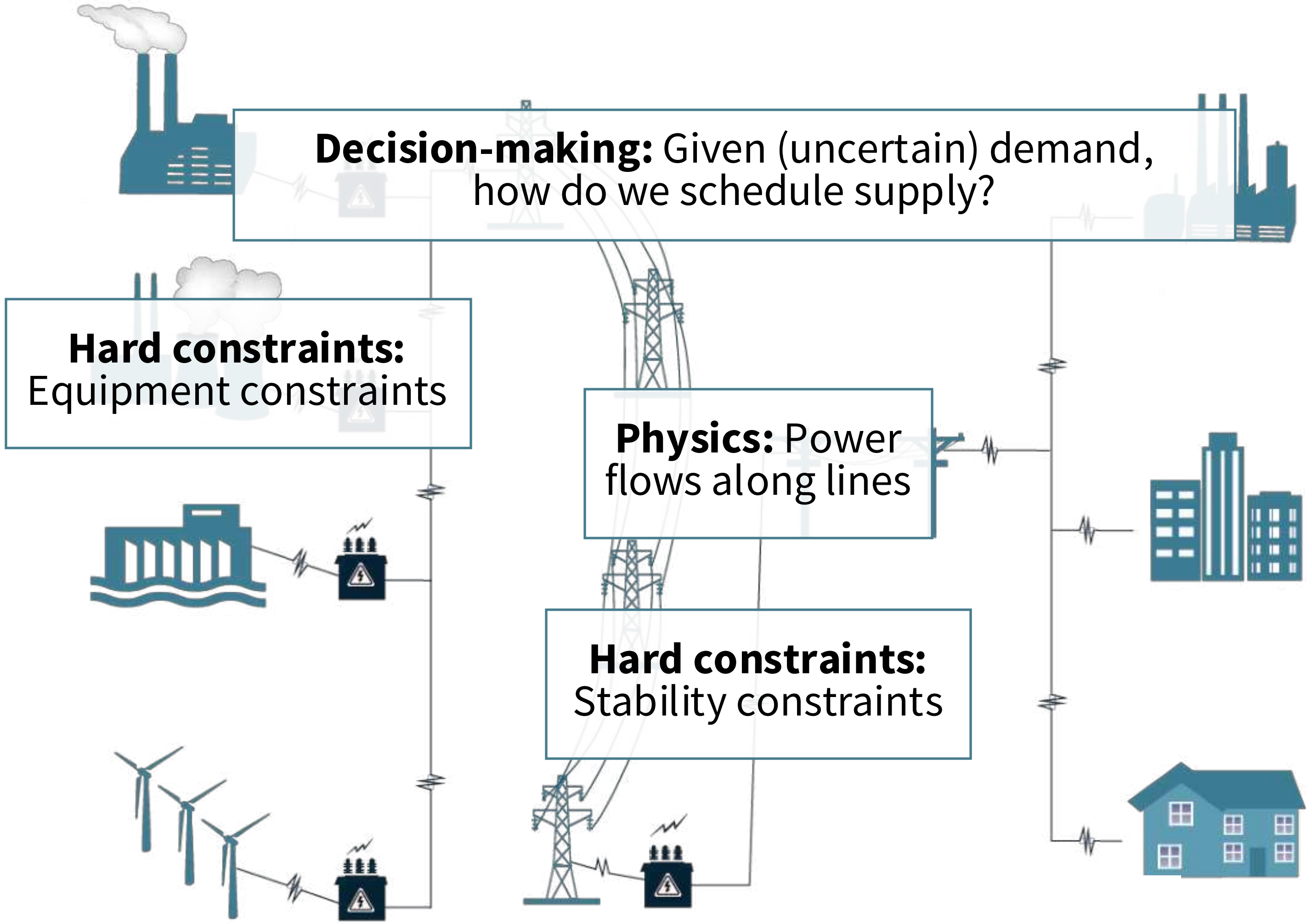
Optimization-in-the-loop ML for power grid operations

Priya L. Donti

Assistant Professor, MIT EECS and LIDS

Reconciling physics and hard constraints with fast and scalable computation



Trad. optimization & control
- Satisfies (many) constraints
- Struggles with speed / scale

