Caltech

Learning for Reliable Control in Dynamical Systems

Yisong Yue

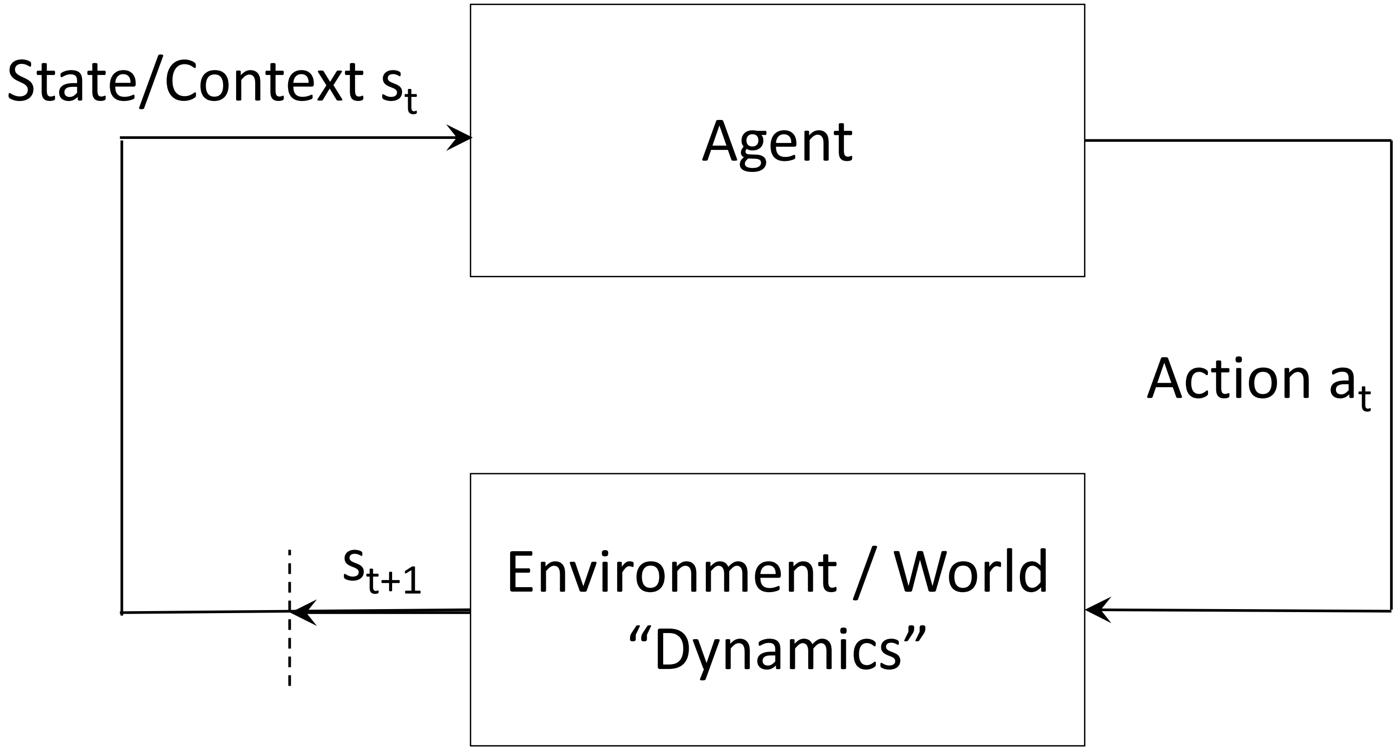
Policy/Controller Learning (Reinforcement & Imitation)

Goal: Find "Optimal" Policy

Imitation Learning:
Optimize imitation loss

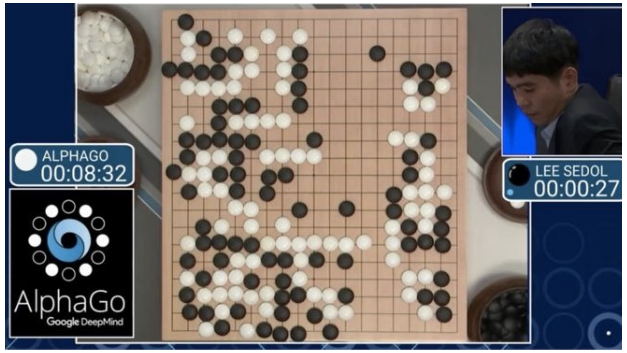
Reinforcement Learning:
Optimize environmental reward

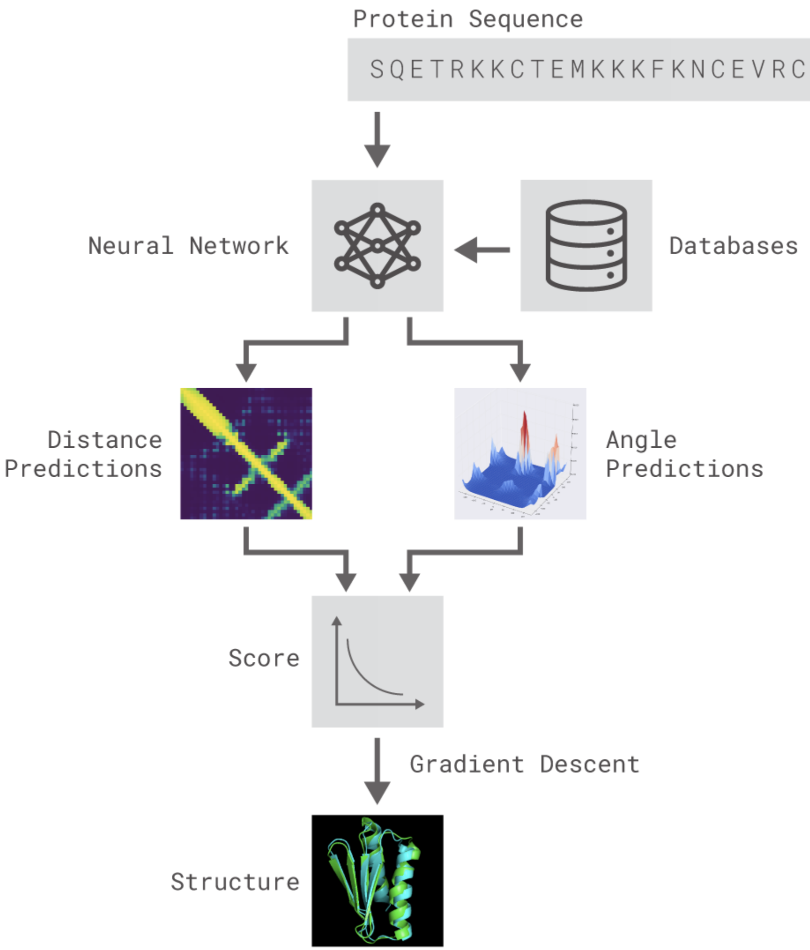
Learning-based Approach for Sequential Decision Making



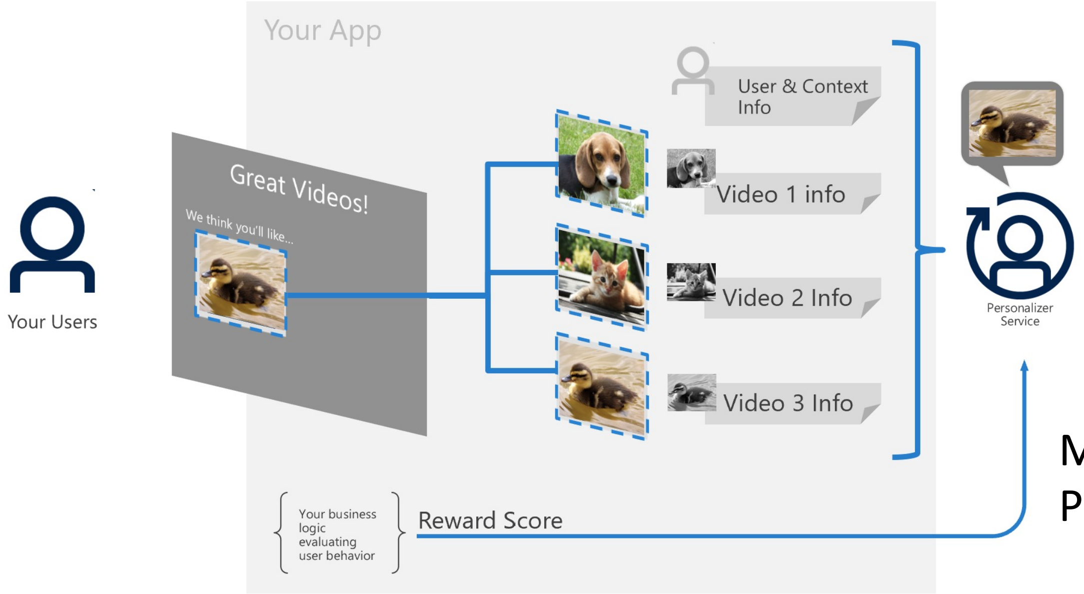
Non-learning approaches include: optimal control, robust control, adaptive control, etc.

Many Exciting Success Stories





AlphaFold



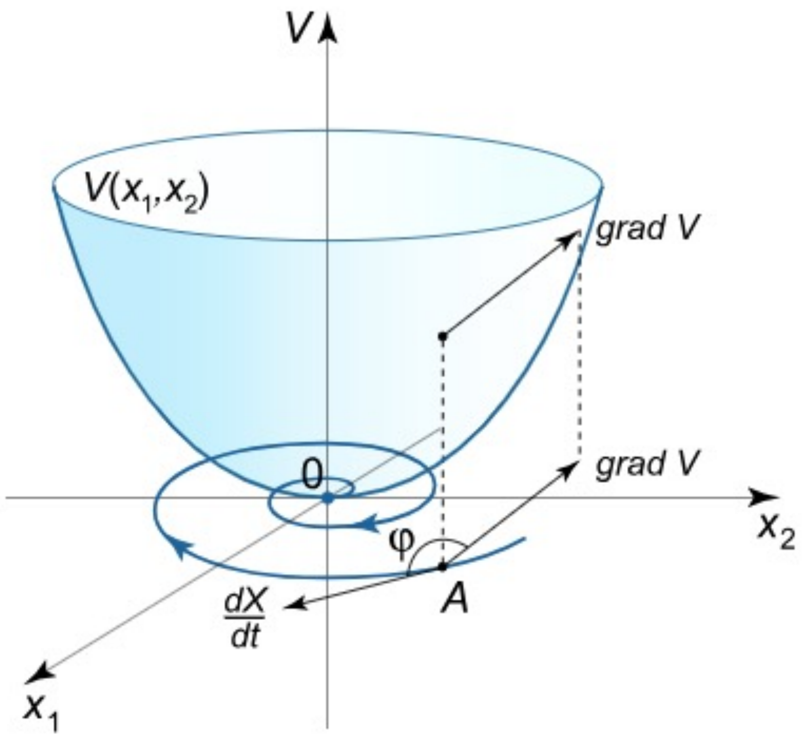
Microsoft Azure Personalizer

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

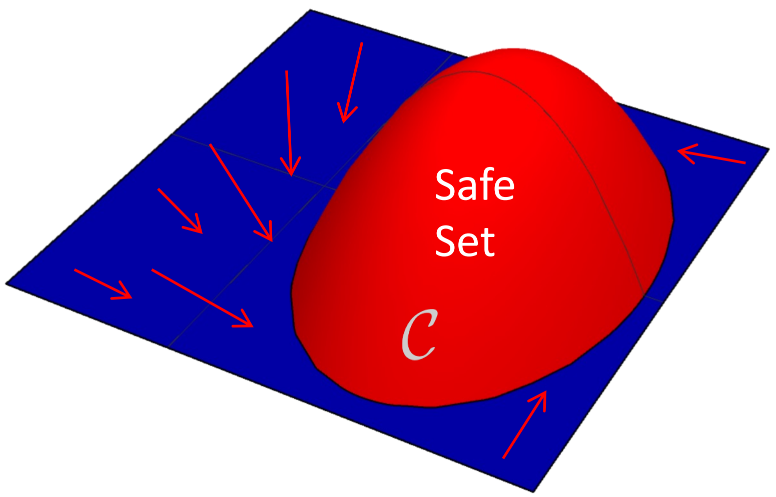
Behavioral Guarantees

Possibly Others:

