Caltech

Learning for Safety-Critical Control in Dynamical Systems

Yisong Yue

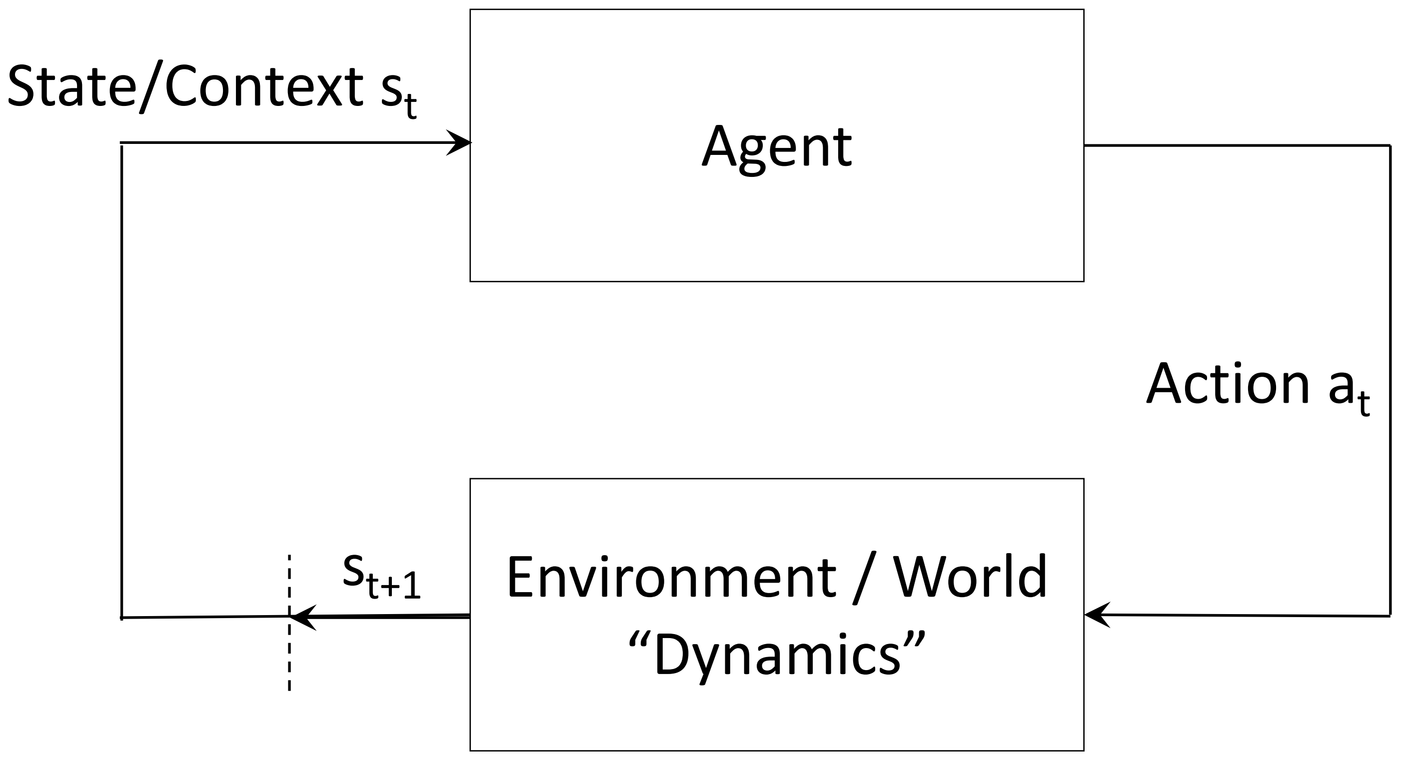
Policy/Controller Learning (Reinforcement & Imitation)

Goal: Find "Optimal" Policy

Imitation Learning:
Optimize imitation loss

Reinforcement Learning:
Optimize environmental reward

Learning-based Approach for Sequential Decision Making



Non-learning approaches include: optimal control, robust control, adaptive control, etc.

Imitation Learning Tutorial

https://sites.google.com/view/icml2018-imitation-learning/

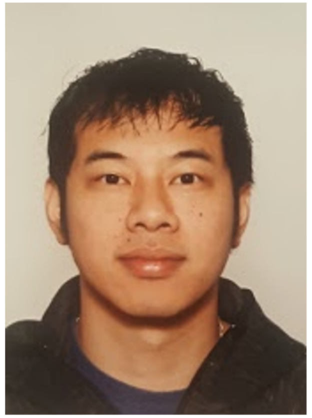
Yisong Yue

yyue@caltech.edu

@YisongYue

yisongyue.com

Hoang M. Le

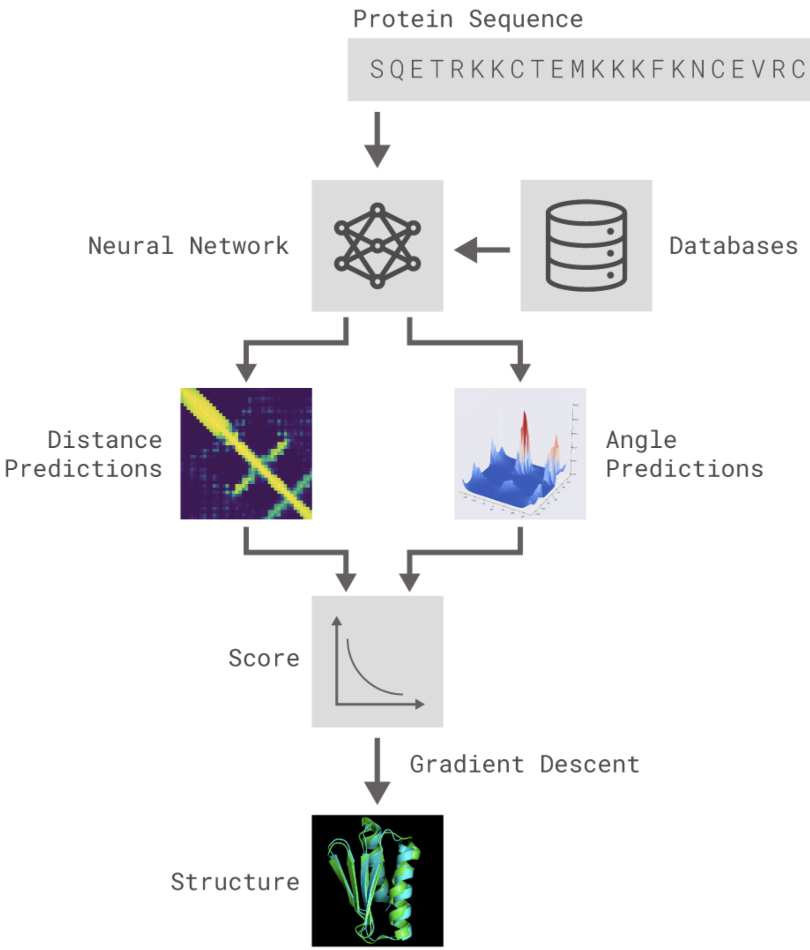


Hoang.Le@Microsoft.com

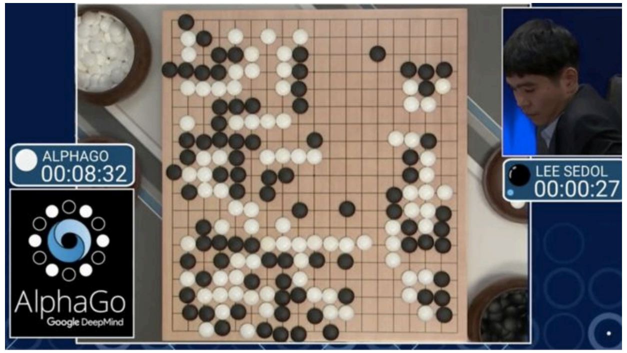
@HoangMinhLe

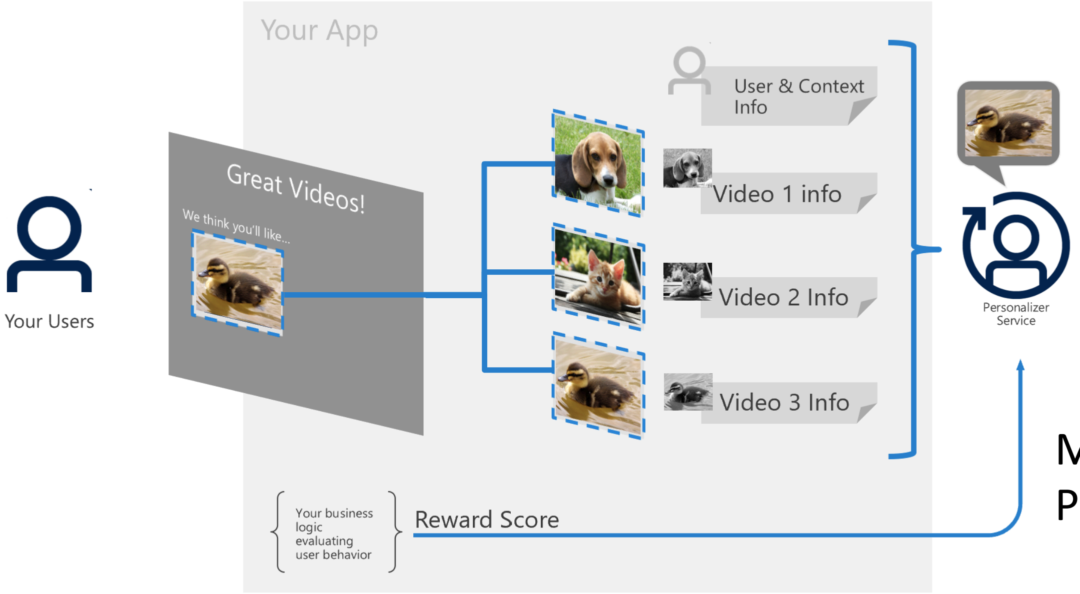
hoangle.info

Many Exciting Success Stories



AlphaFold



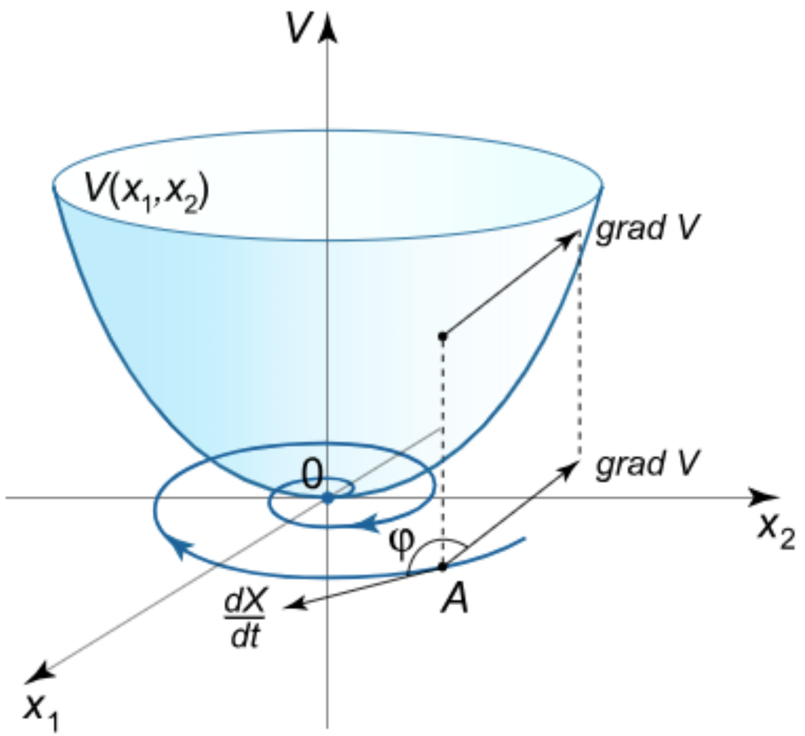


Microsoft Azure Personalizer

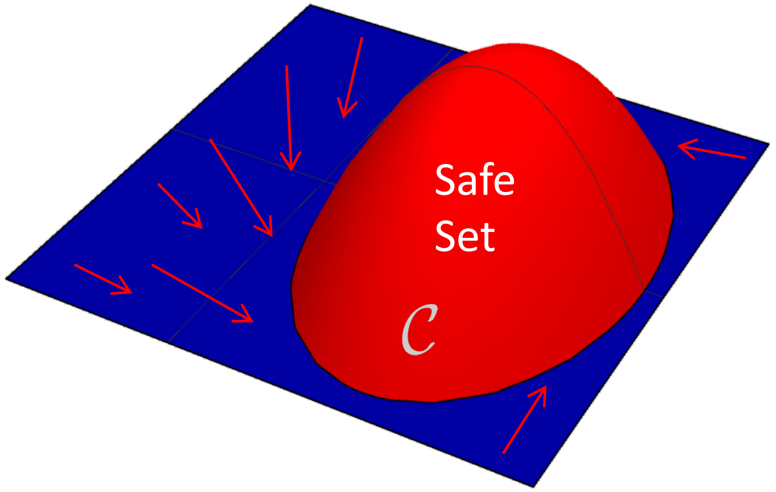
“ I want to use deep learning to optimize the design, manufacturing and operation of our aircrafts. But I need some guarantees. ” -- Aerospace Director

Behavioral Guarantees

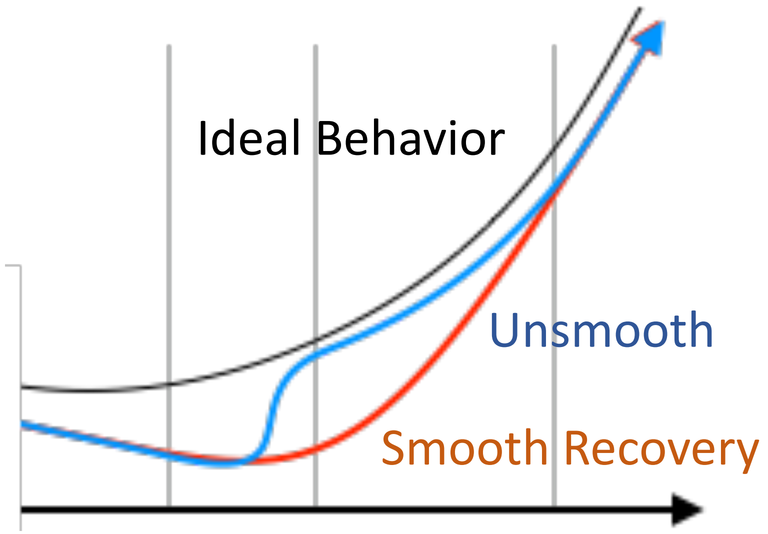
Possibly Others:
- Fairness
- Low-risk
- Temporal logic
- Etc…