Machine learning (ML)
- Fast and scalable
- Struggles with constraints

Figure adapted from: US Congressional Budget Office

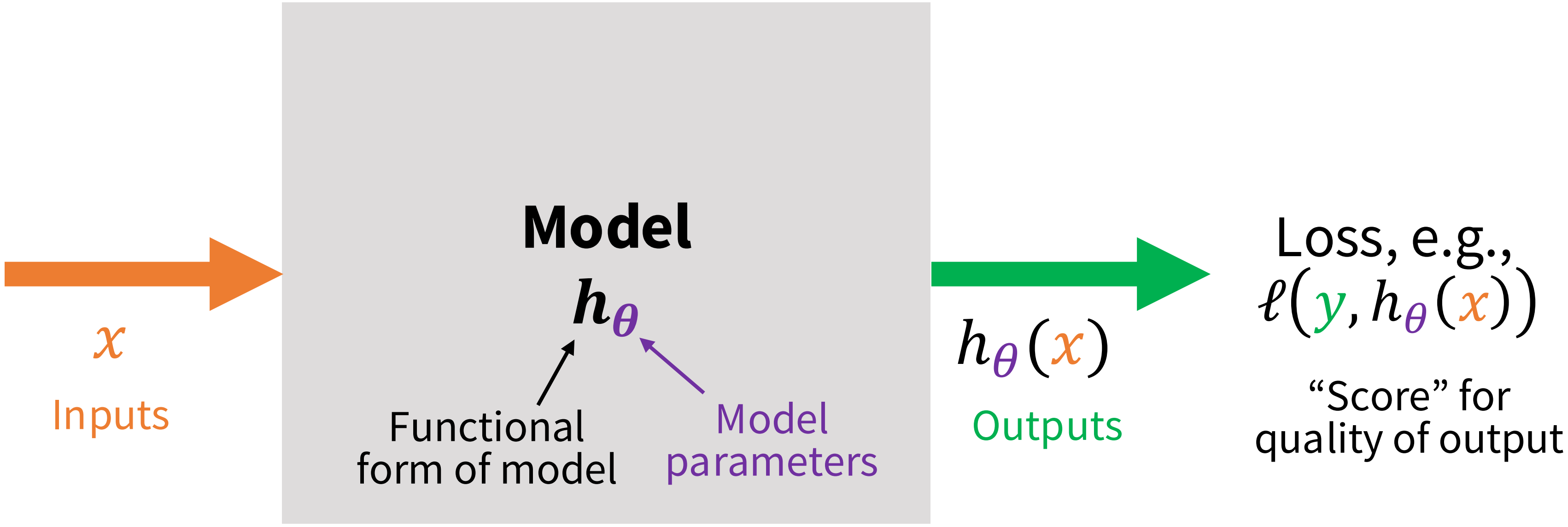
Today's talk

- **Optimization-in-the-loop ML:** Framework for developing ML methods incorporating knowledge of physics/hard constraints, via optimization problems
- **Application to single-agent control:** Optimization-in-the-loop reinforcement learning with enforcement of asymptotic stability or state/action constraints
- **Future directions in distributed control:** Bridging optimization-in-the-loop learning, multi-agent reinforcement learning, and decomposition methods

Today's talk

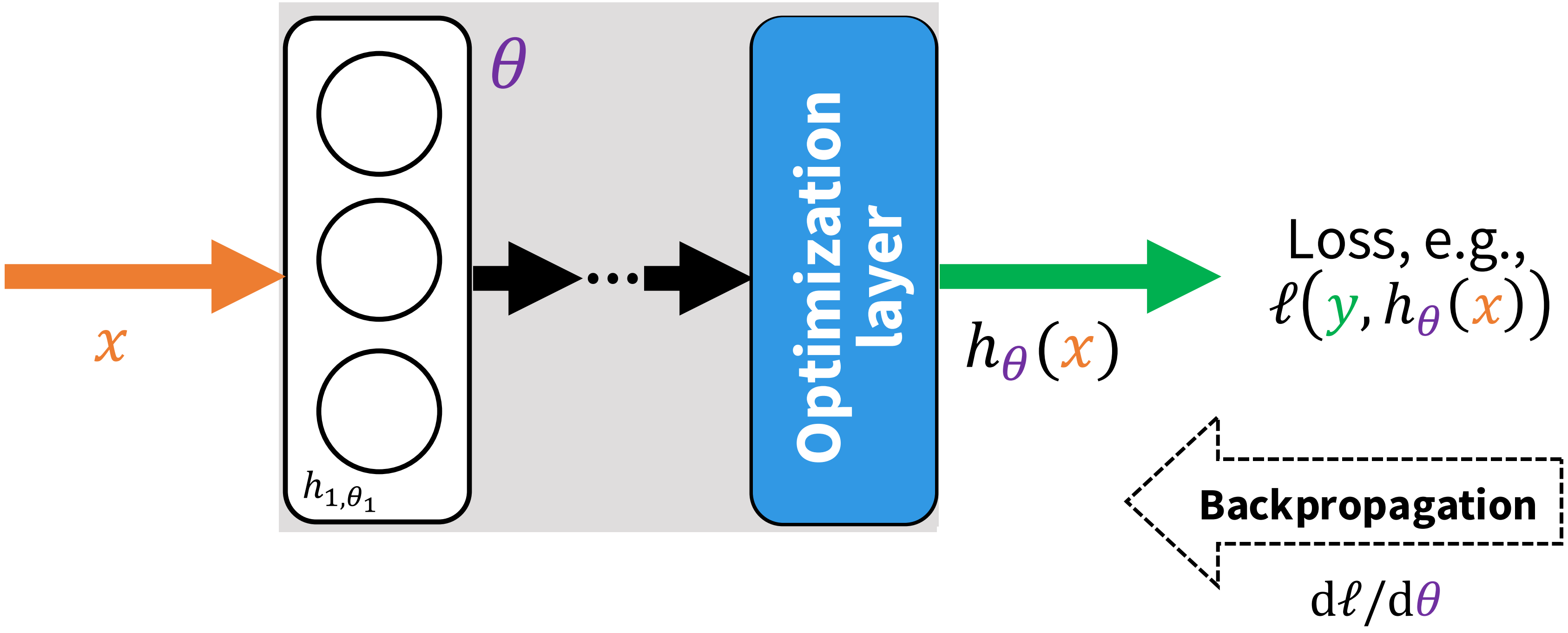
- **Optimization-in-the-loop ML:** Framework for developing ML methods incorporating knowledge of physics/hard constraints, via optimization problems
- **Application to single-agent control:** Optimization-in-the-loop reinforcement learning with enforcement of asymptotic stability or state/action constraints
- **Future directions in distributed control:** Bridging optimization-in-the-loop learning, multi-agent reinforcement learning, and decomposition methods

