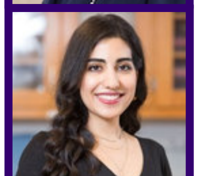
Experiments with sensorimotor games in dynamic human/machine interaction

Electrical & Computer Engineering
University of Washington
Seattle, WA USA


Benjamin Chasnov


Momona Yamagami


Behnoosh Parsa


Lillian Ratliff

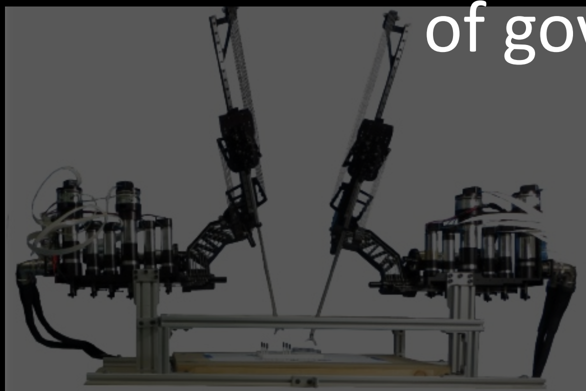

Sam Burden

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

Our physical world is dynamic

Our physical world is dynamic

The state of the world, ***x***, follows a set of governing differential equations.

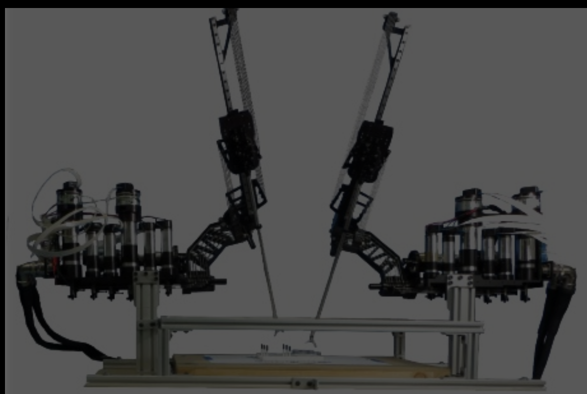


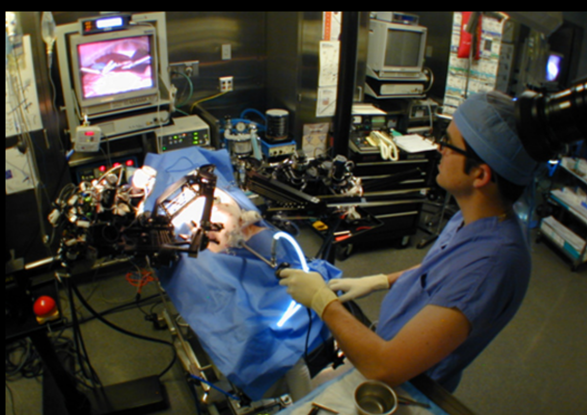
Dynamic environment

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

Our physical world is dynamic

Machine control + Human *teleoperation*



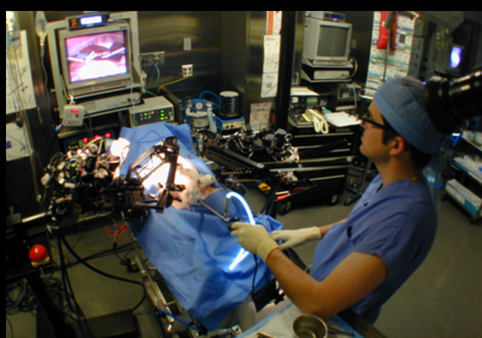


Team control of autonomous systems

Machine actions:



u_M

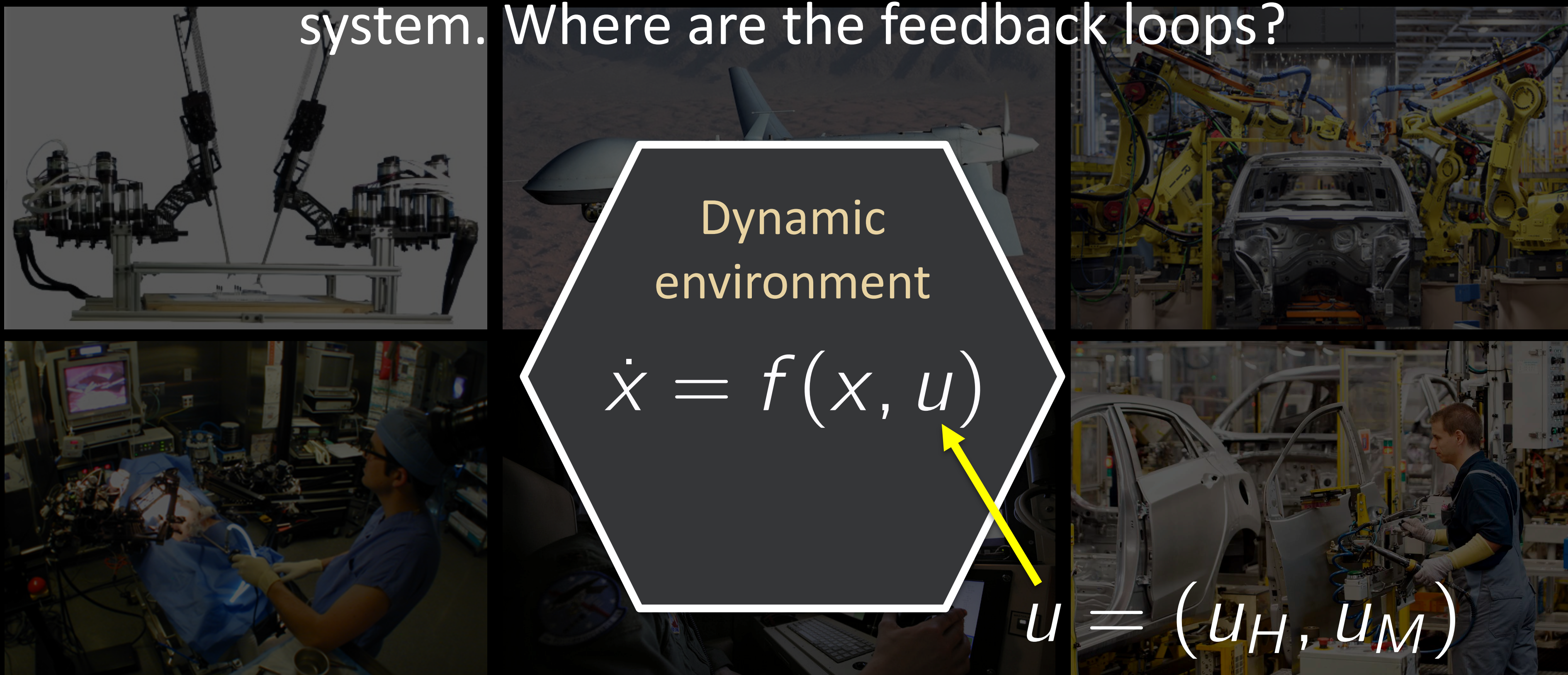


u_H

Human actions:

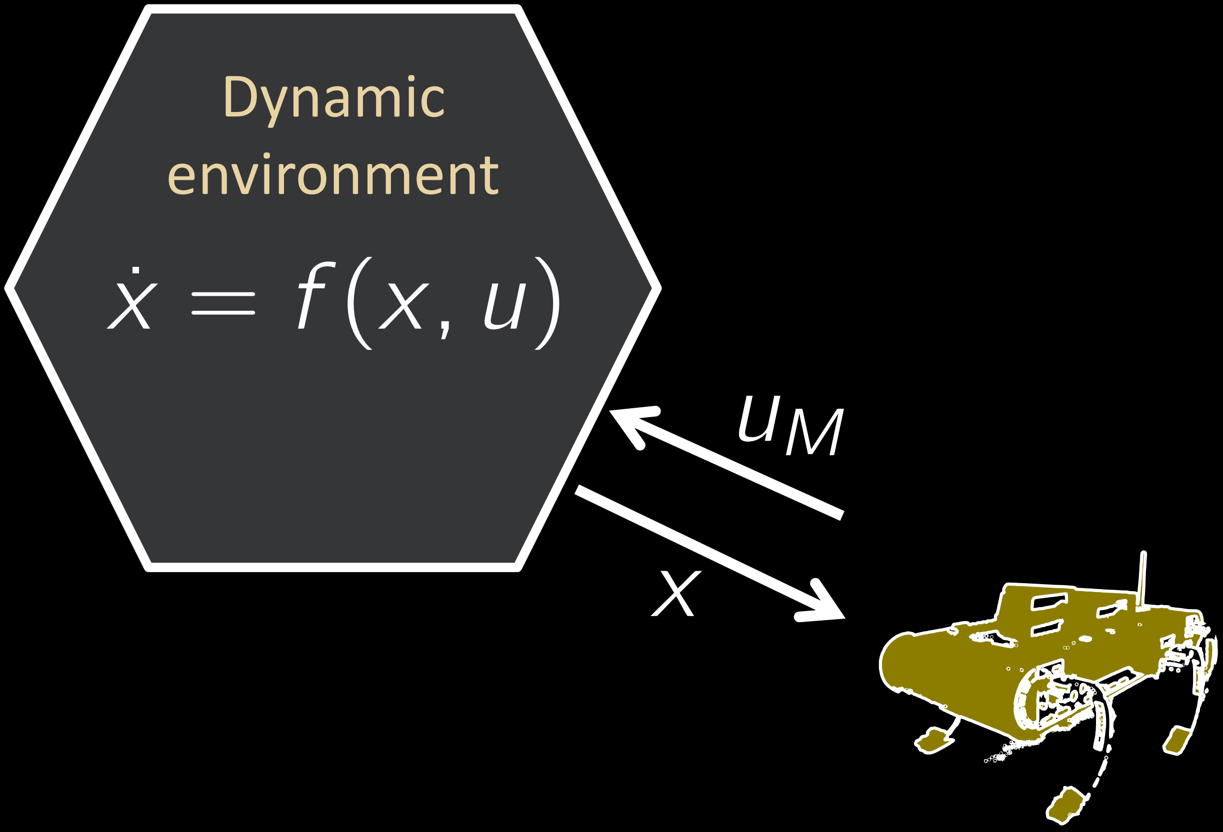
Team control of autonomous systems

The joint action ***u*** drives the state ***x*** in the dynamical system. Where are the feedback loops?



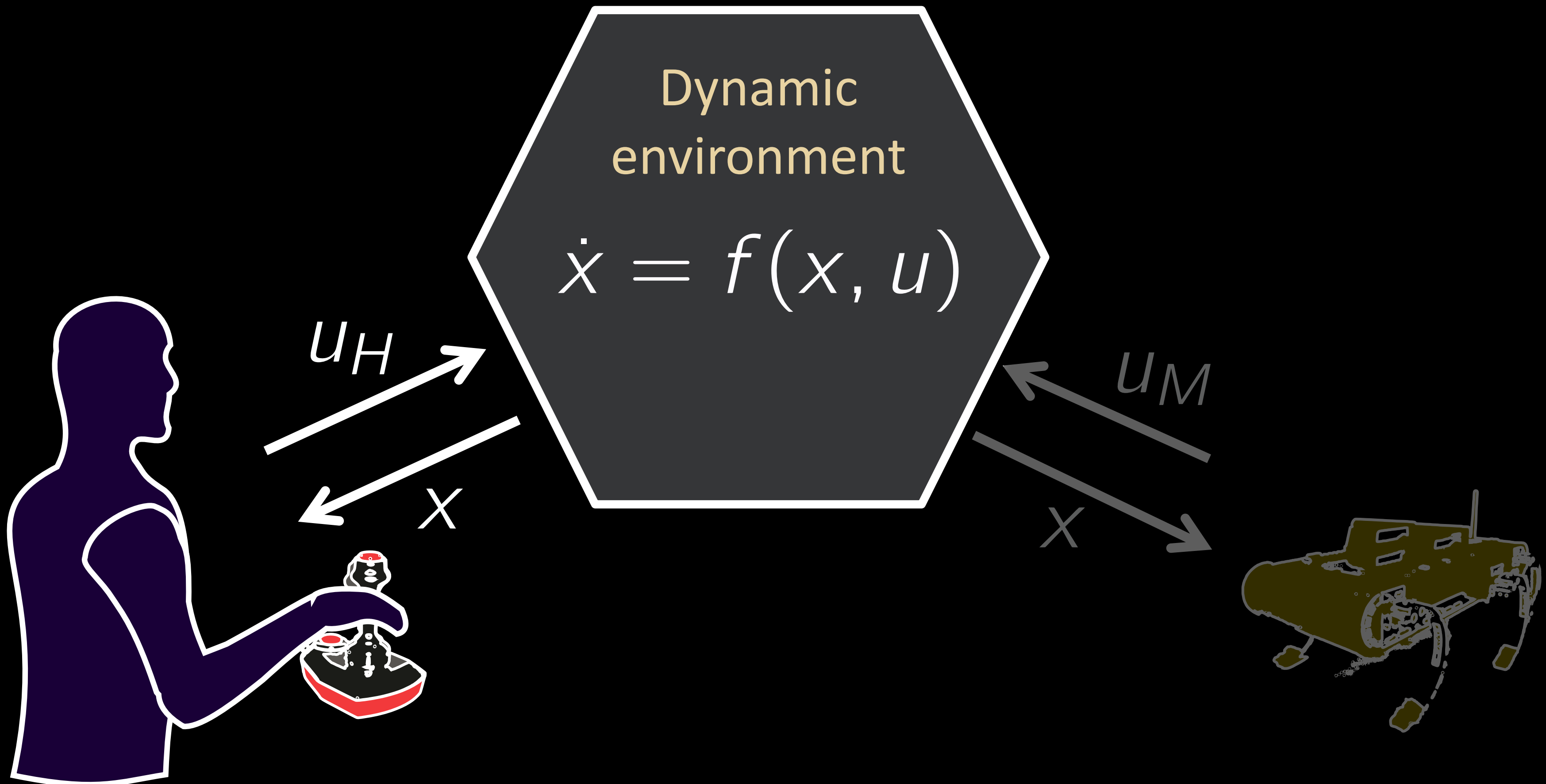
Feedback loops: autonomous controller

The machine controls the system by reacting to state ***x***, and choosing control $\boldsymbol{u_M}$.



