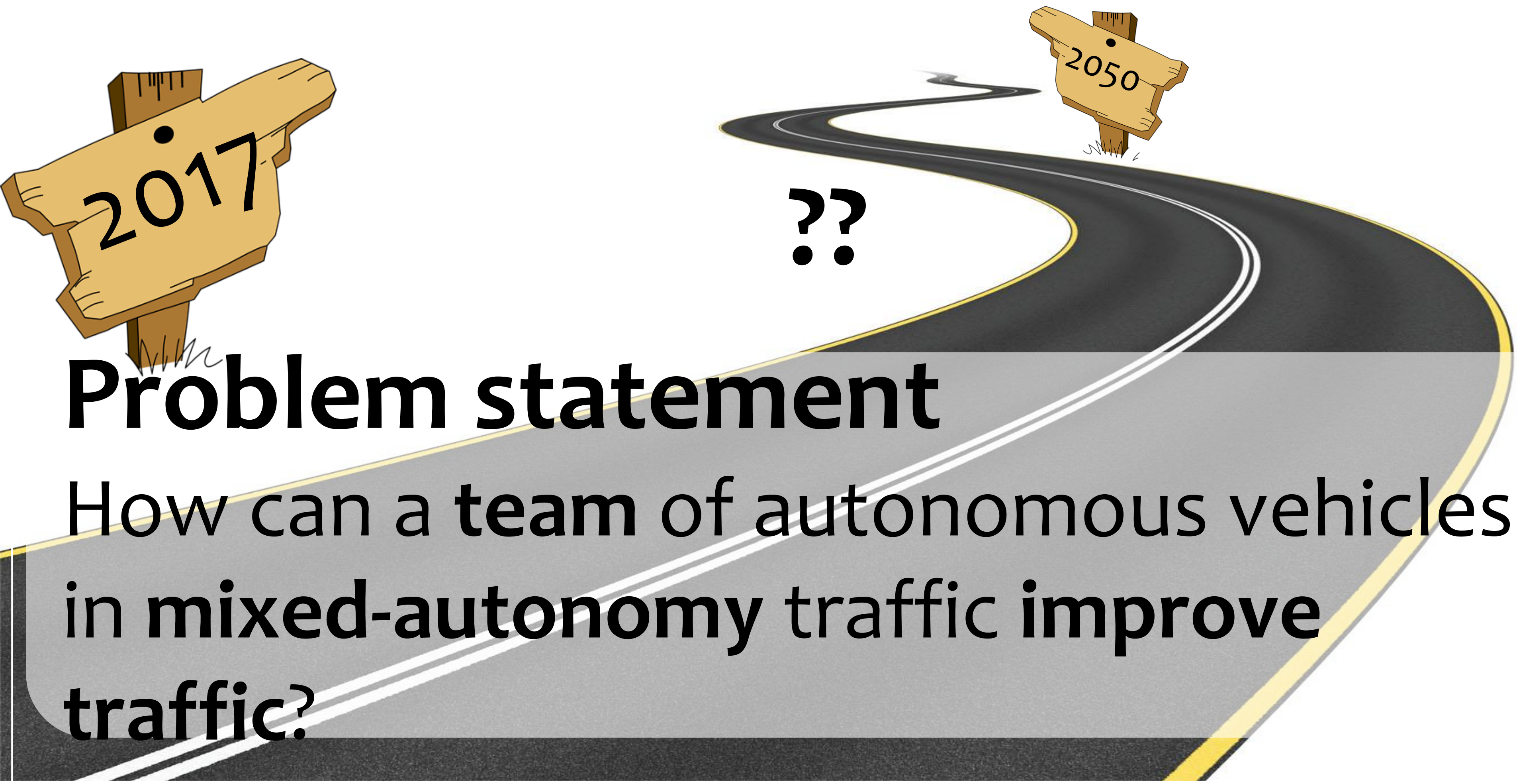
Reinforcement Learning for Mixed-Autonomy Traffic

Cathy Wu, Eugene Vinitsky,
Aboudy Kreidieh, Alexandre Bayen

FORCES Meeting
UC Berkeley
2017-08-23

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

What happens **between 0% and 100%** penetration of autonomous vehicles?

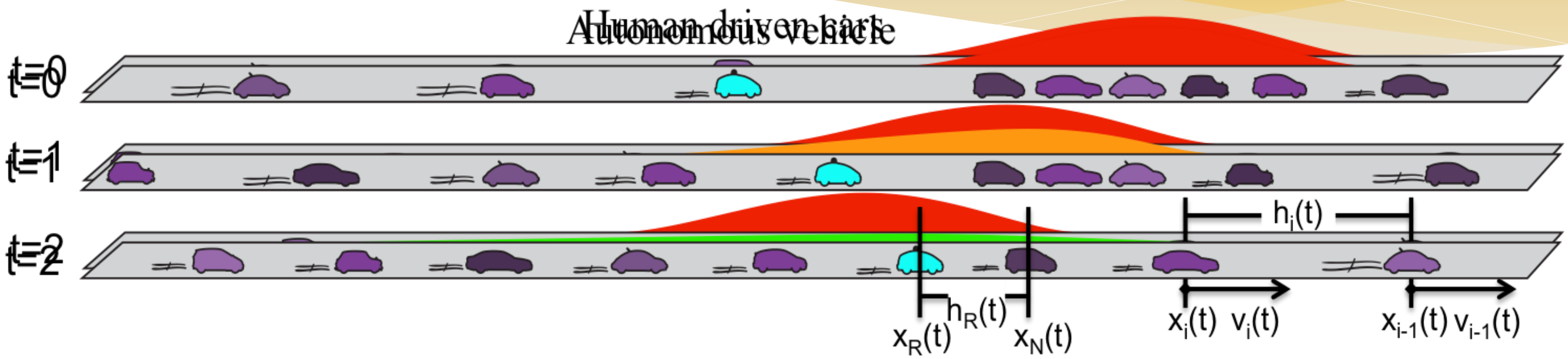


Problem statement

How can a **team** of autonomous vehicles in **mixed-autonomy** traffic **improve traffic**?