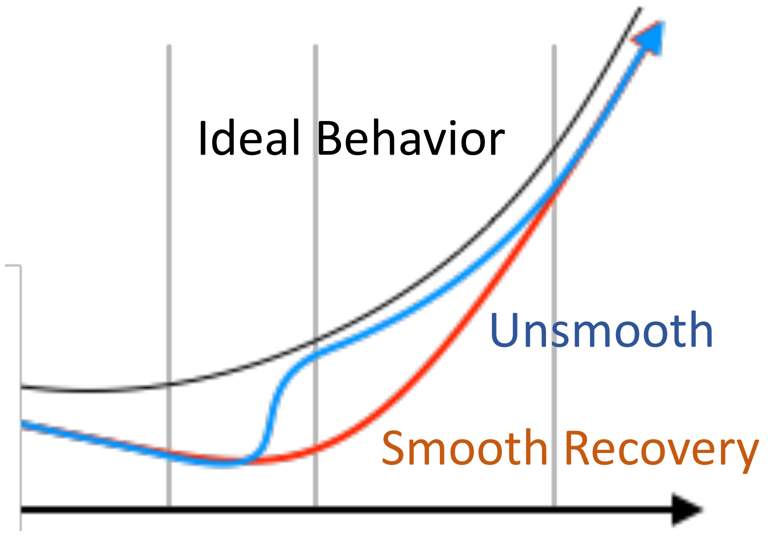
- Fairness
- Low-risk
- Temporal logic
- Etc…



Stability

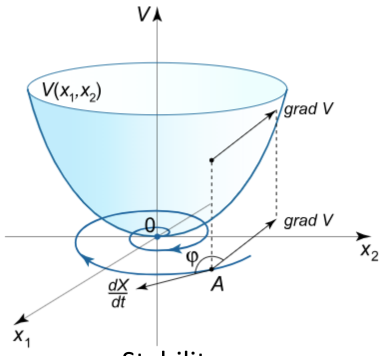


Safety

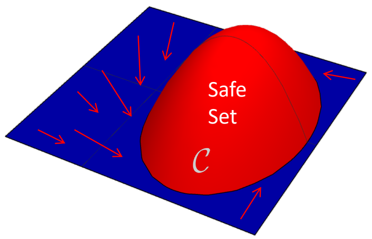


Smoothness

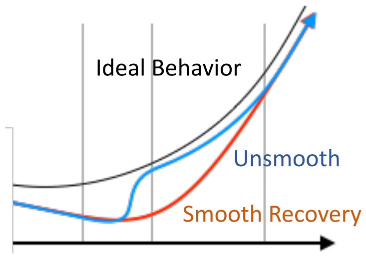
Research Questions



Stability



Safety

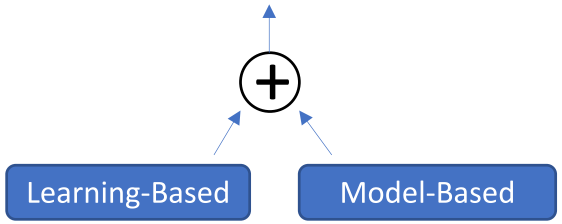


Smoothness

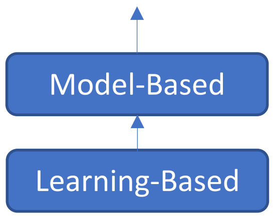
- How to constrain learning to (provably) satisfy guarantees?
- How to integrate domain knowledge from physics & control theory?
    - (Towards) a unified framework?
- How to exploit structure for faster learning?
    - (both computational & statistical)

Integration of Learning at Varying Levels

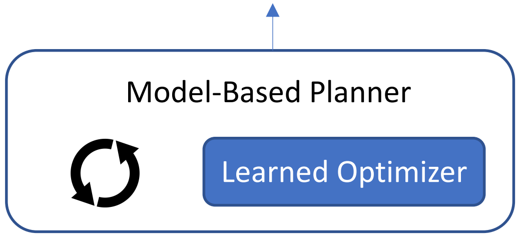
- Integration in control/action



- Integration in dynamics modeling



- Integration in optimization problem



Starting Point

Standard IL/RL Objective

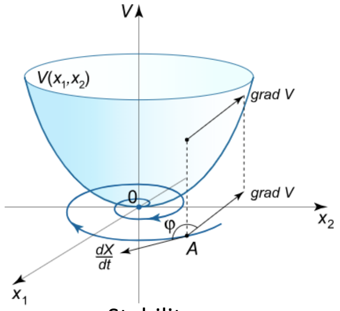
- Model-Based/Free
- On/Off Policy
- Imitation/Reinforcement
- Optimal Control

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

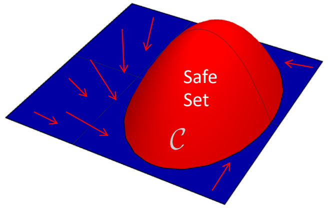
In general, very hard to verify/optimize!

(e.g., for all inputs, h is safe)

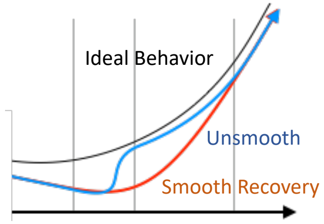
Side Constraint



Stability



Safety



Smoothness

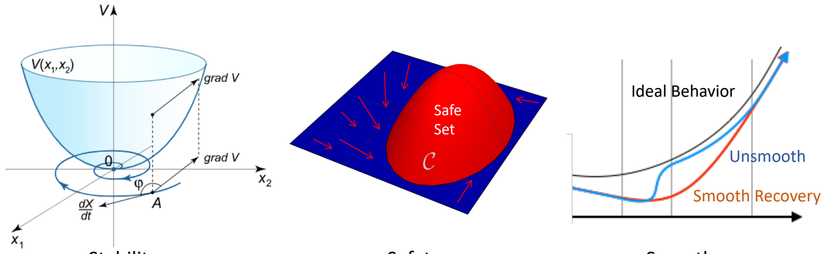
Functional Regularization
(to a certified controller)

$$argmin_h L(h)$$
$$s.t.$$
$$\exists g \in G : \|h - g\|^2 < \kappa$$

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

Intractable?

Model-Based Controllers
(certified by construction)



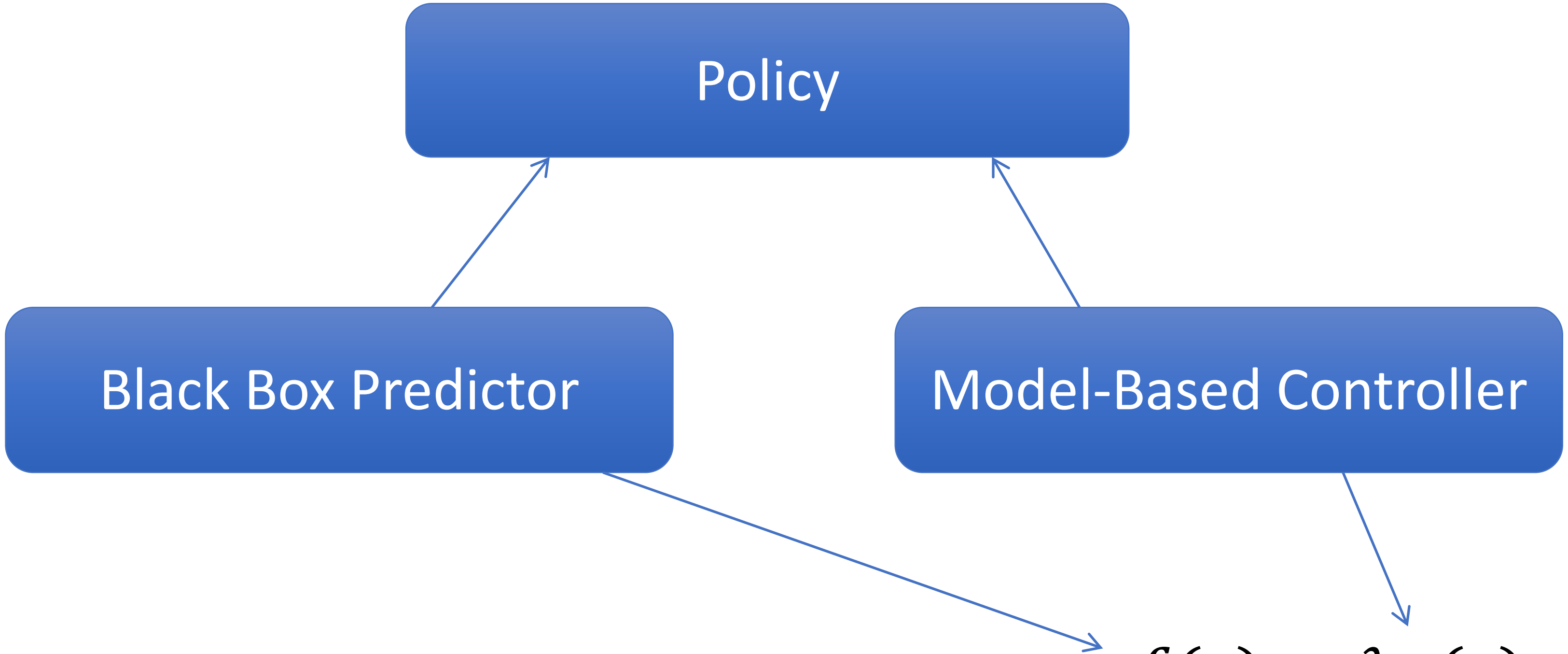
Stability Safety Smoothness

Key idea: G encodes domain knowledge & guarantees

Blended Policy Class (solution concept)

Hoang Le

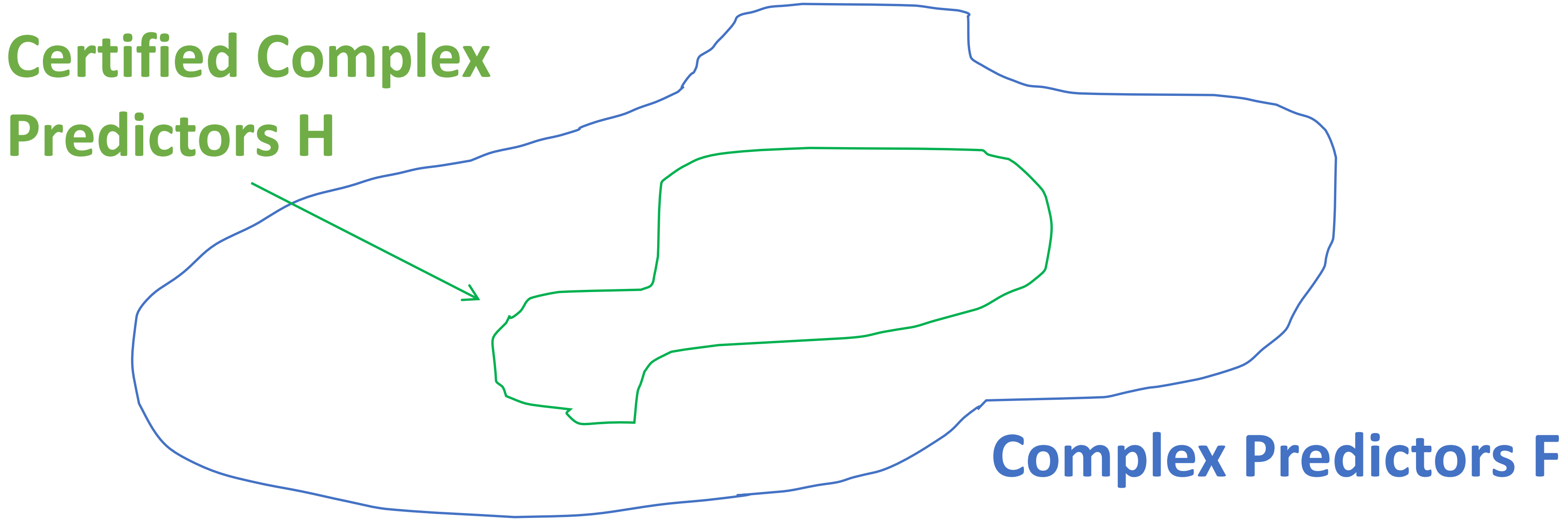
Richard Cheng



$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = \frac{f(s) + \lambda g(s)}{1+\lambda}$$

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

$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = \frac{f(s) + \lambda g(s)}{1 + \lambda}$$

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

Theoretical Guarantees

$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = \frac{f(s) + \lambda g(s)}{1+\lambda}$$

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

Run-time regularization

- Compatible with IL/RL
  - New learning approaches

Convergence analysis

- Very data efficient

Low-Variance Gradients

Comments on Certified by Construction

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

- **Assumption:** all $g \in G$ are certified by construction
  - Robust against disturbances
  - Satisfied for many physical systems
- **Disturbance:** $f \in F$ is a "disturbance" of g
  - Worst-case disturbance depends $\max_{s} f(s)$ and λ
  - Guarantees worsen as λ decreases
- **Note:** local per-state guarantee => global guarantee