Stability

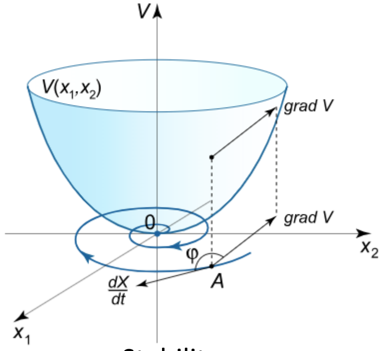


Safety

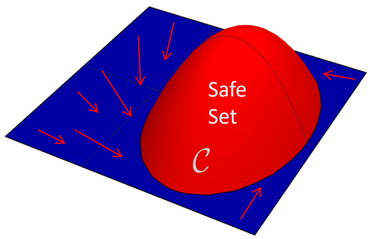


Smoothness

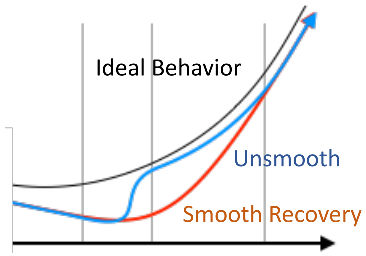
Research Questions



Stability



Safety

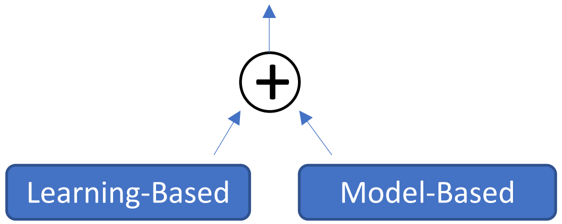


Smoothness

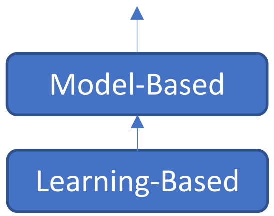
- How to constrain learning to (provably) satisfy guarantees?
- How to integrate domain knowledge from physics & control theory?
    - (Towards) a unified framework?
- How to exploit structure for faster learning?
    - (both computational & statistical)

Integration of Learning at Varying Levels

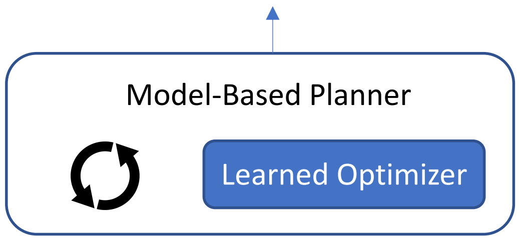
- Integration in control/output



- Integration in dynamics modeling



- Integration in optimization problem



Starting Point

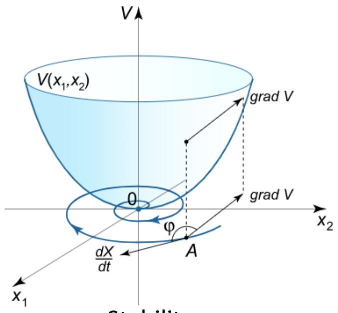
Standard IL/RL Objective

- Model-Based/Free
- On/Off Policy
- Imitation/Reinforcement
- Optimal Control

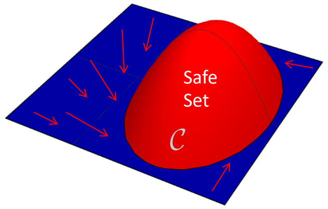
$$argmin_h L(h)$$
$$s.t.$$
$$R(h) < \kappa$$

In general, very hard to verify/optimize!

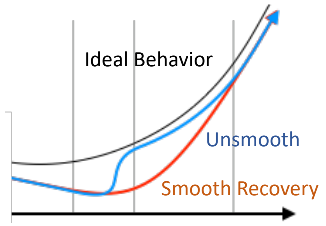
Side Constraint



Stability



Safety



Smoothness

Functional Regularization

(to a certified controller)

$$argmin_h L(h)$$

$$s.t.$$

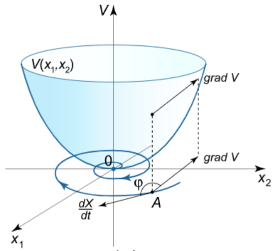
$$\exists g \in G: \|h-g\|^2 < \kappa$$

$$argmin_{h,g} L(h) + \lambda \|h-g\|^2$$

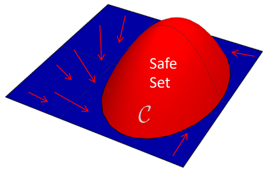
Model-Based Controllers

(certified behavioral properties)

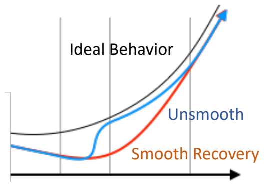
Intractable?



Stability



Safety



Smoothness

Blended Policy Class (solution concept)


Hoang Le


Richard Cheng

Policy

Black Box Predictor

Model-Based Controller

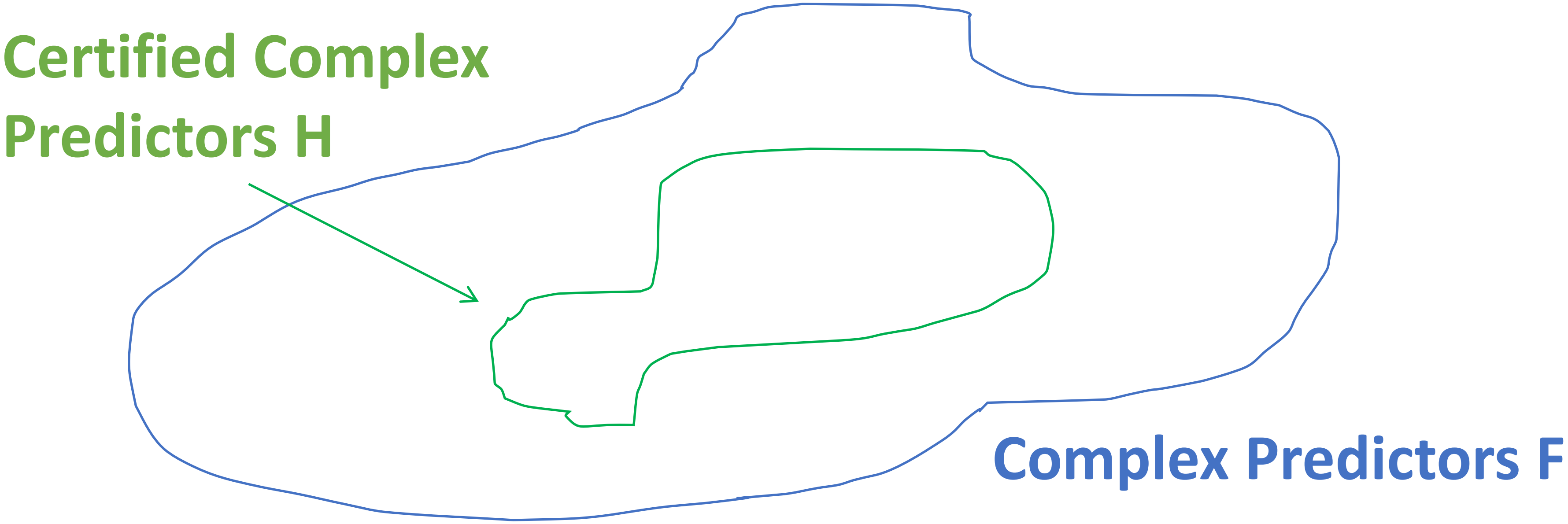
$$\operatorname{argmin}_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = \operatorname{argmin}_{a'} (f(s) - a')^2 + \lambda (g(s) - a')^2$$

$$= \frac{f(s)+\lambda g(s)}{1+\lambda}$$

Smooth Imitation Learning for Online Sequence Prediction, Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016
Control Regularization for Reduced Variance Reinforcement Learning, Richard Cheng, Abhinav Verma, et al. ICML 2019

Test-Time Functional Regularization

Hoang Le



$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = argmin_{a'}(f(s)-a')^2 + \lambda(g(s)-a')^2$$

$$= \frac{f(s)+\lambda g(s)}{1+\lambda}$$

Smooth Imitation Learning for Online Sequence Prediction
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

Comments on Optimization/Learning

$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = argmin_{a'} (f(s) - a')^2 + \lambda (g(s) - a')^2$$

$$= \frac{f(s) + \lambda g(s)}{1 + \lambda}$$

- Often use alternating optimization
  - Hold g fixed, optimize f
  - Hold f fixed, optimize g
  - (see NeurIPS 2019 paper for clean treatment)
- Can also consider fully differentiable learning

Reduces to "standard" approaches

Imitation-Projected Policy Gradient for Programmatic Reinforcement Learning
Abhinav Verma, Hoang Le, Yisong Yue, Swarat Chaudhuri. NeurIPS 2019

Theoretical Guarantees

$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = argmin_{a'} (f(s) - a')^2 + \lambda (g(s) - a')^2$$

$$= \frac{f(s) + \lambda g(s)}{1 + \lambda}$$

- By construction: h "close" to g
  - Certifications on g => (relaxed) certifications on h

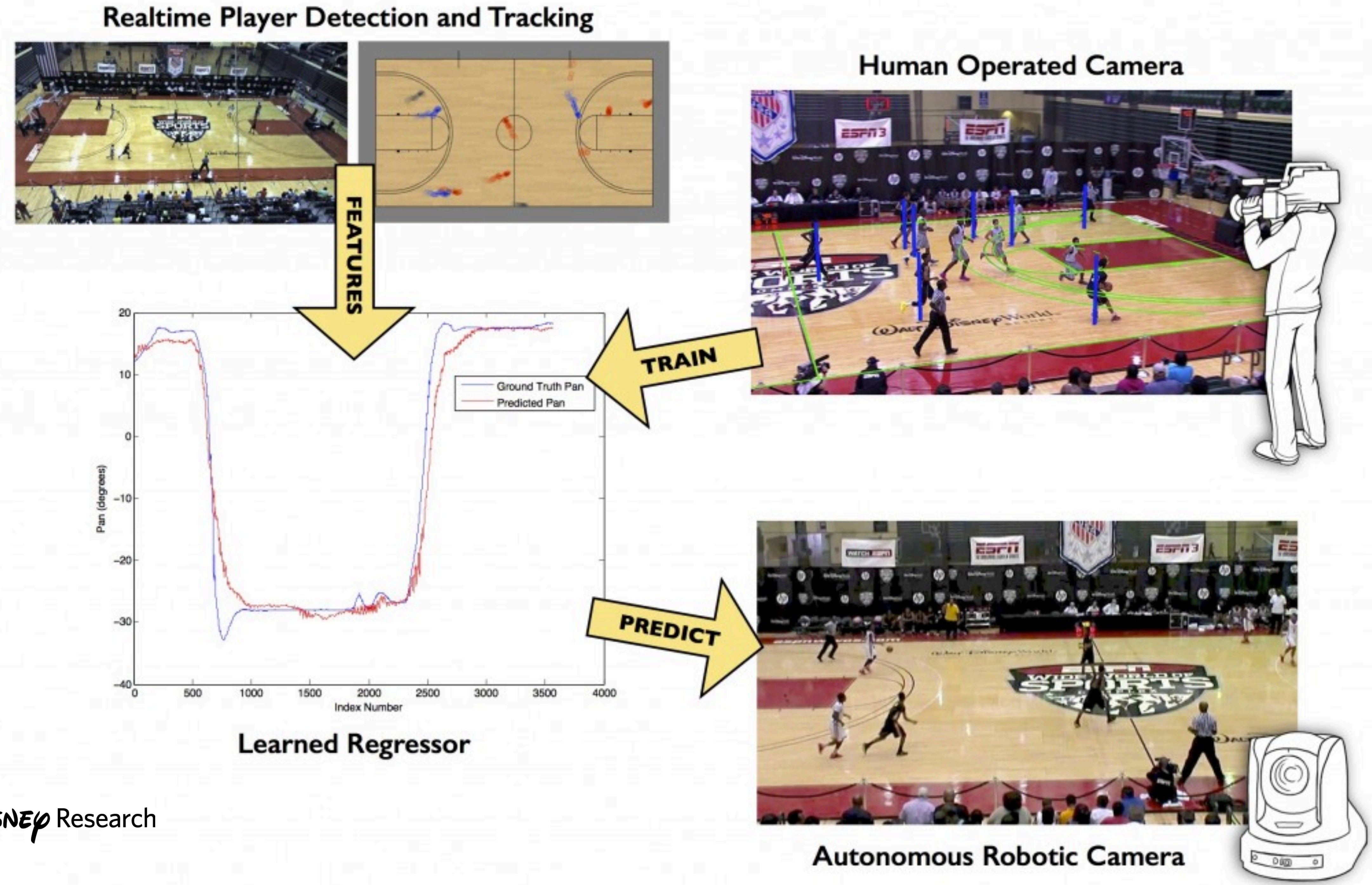
Run-time regularization

- Compatible with many forms of IL/RL
  - Can be exponentially faster than prior work (SEARN)

Adaptive Step Size
Exploits Lipschitz

- Can be very data efficient

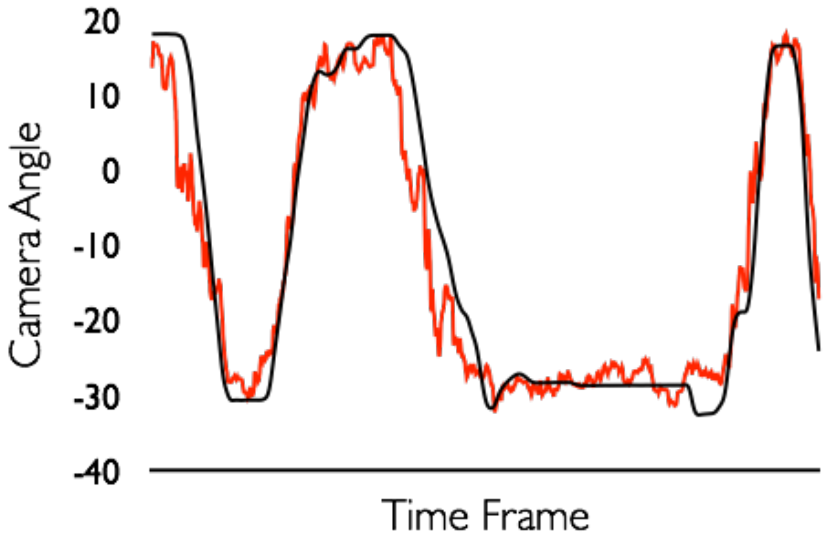
Low-Variance Gradients



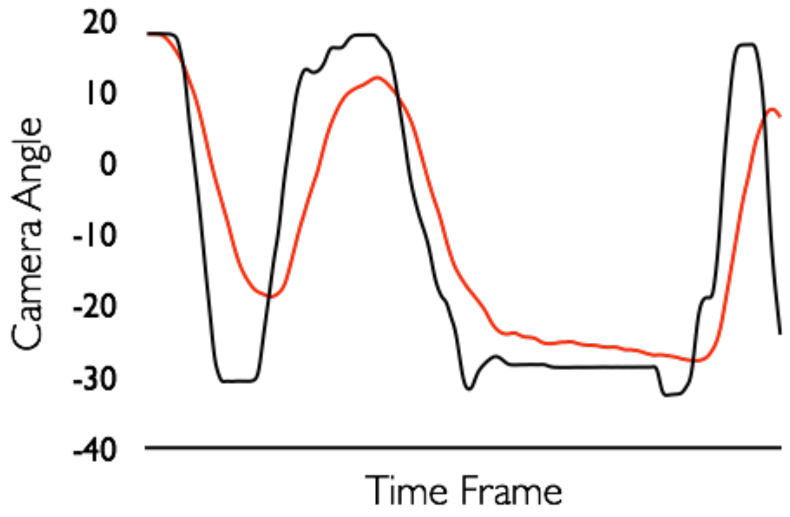
Disney Research

Naïve Approach

- Supervised learning of demonstration data
  - Train predictor per frame
  - Predict per frame