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Problem setup



Linear dynamics: $\ddot{\tilde{x}}_i = k_p(\tilde{x}_{i-1} - \tilde{x}_i) + k_d(\dot{\tilde{x}}_{i-1} - \dot{\tilde{x}}_i) - k_v(\dot{\tilde{x}}_i)$

Robot control: $\ddot{\tilde{x}}_i = k_{p_r}(\tilde{x}_{i-1} - \tilde{x}_i) + k_{d_r}(\dot{\tilde{x}}_{i-1} - \dot{\tilde{x}}_i) - k_{v_r}(\dot{\tilde{x}}_i)$

[Cui, et al., IV, 2017; Wu et al., In submission, 2017]

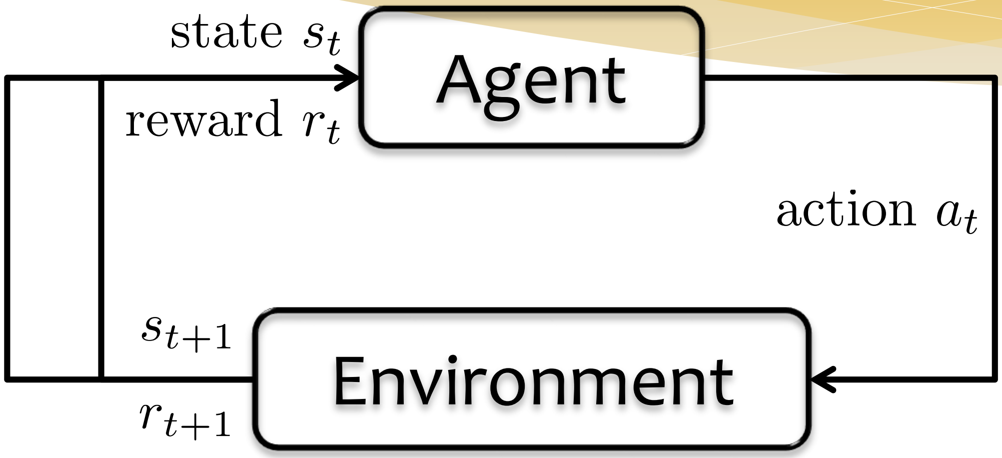
Alternatively: $\ddot{\tilde{x}}_i = f(\tilde{x}_{i-1}, \tilde{x}_i, \tilde{x}_{i-1}, \dot{\tilde{x}}_i)$

[Wu et al., In submission, 2017]

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Image credit: Florian Brown-Altvater

Reinforcement learning



Goal: learn policy $\pi : S \rightarrow A$ to maximize reward

Example:

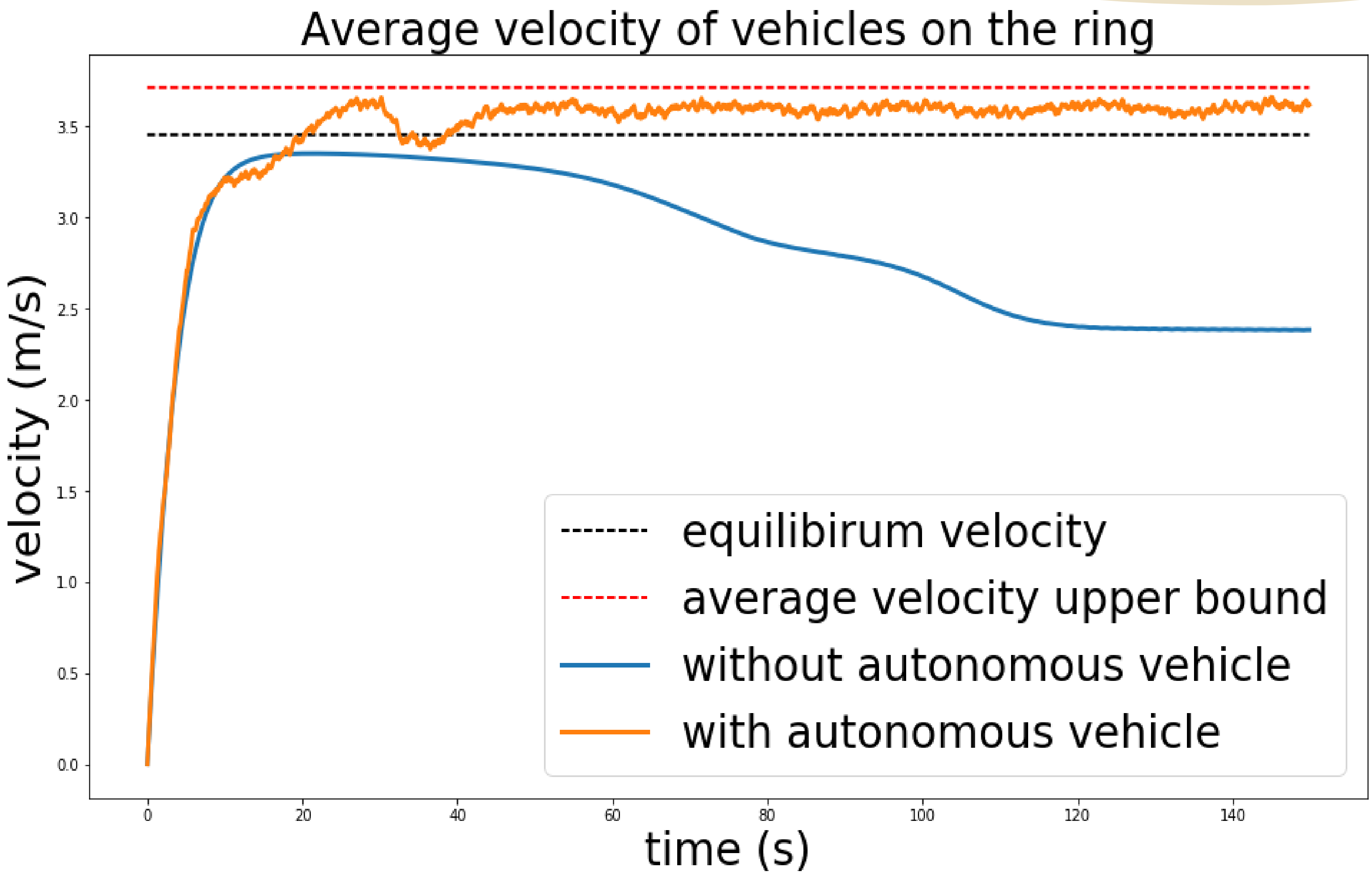
max –(total fuel consumption) +
–λ(total travel time to destination)
s.t. no collisions

[Sutton98]