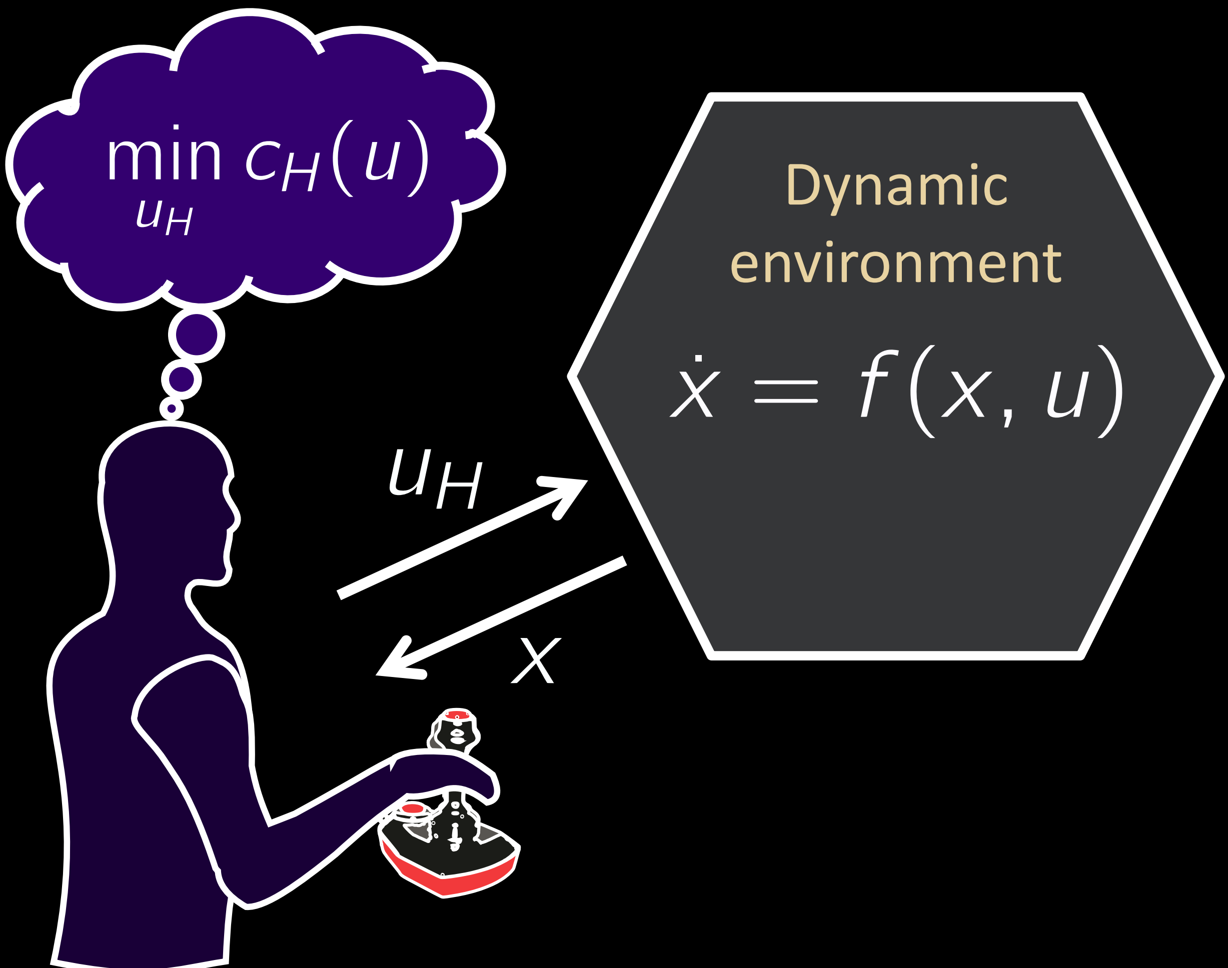
Feedback loops: human operator

Human *teleoperates* a dynamical system by providing control input $\boldsymbol{u_H}$ in feedback with observing $\boldsymbol{x}$.



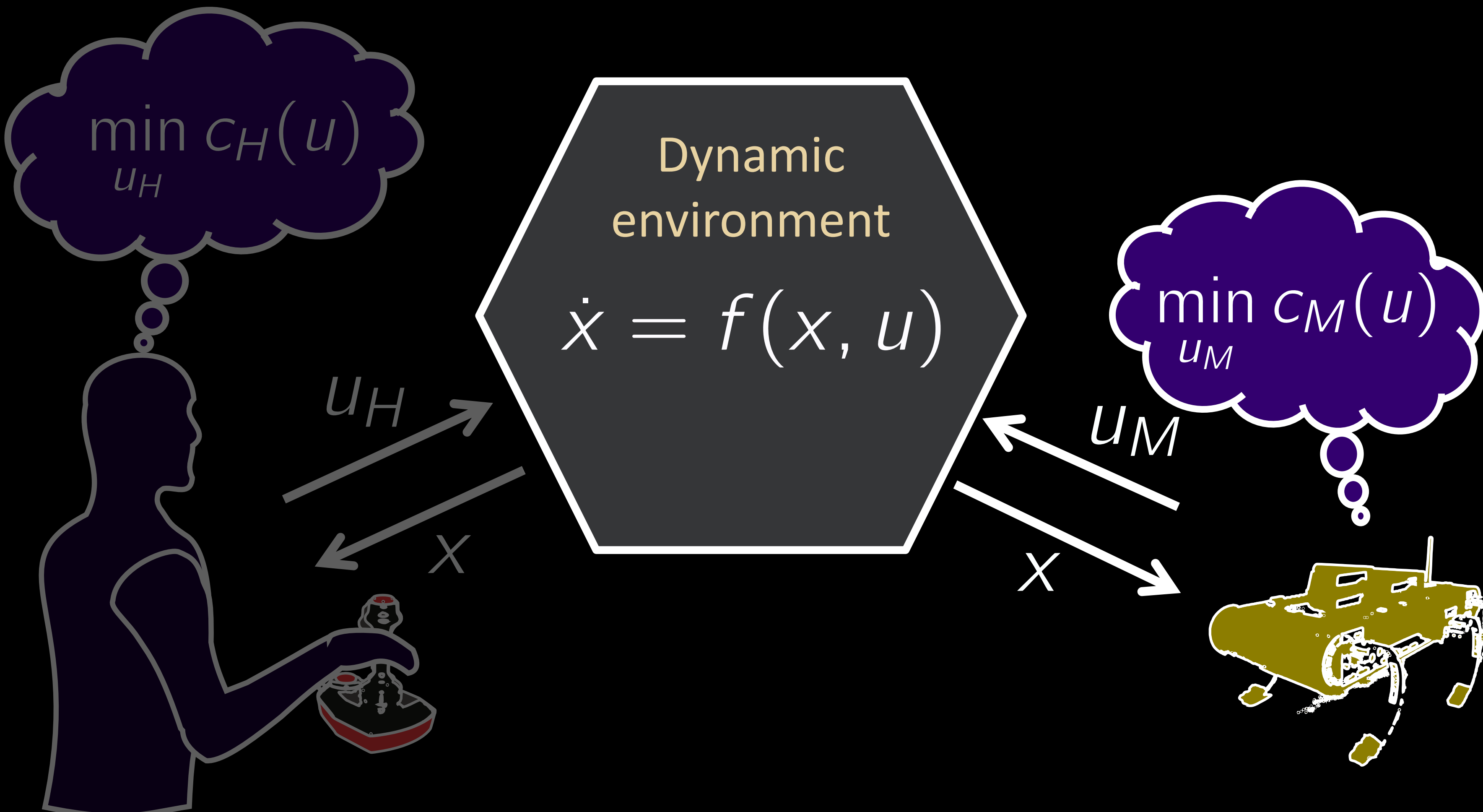
A model of decision-making

A "rational" agent minimizes a cost ***c_H(u)*** subject to dynamical constraints.