Deep learning is differentiable function composition



Deep learning is differentiable function composition

- Neural network h_θ = composition of nonlinear, parameterized functions (*layers*)
- Update parameters θ to minimize loss ℓ using gradients from *backpropagation*
- All components (layers and loss) **must be differentiable**



Differentiating through optimization problems (broader literature: implicit layers)

Insight: Apply implicit function theorem to equilibrium or optimality conditions (and use computational tricks to efficiently compute $\mathrm{d}\ell/\mathrm{d}\theta$ directly)

OptNet: Differentiable Optimization as a Layer in Neural Networks

Task-based End-to-end Model Learning in Stochastic Optimization

SATNet: Bridging deep learning and logical reasoning using a differentiable

laximization

Powerful toolkit for optimization-in-the-loop ML (and differentiable programming more broadly)

End-to-End Diffe
for Learning and Control

Layers

Filipe de A. Belbute-Peres
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
filiped@cs.cmu.edu

Kevin A. Smith
Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139
k2smith@mit.edu

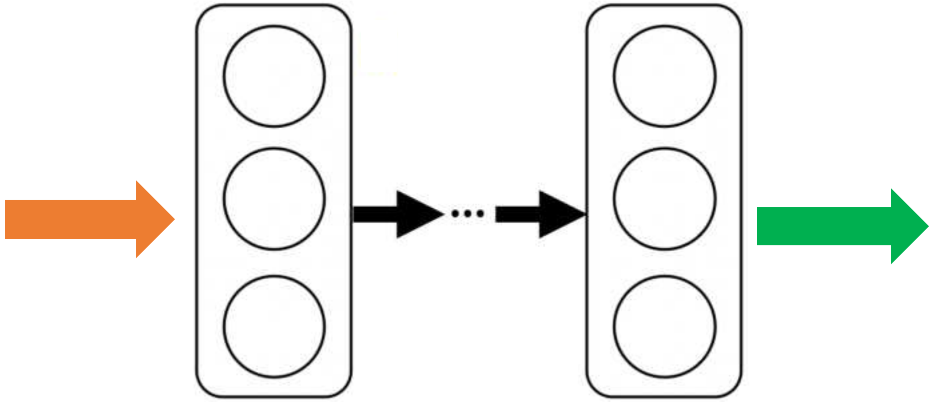
Neural Ordinary Differential Equations

Ricky T. Q. Chen*, Yulia Rubanova*, Jesse Bettencourt*, David Duvenaud
University of Toronto, Vector Institute
{rtqichen, rubanova, jessebett, duvenaud}@cs.toronto.edu

Today's talk

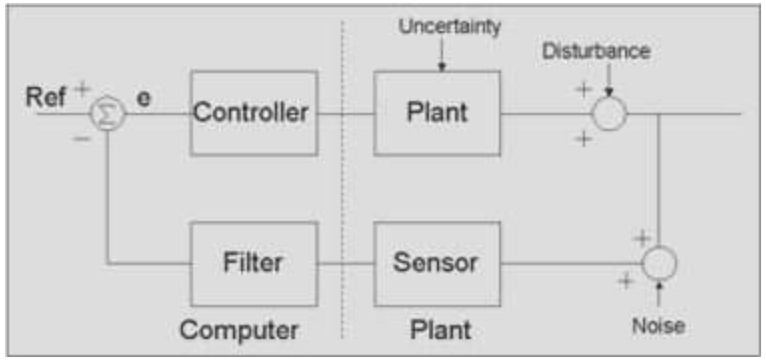
- **Optimization-in-the-loop ML:** Framework for developing ML methods incorporating knowledge of physics/hard constraints, via optimization problems
- **Application to single-agent control:** Optimization-in-the-loop reinforcement learning with enforcement of asymptotic stability or state/action constraints
- **Future directions in distributed control:** Bridging optimization-in-the-loop learning, multi-agent reinforcement learning, and decomposition methods