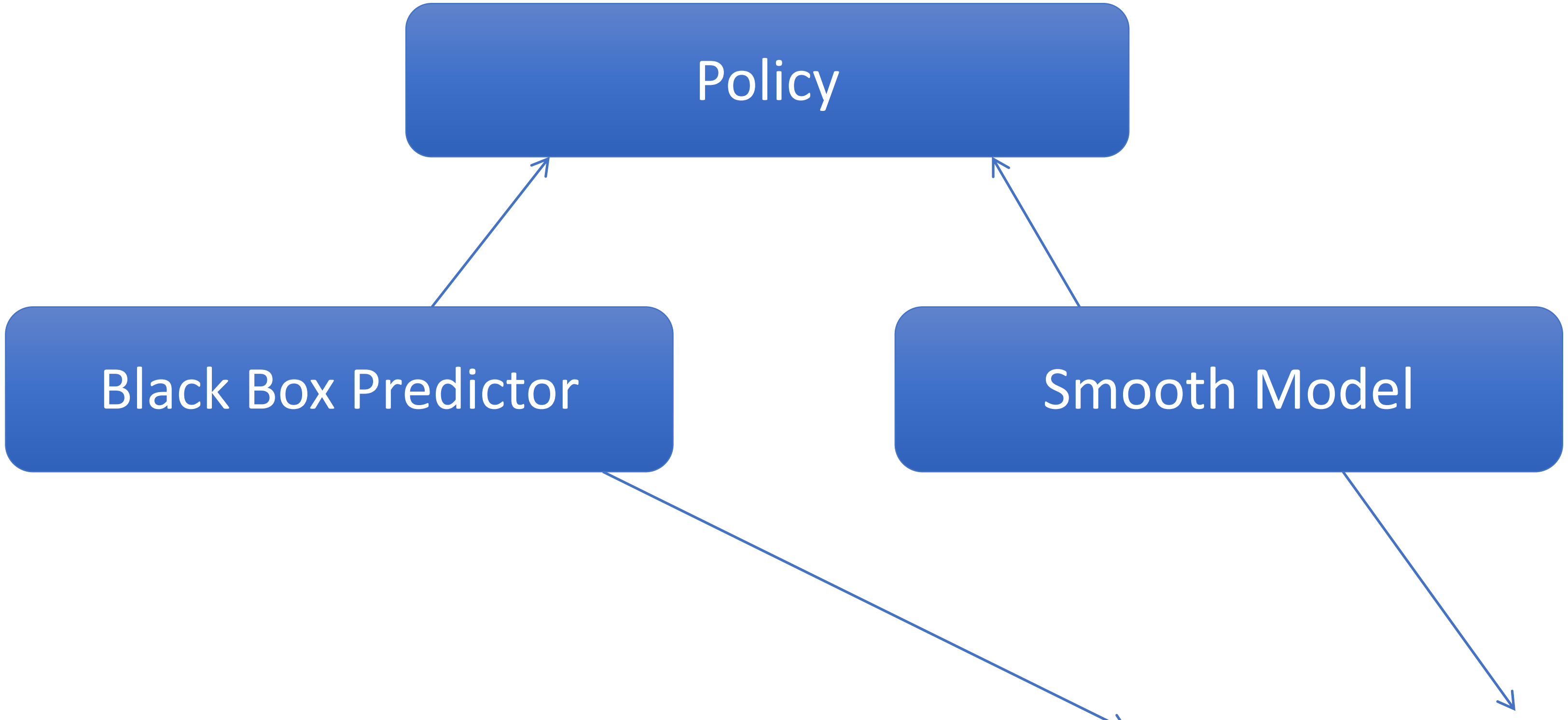


In practice, 2-step smoothing:



Smooth Policy Class

Hoang
Le



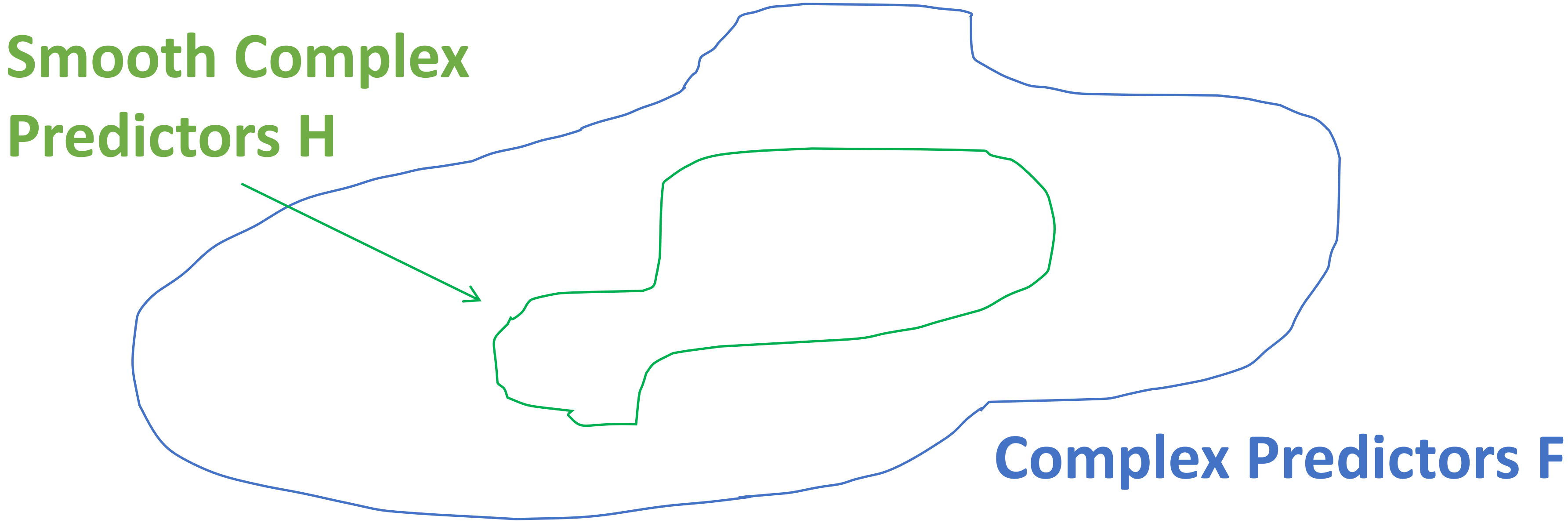
$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = argmin_{a'} (f(s) - a')^2 + \lambda (g(s) - a')^2$$

$$= \frac{f(s) + \lambda g(s)}{1 + \lambda}$$

Smooth Imitation Learning for Online Sequence Prediction
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

Test-Time Functional Regularization

Hoang Le

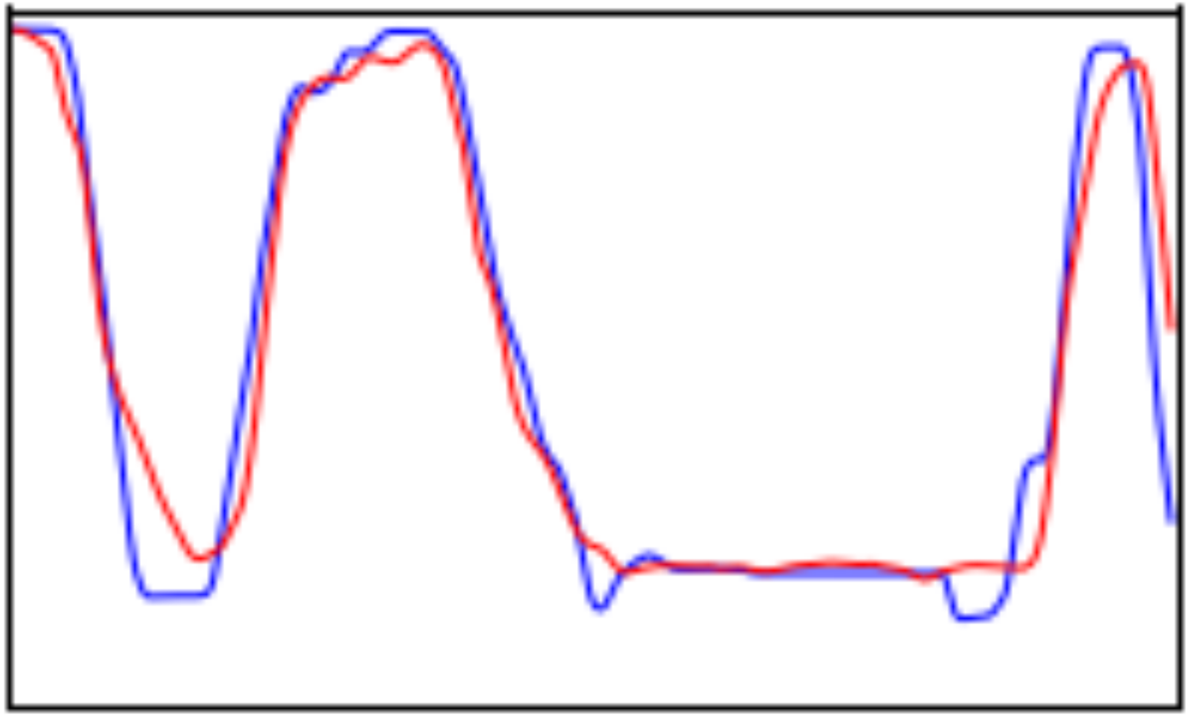


$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = argmin_{a'} (f(s) - a')^2 + \lambda(g(s) - a')^2$$

$$= \frac{f(s)+\lambda g(s)}{1+\lambda}$$

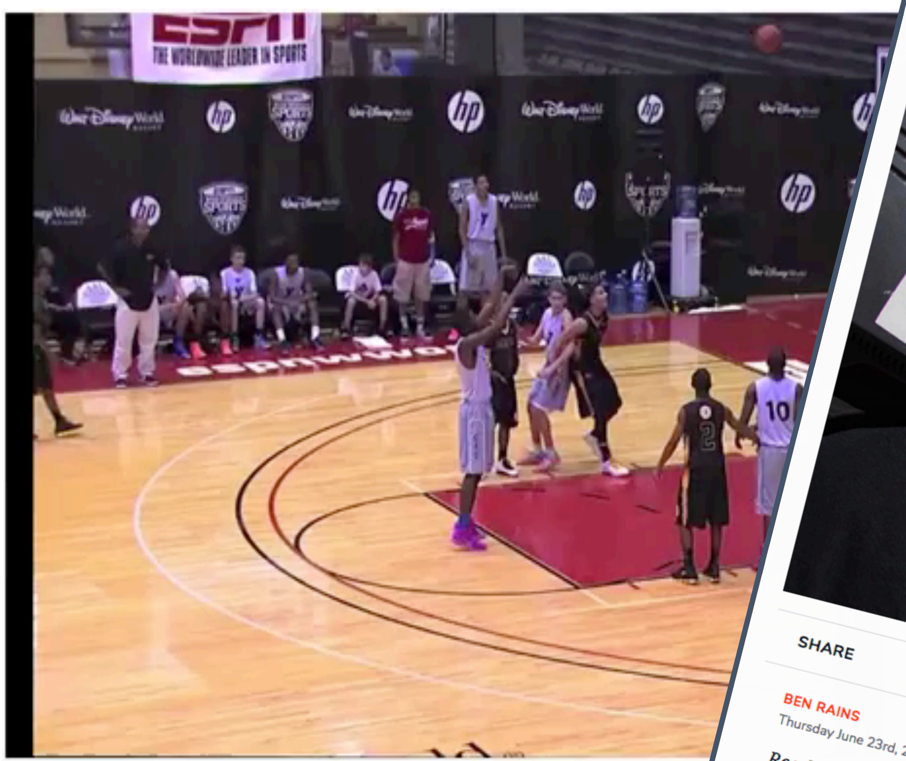
Smooth Imitation Learning for Online Sequence Prediction
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

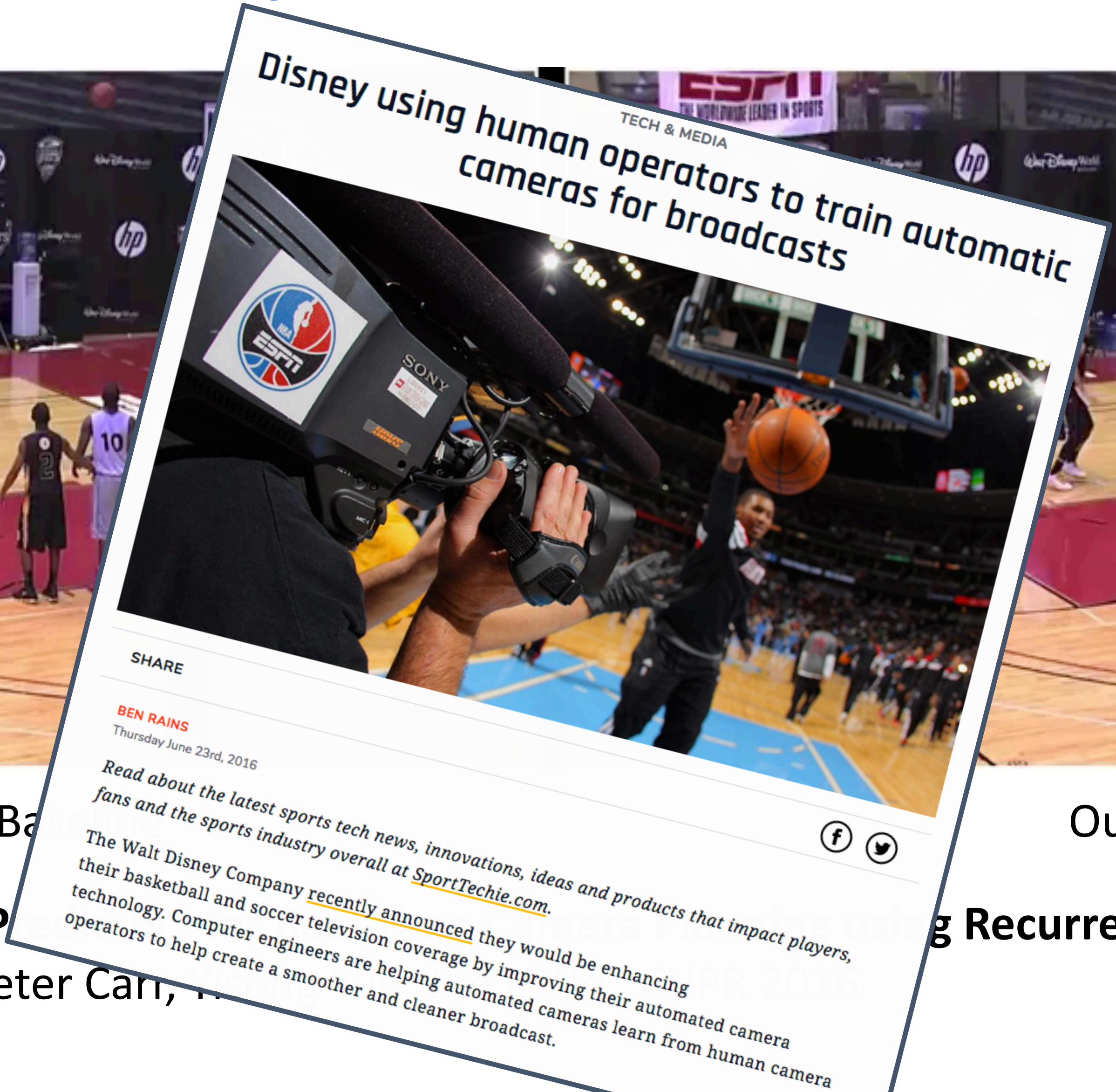
Our Results



Smooth Imitation Learning for Online Sequence Prediction
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

Qualitative Comparison





TECH & MEDIA

Disney using human operators to train automatic cameras for broadcasts

SHARE

BEN RAINS
Thursday June 23rd, 2016

Read about the latest sports tech news, innovations, ideas and products that impact players, fans and the sports industry overall at SportTechie.com.