Single-lane ring road
1 RL, 21 human

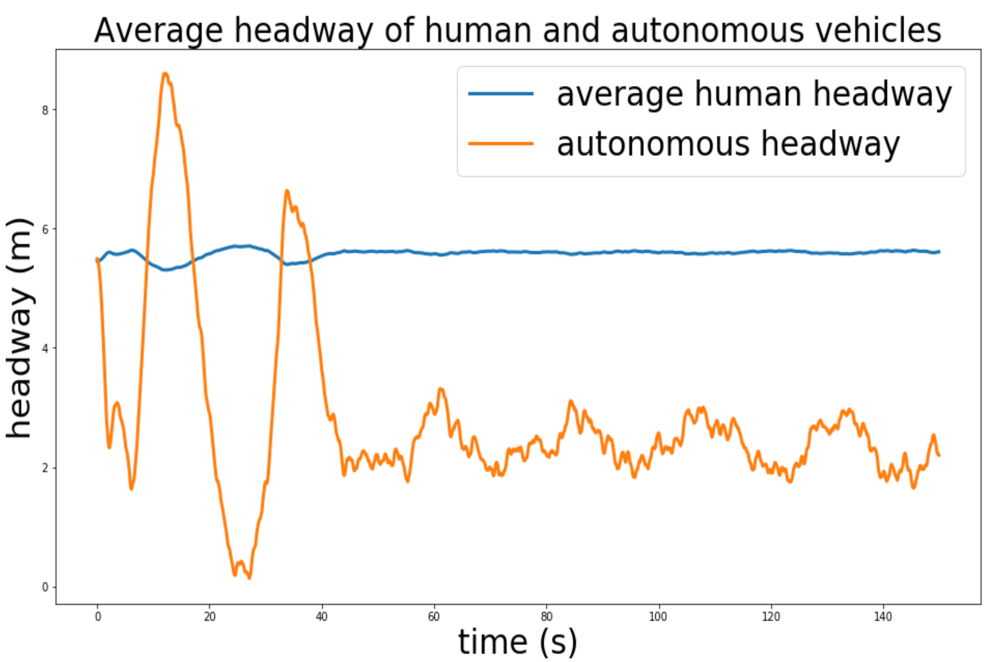
Reward function: deviation from target velocity

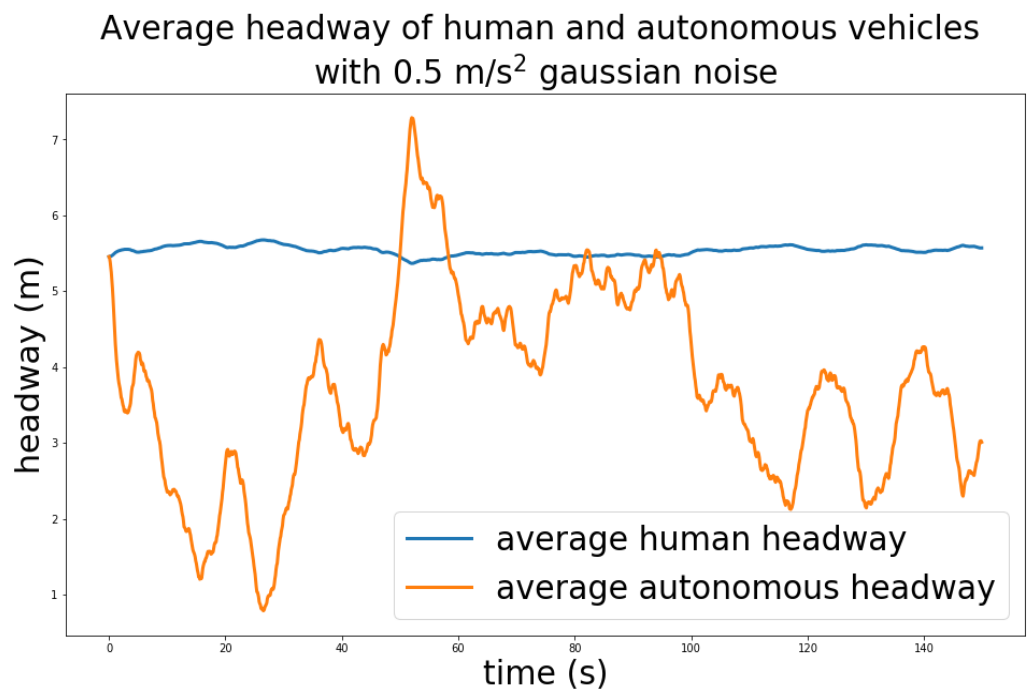
Velocity Improvement!



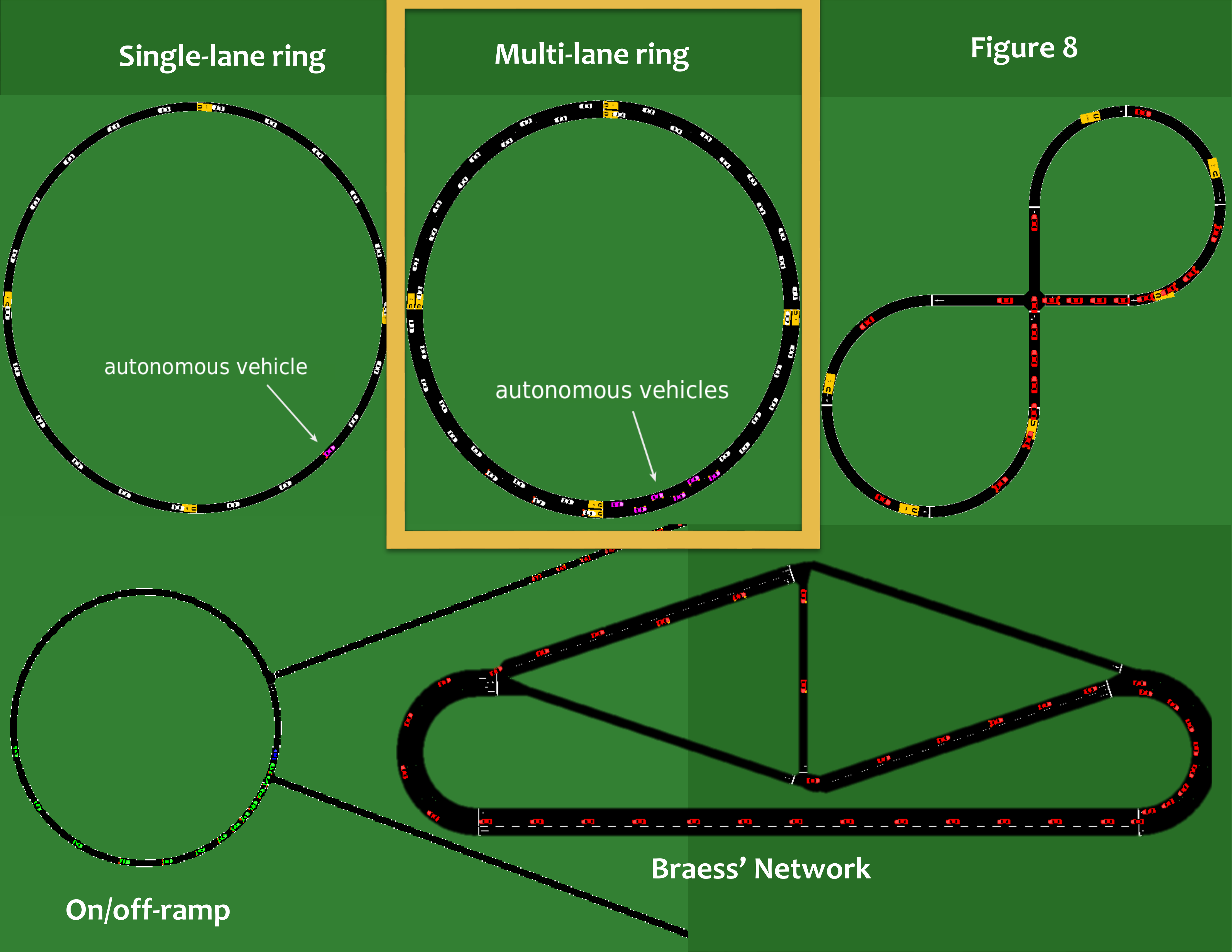
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Tailgating speedup