Deep reinforcement learning vs. robust control



Deep RL

Pro: Expressive, well-performing policies
Con: Potential (catastrophic) failures



Robust control

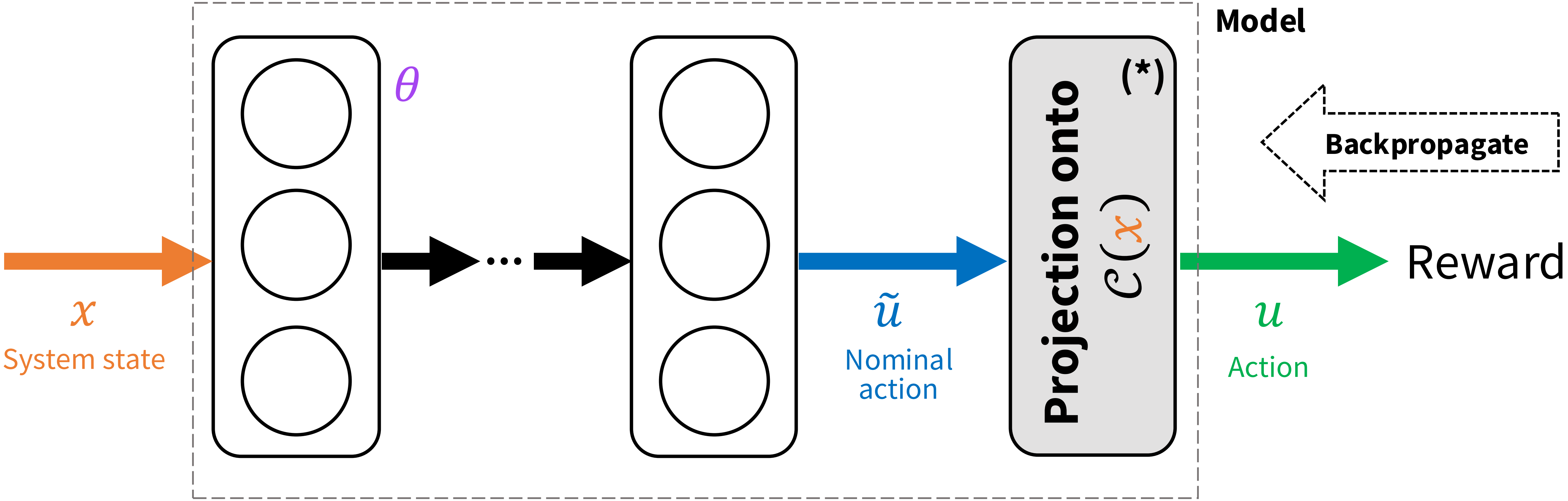
Pro: Provable stability guarantees
Con: Simple policies (e.g., linear)

Can we improve performance while still guaranteeing stability?

Priya L. Donti, Melrose Roderick, Mahyar Fazlyab, and J. Zico Kolter. "Enforcing robust control guarantees within neural network policies." *International Conference on Learning Representations (ICLR) 2021.*

Differentiable projection onto stabilizing actions

Deep learning-based policy with **provable robustness guarantees** (even for a randomly initialized neural network), trainable using reinforcement learning



Priya L. Donti, Melrose Roderick, Mahyar Fazlyab, and J. Zico Kolter. "Enforcing robust control guarantees within neural network policies." *International Conference on Learning Representations (ICLR) 2021.*

Finding a set of stabilizing actions (example)

Insight: Find a set of actions that are guaranteed to satisfy relevant Lyapunov stability criteria at a given state, even under worst-case conditions

Given the following (from robust control)**:**

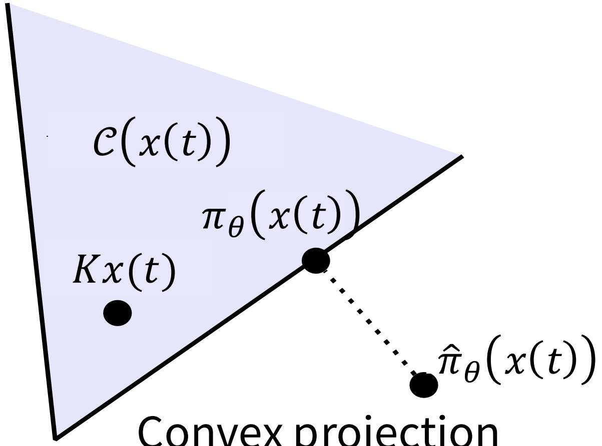
- Uncertainty model: e.g., $\dot{x}(t) \in Ax(t) + Bu(t) + Gw(t)$ s.t. $\|w(t)\|_2 \leq \|Cx(t) + Du(t)\|_2$
- Lyapunov function V obtained via robust control synthesis
- Exponential stability criterion: $\dot{V}\big(x(t)\big) \leq -\alpha V\big(x(t)\big), \forall x \neq 0$

Find: For given x, set of actions satisfying exponential stability criterion even in worst case

$$\mathcal{C}(x) \equiv \left\{ u : \left(\sup_{w\,:\,\|w\|_2 \leq \|Cx+Du\|_2} \dot{V}(x) \right) \leq -\alpha V(x) \right\}$$