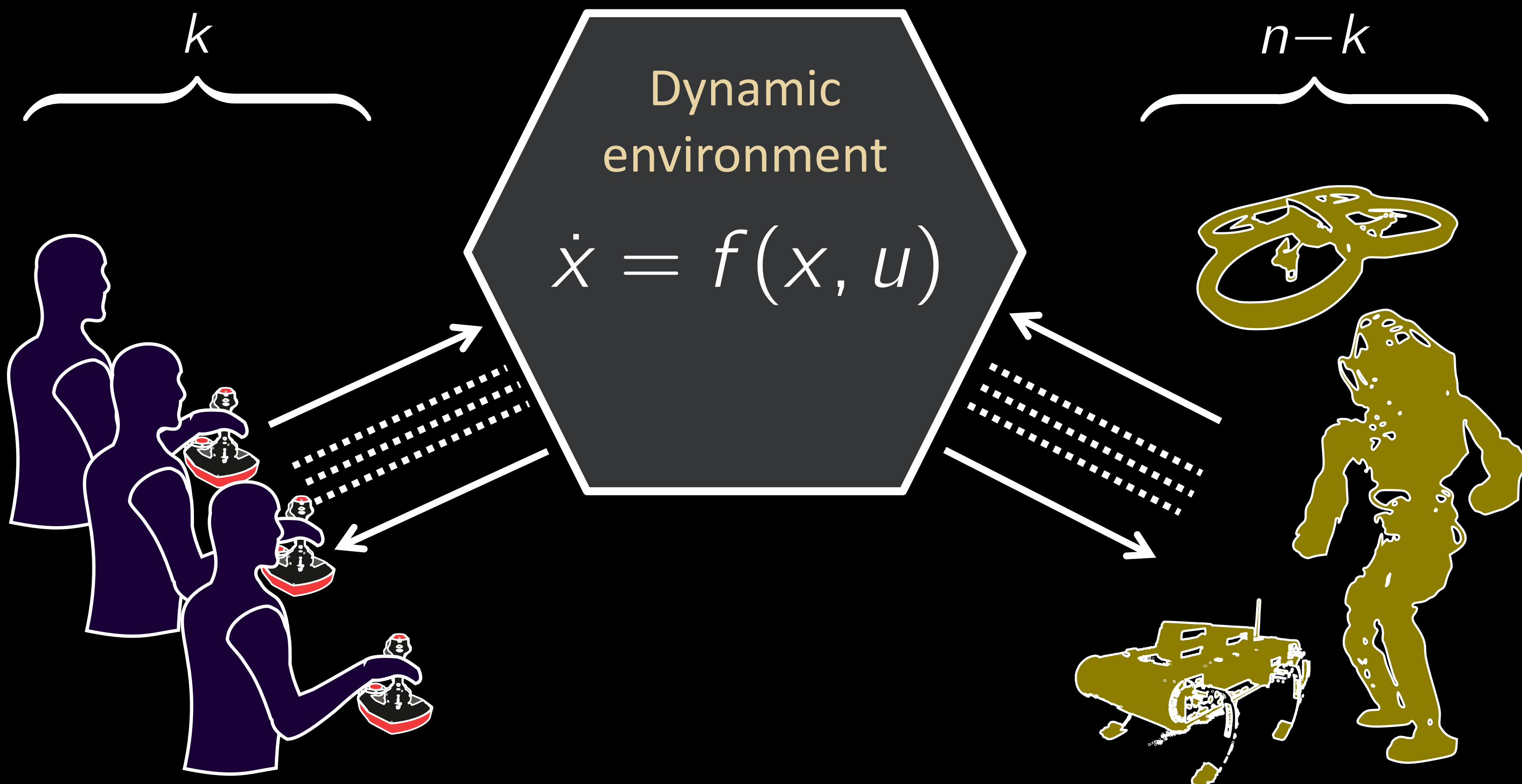
A model of decision-making

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



A model of *team* decision-making

Human and machines play a *sensorimotor game.*



(games are *non-cooperative*)

Cooperative

$$\boxed{\min_u \sum_{i=1}^n c_i(u)}$$

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

Trust and communication

$$\neq$$

Non-cooperative

$$\boxed{\begin{array}{c} \min_{u_1} \ c_1(u) \\ \vdots \\ \min_{u_n} \ c_n(u) \end{array}}$$

Stationary conditions:

Pareto optimum

Nash equilibrium

Learning to make decisions by optimization

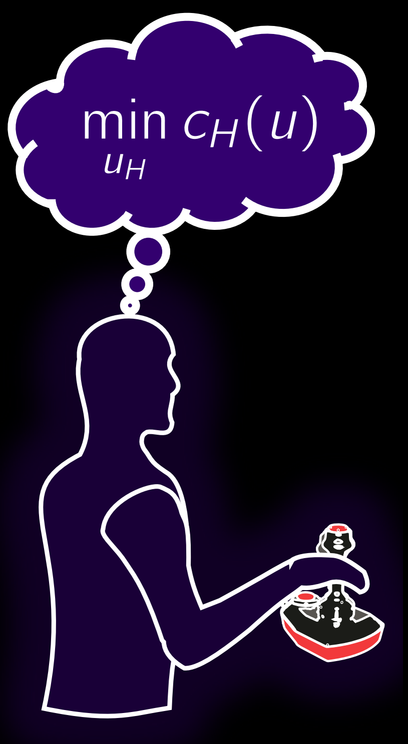
A “rational” human minimizes its cost $c_H(u)$ by descending its steepest gradient, $D_H c_H(u)$.

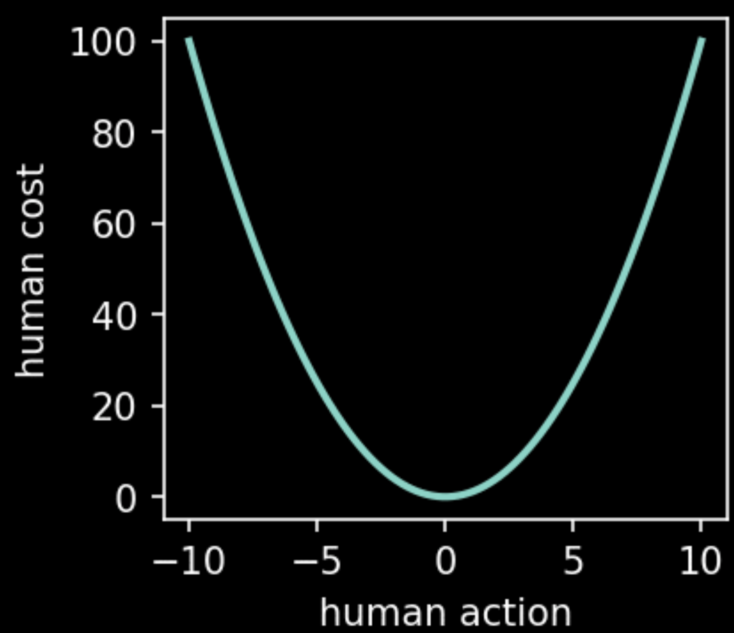
$$u_H^+ = u_H - \gamma D_{u_H} c_H(u)$$

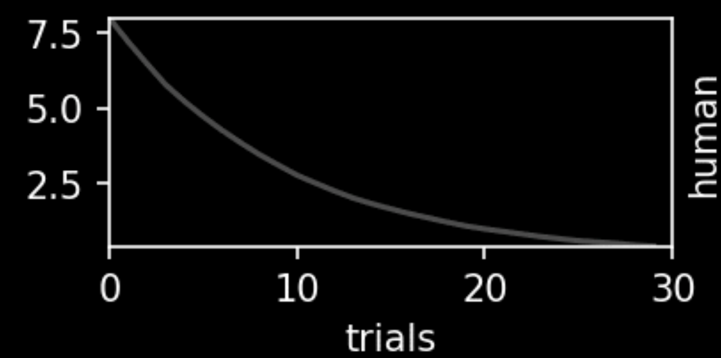
Learning to make decisions by optimization