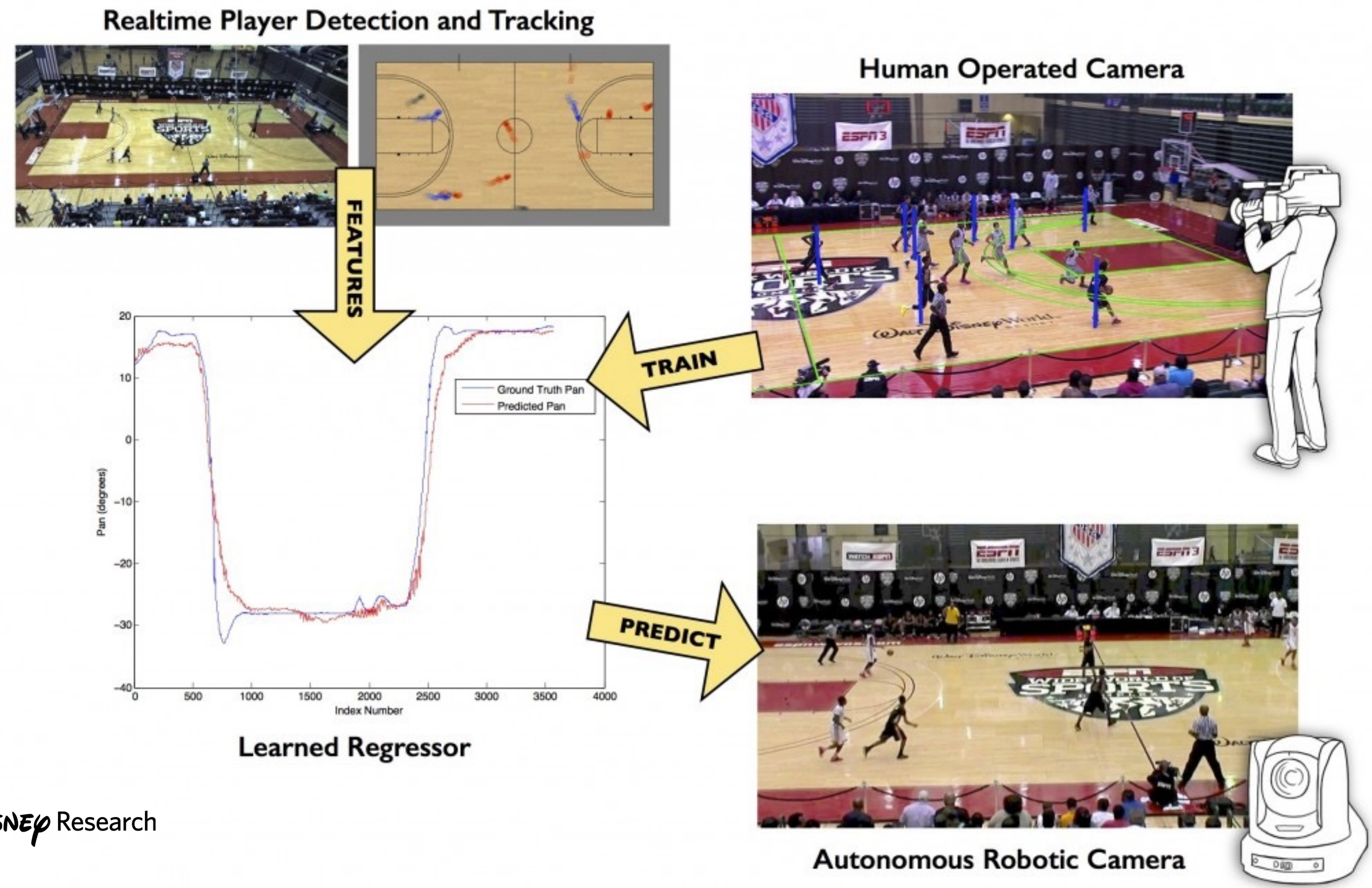
Comments on Optimization/Learning

$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = \frac{f(s) + \lambda g(s)}{1 + \lambda}$$

- Alternating optimization
  - Hold g fixed, optimize f
  - Hold h fixed, optimize g
  - (see NeurIPS 2019 paper for clean treatment)

Reduces to "standard" approaches

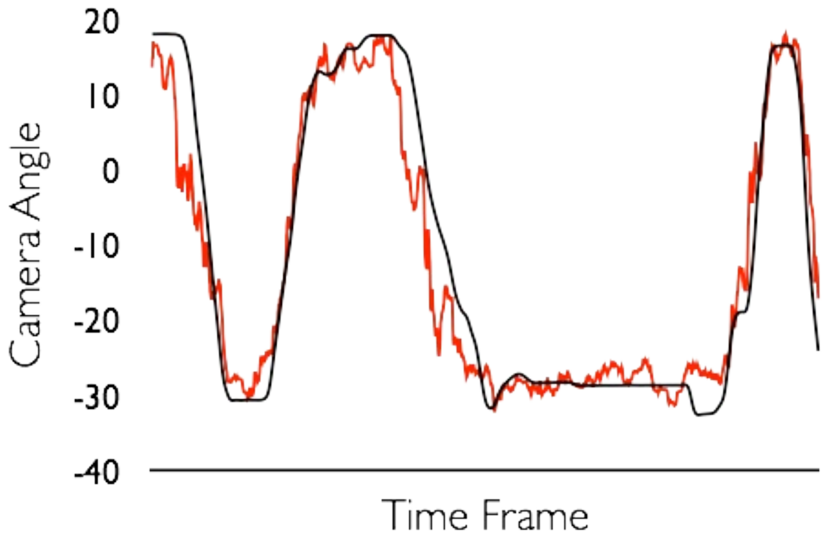
Imitation-Projected Programmatic Reinforcement Learning
Abhinav Verma, Hoang Le, Yisong Yue, Swarat Chaudhuri. NeurIPS 2019



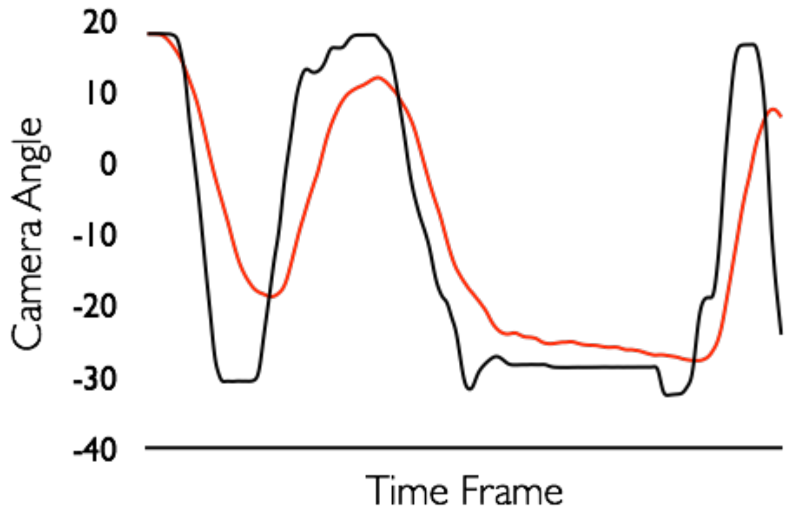
Disney Research

Naïve Approach

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

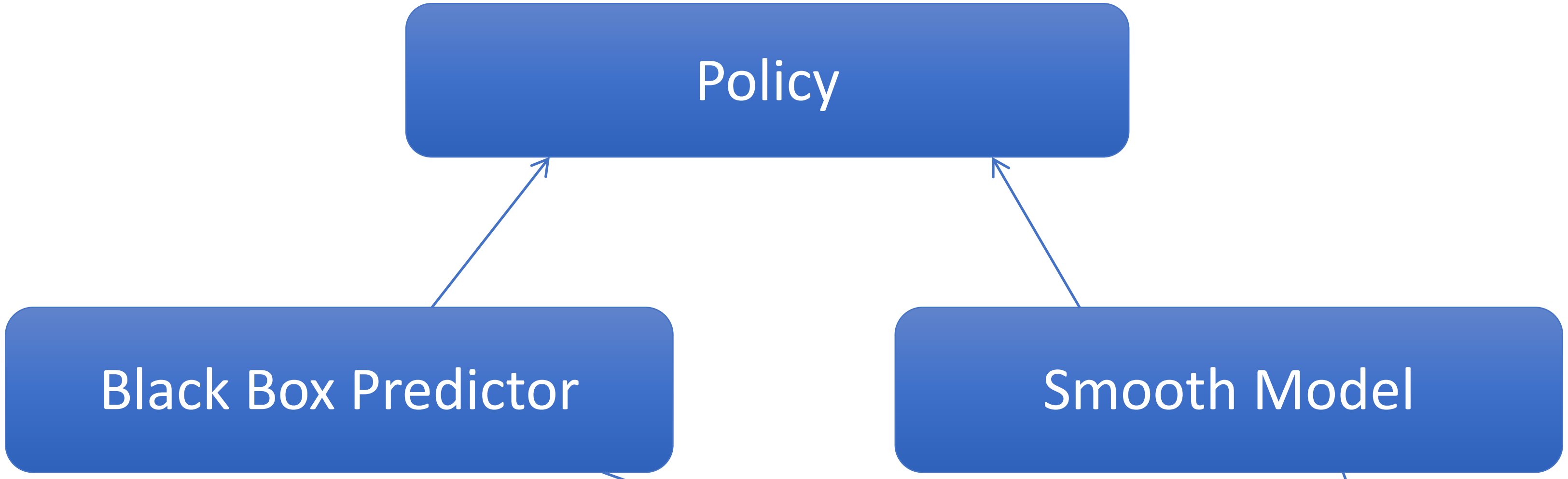


In practice, 2-step smoothing:



Smooth Policy Class

Hoang
Le

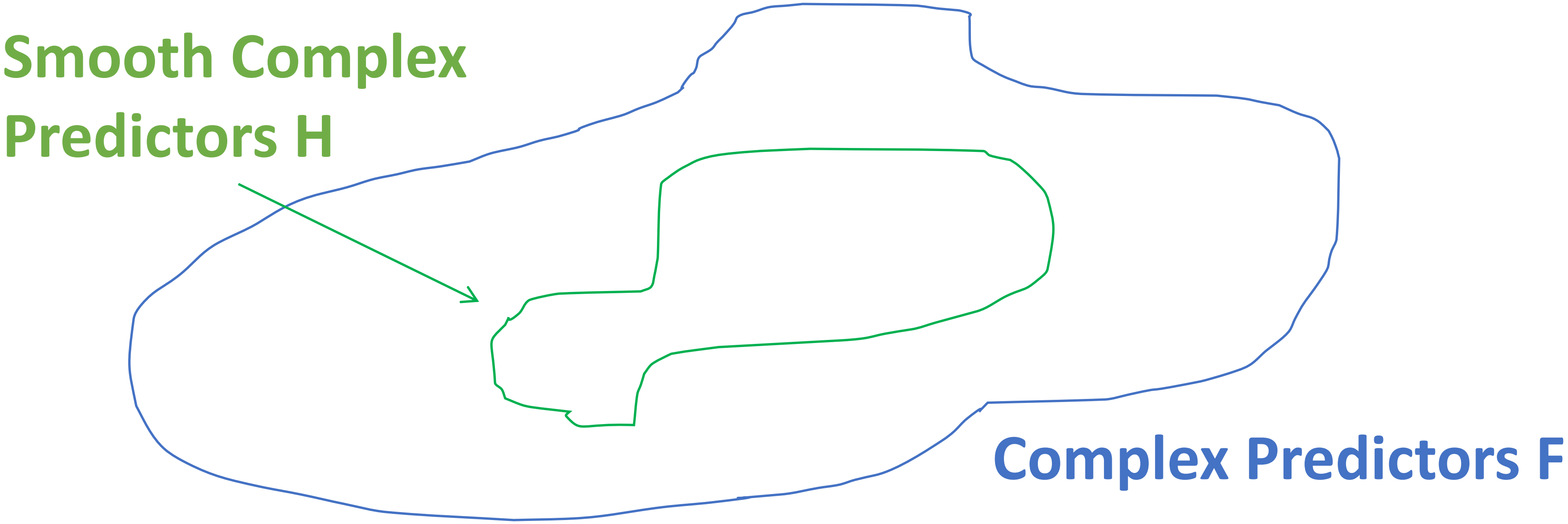


$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = \frac{f(s) + \lambda g(s)}{1 + \lambda}$$

