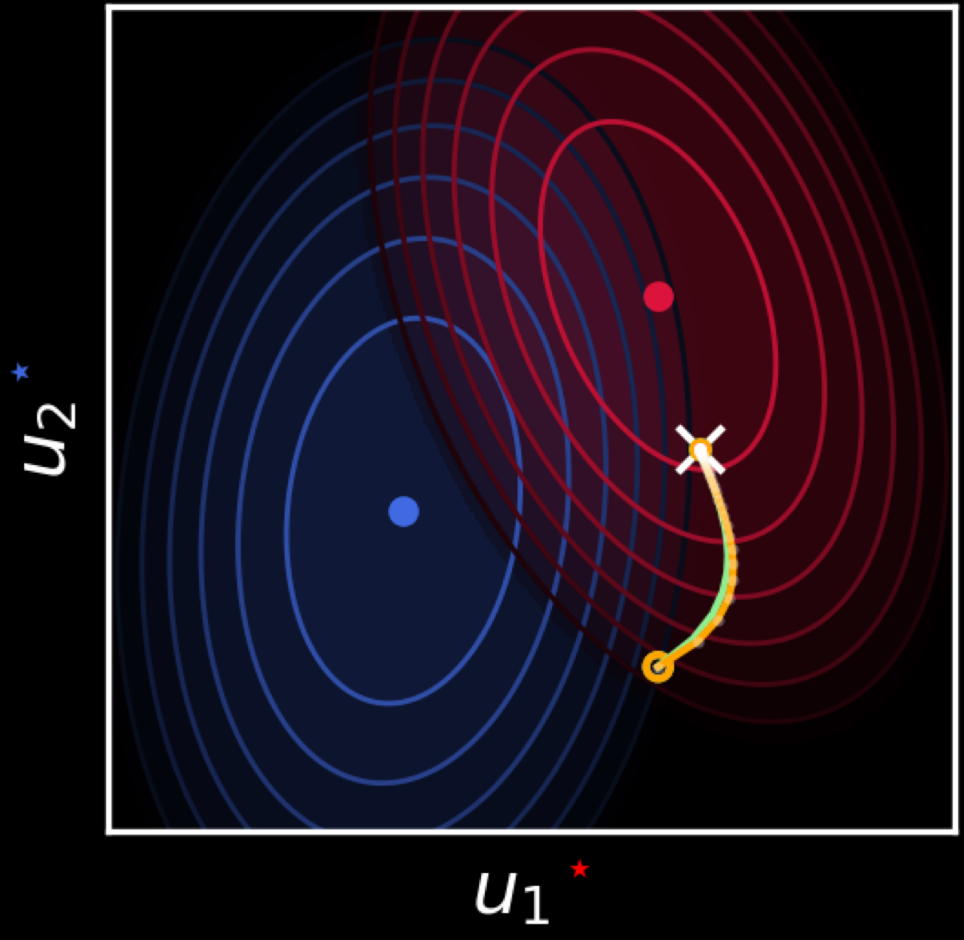
- Modify learning rule to avoid stable non-Nash and limit cycles

Dynamical systems perspective

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

(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).