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

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





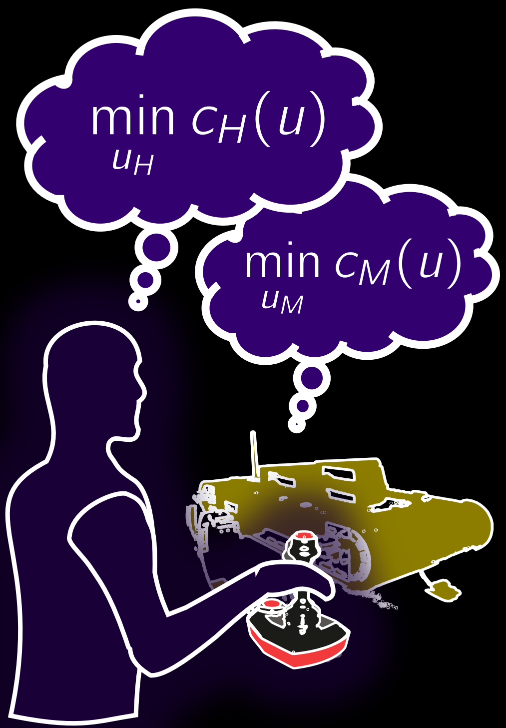


Learning as a team: coupled optimization

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

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

$$u = (u_H, u_M)$$



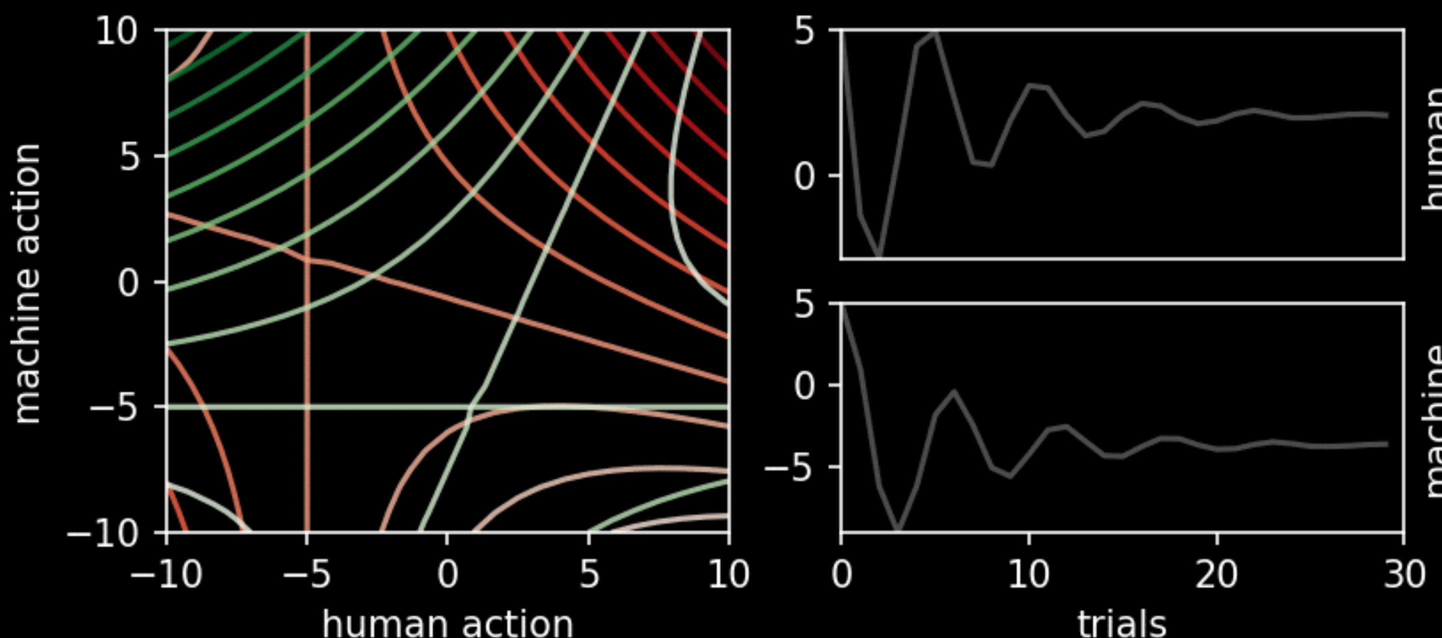
Learning as a team: coupled optimization

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

$$u = (u_H, u_M)$$

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

Stable attractor:



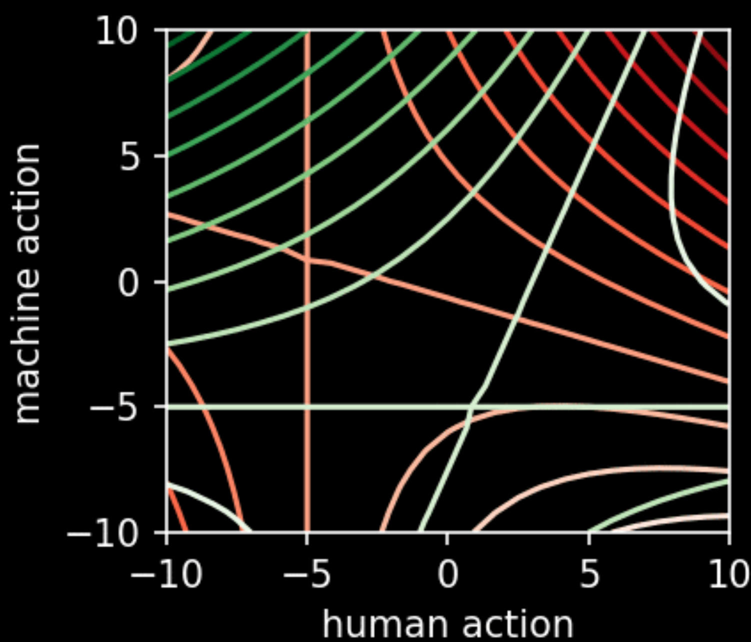
Prediction 1: *periodic orbits*

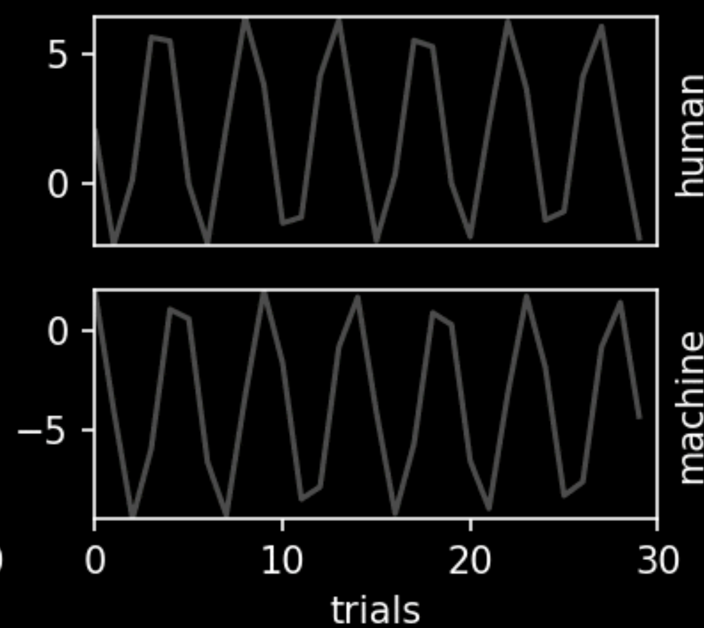
A group of optimization agents minimize their *own* cost with respect to their *own* action

$$u = (u_H, u_M)$$

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

Periodic orbit:



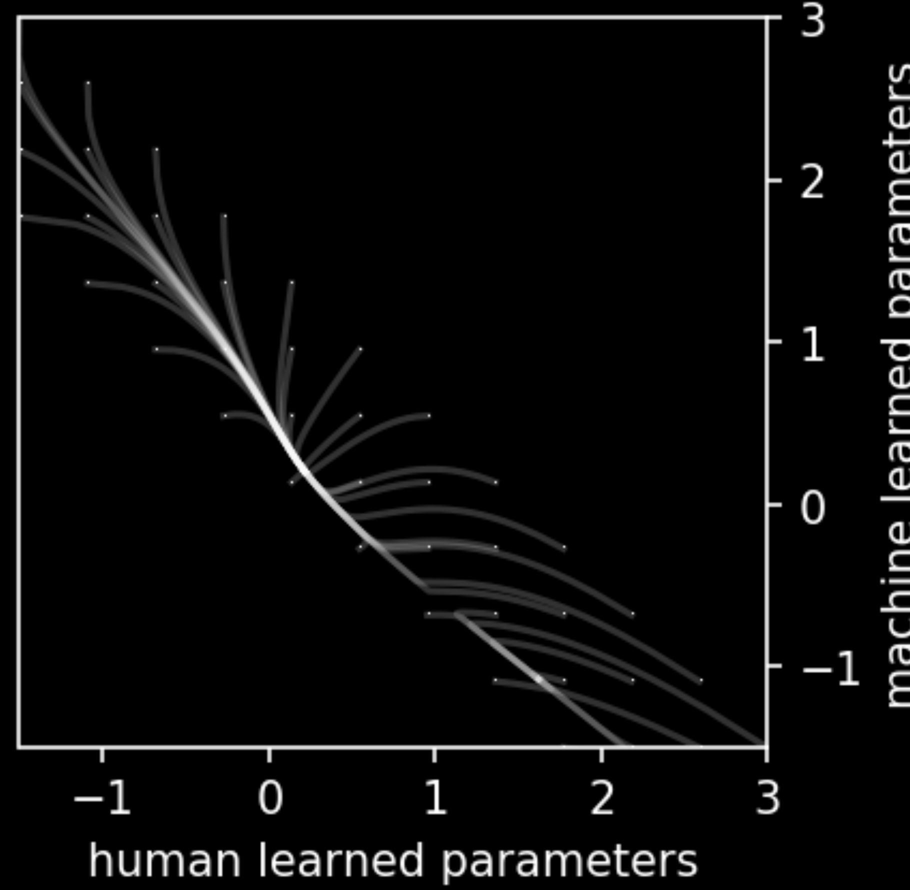


Prediction 2: *Spurious attractors*

Gradient learning dynamics do not guarantee convergence to (Nash) optimal solution.