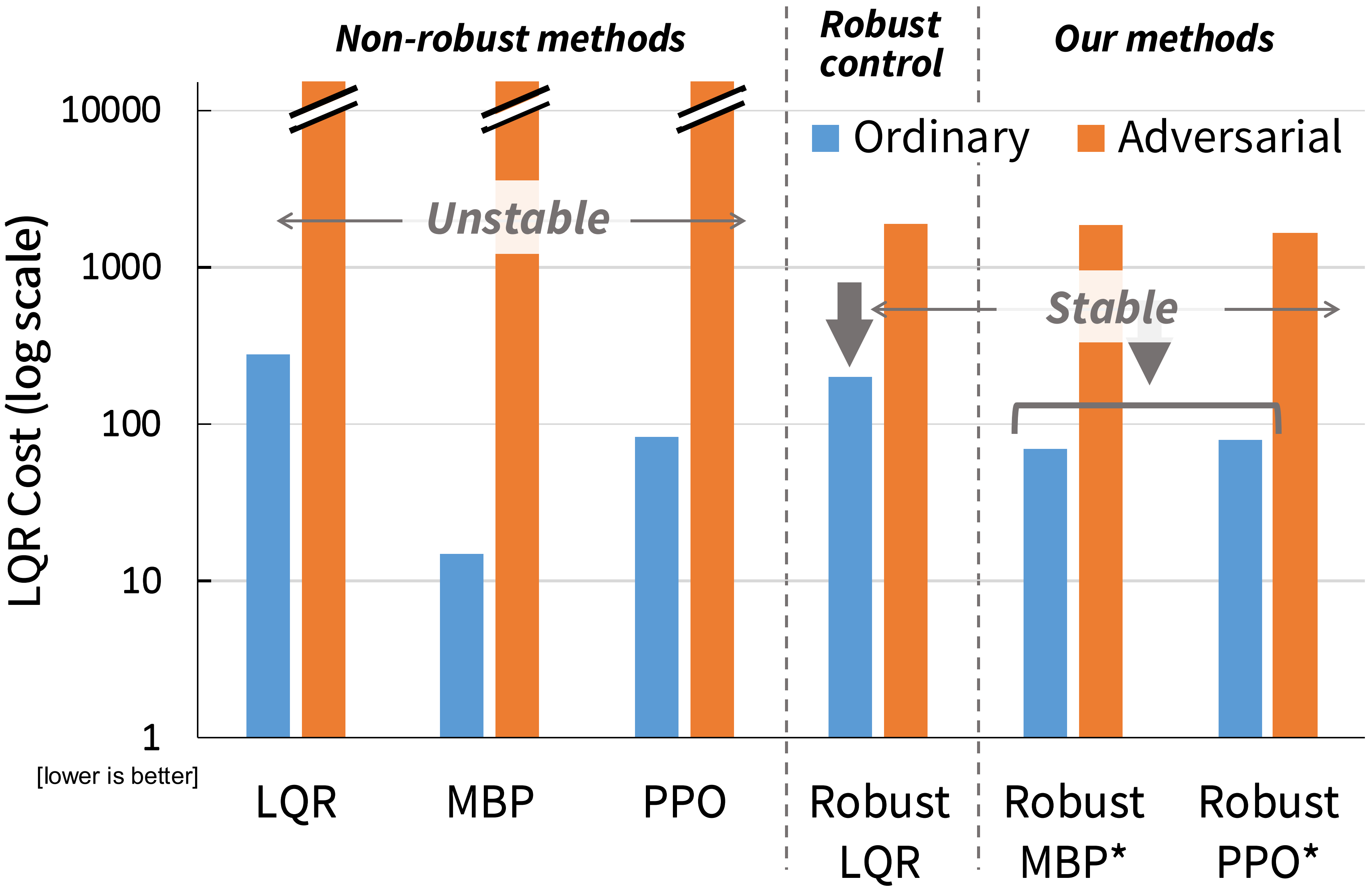
$$\Rightarrow \{u : \|k_1(x) + Du\|_2 \leq k_2(x) + k_3(x)^T u\}$$

Convex (non-empty) set in $u(t)$



Note: t-dependence has been dropped for brevity

Illustrative results: Synthetic NLDI system



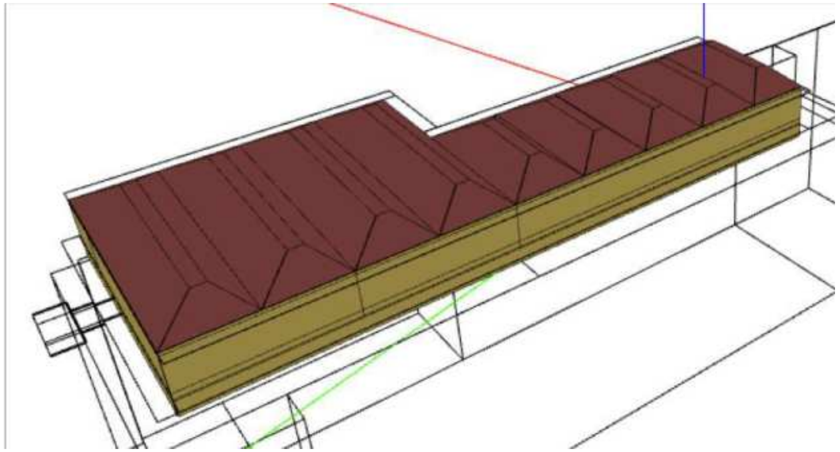
Improved "average-case" performance over robust baselines

Provably stable under "worst-case" dynamics (unlike non-robust baselines)