The Walt Disney Company recently announced they would be enhancing their basketball and soccer television coverage by improving their automated camera technology. Computer engineers are helping automated cameras learn from human camera operators to help create a smoother and cleaner broadcast.

2-Step Ba…

Our Approach

Learning Online Smooth P… g Recurrent Decision Trees

Jianhui Chen, Hoang Le, Peter Carr, …

Control Regularization

Richard Cheng

$$h(s) = \frac{f(s)+\lambda g(s)}{1+\lambda}$$

- f is black box learning
- g is "control prior" (e.g., H-infinity controller)

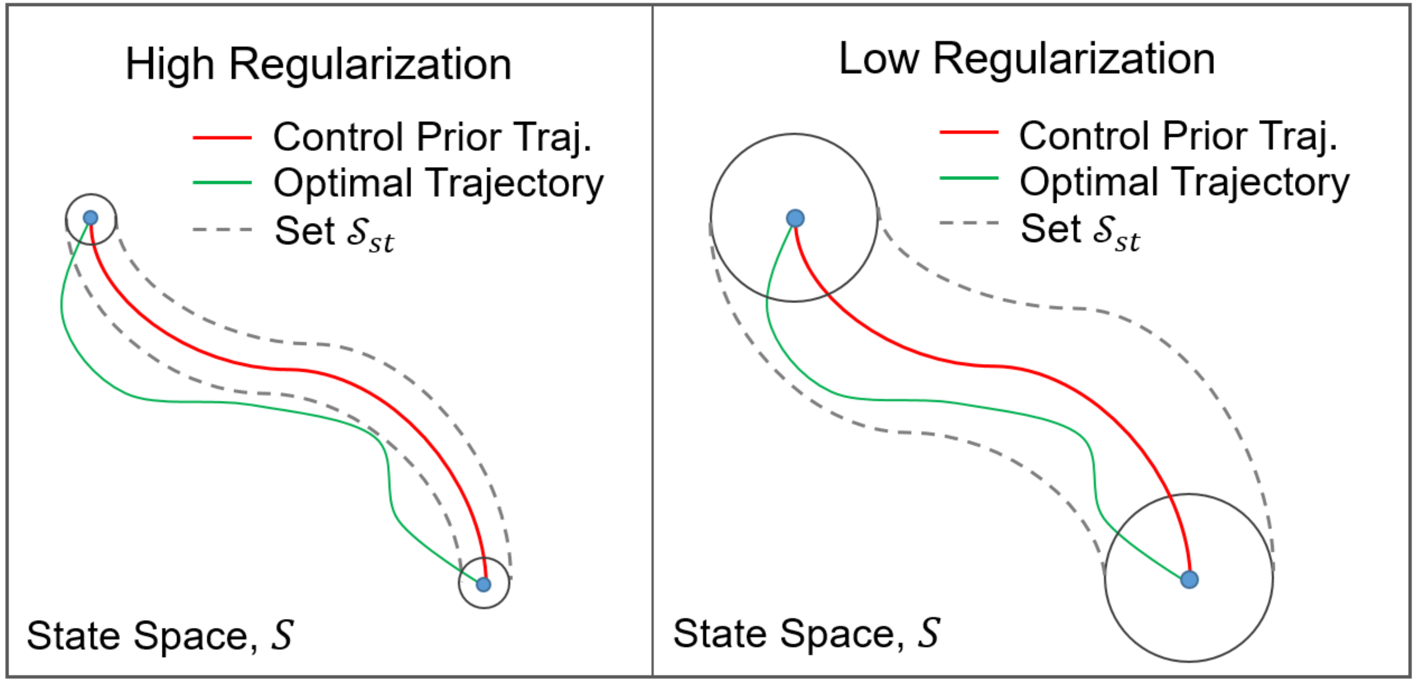
- Learn f using policy gradient using any standard RL method

Control Regularization for Reduced Variance Reinforcement Learning
Richard Cheng, Abhinav Verma, Gabor Orosz, Swarat Chaudhuri, Yisong Yue, Joel Burdick. ICML 2019

Control Regularization

Richard
Cheng

- (Relaxed) Lyapunov stability bounds:



Control Regularization for Reduced Variance Reinforcement Learning
Richard Cheng, Abhinav Verma, Gabor Orosz, Swarat Chaudhuri, Yisong Yue, Joel Burdick. ICML 2019

Control Regularization

Richard Cheng

$$h(s) = \frac{f(s)+\lambda g(s)}{1+\lambda}$$

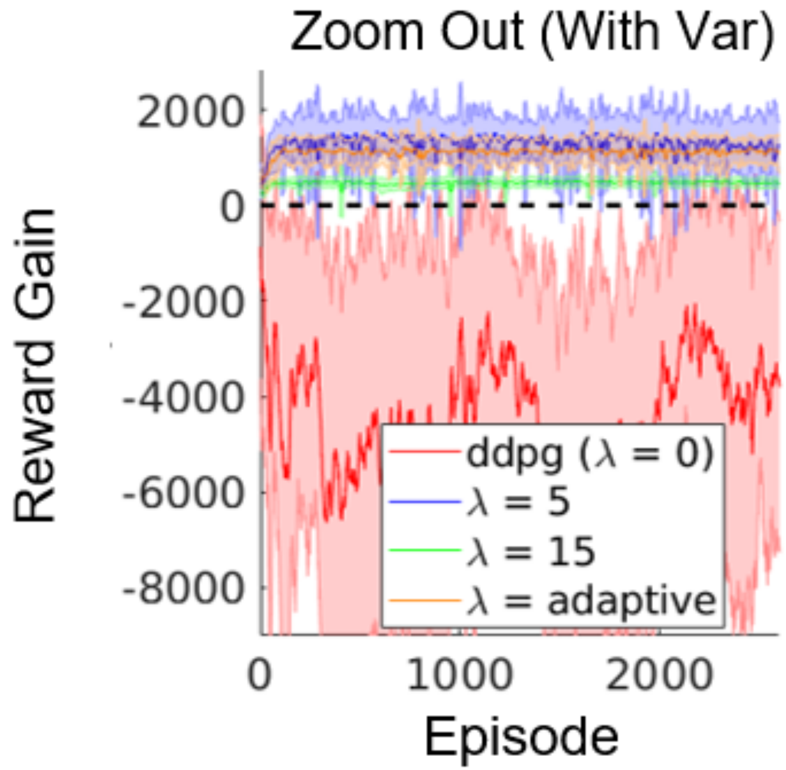
- Theorem (informal):
  - Variance of policy gradient decreases by factor of: $\left(\frac{1}{1+\lambda}\right)^2$

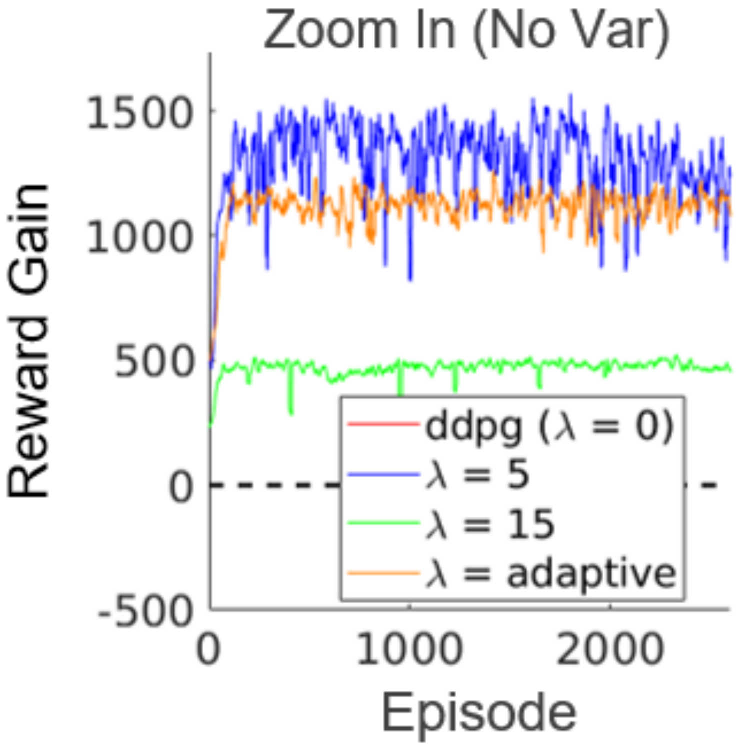
Implies much faster learning!

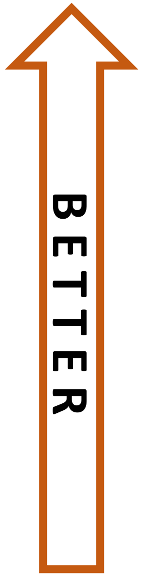
  - Bias converges to: $D_{TV}(h^*, g)$

Control Regularization for Reduced Variance Reinforcement Learning
Richard Cheng, Abhinav Verma, Gabor Orosz, Swarat Chaudhuri, Yisong Yue, Joel Burdick. ICML 2019

Generalized Control Regularization







Richard Cheng

Control Regularization for Reduced Variance Reinforcement Learning
Richard Cheng, Abhinav Verma, Gabor Orosz, Swarat Chaudhuri, Yisong Yue, Joel Burdick. ICML 2019


Ready
3 4 5 6 7
120 180 240
Ready
3 4 5 6 7
120 180 240

Summary: Functional Regularization

Equivalence Between Regularization & Constrained Learning

$$\begin{aligned} &argmin_h L(h) \\ s.t.\ &\\ &R(h) < \kappa \end{aligned} \quad \Longleftrightarrow \quad argmin_h L(h) + \lambda R(h)$$

IL/RL Objective

Side Guarantees

Hybrid Policy Solution Concept

$$h(s) = argmin_{a'} (f(s) - a')^2 + \lambda (g(s) - a')^2$$

$$= \frac{f(s) + \lambda g(s)}{1 + \lambda}$$

Summary: Functional Regularization (cont.)

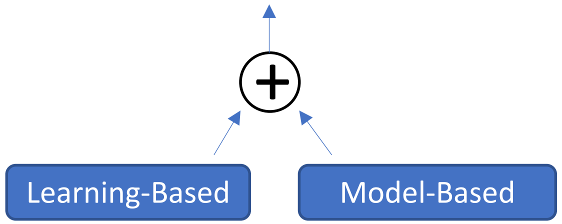
- Control methods => analytic guarantees — (side guarantees)
- Blend w/ learning => improve precision/flexibility — (real-world improvements)
- Preserve side guarantees — (possibly relaxed)
- Interpret as functional regularization — (speeds up learning)
- Other directions:

Batch Policy Learning under Constraints
Hoang Le, Cameron Voloshin, Yisong Yue. ICML 2019 — (offline learning)

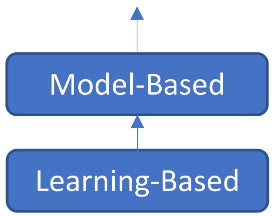
Imitation-Projected Policy Gradient for Programmatic Reinforcement Learning
Abhinav Verma, Hoang Le, Yisong Yue, Swarat Chaudhuri. NeurIPS 2019 — (programmatic controllers)

Integration of Learning at Varying Levels

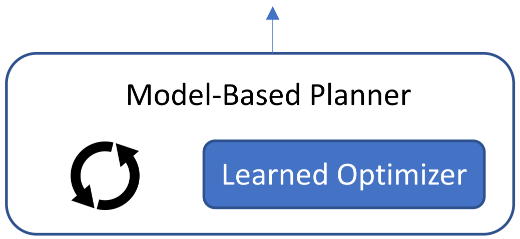
- Integration in control/output



- Integration in dynamics modeling



- Integration in optimization problem



Model-Based Control

New State

Current Action (aka control input)

$$s_{t+1} = F(s_t, u_t) + \epsilon$$

Current State

Unmodeled Disturbance / Error

(Value Iteration is also contraction mapping)

Robust Control (fancy contraction mappings)

- Stability guarantees (e.g., Lyapunov)
- Precision/optimality depends on error

Learning Residual Dynamics

F = nominal dynamics
$\tilde{F}$ = learned dynamics

New State

Current Action (aka control input)

$$s_{t+1} = F(s_t, u_t) + \tilde{F}(s_t, u_t) + \epsilon(s_t, u_t)$$

Current State

Unmodeled Disturbance / Error

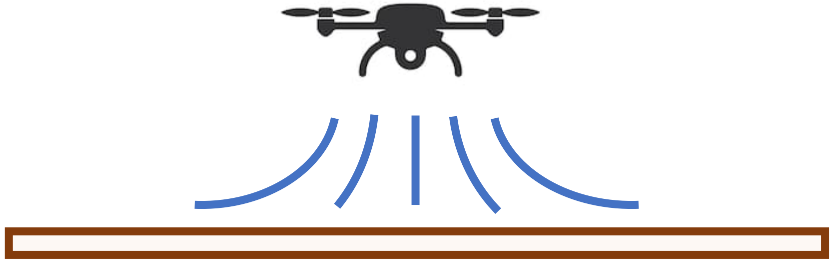
Leverage robust control (fancy contraction mappings)

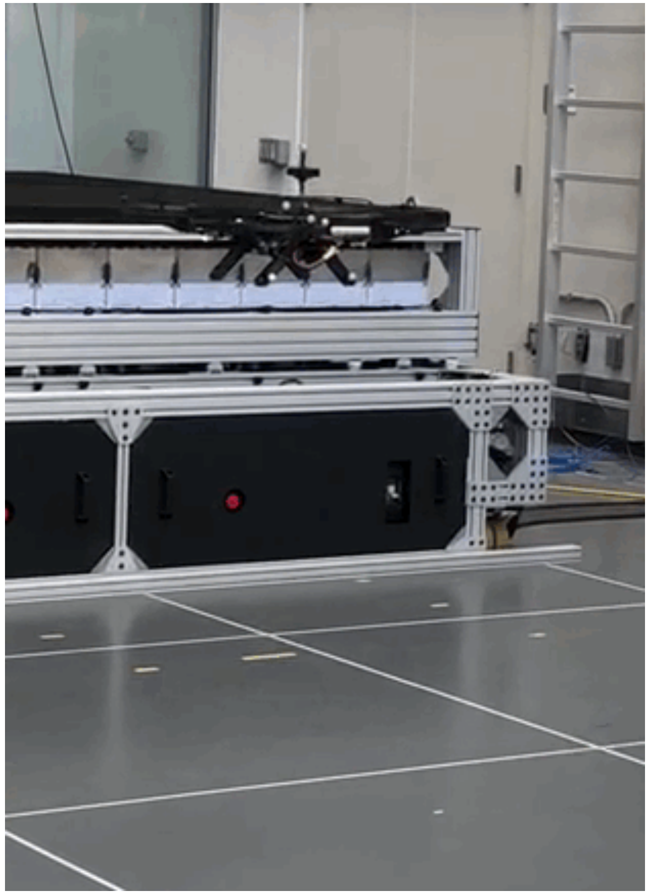
- Preserve stability (even using deep learning)
- Requires $\tilde{F}$ Lipschitz & bounded error

Stable Drone Landing

Guanya Shi

Ground effect





Neural Lander: Stable Drone Landing Control using Learned Dynamics

Guanya Shi, Xichen Shi, Michael O'Connell, Rose Yu, Kamyar Azizzadenesheli, Anima Anandkumar, Yisong Yue, Soon-Jo Chung. ICRA 2019