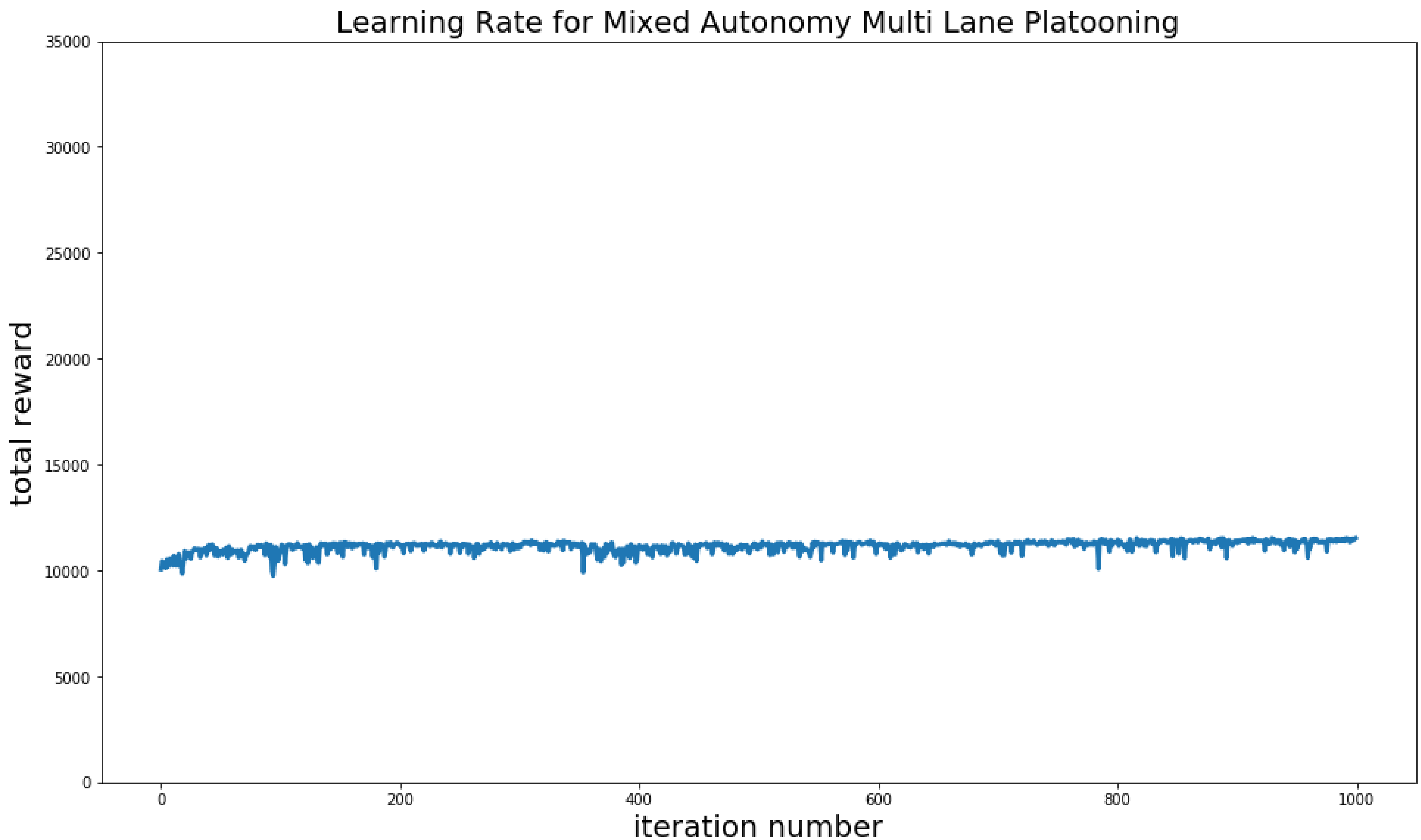


FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS



Single-lane ring
Multi-lane ring
Figure 8
autonomous vehicle
autonomous vehicles
On/off-ramp
Braess' Network