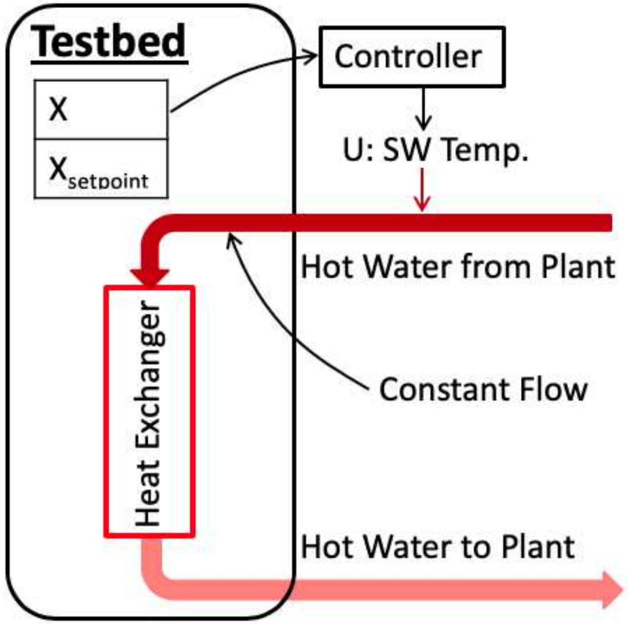
Needs improvement: Computational cost

Energy-efficient heating and cooling

Goal: Control the HVAC supply water temperature to minimize energy use, while respecting equipment constraints and maintaining thermal comfort



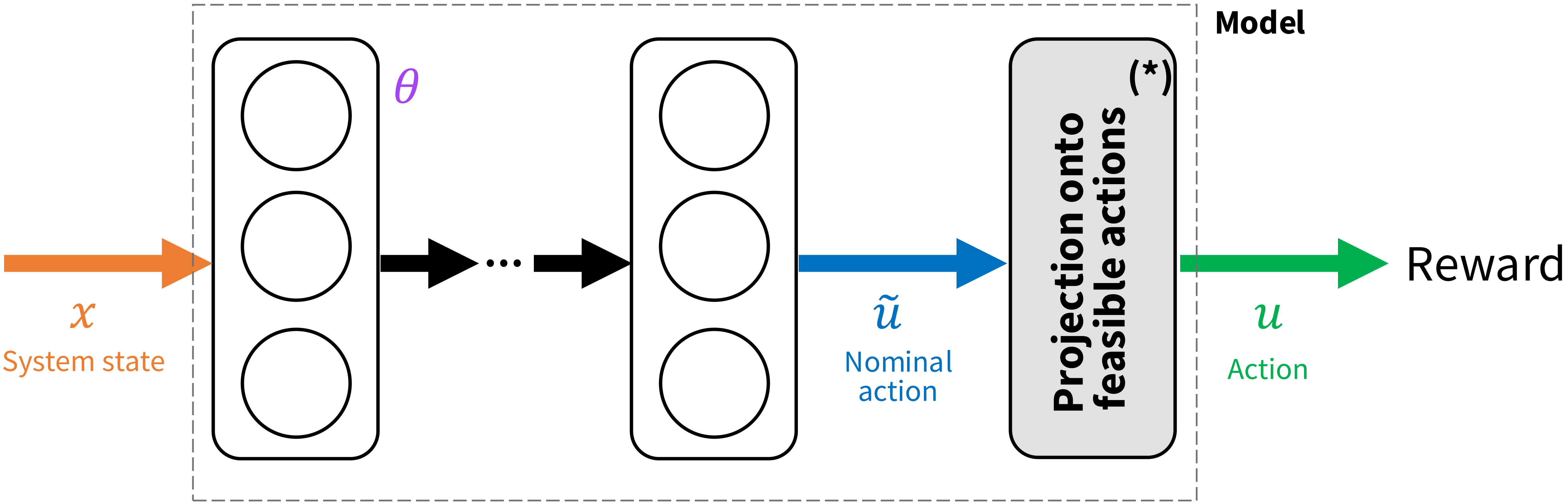
Intelligent Workplace
Margaret Morrison Hall, 4th Floor
(✤ Zhang & Lam, 2018)



HVAC Schematic

Bingqing Chen*, Priya L. Donti*, Kyri Baker, J. Zico Kolter, and Mario Berges. "Enforcing Policy Feasibility Constraints through Differentiable Projection for Energy Optimization." *ACM International Conference on Future Energy Systems (ACM e-Energy) 2021*.