Control System Formulation

Learn the Residual

- Dynamics:

$$\begin{cases} \dot{\mathbf{p}} = \mathbf{v}, & m\dot{\mathbf{v}} = m\mathbf{g} + R\mathbf{f}_u + \mathbf{f}_a \\ \dot{R} = RS(\boldsymbol{\omega}), & J\dot{\boldsymbol{\omega}} = J\boldsymbol{\omega} \times \boldsymbol{\omega} + \boldsymbol{\tau}_u + \boldsymbol{\tau}_a \end{cases}$$

- Control:

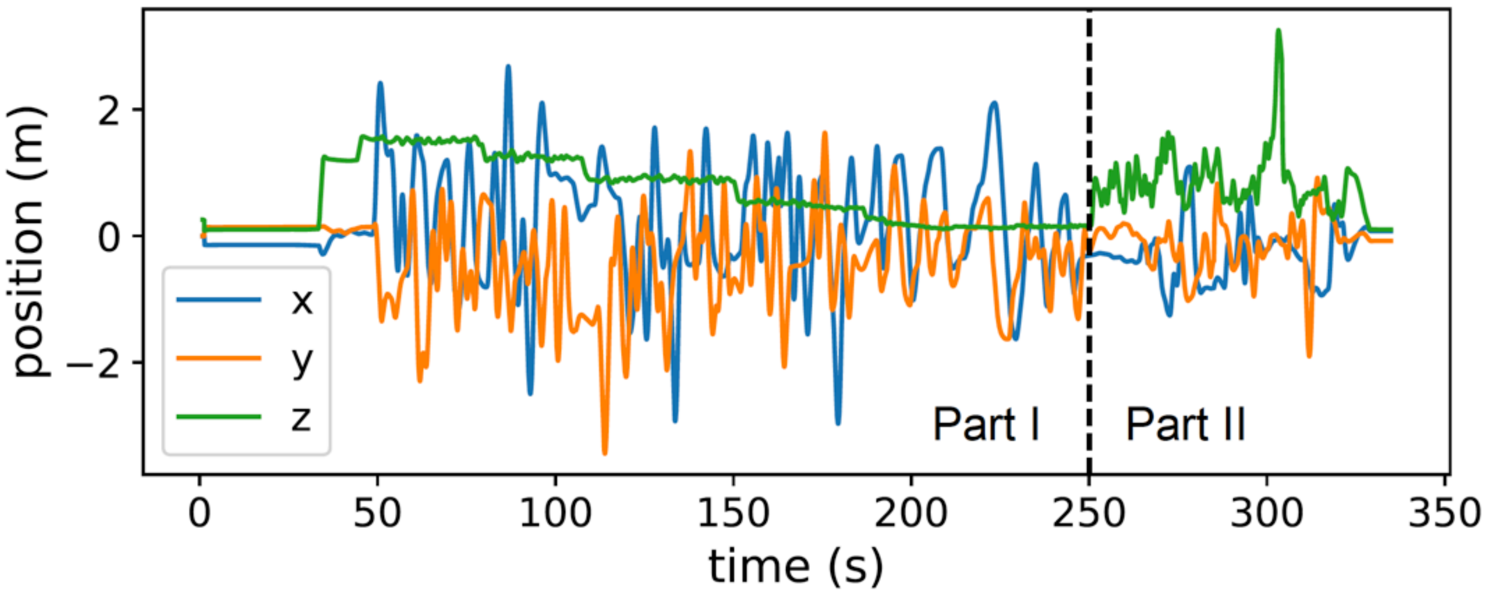
$$\begin{cases} \mathbf{f}_u = [0, 0, T]^\top \\ \boldsymbol{\tau}_u = [\tau_x, \tau_y, \tau_z]^\top \\ \begin{bmatrix} T \\ \tau_x \\ \tau_y \\ \tau_z \end{bmatrix} = \begin{bmatrix} c_T & c_T & c_T & c_T \\ 0 & c_T l_{\text{arm}} & 0 & -c_T l_{\text{arm}} \\ -c_T l_{\text{arm}} & 0 & c_T l_{\text{arm}} & 0 \\ -c_Q & c_Q & -c_Q & c_Q \end{bmatrix} \begin{bmatrix} n_1^2 \\ n_2^2 \\ n_3^2 \\ n_4^2 \end{bmatrix} \end{cases}$$

- Unknown forces & moments:

$$\begin{cases} \mathbf{f}_a = [f_{a,x}, f_{a,y}, f_{a,z}]^\top \\ \boldsymbol{\tau}_a = [\tau_{a,x}, \tau_{a,y}, \tau_{a,z}]^\top \end{cases}$$

Learn the Residual

Data Collection (Manual Exploration)



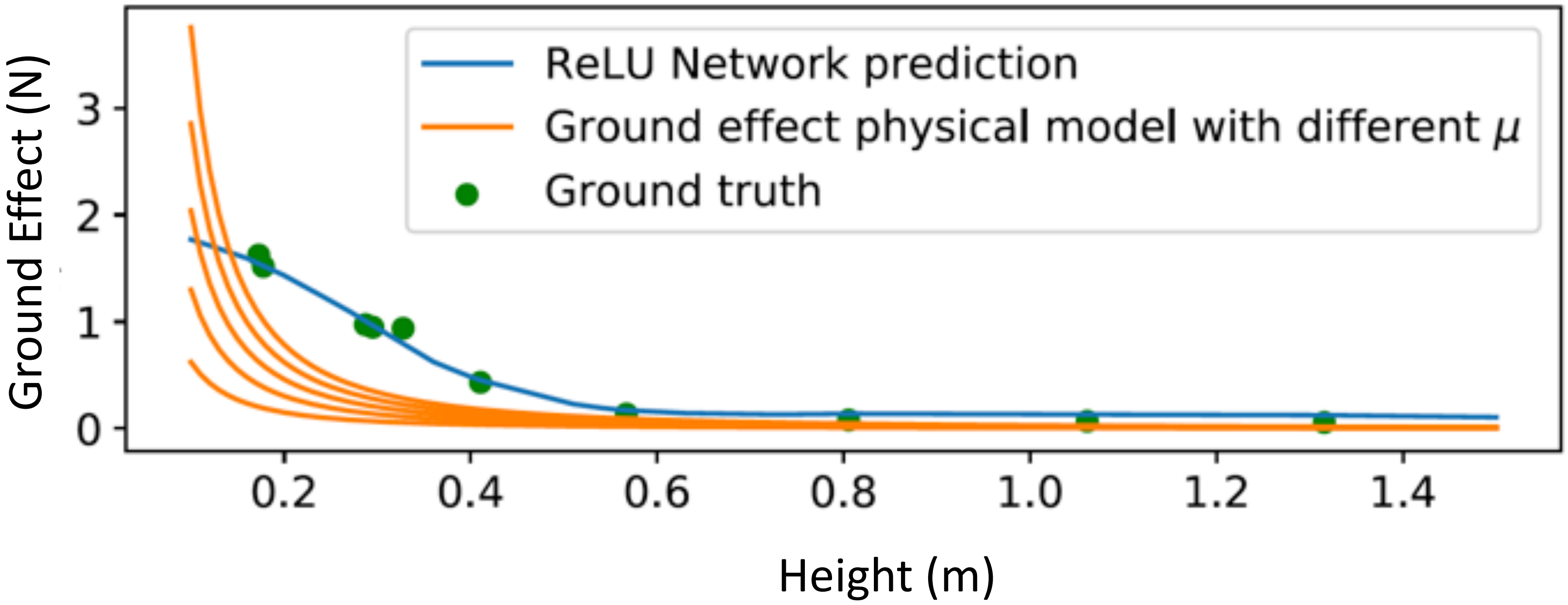
- Learn ground effect: $\tilde{F}(s,u) \rightarrow \mathbf{f}_a = [f_{a,x}, f_{a,y}, f_{a,z}]^\top$
- (s,u): height, velocity, attitude and four control inputs

Notable Extension:
Safe Exploration

Ensures $\widetilde{F}$ is Lipshitz
[Bartlett et al., NeurIPS 2017]
[Miyato et al., ICLR 2018]

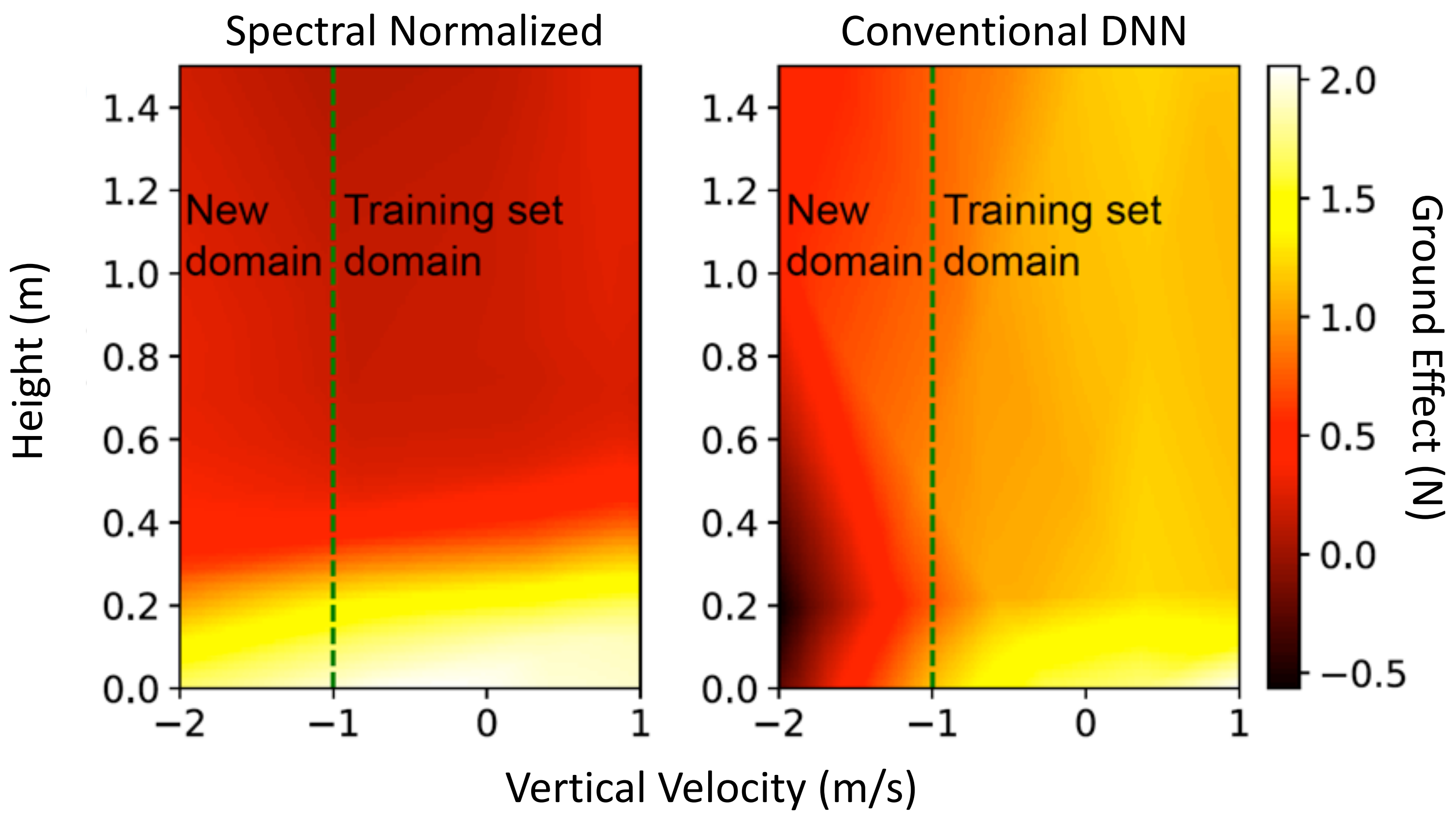
Spectral-Normalized 4-Layer Feed-Forward

Prediction Results



Neural Lander: Stable Drone Landing Control using Learned Dynamics
Guanya Shi, Xichen Shi, Michael O'Connell, Rose Yu, Kamyar Azizzadenesheli, Anima Anandkumar, Yisong Yue, Soon-Jo Chung. ICRA 2019.

Prediction Results



Neural Lander: Stable Drone Landing Control using Learned Dynamics
Guanya Shi, Xichen Shi, Michael O'Connell, Rose Yu, Kamyar Azizzadenesheli, Anima Anandkumar, Yisong Yue, Soon-Jo Chung. ICRA 2019.

Controller Design (simplified)

Guanya Shi

- Nonlinear Feedback Linearization:

$$u_{nominal} = K_s \eta \qquad \eta = \begin{bmatrix} p - p^* \\ v - v^* \end{bmatrix}$$

Desired Trajectory
(tracking error)

Feedback Linearization (PD control)

- Cancel out ground effect $\tilde{F}(s, u_{old})$: $u = u_{nominal} + u_{residual}$

Requires Lipschitz & small time delay

Controller Design (simplified)

Guanya Shi

- Nonlinear Feedback Linearization:

$$u_{nominal} = K_s \eta \qquad \eta = \begin{bmatrix} p - p^* \\ v - v^* \end{bmatrix}$$

Desired Trajectory
(tracking error)

Stability Guarantee:
(simplified)