Simultaneous learning:

$$\left.\begin{aligned} k_H^+ &= k_H - \gamma D_H c_H(k) \\ k_M^+ &= k_M - \gamma D_M c_M(k) \end{aligned}\right\}$$

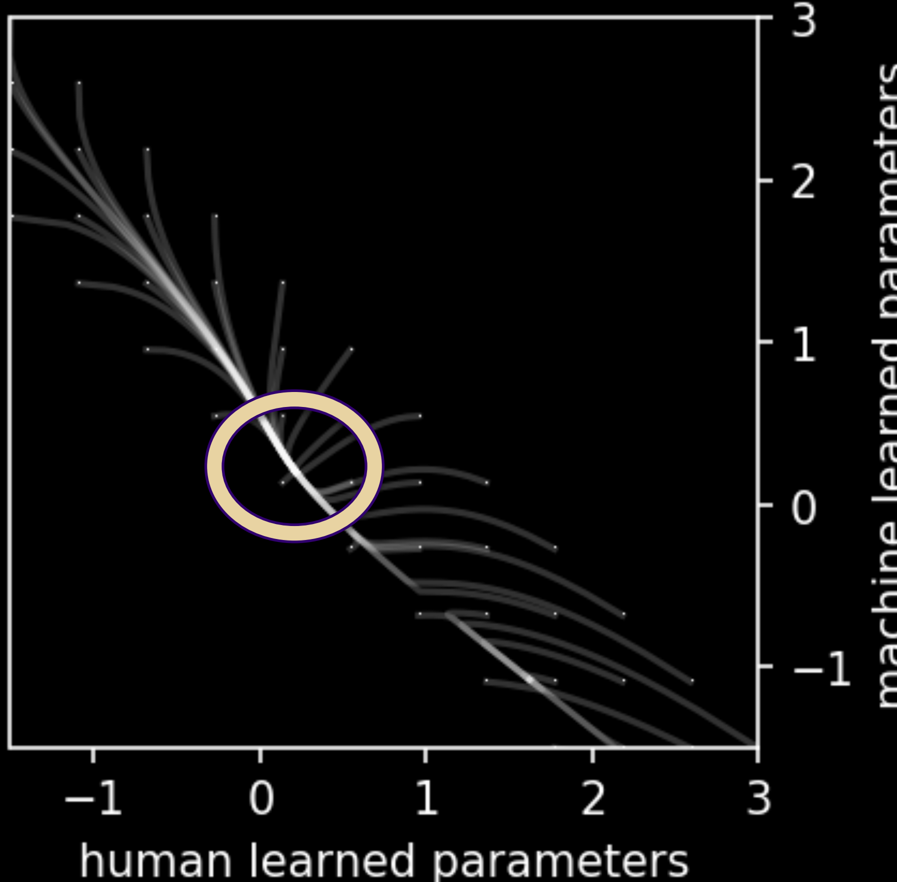


Prediction 2: *Spurious attractors*

Gradient learning dynamics do not guarantee convergence to (Nash) optimal solution.

Simultaneous learning:

$$\left.\begin{aligned} k_H^+ &= k_H - \gamma D_H c_H(k) \\ k_M^+ &= k_M - \gamma D_M c_M(k) \end{aligned}\right\}$$

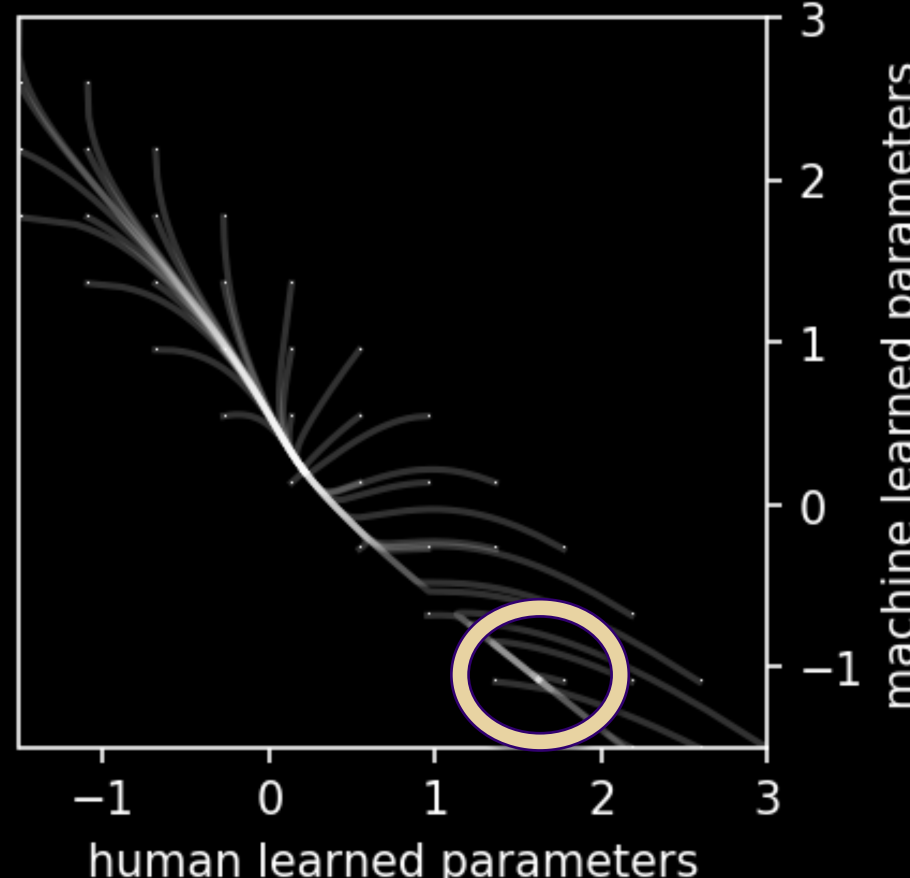


Prediction 2: *Spurious attractors*

Gradient learning dynamics do not guarantee convergence to (Nash) optimal solution.

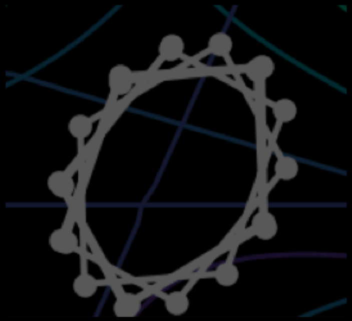
Simultaneous learning:

$$\left.\begin{aligned} k_H^+ &= k_H - \gamma D_H c_H(k) \\ k_M^+ &= k_M - \gamma D_M c_M(k) \end{aligned}\right\}$$

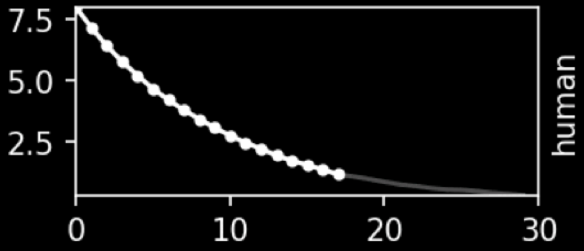


Team learning for human and machine systems

Simulation: How do the **coupling effects** of states and actions make learning difficult in *team settings*?



Experiment: How do humans effectively **learn to control** dynamical systems (individually + team settings)?