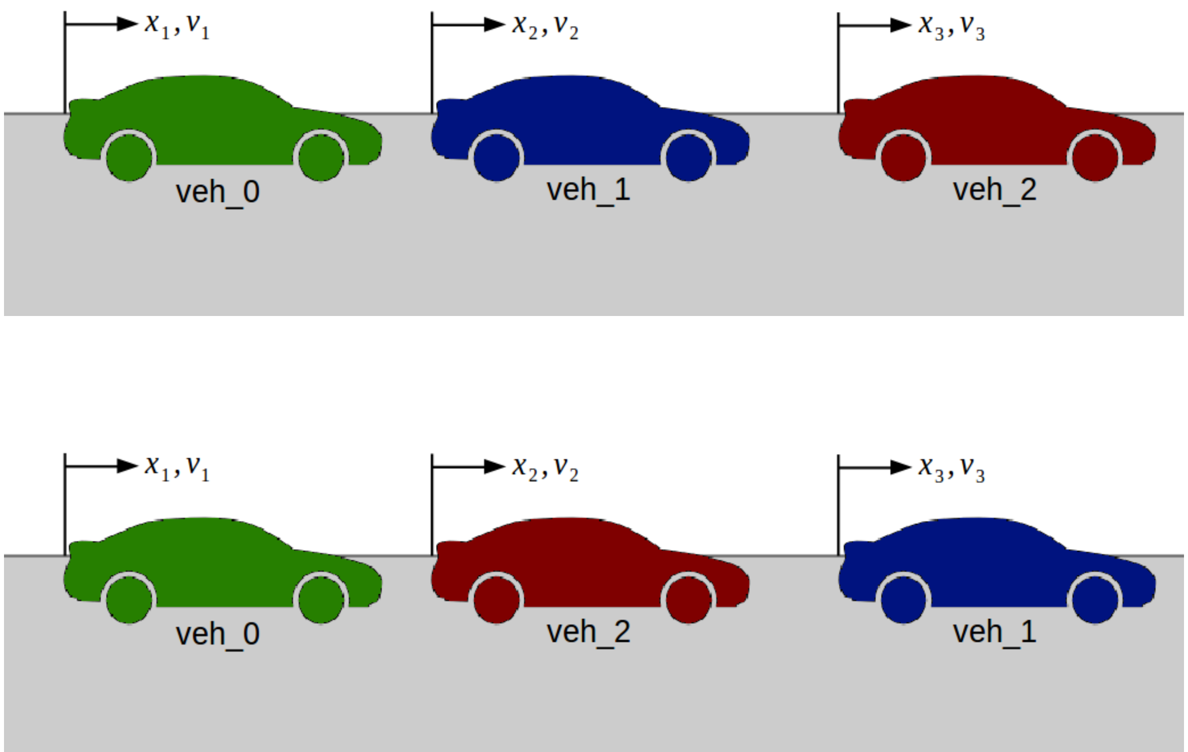
Multi-lane Platooning



FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

MDPs with symmetry

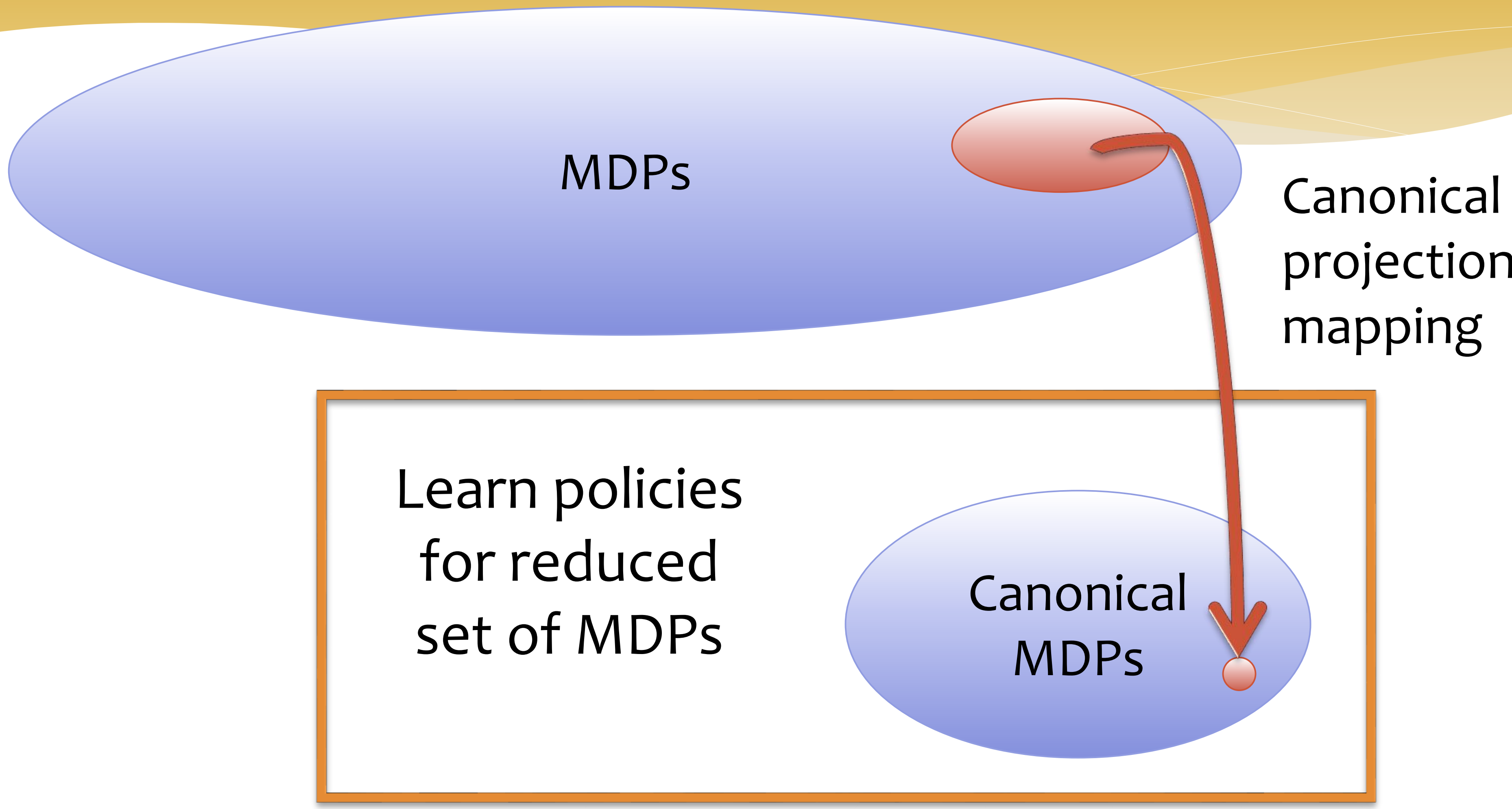
Challenge: poor sample efficiency due to combinatorial explosion in state and action spaces.