Time delay

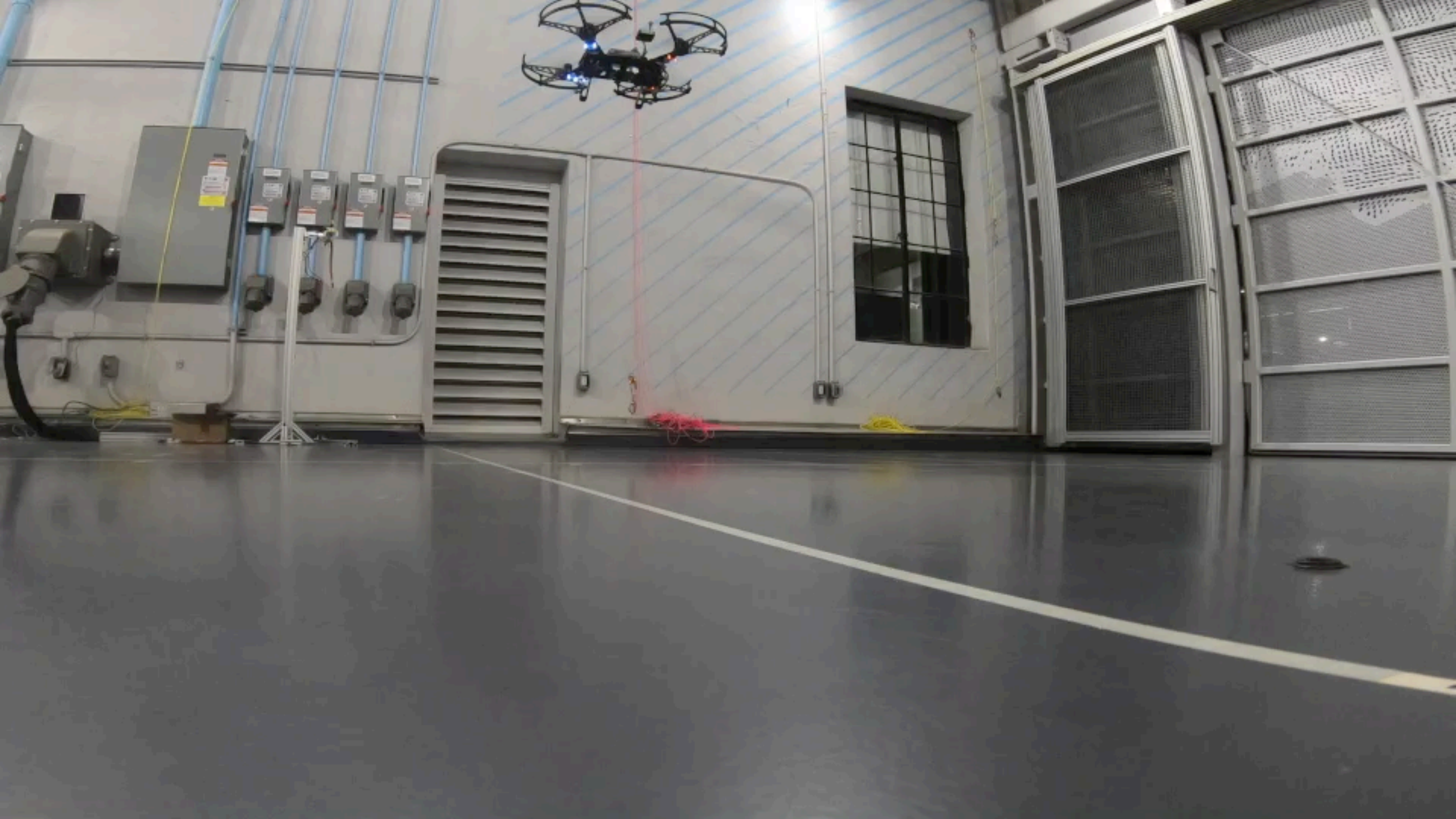
Unmodeled disturbance

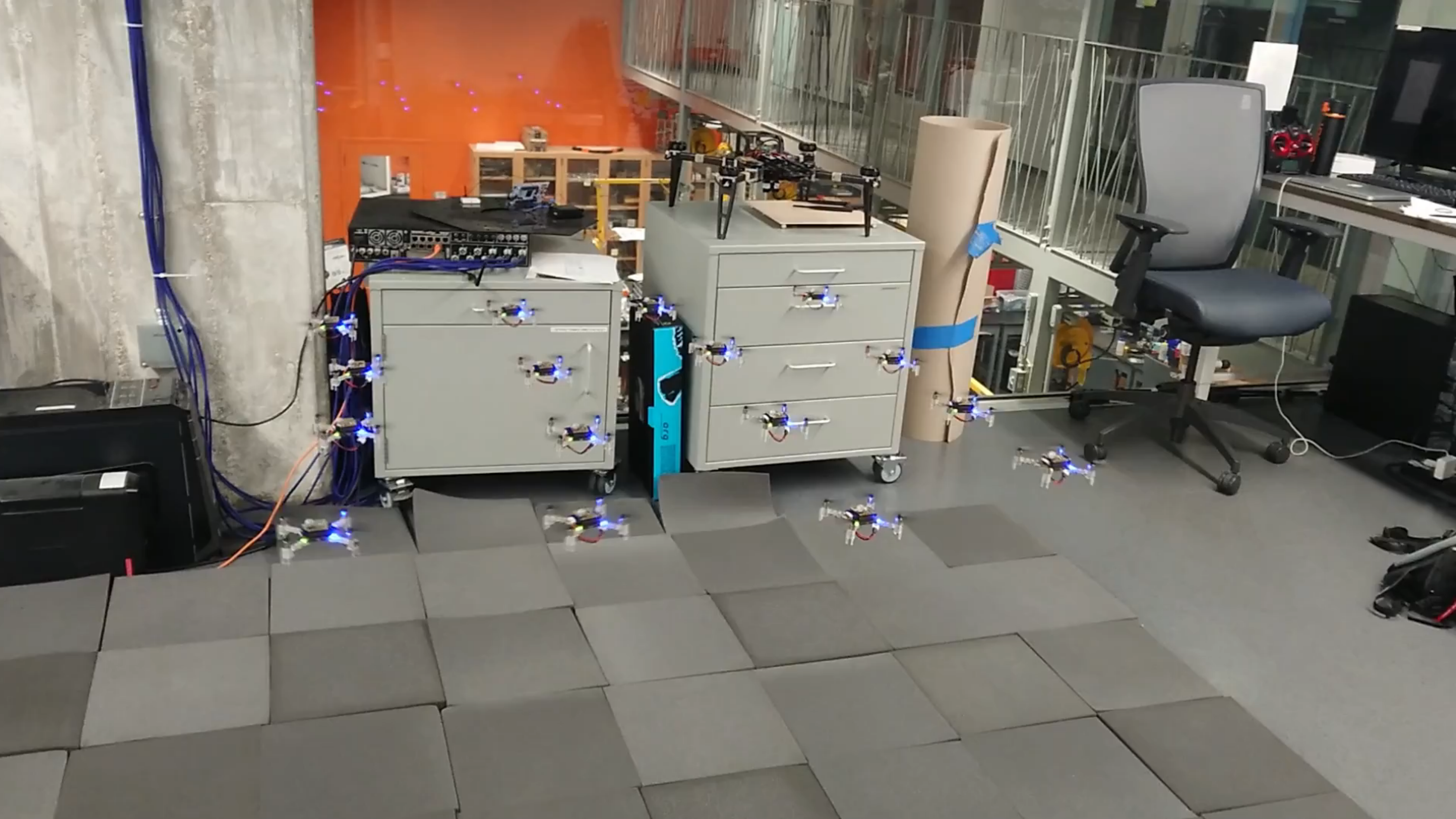
$$\|\eta(\mathrm{t})\| \leq \|\eta(0)\| \exp\left\{\frac{\lambda_{min}(K) - \tilde{L}\rho}{C} t\right\} + \frac{\epsilon}{\lambda_{min}(K) - \tilde{L}\rho}$$

Lipschitz of NN

$$\Longrightarrow \|\eta(\mathrm{t})\| \rightarrow \frac{\epsilon}{\lambda_{min}(K) - \tilde{L}\rho}$$

Exponentially fast

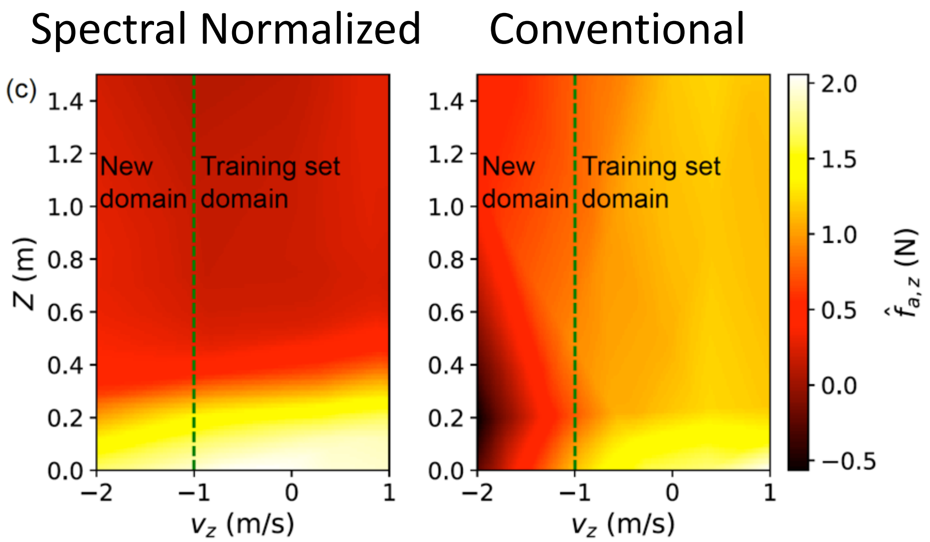


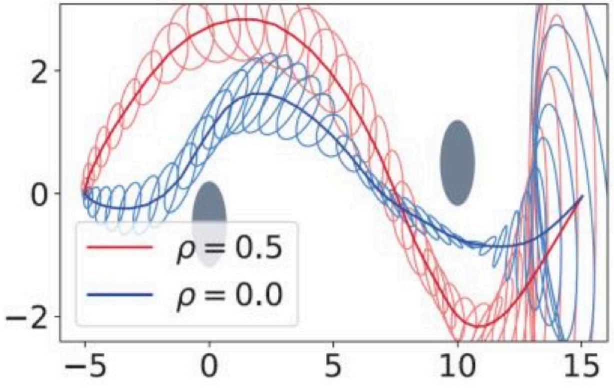


Aside: Robust Regression for Safe Exploration

- Robust regression for provable extrapolation => Safe Exploration!

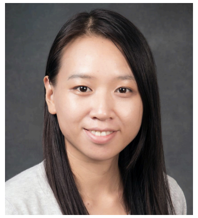
Robust regression guarantees low error!





Provably safe trajectory planning for exploration!

Robust Regression for Safe Exploration in Control,
Angie Liu, Guanya Shi, et al., L4DC 2020
Chance-Constrained Trajectory Optimization for Safe Exploration and Learning of Nonlinear Systems,
Yashwanth Kumar Nakka, et al. arXiv


Angie Liu


Yashwanth Nakka

Aside: Learning Control Lyapunov/Barrier Functions

- CLFs & CBFs encode low-dimensional projection of dynamics
  - DOF of action space rather than state space
  - Can be easier to learn than full dimensional dynamics
- How to learn CLF/CBF for controller design?
- How to analyze stability/safety under uncertainty?

Andrew Taylor

Victor Dorobantu

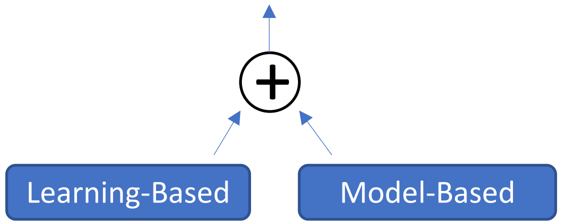
Episodic Learning with Control Lyapunov Functions for Uncertain Robotic Systems
Andrew J. Taylor, Victor D. Dorobantu, Hoang M. Le, Yisong Yue, Aaron D. Ames. IROS 2019.
A Control Lyapunov Perspective on Episodic Learning via Projection to State Stability
Andrew J. Taylor, Victor D. Dorobantu, Meera Krishnamoorthy, Hoang M. Le, Yisong Yue, Aaron D. Ames. CDC 2019.
Learning for Safety-Critical Control with Control Barrier Functions
Andrew Taylor, Andrew Singletary, Yisong Yue, Aaron Ames. L4DC 2020.
A Control Barrier Perspective on Episodic Learning via Projection-to-State Safety
Andrew J. Taylor, Andrew Singletary, Yisong Yue, Aaron D. Ames. L-CSS 2020.

Summary: Dynamics Learning

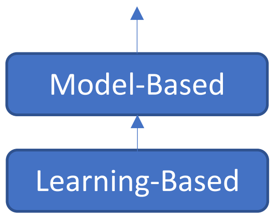
- Learn residual dynamics (data efficient)
- Control Lipschitz constant (imposes compatible structure)
- Standard controller design (inherits guarantees)
- Robust regression for safe exploration (provable limited extrapolation)

Integration of Learning at Varying Levels

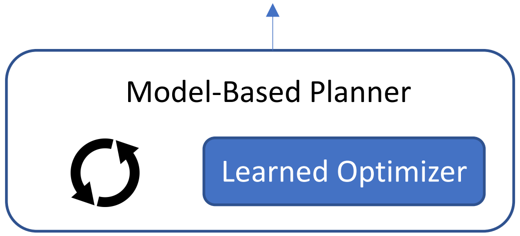
- Integration in control/output



- Integration in dynamics modeling



- Integration in optimization problem

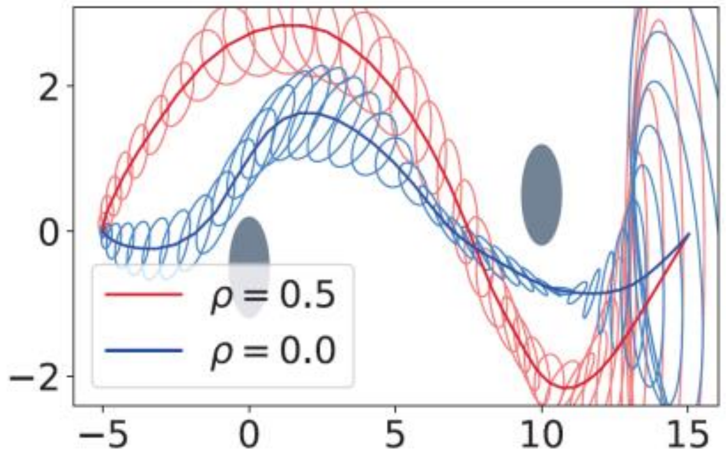


Model-Based Planning

- Environment model is given

$$s_{t+1} = F(s_t, u_t) + \epsilon$$

- Design global plan (aka trajectory)



- Satisfy global constraints
  - Previous topics only ensured local constraints
  - E.g., Lyapunov stability, smoothness

- **NP-Hard optimization problem!**

Optimization as Sequential Decision Making

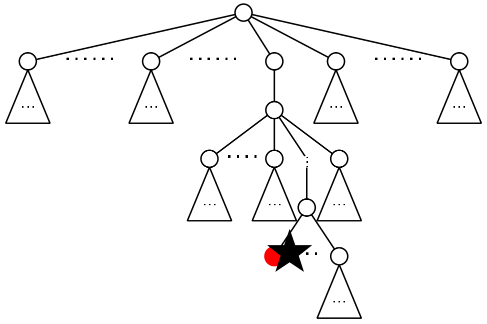
- Many Solvers are Sequential
  - Tree-Search
  - Greedy
  - Gradient Descent

- Can view solver as “agent” or “policy”
  - State = intermediate solution
  - Find a state with high reward (solution)
  - **Learn better local decision making**

Optimization as Sequential Decision Making

Learning to Search/Plan

- Discrete Optimization (Tree Search), Sparse Rewards
- **Learning to Search via Retrospective Imitation** [arXiv]
- **Co-training for Policy Learning** [UAI 2019]
- **GLAS: Global-to-Local Safe Autonomy Synthesis** [RA-L 2020]
- **A General Large Neighborhood Search Framework for Solving Integer Programs** [arXiv]

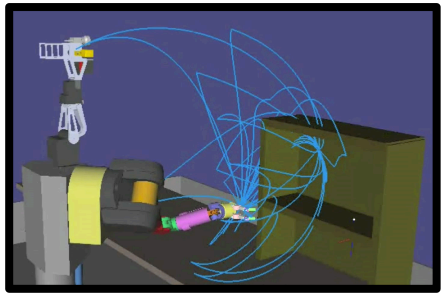


Jialin Song

Ben Riviere

Contextual Submodular Maximization

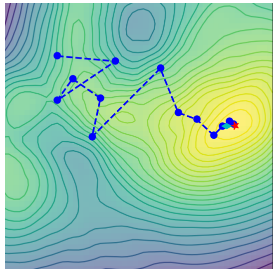
- Discrete Optimization (Greedy), Dense Rewards
- **Learning Policies for Contextual Submodular Prediction** [ICML 2013]

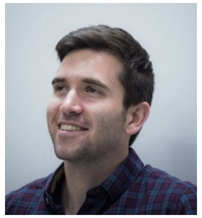


Stephane Ross

Learning to Infer

- Continuous Optimization (Gradient-style), Dense Rewards
- **Iterative Amortized Inference** [ICML 2018]
- **A General Method for Amortizing Variational Filtering** [NeurIPS 2018]



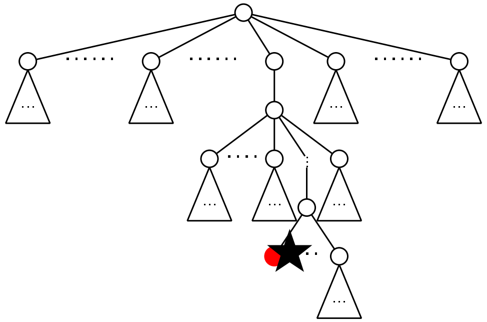


Joe Marino

Optimization as Sequential Decision Making

Learning to Search/Plan

- Discrete Optimization (Tree Search), Sparse Rewards
- **Learning to Search via Retrospective Imitation** [arXiv]
- **Co-training for Policy Learning** [UAI 2019]
- **GLAS: Global-to-Local Safe Autonomy Synthesis** [RA-L 2020]
- **A General Large Neighborhood Search Framework for Solving Integer Programs** [arXiv]