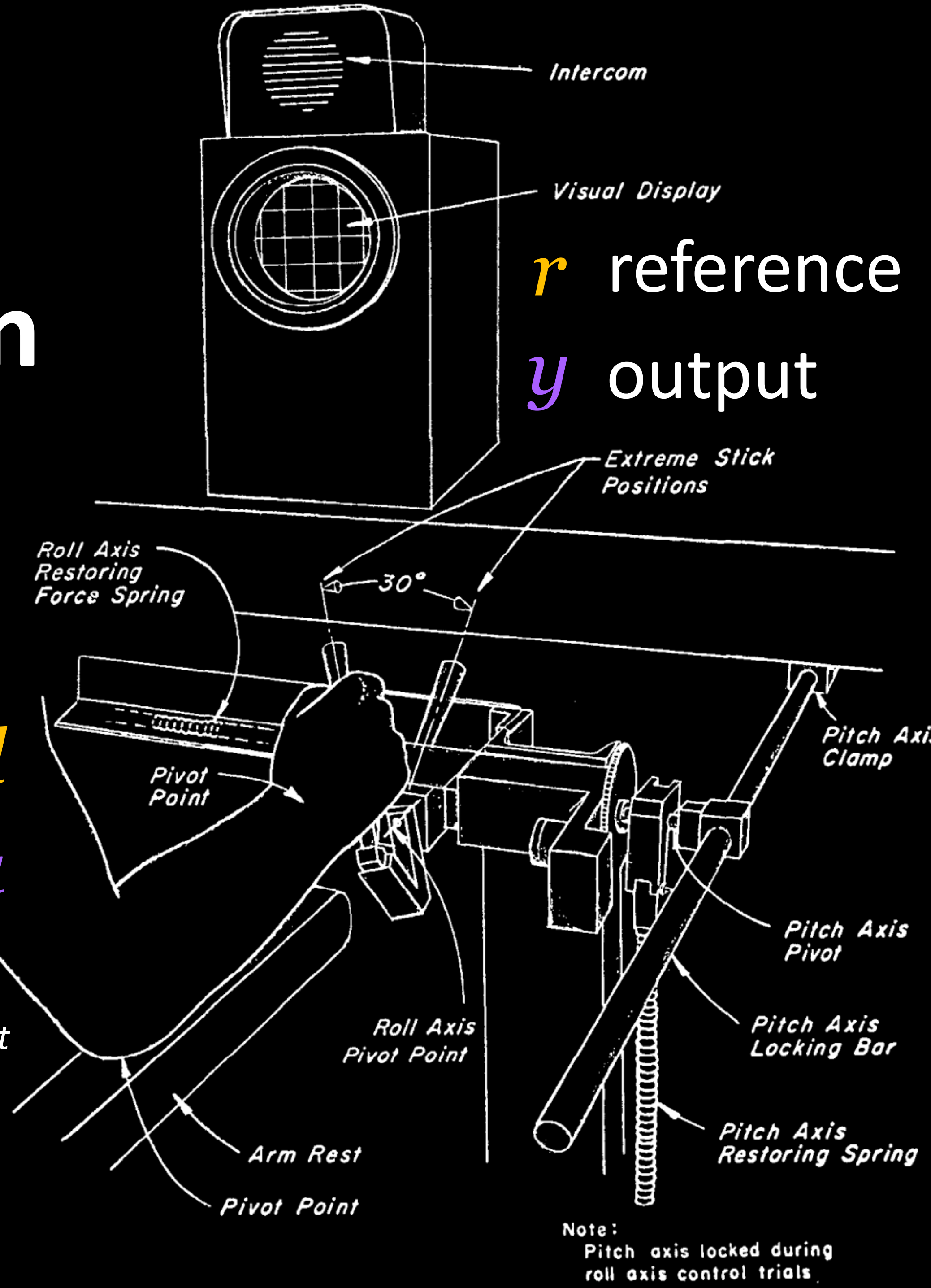
Experiment: Human teleoperation

r reference

y output

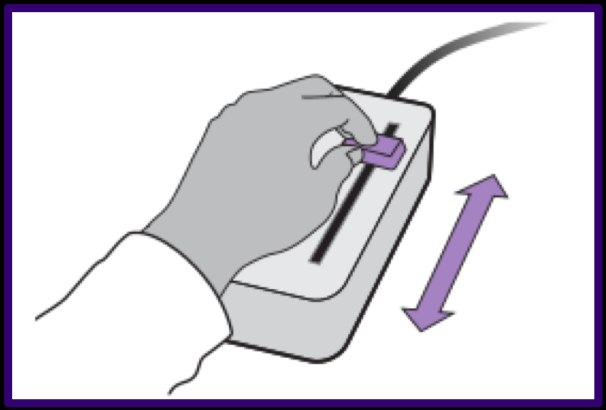
disturbance d

input u

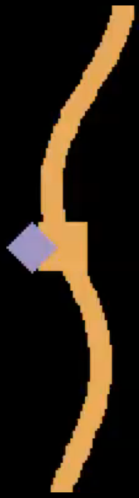


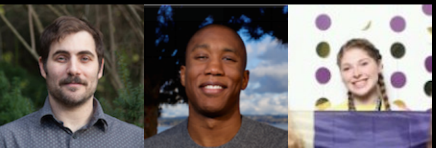
McRuer, Krendel *J Franklin Institute* 1959
The Human Operator as a Servo System Element
McRuer *Automatica* 1980
Human *Dynamics in Man-Machine Systems*

Human sensorimotor learning

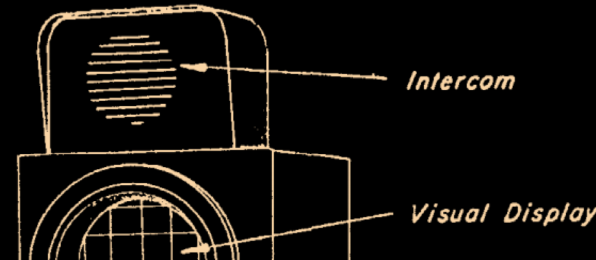


– subjects use 1-dimensional input device to control **cursor motion** to track **specified reference**

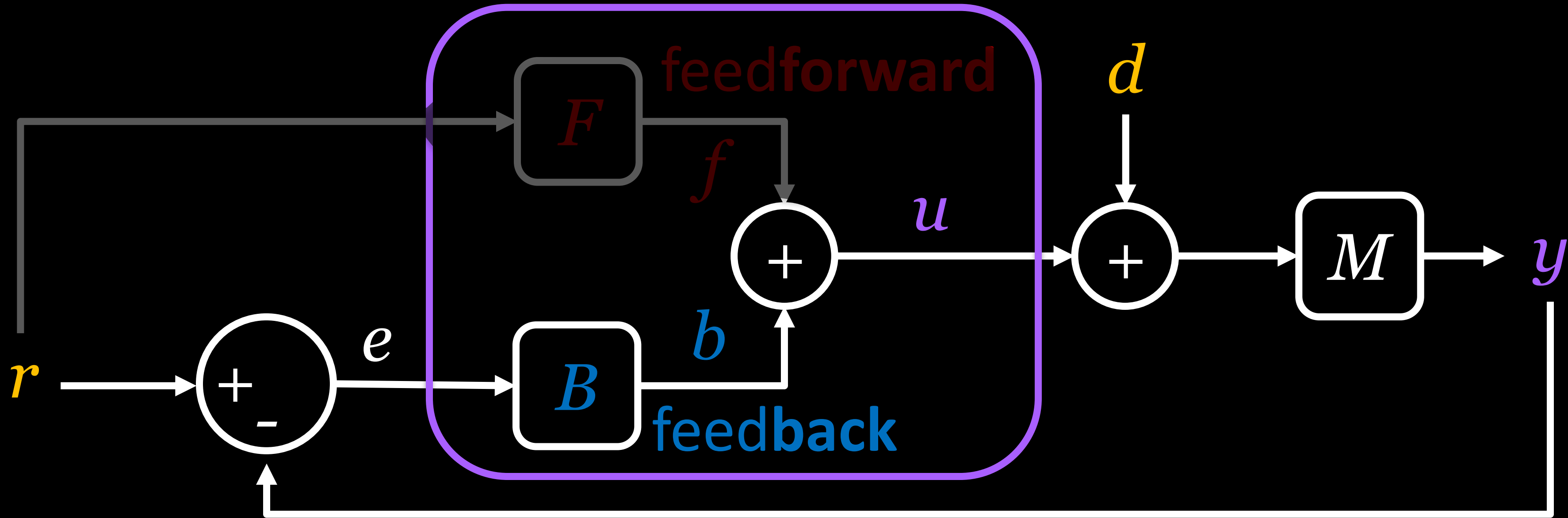




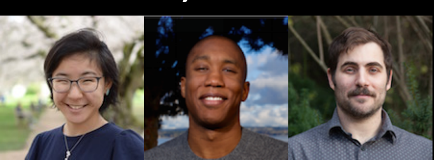
Roth, Howell, Beckwith, Burden *SPIE 2017*
Toward experimental validation of a model for human sensorimotor learning and control in teleoperation



human/machine system



McRuer, Krendel *J Franklin Institute* 1959
The Human Operator as a Servo System Element
McRuer *Automatica* 1980
Human Dynamics in Man-Machine Systems