$$s = (x_1, v_1, \underline{x_2, v_2}, \underline{x_3, v_3})$$

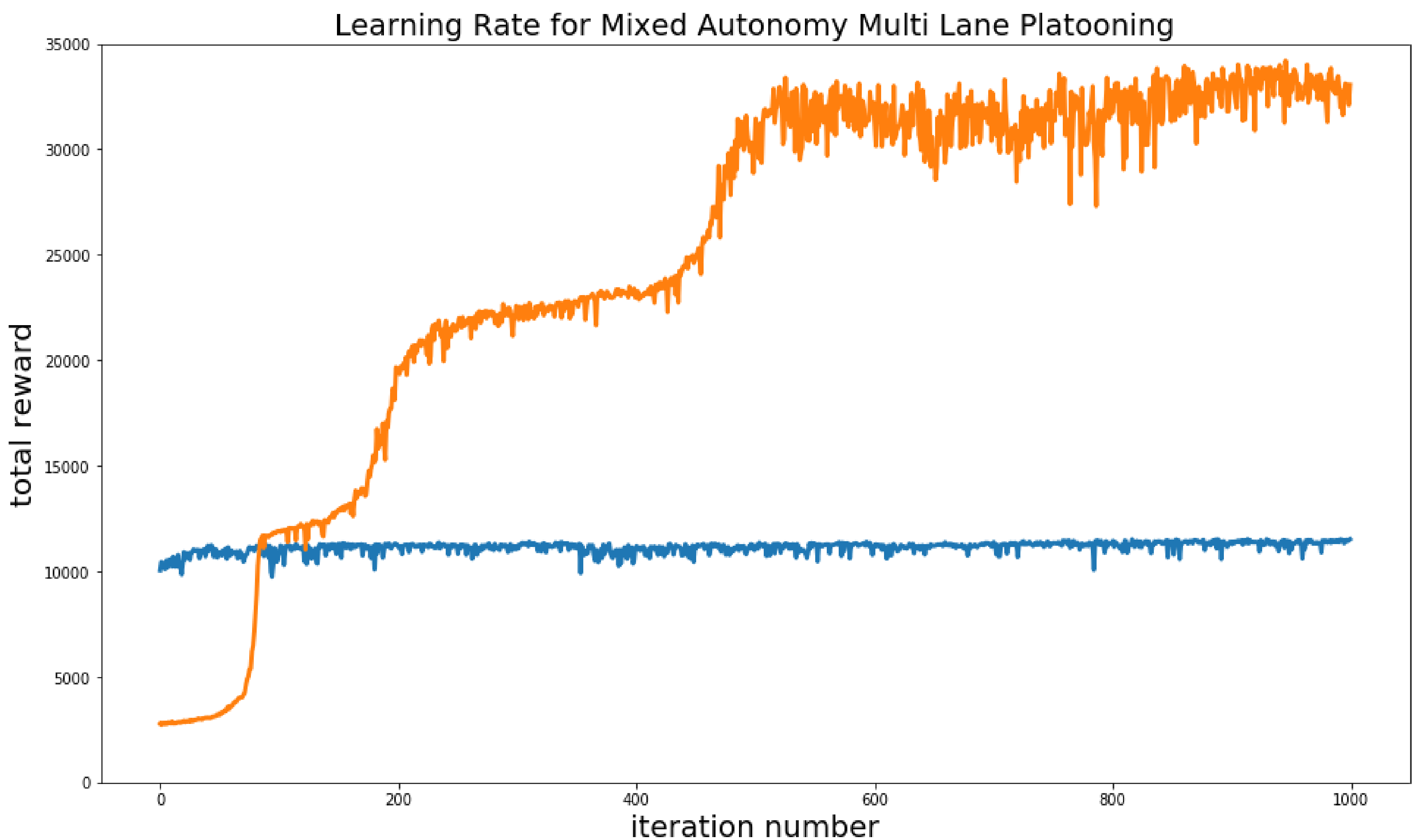
$$s = (x_1, v_1, x_3, v_3, x_2, v_2)$$

MDPs with symmetry



Proposed canonical projection mapping:
Selective sorting of the state space

Canonical Projection Result



FORCES
FOUNDATIONS OF RESILIENT
CYBER-PHYSICAL SYSTEMS

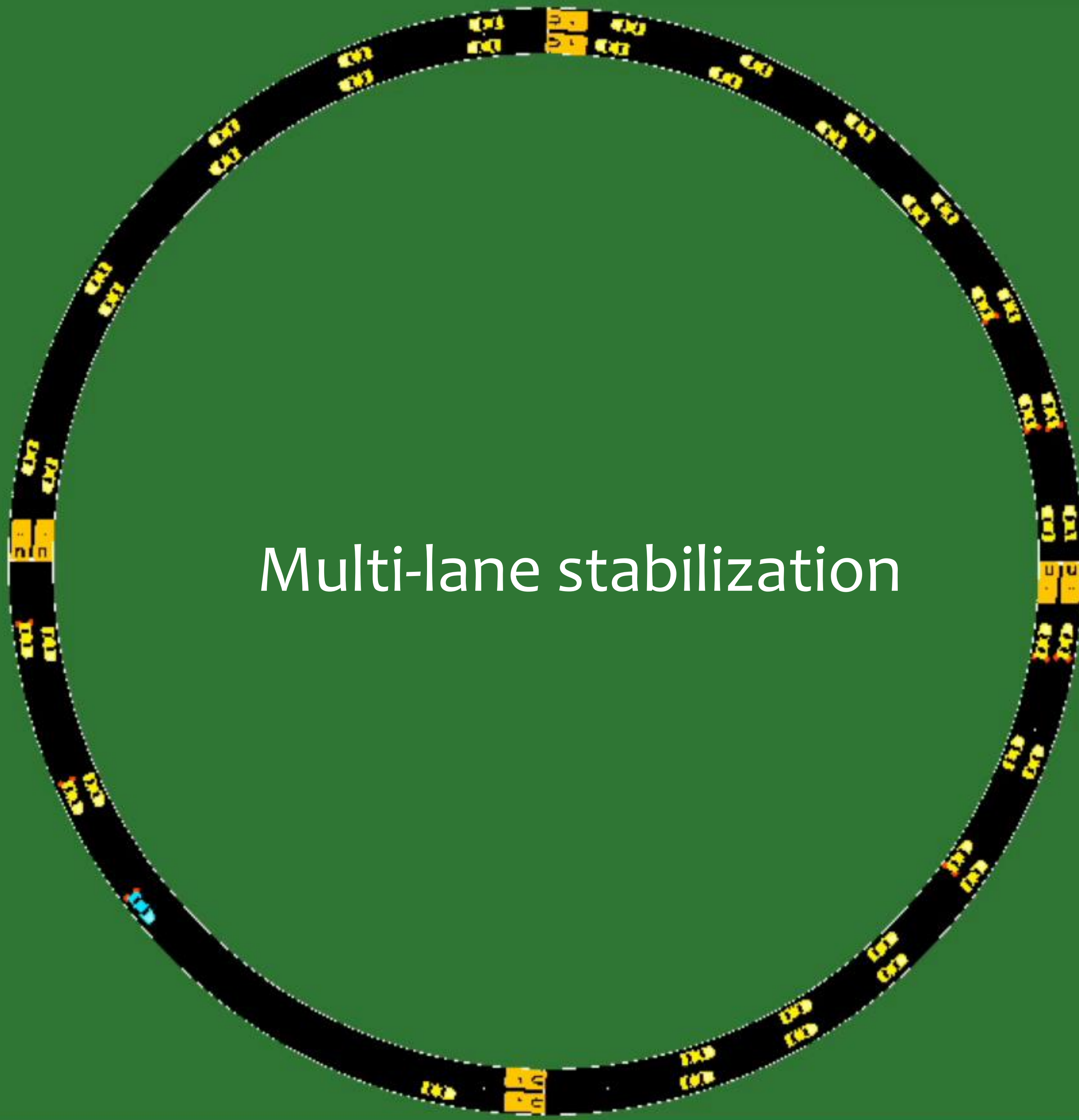

Multi-lane stabilization

Figure 8 (1 RL)