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

Test-Time Functional Regularization

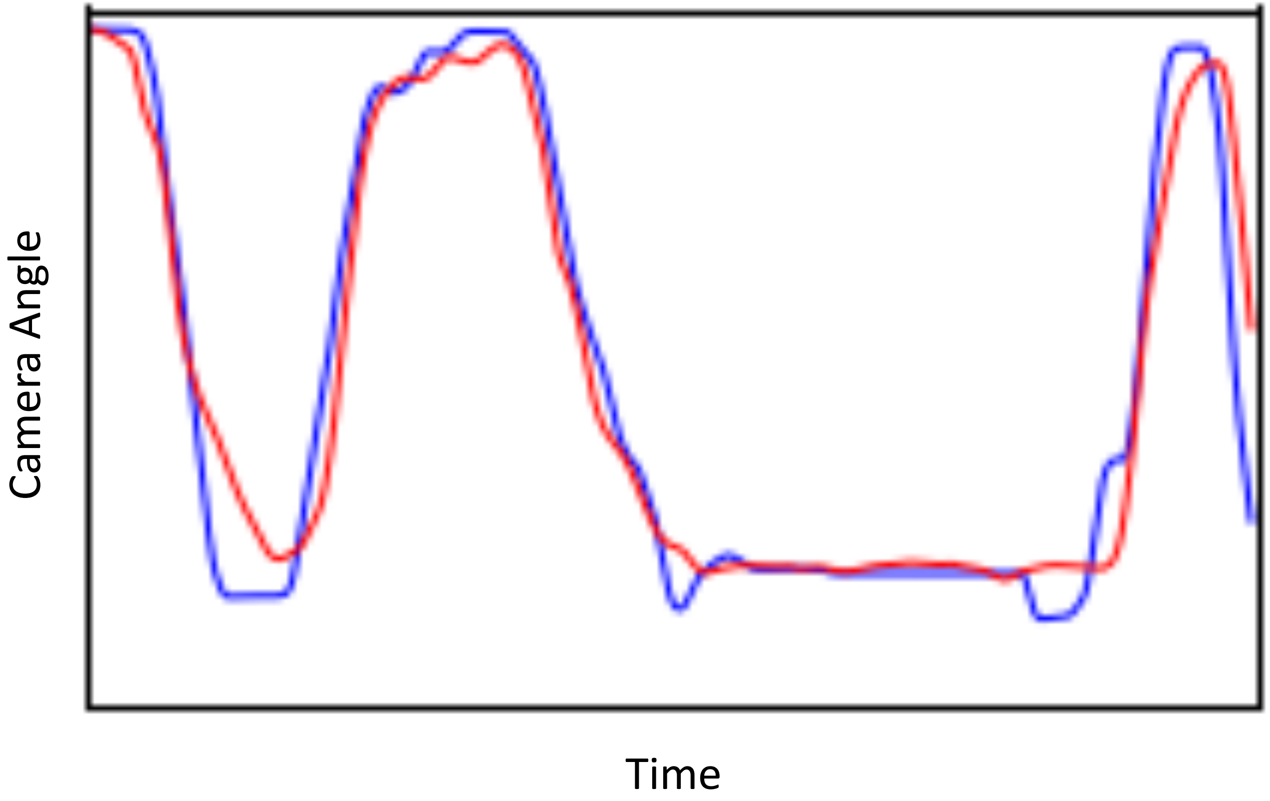
Hoang Le



$$argmin_{h=(f,g)} L(h) \quad \text{s.t.} \quad h(s) = \frac{f(s) + \lambda g(s)}{1 + \lambda}$$

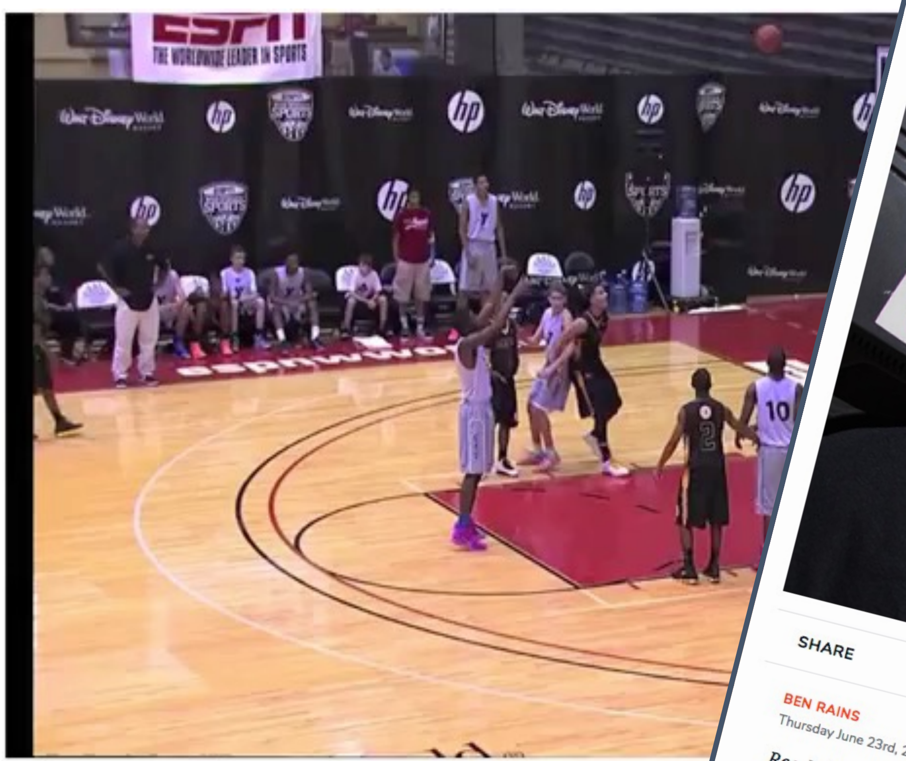
Smooth Imitation Learning for Online Sequence Prediction
Hoang Le, Andrew Kang, Yisong Yue, Peter Carr. ICML 2016

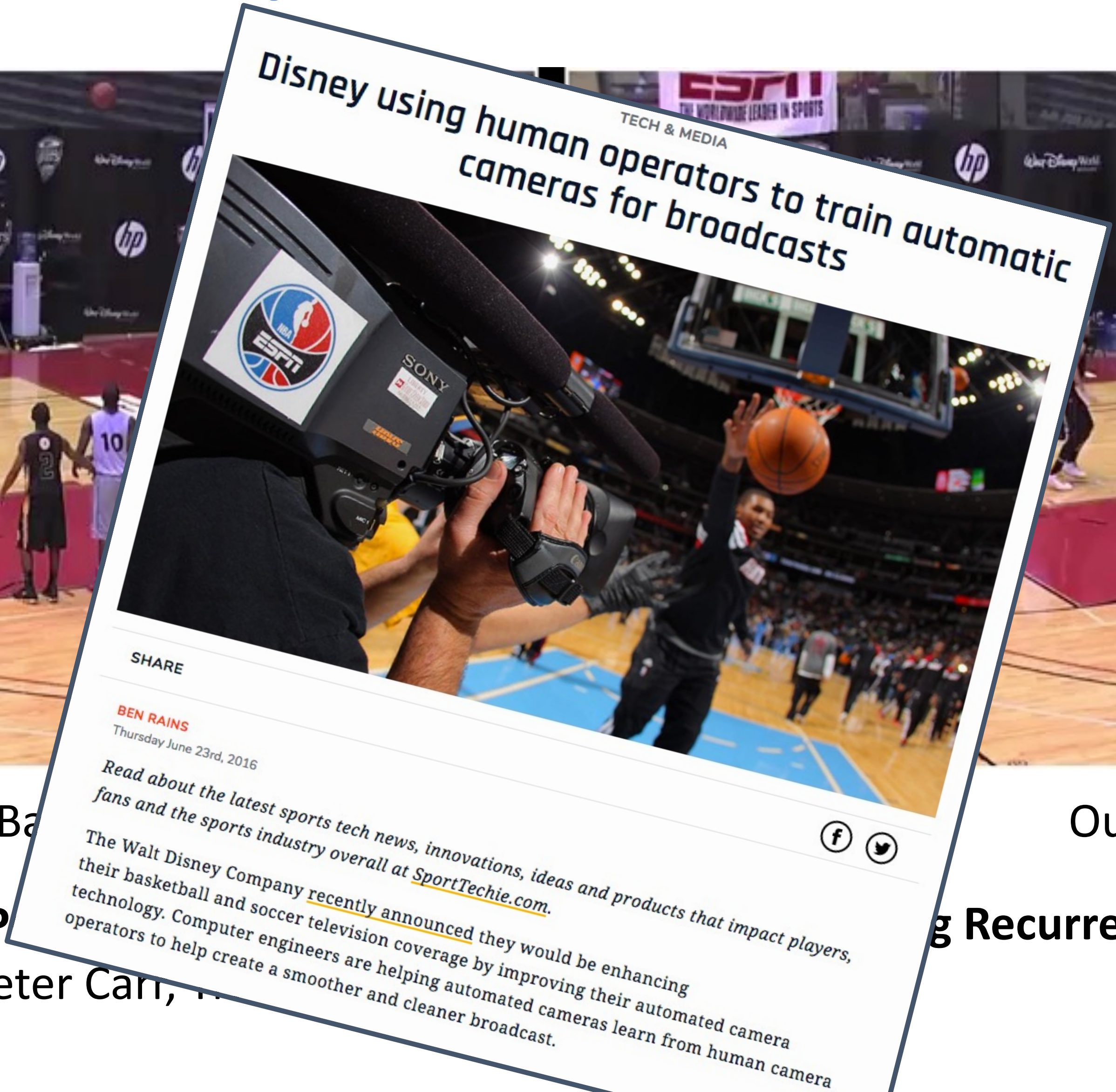
Our Results



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.

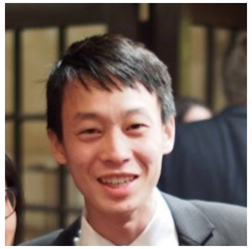
2-Step Ba

Our Approach

Learning Online Smooth P **g Recurrent Decision Trees**

Jianhui Chen, Hoang Le, Peter Carr,

Control Regularization



Richard Cheng

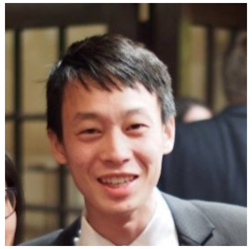
$$h(s) = \frac{f(s)+\lambda g(s)}{1+\lambda}$$

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

- Learn f using any RL method

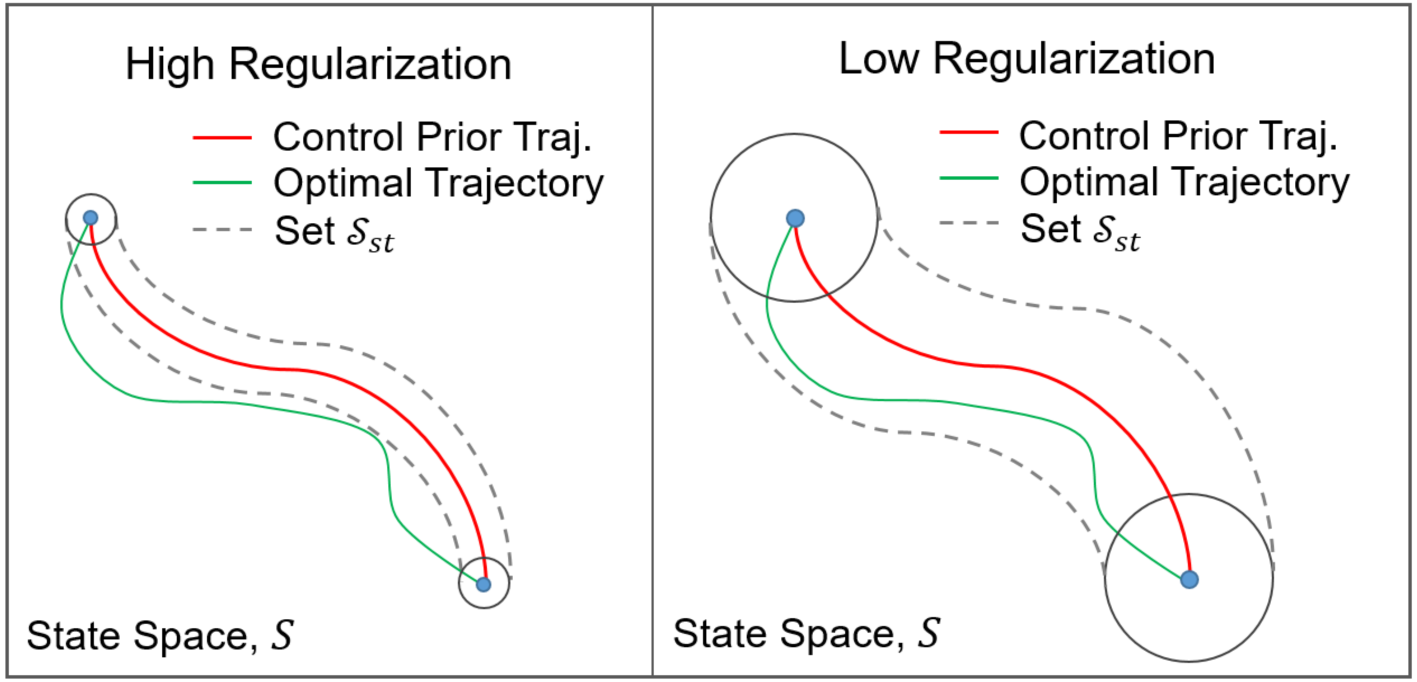
Control Regularization for Reduced Variance Reinforcement Learning
Richard Cheng, Abhinav Verma, Gabor Orosz, Swarat Chaudhuri, Yisong Yue, Joel Burdick. ICML 2019

Control Regularization



Richard Cheng

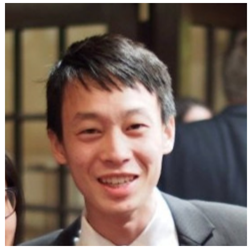
- (Relaxed) Lyapunov stability bounds:



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

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

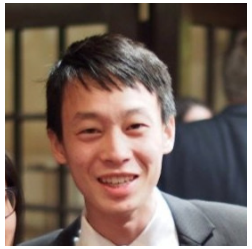
- Theorem (informal):
  - Variance of policy gradient decreases by factor of: $\left(\frac{1}{1+\lambda}\right)^2$

Implies much faster learning!