Differentiable projection onto feasible actions

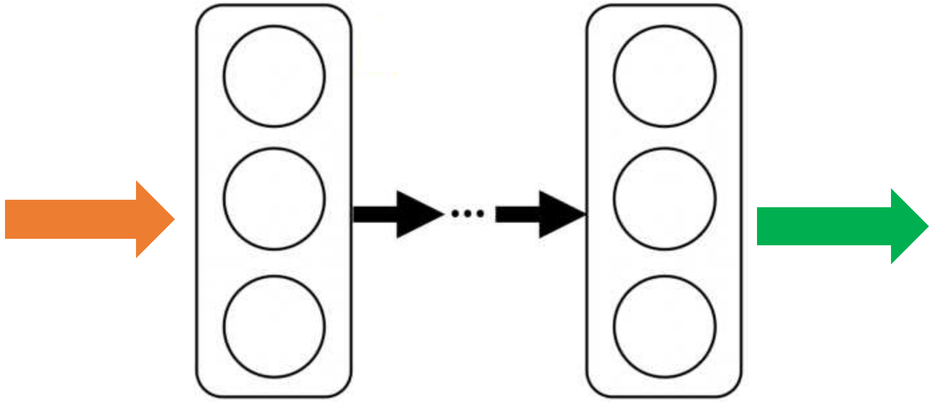


Results on realistic-scale building simulator

Improved energy efficiency (4-24%) | **Comparable thermal comfort**

	Total Heating Demand (kWh)	Predicted Percentage Dissatisfied Mean (%)	Predicted Percentage Dissatisfied Std (%)
Existing controller	43709	**9.45**	5.59
Agent #6 (Zhang & Lam, 2018)	37131	11.71	3.76
Gnu-RL (Chen et al., 2019)	34678	9.56	6.39
PROF (ours)	**33271**	9.68	**3.66**

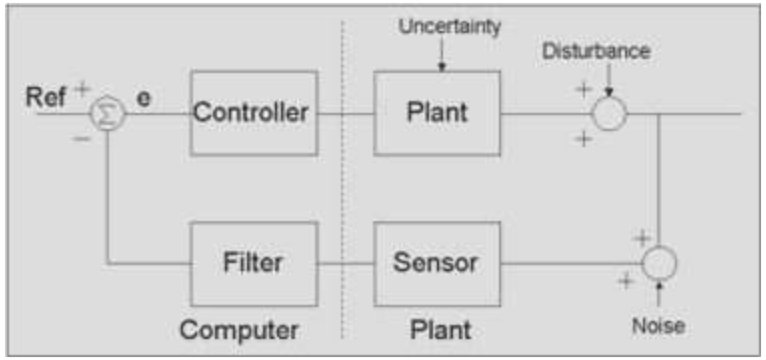
Deep reinforcement learning vs. robust control



Deep RL

Pro: Expressive, well-performing policies

Con: Potential (catastrophic) failures



Robust control

Pro: Provable stability guarantees

Con: Simple policies (e.g., linear)

Can we improve performance while still guaranteeing stability?

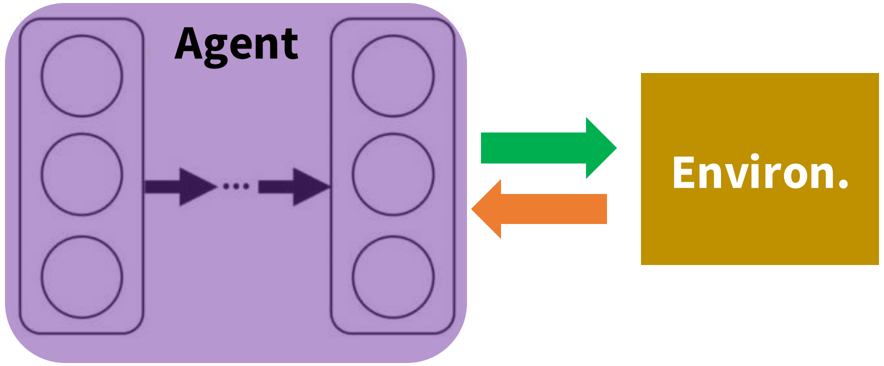
Today's talk

- **Optimization-in-the-loop ML:** Framework for developing ML methods incorporating knowledge of physics/hard constraints, via optimization problems
- **Application to single-agent control:** Optimization-in-the-loop reinforcement learning with enforcement of asymptotic stability or state/action constraints
- **Future directions in distributed control:** Bridging optimization-in-the-loop learning, multi-agent reinforcement learning, and decomposition methods