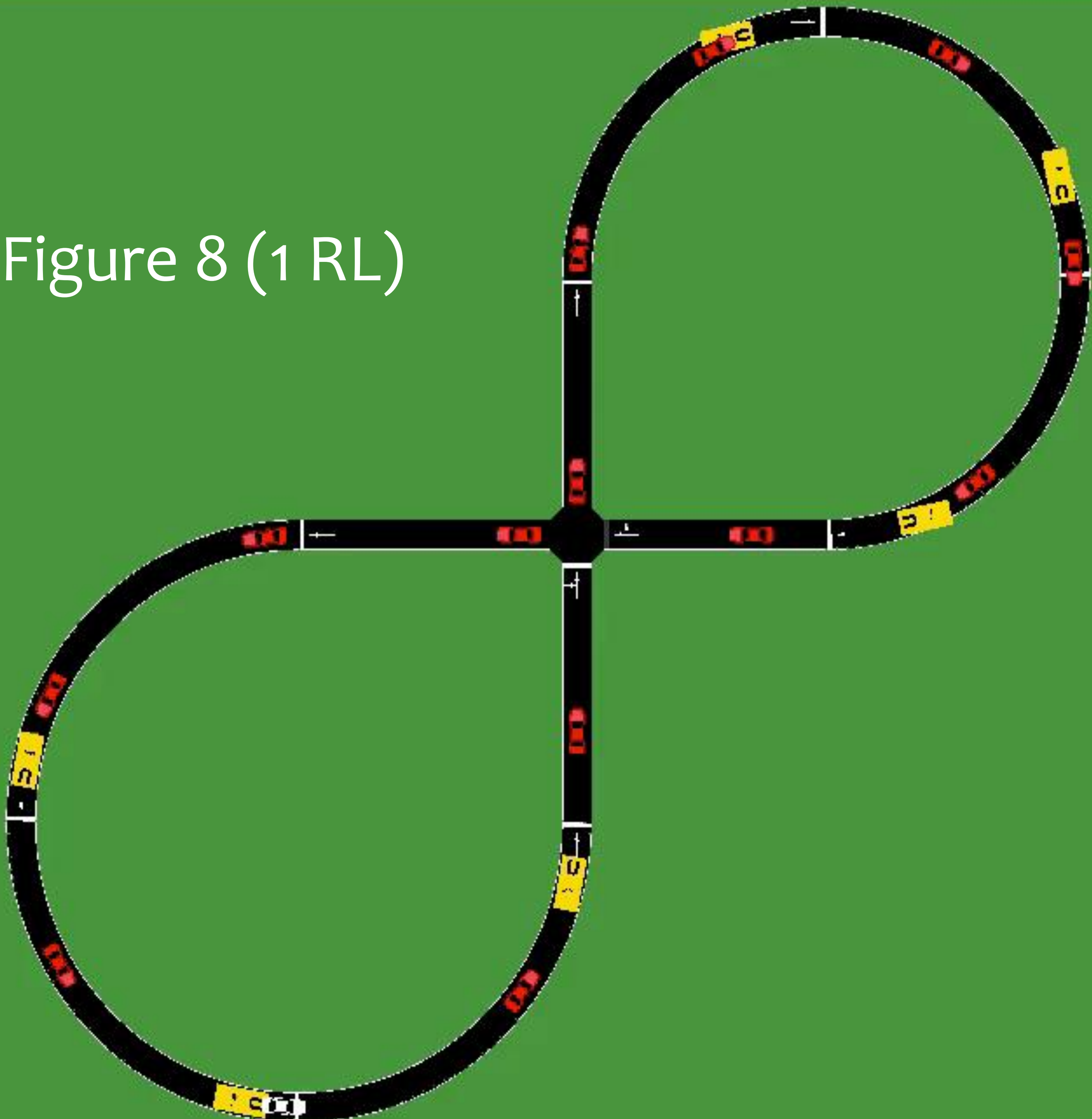
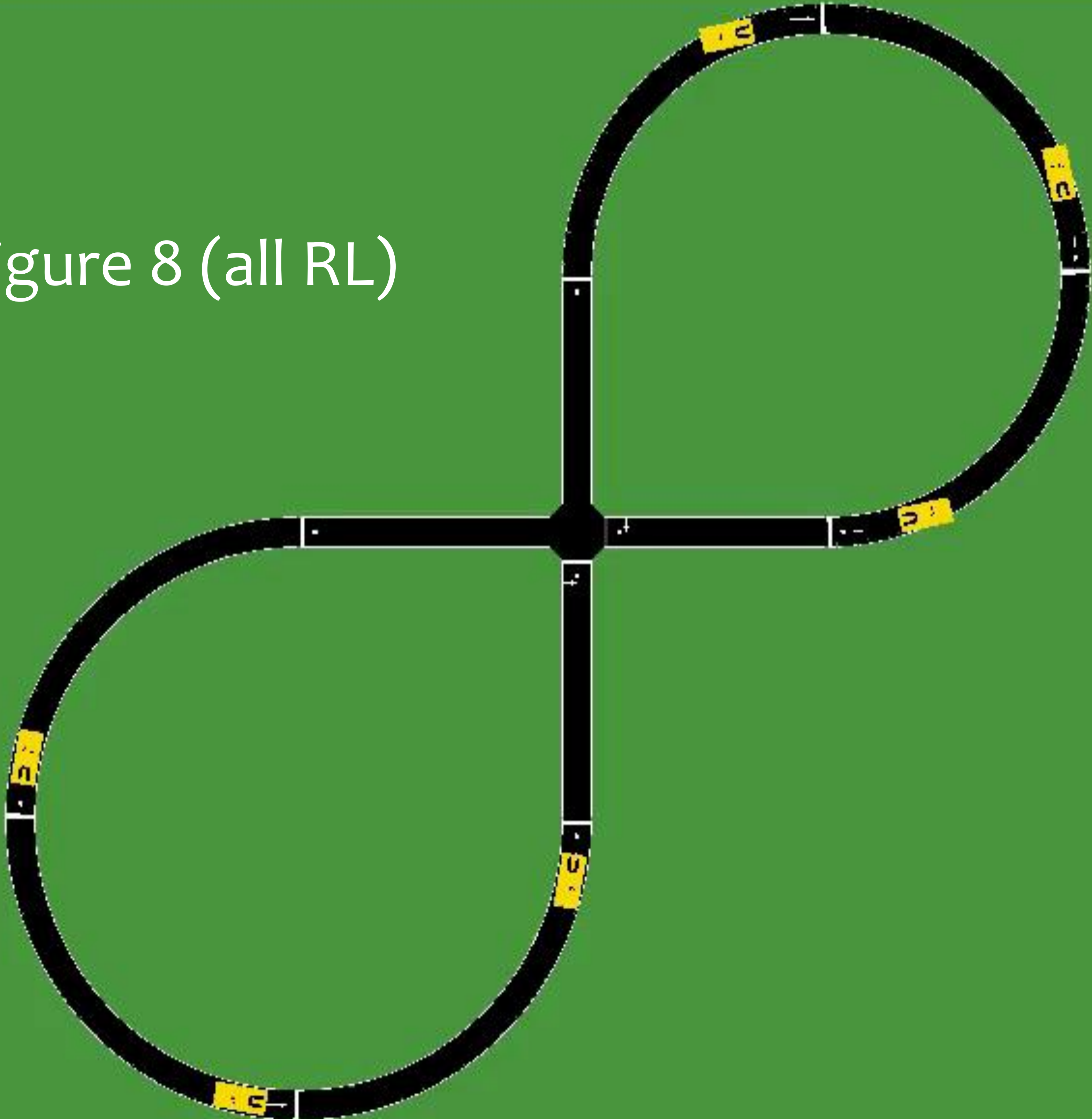
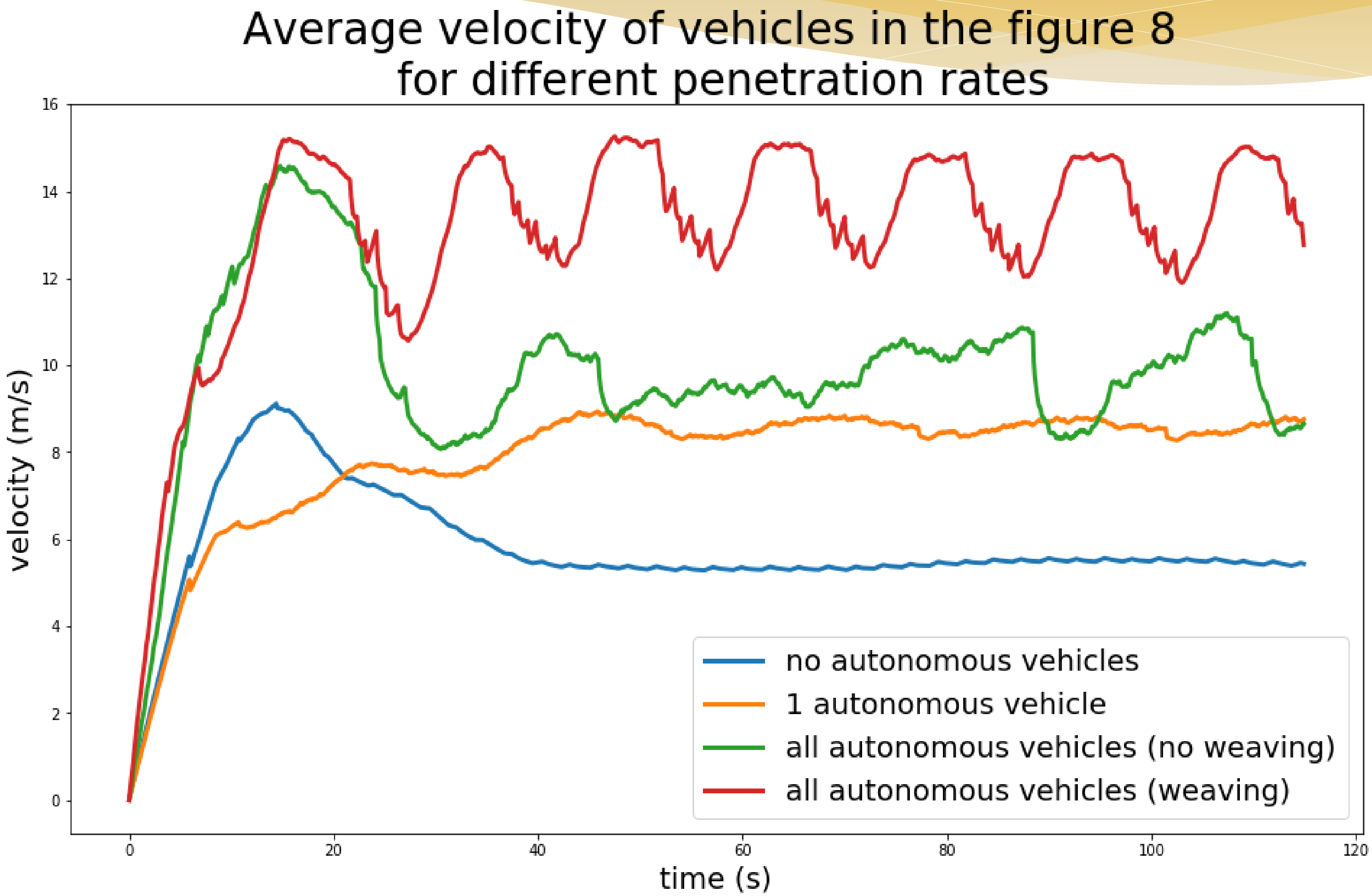