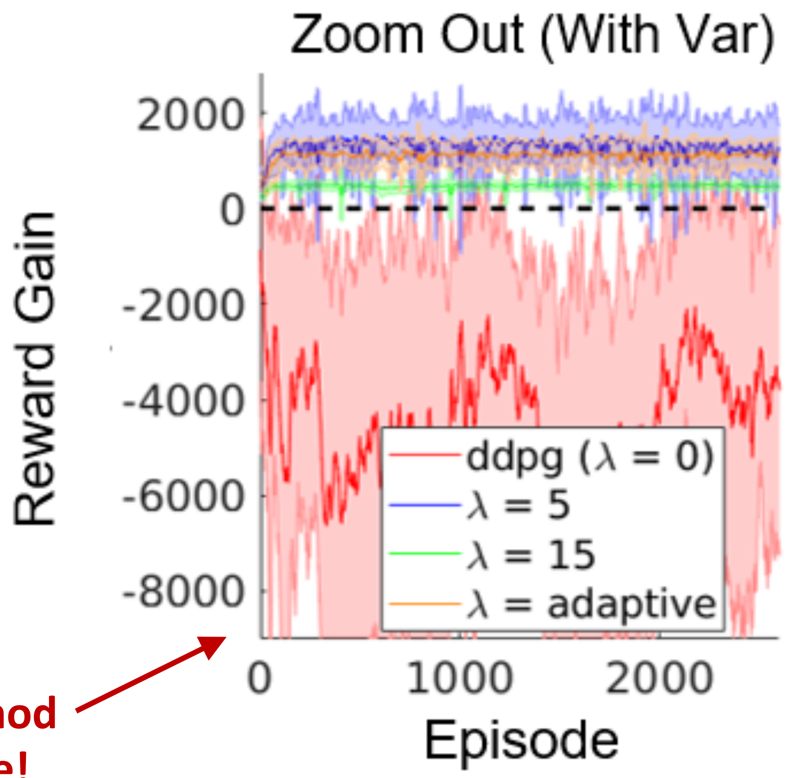
  - Bias converges to: $\left(\frac{\lambda}{1+\lambda}\right) D_{TV}(h^*, g)$

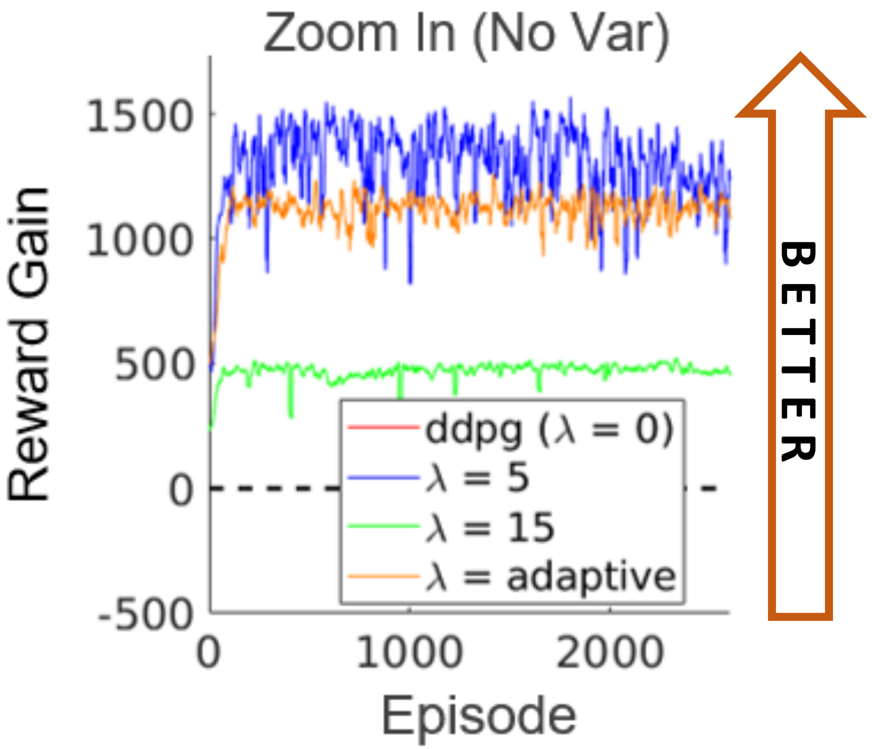
Control Regularization for Reduced Variance Reinforcement Learning
Richard Cheng, Abhinav Verma, Gabor Orosz, Swarat Chaudhuri, Yisong Yue, Joel Burdick. ICML 2019

Control Regularization



Richard Cheng





Baseline RL Method has High Variance!

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

Regularization ↔ Constrained Learning

IL/RL Objective

$$argmin_h L(h) \quad s.t. \quad R(h) < \kappa \quad \Longleftrightarrow \quad argmin_h L(h) + \lambda R(h)$$

Side Guarantees

Hybrid Policy Solution Concept

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

Summary: Functional Regularization (cont.)

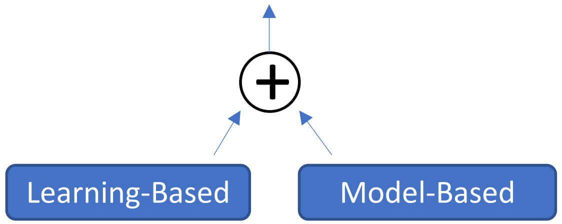
- Control methods => analytic guarantees (side guarantees)
- Blend w/ learning => improve precision/flexibility (real-world improvements)
- Preserve behavioral guarantees (possibly relaxed)
- Interpret as functional regularization (speeds up learning)
- Other directions:

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

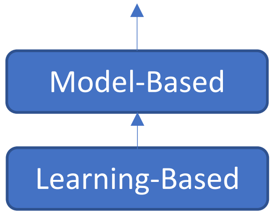
Imitation-Projected Programmatic Reinforcement Learning
Abhinav Verma, Hoang Le, Yisong Yue, Swarat Chaudhuri. NeurIPS 2019 (neurosymbolic policies)

Integration of Learning at Varying Levels

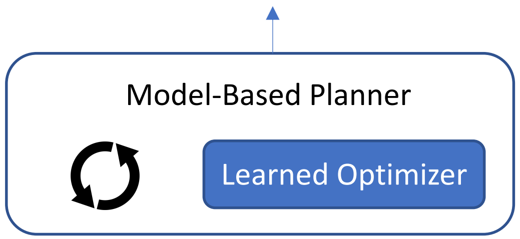
- Integration in control/action



- 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/Optimal 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)

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

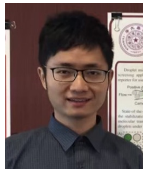
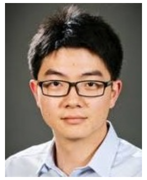
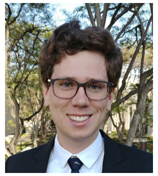
Current State

Unmodeled Disturbance / Error

Leverage robust/optimal control (fancy contraction mappings)

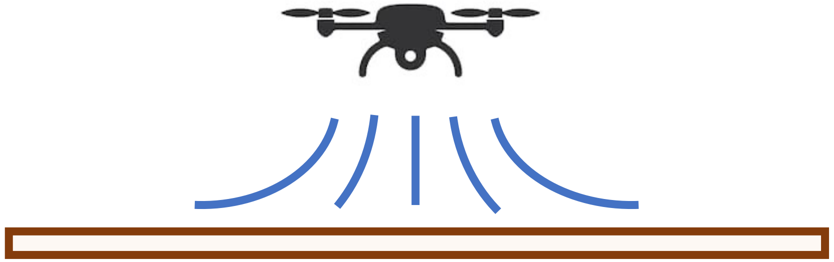
- Preserve stability (even using deep learning)
- Requires $\tilde{F}$ Lipschitz & bounded error

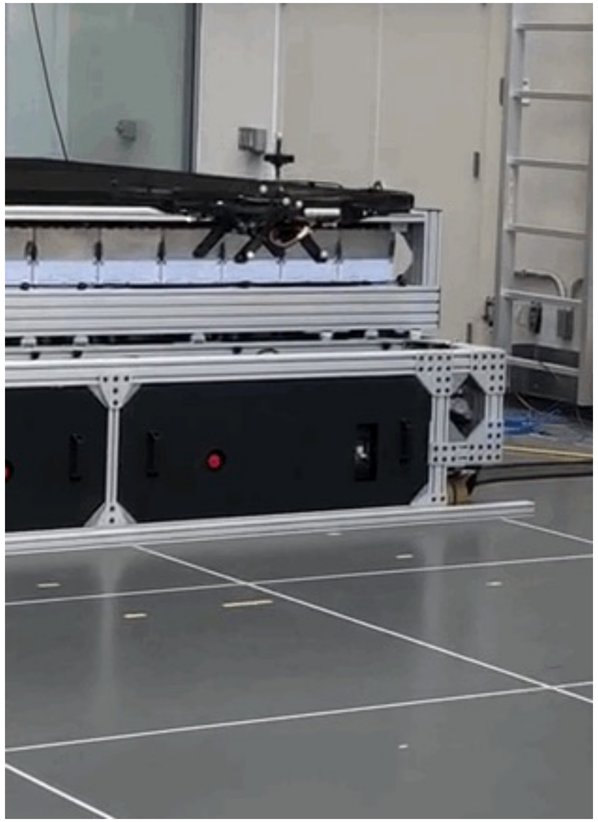
Boundary Conditions

Guanya Shi | Xichen Shi | Michael O'Connell

Ground effect





Neural Lander: Stable Drone Landing Control using Learned Dynamics, Guanya Shi, Xichen Shi, Michael O'Connell, et al. ICRA 2019
Neural-Swarm: Decentralized Close-Proximity Multirotor Control Using Learned Interactions, Guanya Shi et al., ICRA 2020
Neural-Swarm2: Planning and Control of Heterogeneous Multirotor Swarms using Learned Interactions, Guanya Shi et al., T-RO 1021

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:

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

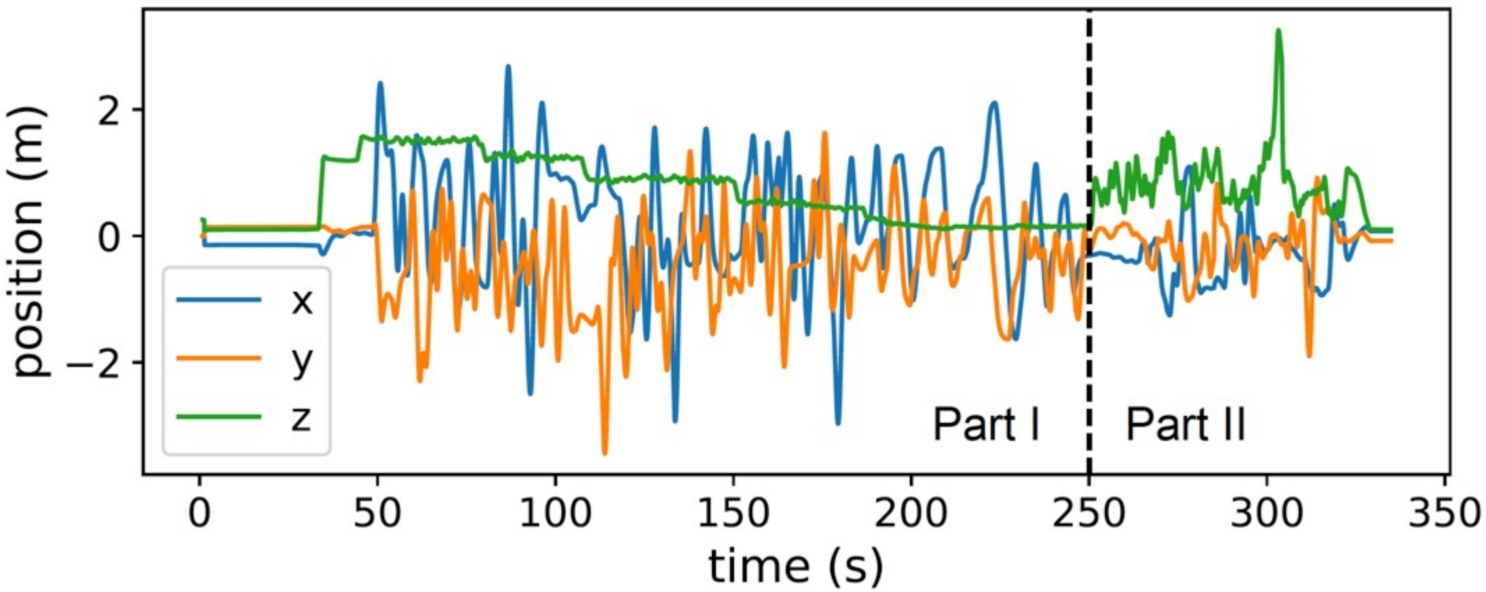
- Unknown forces & moments:

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

Learn the Residual

Data Collection (Manual Exploration)

Notable Extension:
Safe Exploration

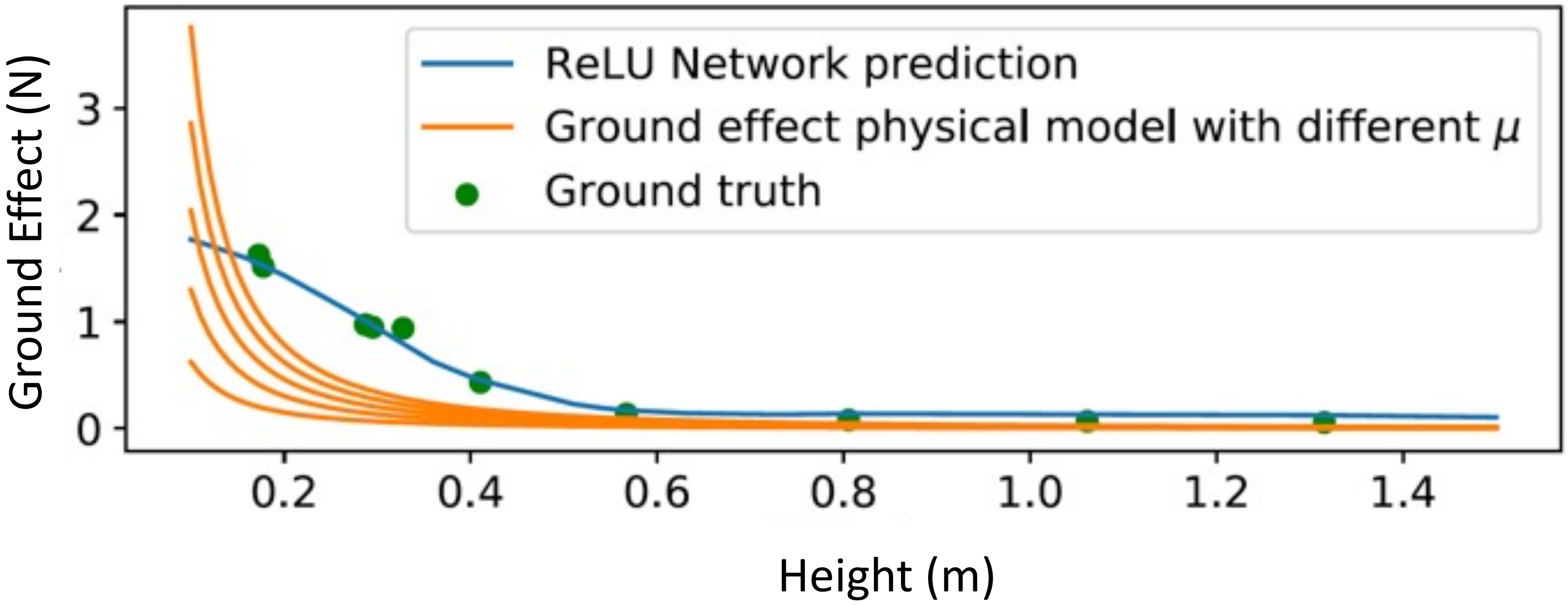


Ensures $\widetilde{F}$ is Lipshitz
[Bartlett et al., NeurIPS 2017]
[Miyato et al., ICLR 2018]

Spectral-Normalized 4-Layer Feed-Forward

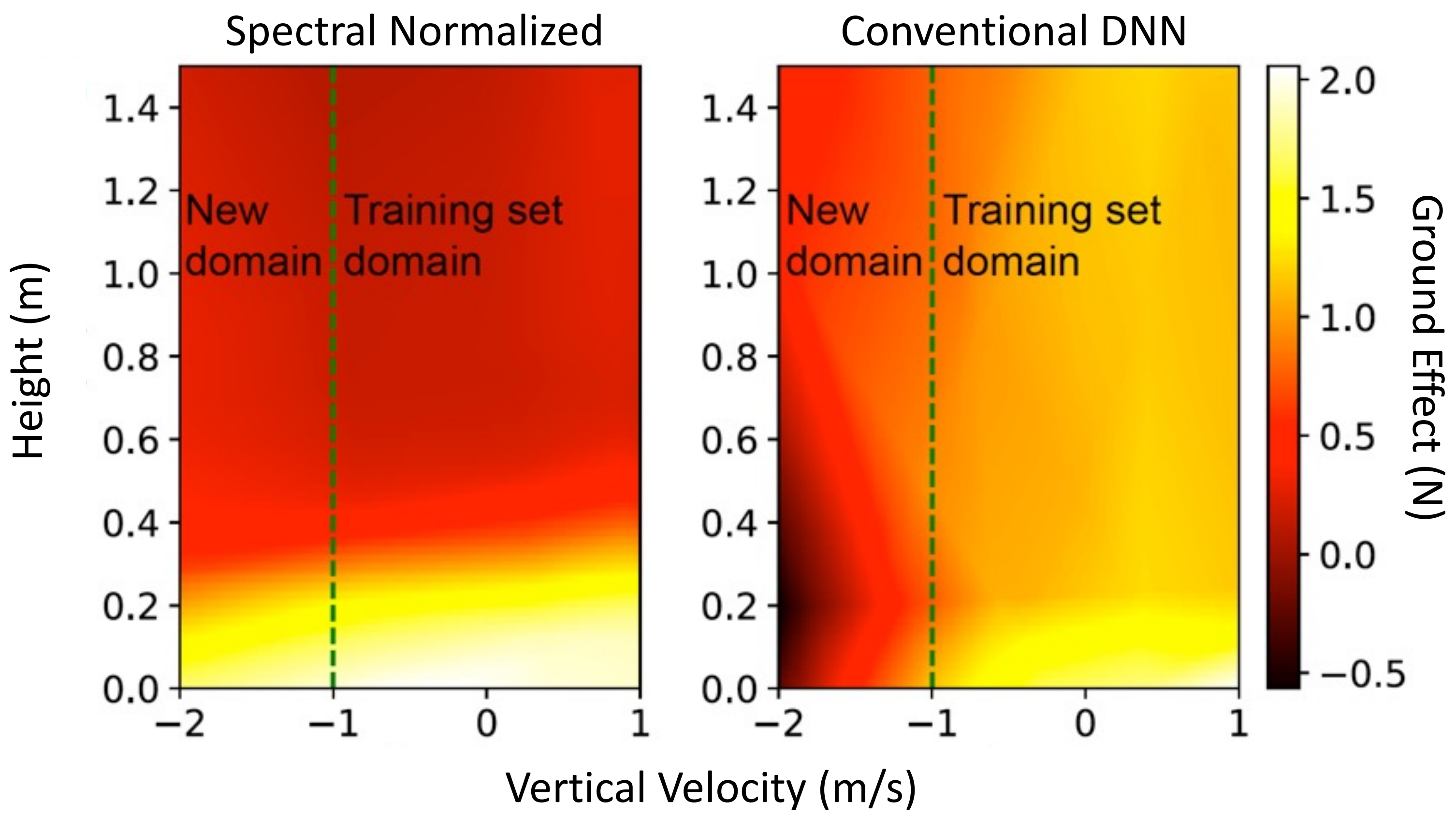
- Learn ground effect: $\tilde{F}(s,u) \rightarrow \mathbf{f}_a = [f_{a,x}, f_{a,y}, f_{a,z}]^\top$
- (s,u): height, velocity, attitude and four control inputs

Prediction Results



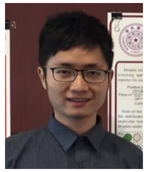
Neural Lander: Stable Drone Landing Control using Learned Dynamics
Guanya Shi, Xichen Shi, Michael O'Connell, Rose Yu, Kamyar Azizzadenesheli, Anima Anandkumar, Yisong Yue, Soon-Jo Chung. ICRA 2019.

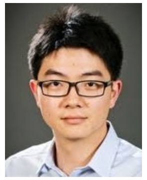
Prediction Results

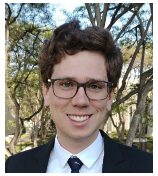


Neural Lander: Stable Drone Landing Control using Learned Dynamics
Guanya Shi, Xichen Shi, Michael O'Connell, Rose Yu, Kamyar Azizzadenesheli, Anima Anandkumar, Yisong Yue, Soon-Jo Chung. ICRA 2019.

Controller Design (simplified)


Guanya Shi


Xichen Shi


Michael O'Connell

- Nonlinear Feedback Linearization:

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

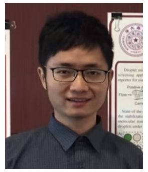
Desired Trajectory
(tracking error)

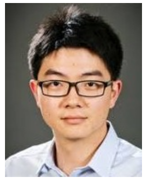
Feedback Linearization (PD control)

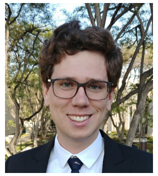
- Cancel out ground effect $\tilde{F}(s, u_{old})$: $u = u_{nominal} + u_{residual}$

Requires Lipschitz & small time delay

Controller Design (simplified)


Guanya Shi


Xichen Shi


Michael O'Connell

- 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)