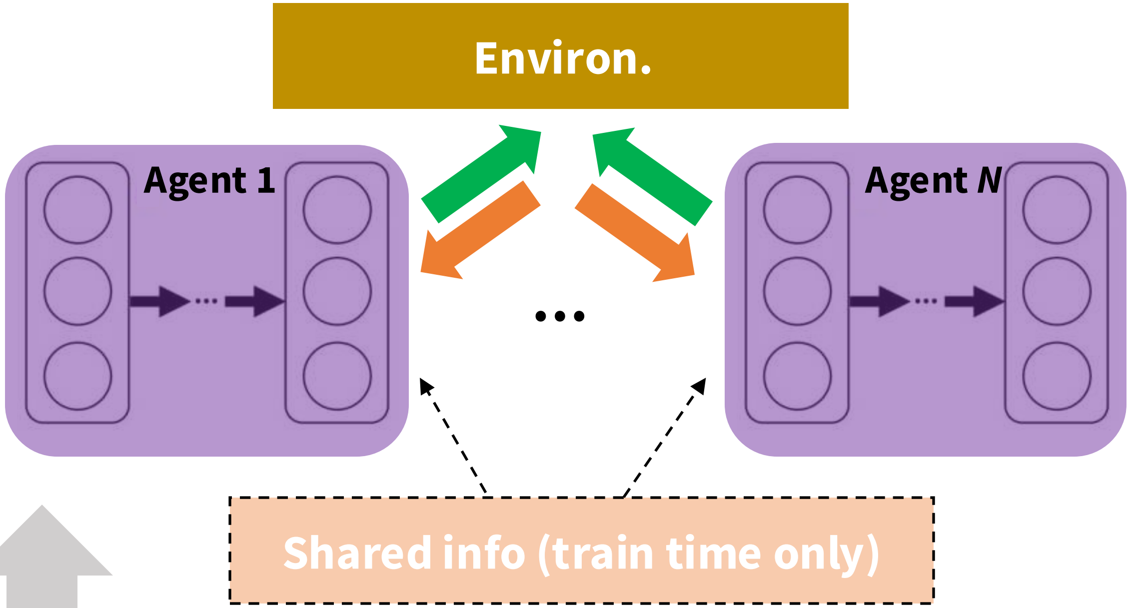
Single-agent → Multi-agent

Reinforcement learning

Standard reinforcement learning training



Centralized Training with Decentralized Execution (CTDE)



Optimization

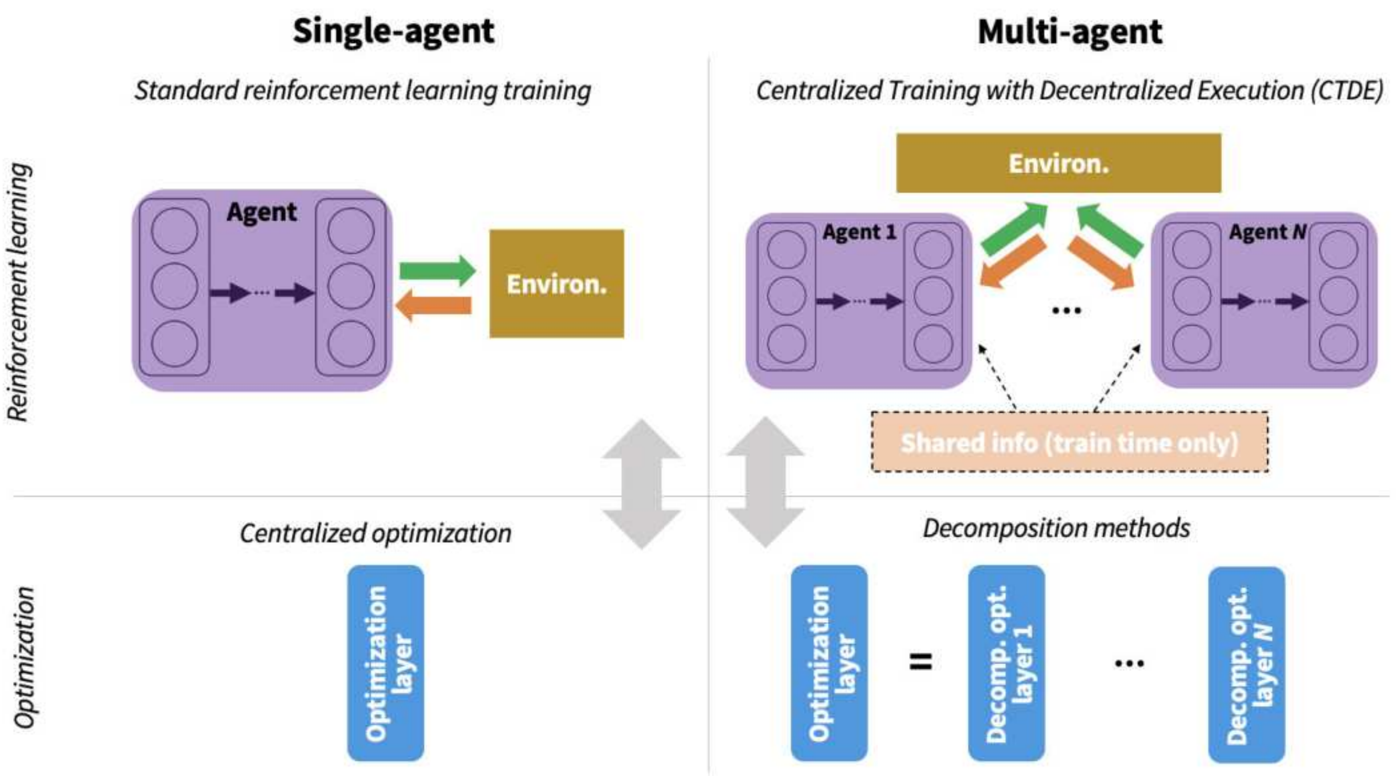
Centralized optimization

Optimization layer

Decomposition methods

Optimization layer = **Decomp. opt. layer 1** ... **Decomp. opt. layer N**

Many questions – seeking collaborators!



Addressing the "right" settings

- Distributed vs. decentralized?
- Competitive vs. cooperative?
- Communication constraints?
- Stability and safety reqs.?

Bridging the "best" methods

- Robust control formulations and synthesis techniques
- Decomposition approaches

Priya L. Donti: donti@mit.edu