Figure 8 (all RL)



Mixed-Autonomous Comparison



FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Current work and opportunities

Control ←→ reinforcement learning

* Learning minimal controllers
* Proving optimality for learned controllers

Advancing reinforcement learning research

* Scaling up multi-agent learning algorithms
* Representation learning for sample efficient learning in MDPs with symmetr
  and locality

Advancing intelligent transportation research

* Supporting large-scale networks
* Learning for mixed-control (intersections, speed limits, AVs).

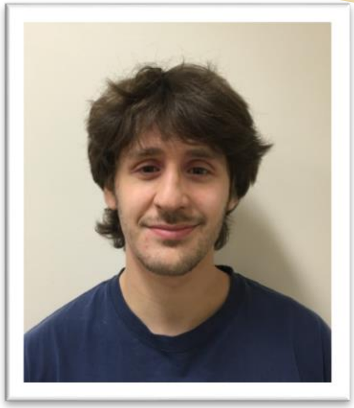
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Learning-traffic team

Graduate researchers

PI

Cathy Wu
Team lead, EECS

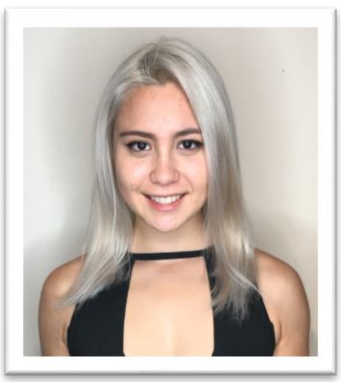


Eugene Vinitsky
MechE

Aboudy Kreidieh
CEE

Alex Bayen
EECS/CEE

Undergraduate researchers



Leah Dickstein
EECS

Kanaad Parvate
EECS

Nishant Kheterpal
EECS

Acknowledgements:
Joseph Wu

FORCES
FOUNDATIONS OF RESILIENT
CYBER-PHYSICAL SYSTEMS