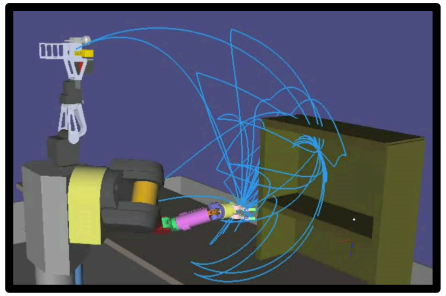


Jialin Song

Ben Riviere

Contextual Submodular Maximization

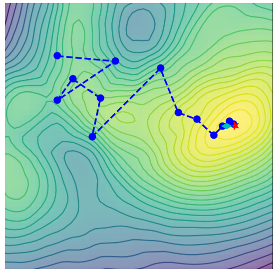
- Discrete Optimization (Greedy), Dense Rewards
- **Learning Policies for Contextual Submodular Prediction** [ICML 2013]

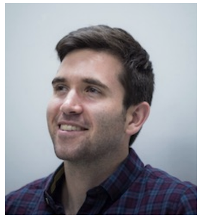


Stephane Ross

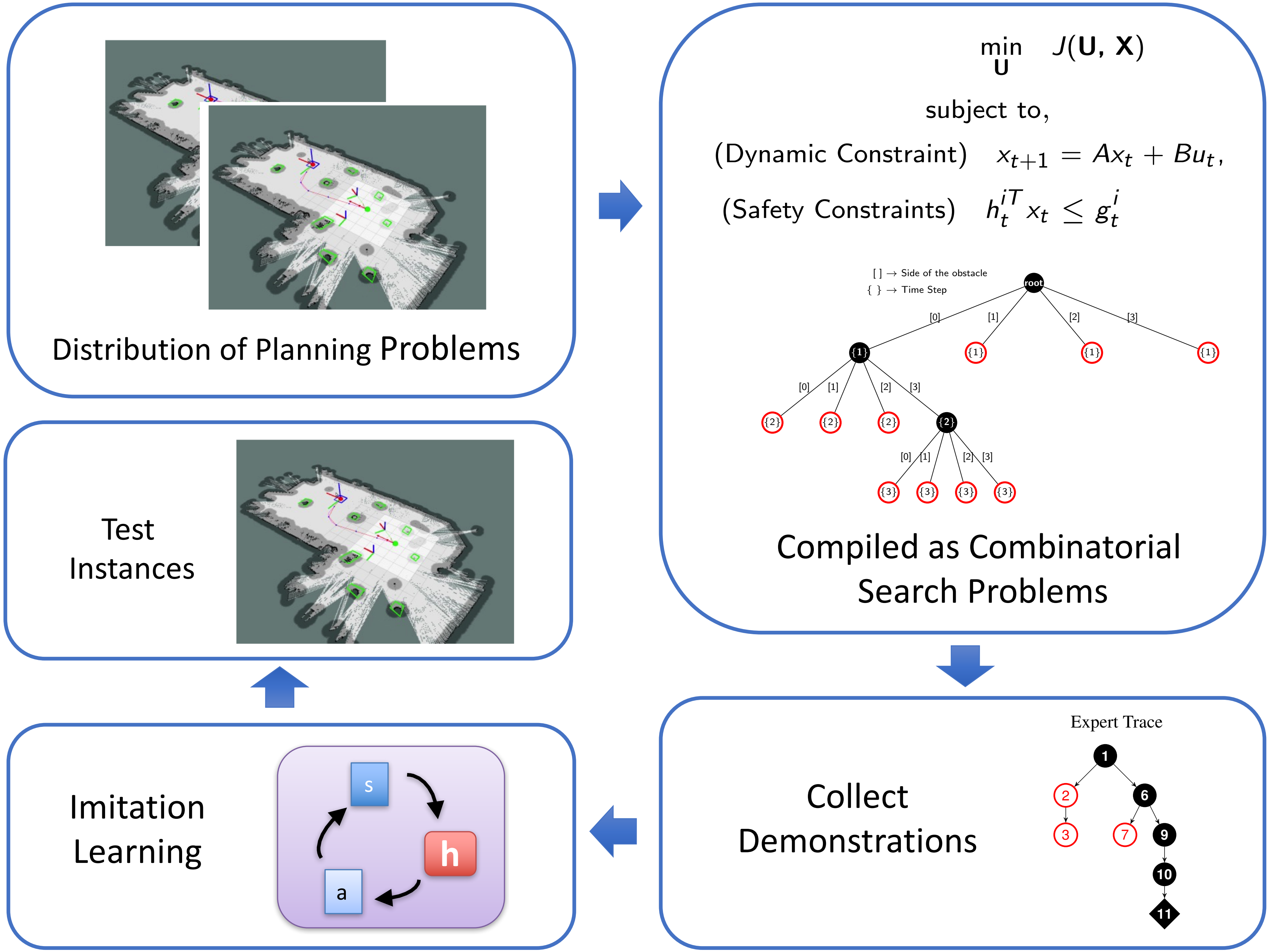
Learning to Infer

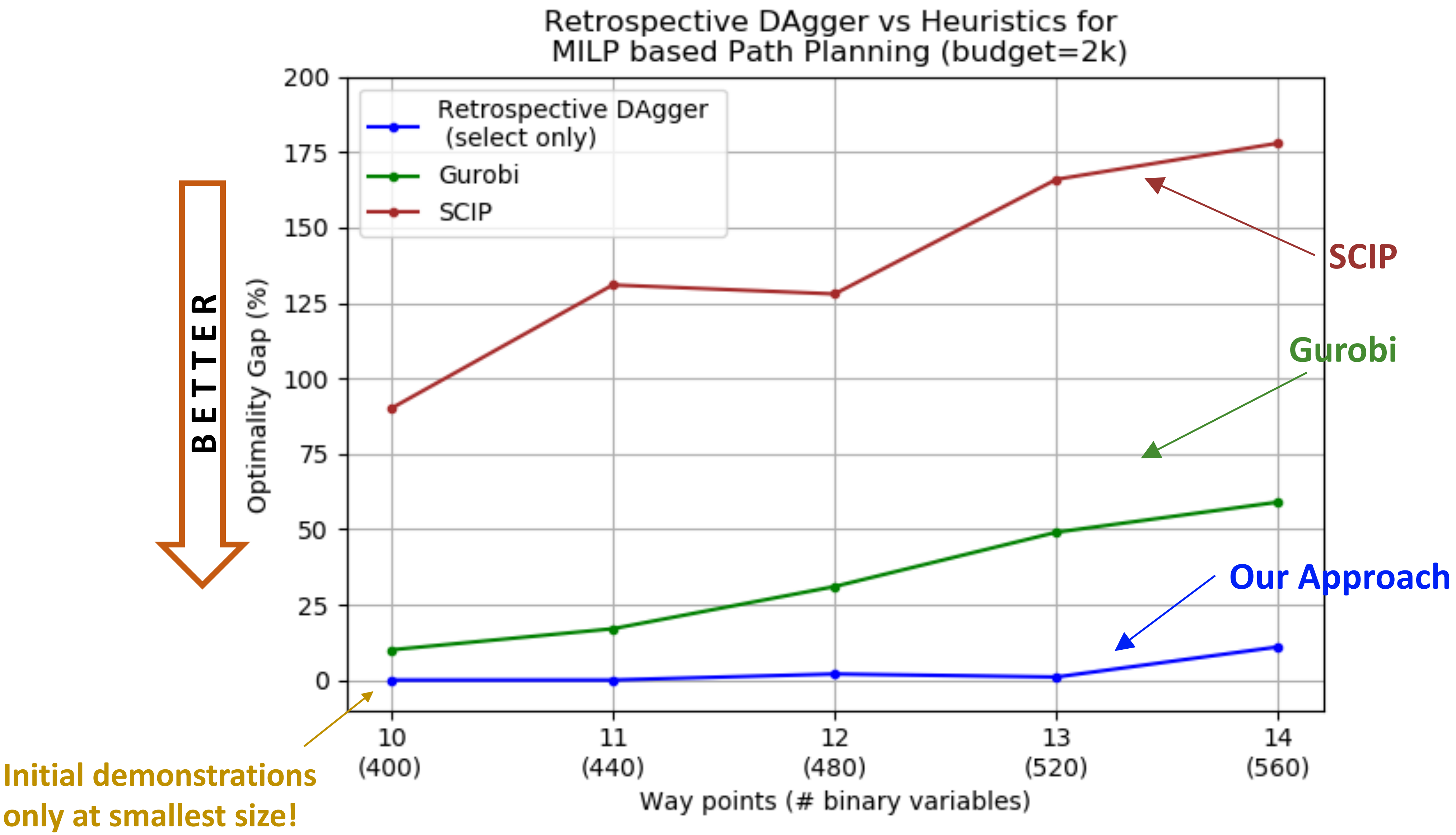
- Continuous Optimization (Gradient-style), Dense Rewards
- **Iterative Amortized Inference** [ICML 2018]
- **A General Method for Amortizing Variational Filtering** [NeurIPS 2018]





Joe Marino

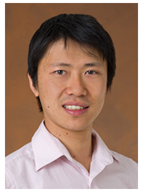

Distribution of Planning Problems
$\min_{\mathbf{U}} \quad J(\mathbf{U}, \mathbf{X})$
subject to,
(Dynamic Constraint) $x_{t+1} = Ax_t + Bu_t,$
(Safety Constraints) $h_t^{iT} x_t \leq g_t^i$
[ ] → Side of the obstacle
{ } → Time Step
root
[0]
[1]
[2]
[3]
{1}
{2}
{3}
Compiled as Combinatorial Search Problems
Collect Demonstrations
Expert Trace
1
2
3
6
7
9
10
11
Imitation Learning
s
h
a
Test Instances



Learning to Search via Retrospective Imitation, Jialin Song, Ravi Lanka, et al., arXiv