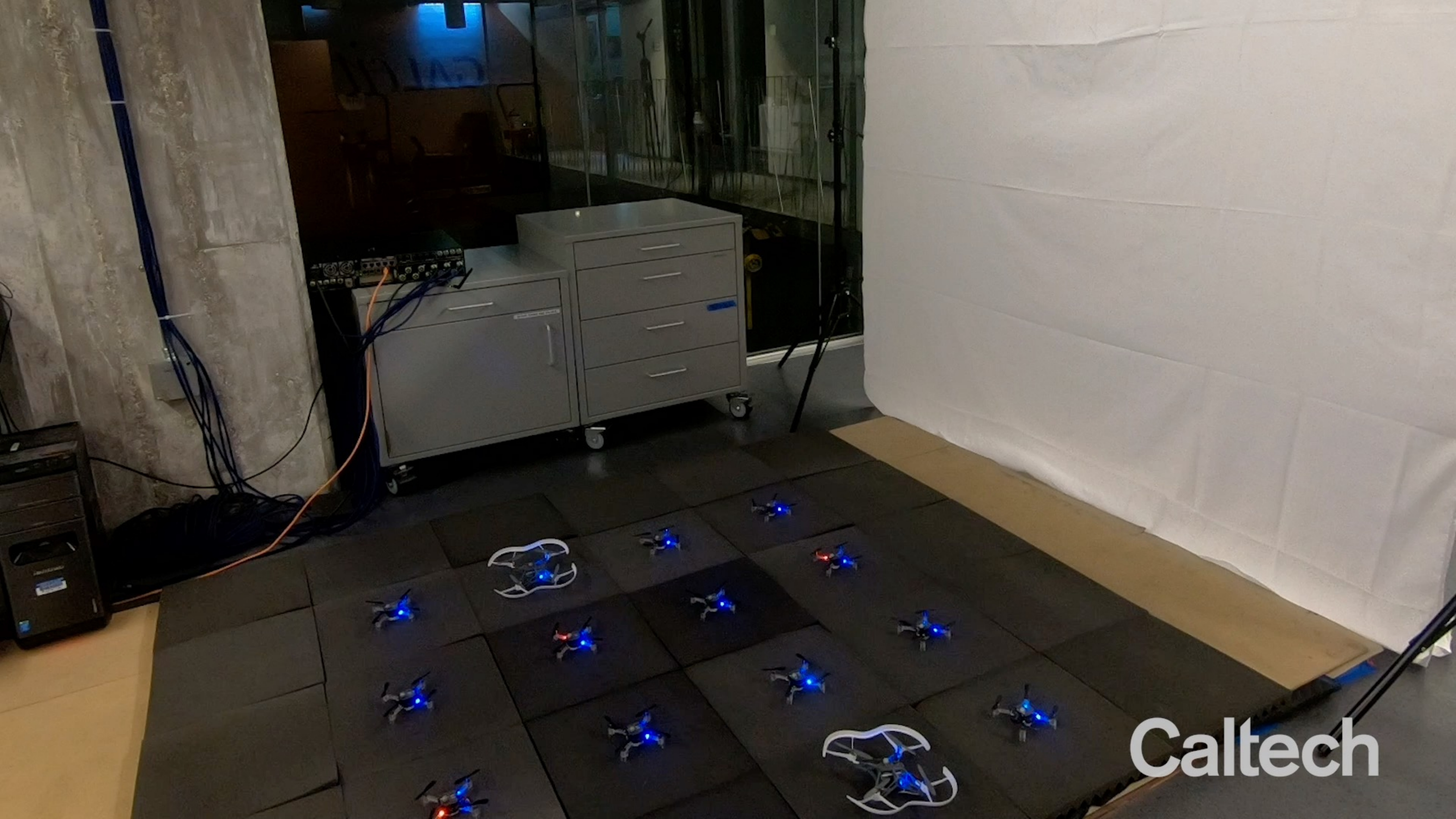
Time delay

Unmodeled disturbance

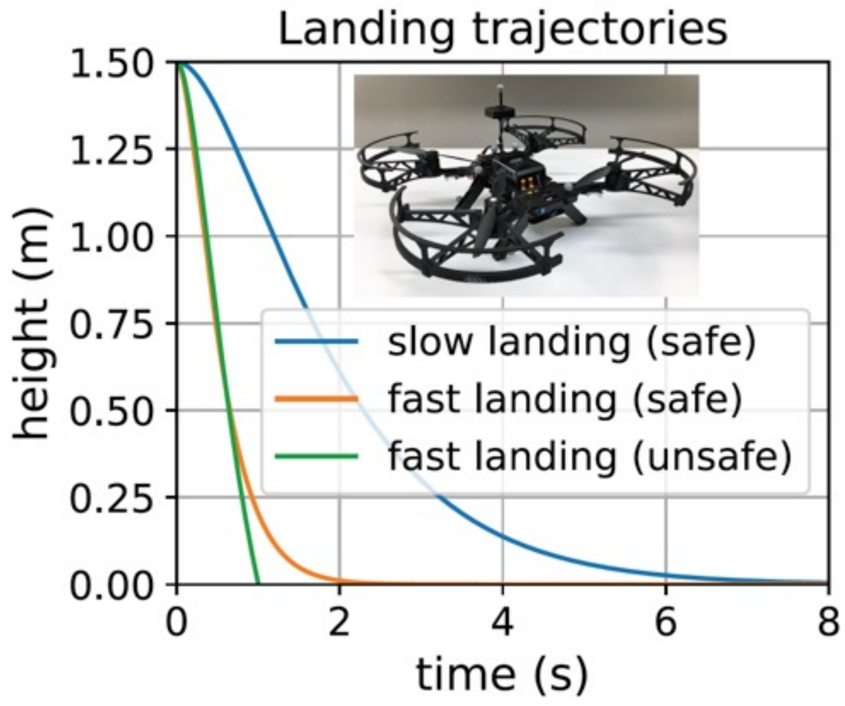
$$\|\eta(t)\| \leq \|\eta(0)\| \exp\left\{-\frac{\lambda_{min}(K) - \tilde{L}\rho}{C} t\right\} + \frac{\epsilon}{\lambda_{min}(K) - \tilde{L}\rho}$$

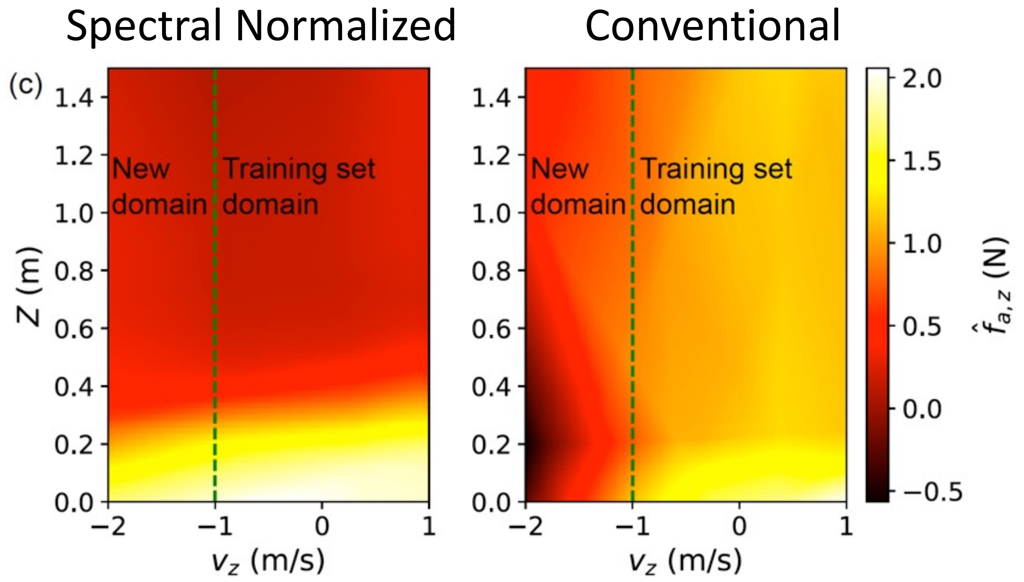
Lipschitz of NN

$$\Longrightarrow \|\eta(t)\| \to \frac{\epsilon}{\lambda_{min}(K) - \tilde{L}\rho}$$ Exponentially fast

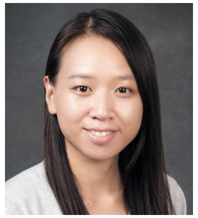

Caltech

Aside: Robust Regression for Safe Exploration





Robust regression guarantees extrapolation!
Enables safe exploration!



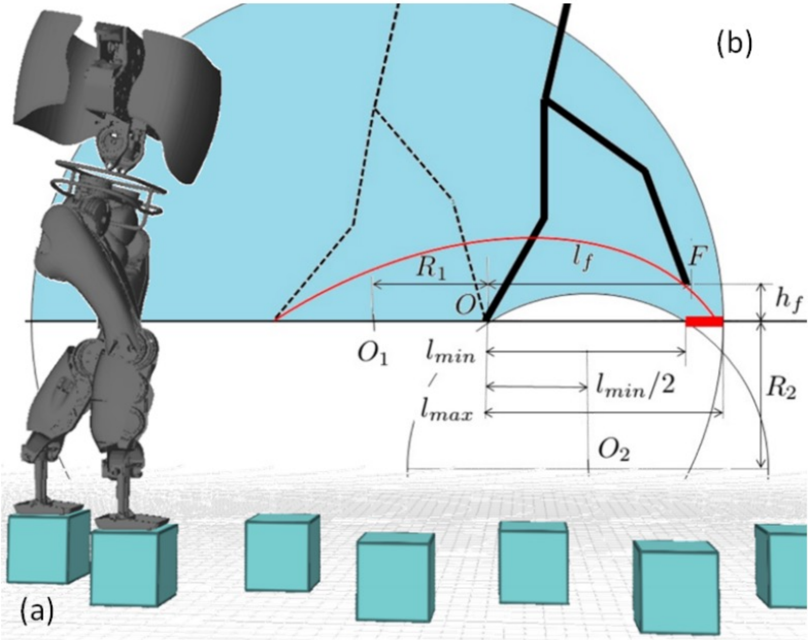
Angie Liu

Yashwanth Nakka

Robust Regression for Safe Exploration in Control,
Angie Liu, Guanya Shi, Soon-Jo Chung, Anima Anandkumar, Yisong Yue, L4DC 2020
Chance-Constrained Trajectory Optimization for Safe Exploration and Learning of Nonlinear Systems,
Yashwanth Kumar Nakka, Angie Liu, Guanya Shi, Anima Anandkumar, Yisong Yue, Soon-Jo Chung, R-AL 2021

Aside: Learning Control Lyapunov/Barrier Functions

- CLFs & CBFs encode low-dim projection of dynamics
- Learn CLF/CBFs?
- Stability/safety under uncertainty?

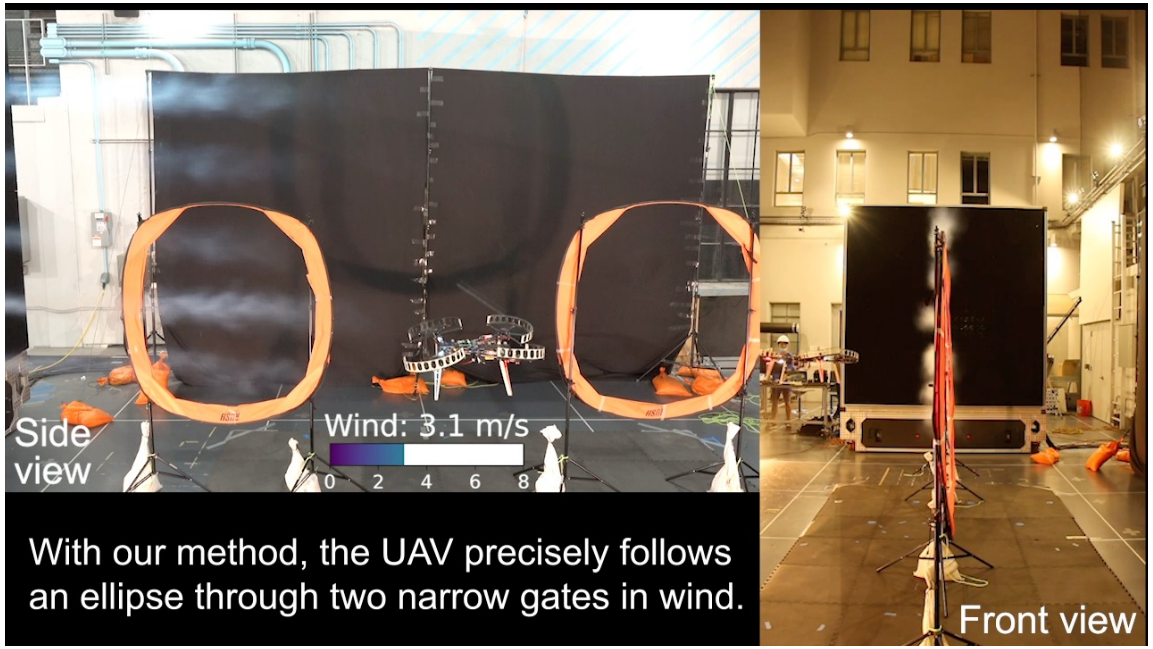


Img: https://arxiv.org/abs/1903.11199

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.