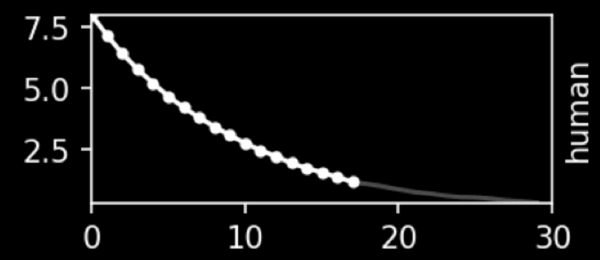


Yamagami, Howell, Roth, Burden *CPHS 2018*
Contributions of feedforward and feedback control in a manual trajectory-tracking task

Preliminary results: sensorimotor learning

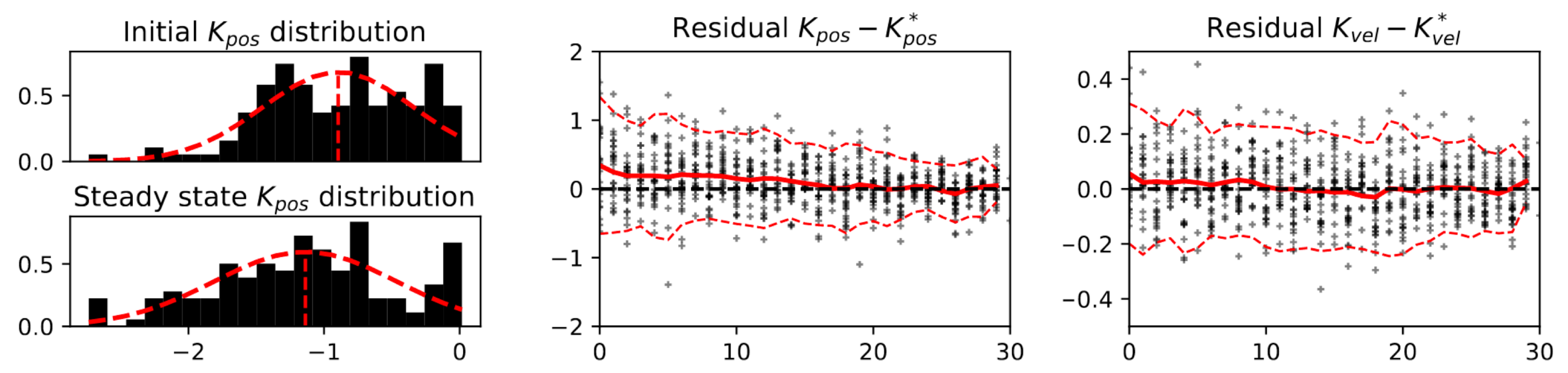
Prediction

Convergence to stationary policy



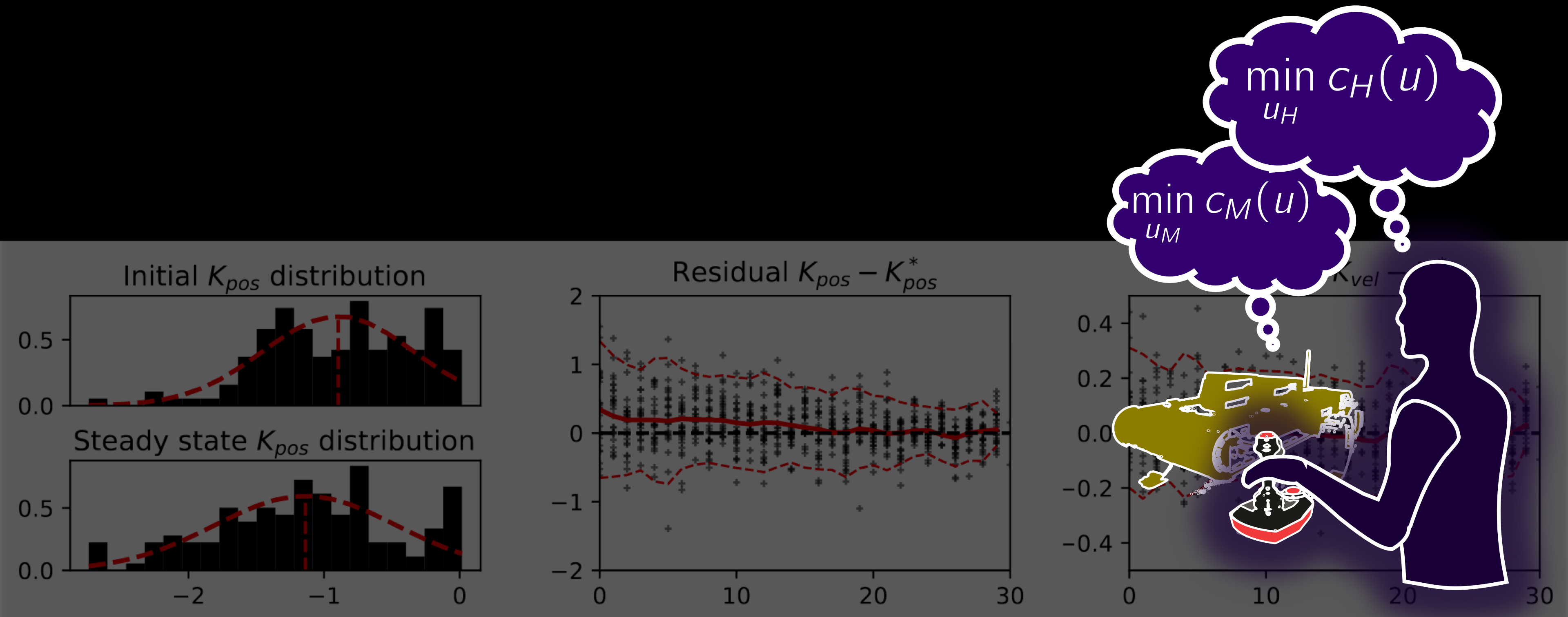
Results

Feedback gains of a second order system



Future work: sensorimotor *games*

- Coupled dynamic system via haptics
- Full information/limited information games
- When do agents play Nash?



thank you!

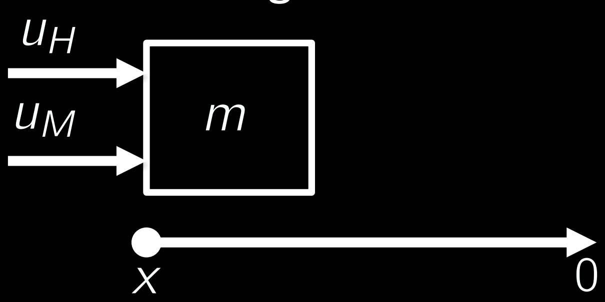
bchasnov@uw.edu
http://students.washington.edu/bchasnov/

COMPUTATIONAL
NEUROSCIENCE
CENTER

Backup slides

Simulation: learning to control a scalar system

First order integrator with two agents:



Simultaneous learning:

$$k_H^+ = k_H - \gamma D_H c_H(k)$$
$$k_M^+ = k_M - \gamma D_M c_M(k)$$

Discrete time system:

$$\begin{aligned} x^+ &= x + u_H + u_M \\ &= x + k_H x + k_M x \end{aligned}$$

With non-cooperative costs:

$$c_i(x, u) = x^2 + R_{i,H} u_H^2 + R_{i,M} u_M^2$$

A model of decision-making

A "rational" agent minimizes a cost ***c(u)*** subject to dynamical constraints

A cost can be decomposed into the two components: one that encodes the ***goal/stability***, and the other ***effort/energy***.

$$c_H(u) \equiv \underbrace{c_{H,Q}(x)}_{\text{state}} + \underbrace{c_{H,R}(u)}_{\text{control}}$$