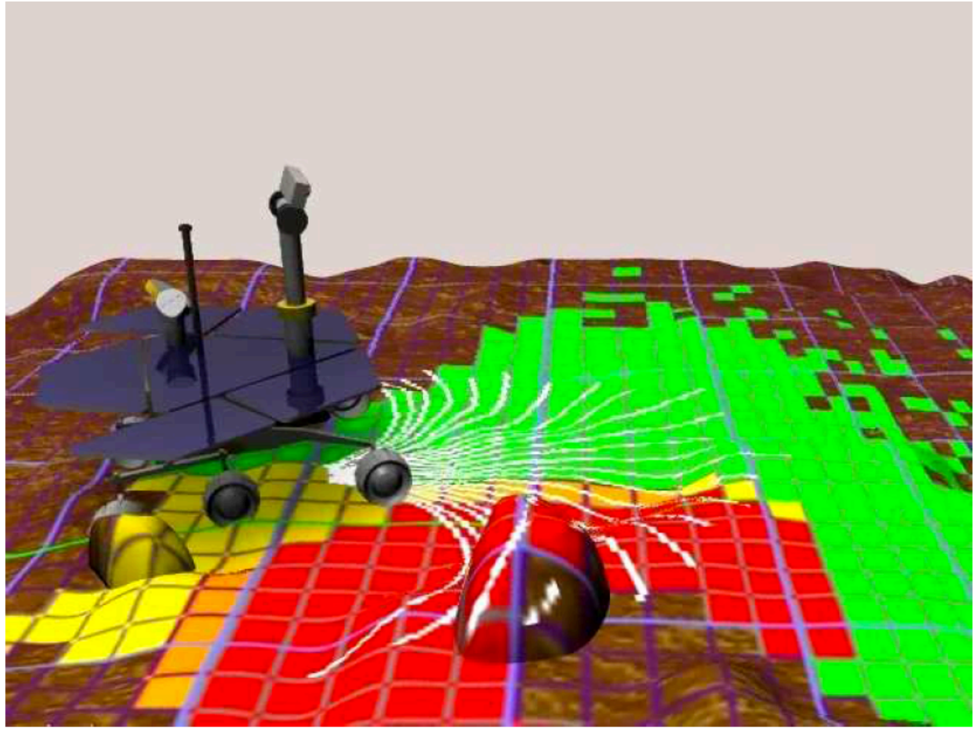
Ongoing: Integration with ENav

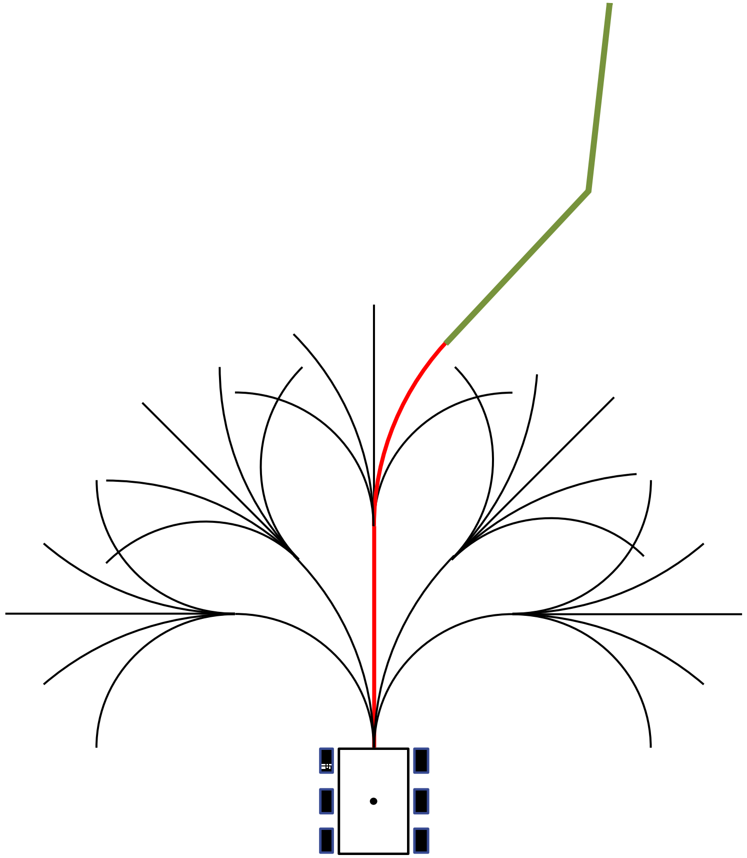

Shreyansh Daftry


Hiro Ono


Olivier Toupet


Neil Abcouwer



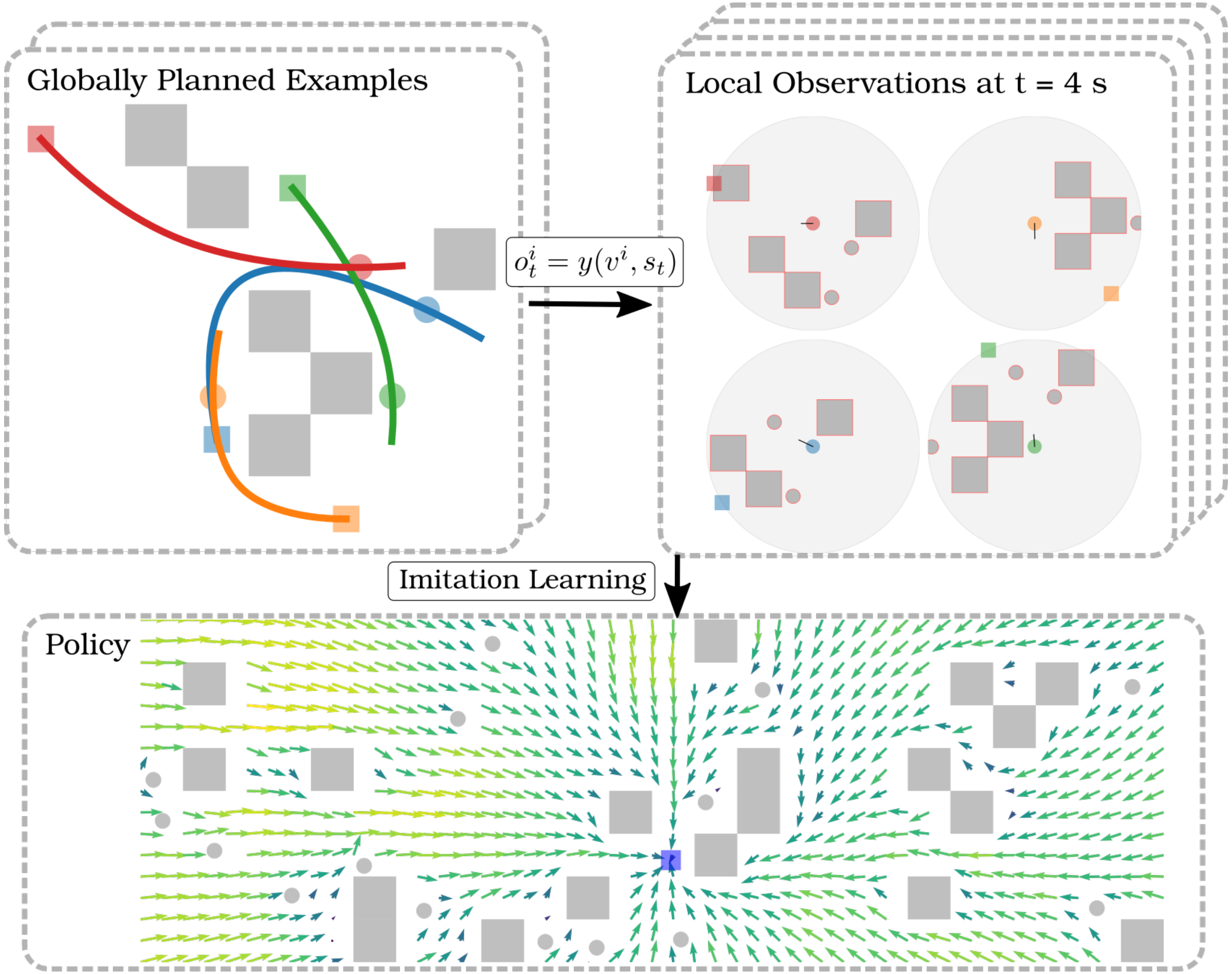


Learned Decentralized Planner

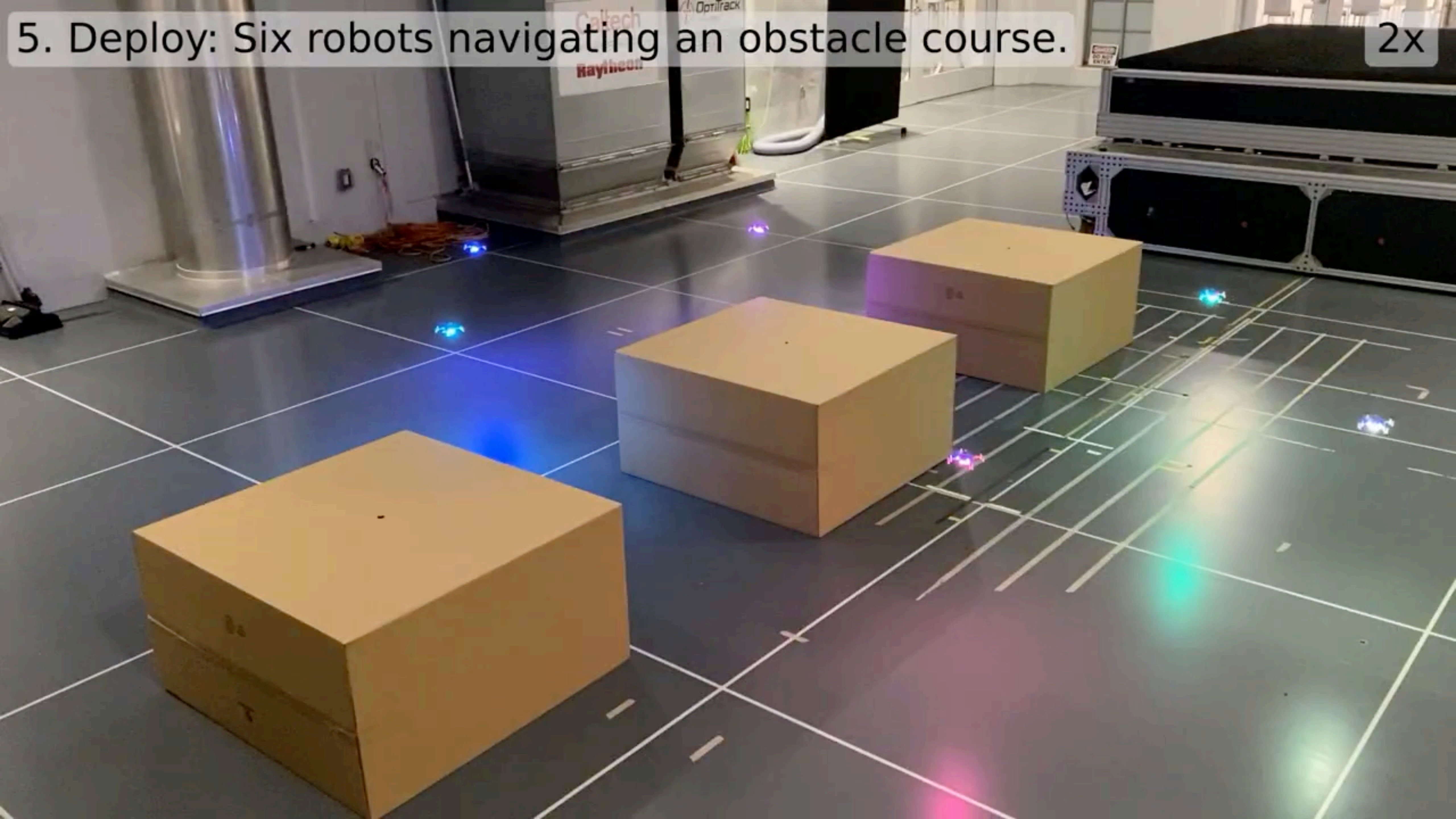
(enforcing safety)

Ben Riviere

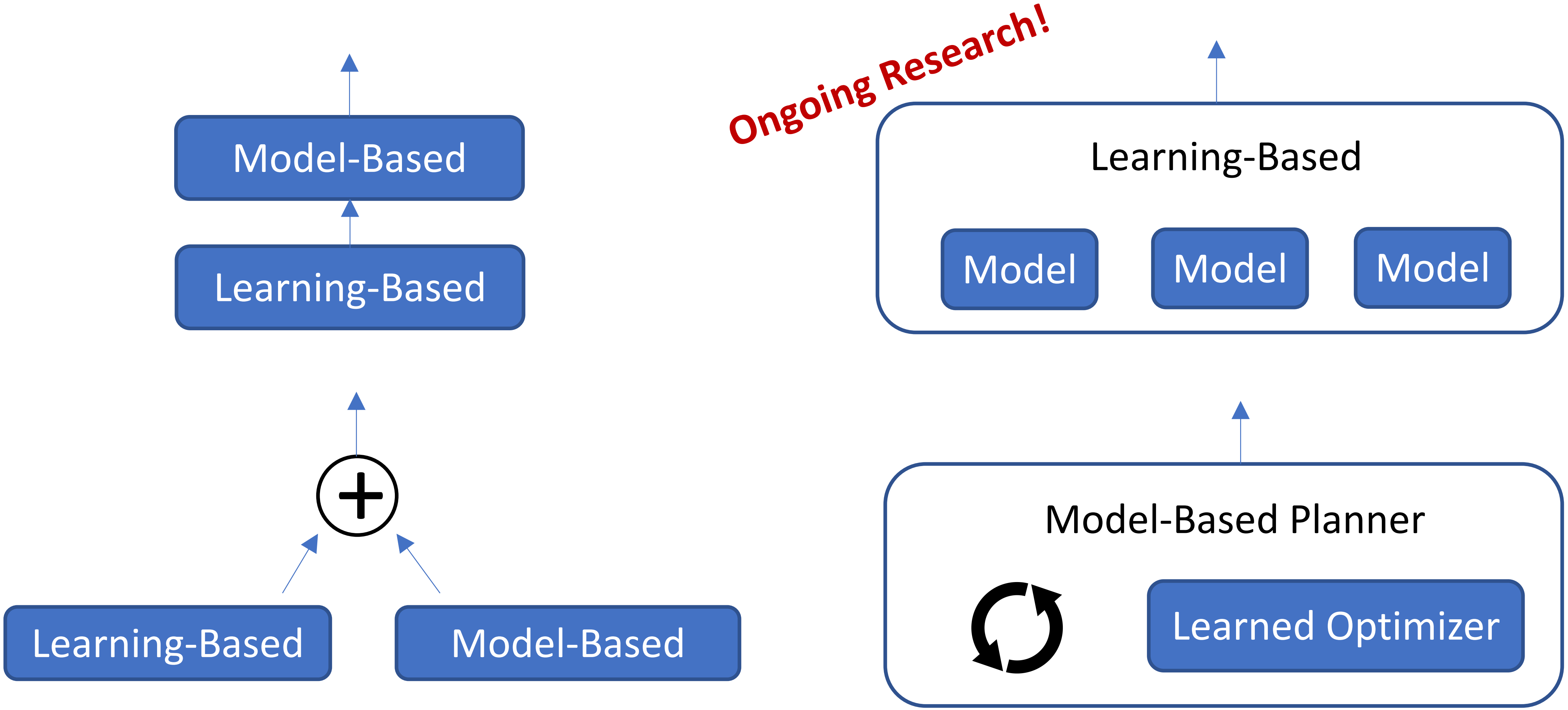
Wolfgang Hoenig



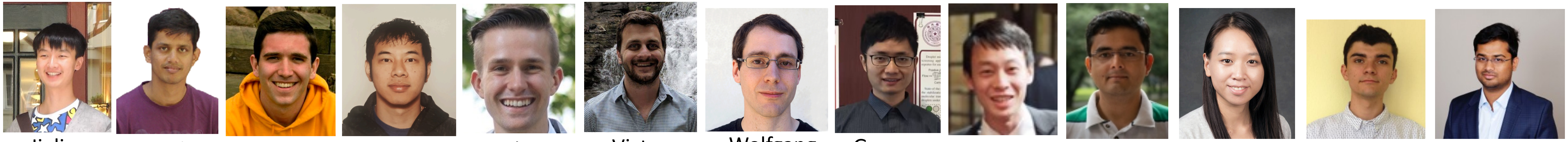
GLAS: Global-to-Local Safe Autonomy Synthesis for Multi-Robot Motion Planning with End-to-End Learning,
Benjamin Rivière, et al., R-AL 2020


5. Deploy: Six robots navigating an obstacle course.
2x

Blending Models/Rules & Black-Box Learning

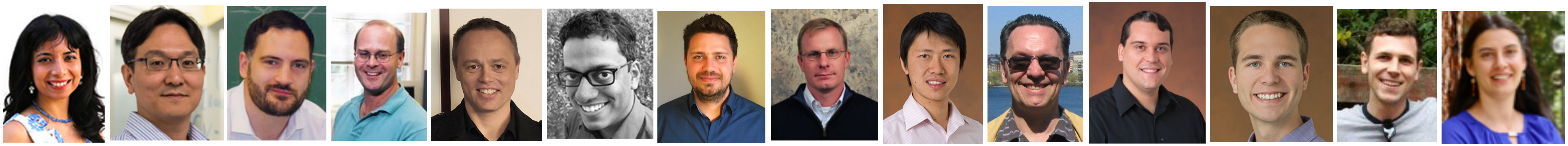


Collaborators



Jialin Song, Ravi Lanka, Joe Marino, Hoang Le, Andrew Taylor, Victor Dorobantu, Wolfgang Hoenig, Guanya Shi, Richard Cheng, Abhinav Verma, Angie Liu, Ben Riviere, Yashwanth Nakka

Cameron Voloshin, Robin Zhou, Jimmy Chen, Andrew Kang, Milan Cvitkovic, Kamyar Azizzadenesheli, Michael O'Connell, Aadyot Bhatnagar, Albert Zhao, Meera Krishnamoorthy, Debadeepta Dey, Shreyansh Daftry, Stephane Ross



Anima Anandkumar, Soon-Jo Chung, Aaron Ames, Joel Burdick, Gabor Orosz, Swarat Chaudhuri, Stephan Mandt, Drew Bagnell, Hiro Ono, Jim Little, Olivier Toupet, Neil Abcouwer, Peter Carr, Bistra Dilkina

References

Smooth Imitation Learning for Online Sequence Prediction, Hoang Le, et al., ICML 2016

Control Regularization for Reduced Variance Reinforcement Learning, Richard Cheng et al. ICML 2019

Batch Policy Learning under Constraints, Hoang Le, et al. ICML 2019

Learning Smooth Online Predictors for Real-Time Camera Planning using Recurrent Decision Trees, Jianhui Chen, et al., CVPR 2016

Imitation-Projected for Programmatic Reinforcement Learning, Abhinav Verma, Hoang Le, et al., NeurIPS 2019

Neural Lander: Stable Drone Landing Control using Learned Dynamics, Guanya Shi, et al., ICRA 2019

Neural-Swarm: Decentralized Close-Proximity Multirotor Control Using Learned Interactions, Guanya Shi et al., ICRA 2020

Robust Regression for Safe Exploration in Control, Angie Liu, Guanya Shi, et al., L4DC 2020

Chance-Constrained Trajectory Optimization for Safe Exploration and Learning of Nonlinear Systems, Yashwanth Kumar Nakka, et al. arXiv

Episodic Learning with Control Lyapunov Functions for Uncertain Robotic Systems, Andrew Taylor, Victor Dorobantu, et al., IROS 2019

A Control Lyapunov Perspective on Episodic Learning via Projection to State Stability, Andrew Taylor, Victor Dorobantu, et al., CDC 2019

Learning for Safety-Critical Control with Control Barrier Functions, Andrew Taylor, et al., L4DC 2020

A Control Barrier Perspective on Episodic Learning via Projection-to-State Safety, Andrew Taylor, et al., L-CSS 2020

Learning to Search via Retrospective Imitation, Jialin Song, Ravi Lanka, et al., arXiv

Co-Training for Policy Learning, Jialin Song, Ravi Lanka, et al., UAI 2019

A General Large Neighborhood Search Framework for Solving Integer Programs, Jialin Song, Ravi Lanka, et al., arXiv

GLAS: Global-to-Local Safe Autonomy Synthesis for Multi-Robot Motion Planning with End-to-End Learning, Benjamin Rivière, et al., R-AL 2020

Learning Policies for Contextual Submodular Optimization, Stephane Ross et al., ICML 2013

Iterative Amortized Inference, Joe Marino et al., ICML 2018

A General Framework for Amortizing Variational Filtering, Joe Marino et al, NeurIPS 2018