Andrew Taylor

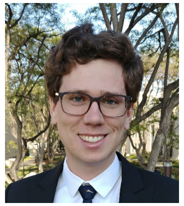
Victor Dorobantu



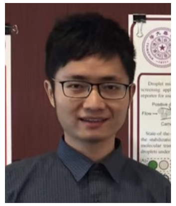
Dynamic Environments

[Neural-Fly, Science Robotics 2022]

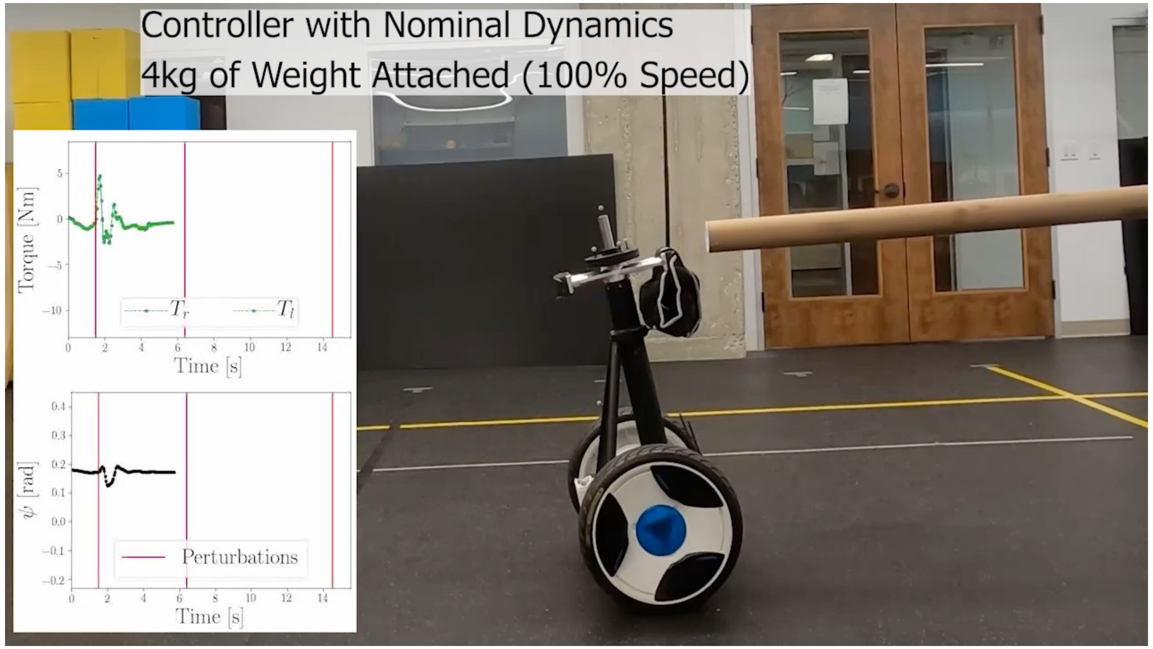
Meta-learning + Adaptive Control



Michael O'Connell



Guanya Shi



Sharp Non-Linearities

https://arxiv.org/abs/2103.04548 [IROS 2021]

Learn Continuous-time Models

Ivan Jimenez Rodriguez

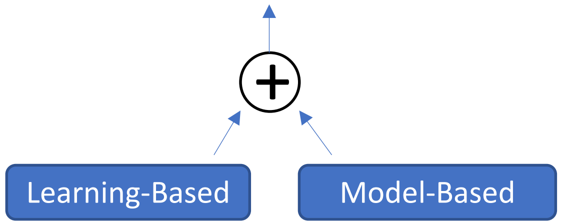
Ugo Rosolia

Summary: Dynamics Learning

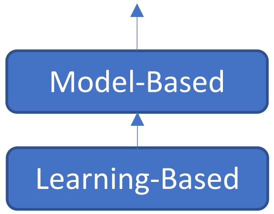
- Learn residual dynamics (data efficient)
- Control Lipschitz constant (imposes compatible structure)
- Standard controller design (inherits guarantees)
- Extend to complex settings (multi-agent, meta-learning, continuous-time, etc.)
- Robust regression for safe exploration (provable limited extrapolation)

Integration of Learning at Varying Levels

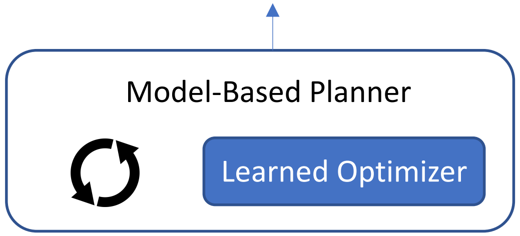
- Integration in control/action



- Integration in dynamics modeling



- Integration in optimization problem

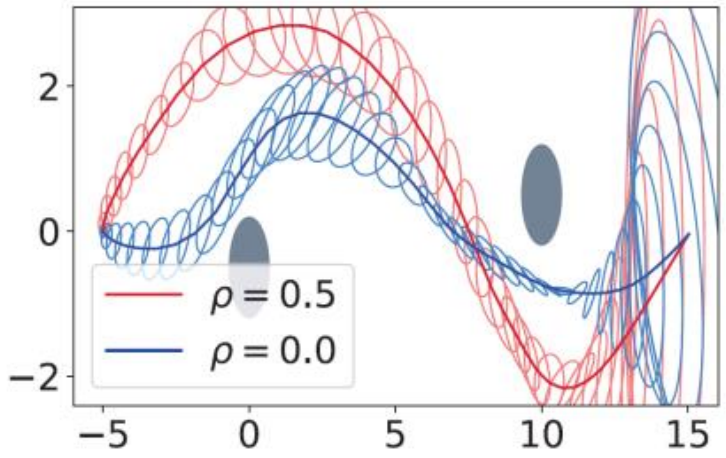


Model-Based Planning

- 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

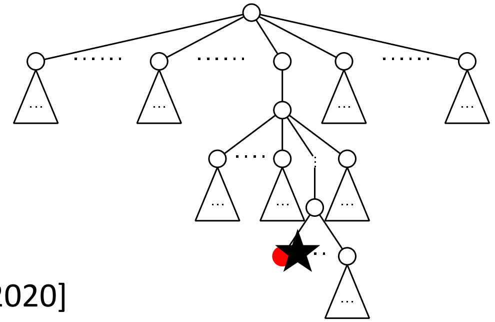
- Many Solvers are Sequential
  - Tree-Search
  - Greedy
  - Gradient Descent

- 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 Search Policy

- **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** [NeurIPS 2020]

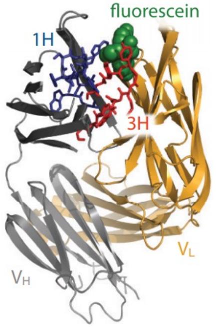


Jialin Song

Ben Riviere

Learning Value Function

- **MLNav: Learning to Safely Navigate on Martian Terrains** [R-AL 2022]
- **Learning to Make Decisions via Submodular Regularization** [ICLR 2021]

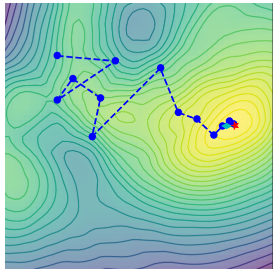


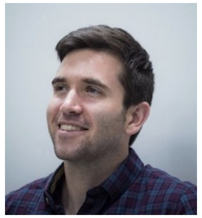
Ayya Alieva

Shreyansh Daftry

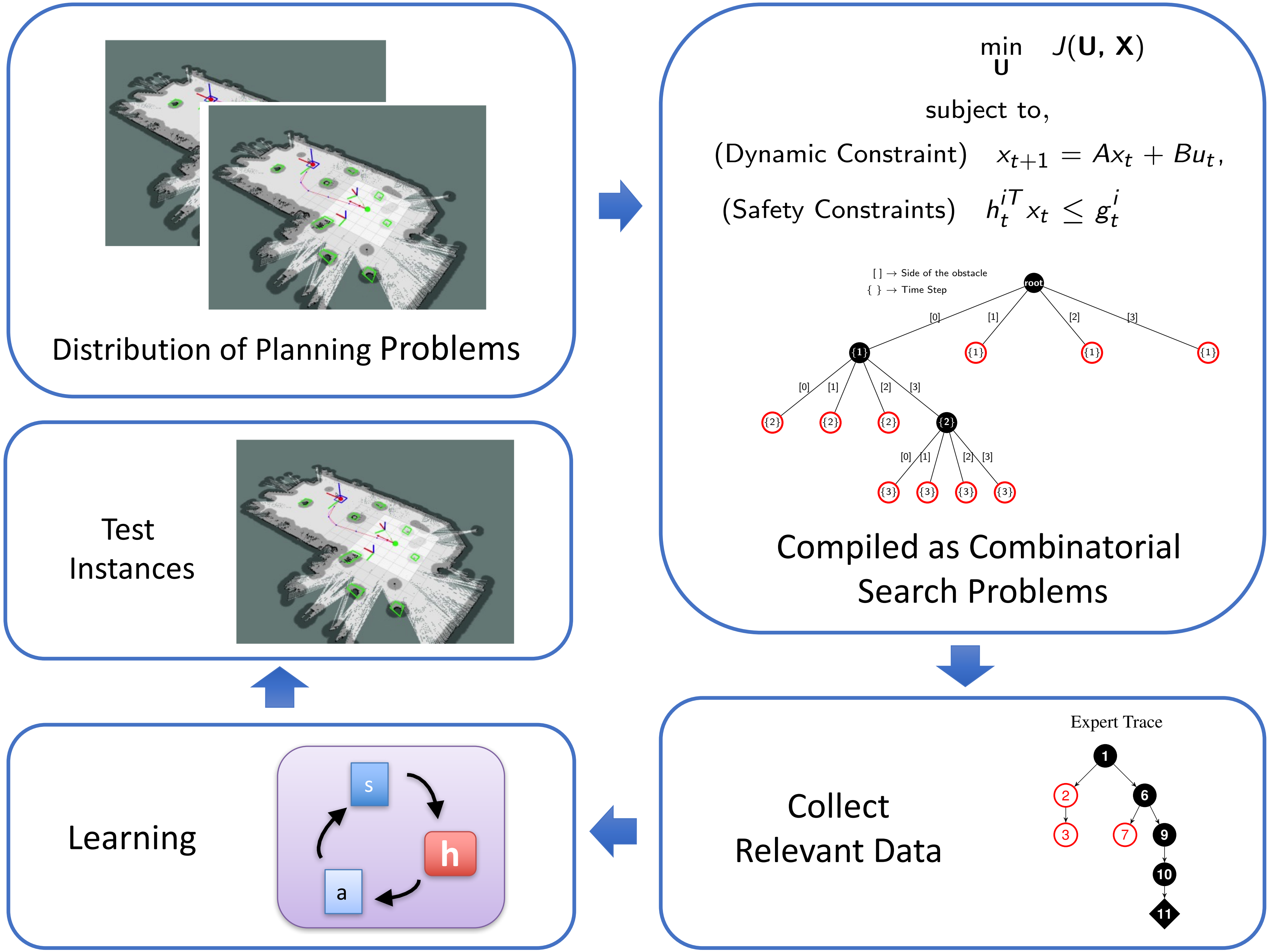
Learning to Infer

- **Iterative Amortized Inference** [ICML 2018]
- **A General Method for Amortizing Variational Filtering** [NeurIPS 2018]
- **Iterative Amortized Policy Optimization** [NeurIPS 2021]





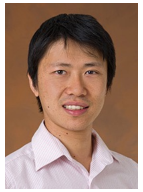
Joe Marino



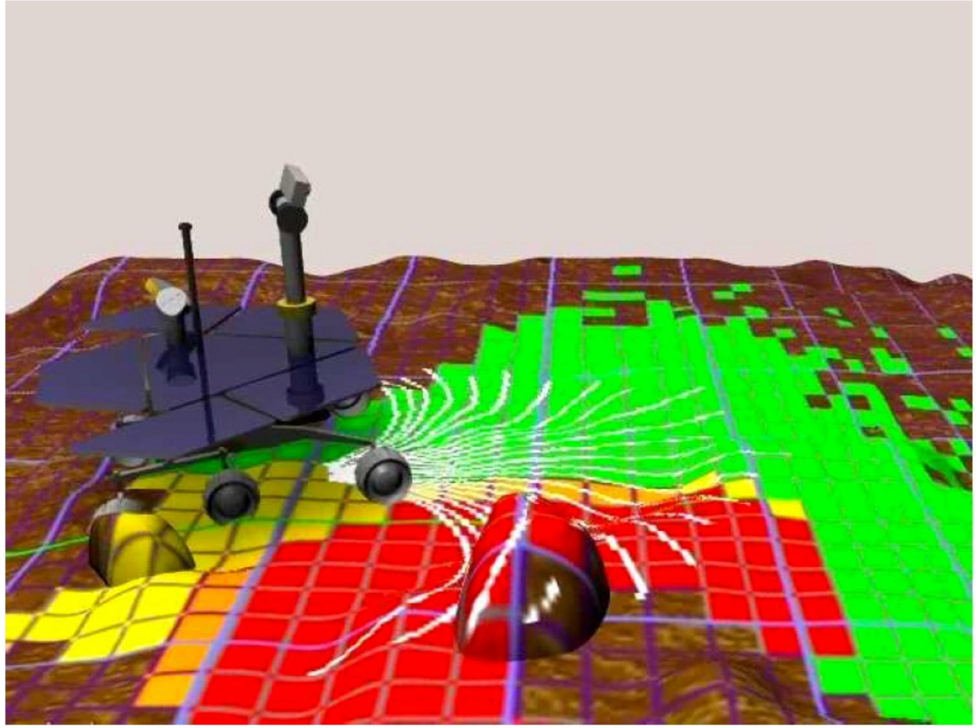
Distribution of Planning Problems
$\min_{\mathbf{U}} \quad J(\mathbf{U}, \mathbf{X})$
subject to,
(Dynamic Constraint) $x_{t+1} = Ax_t + Bu_t,$
(Safety Constraints) $h_t^{iT} x_t \leq g_t^i$
[ ] → Side of the obstacle
{ } → Time Step
root
Compiled as Combinatorial Search Problems
Expert Trace
Collect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

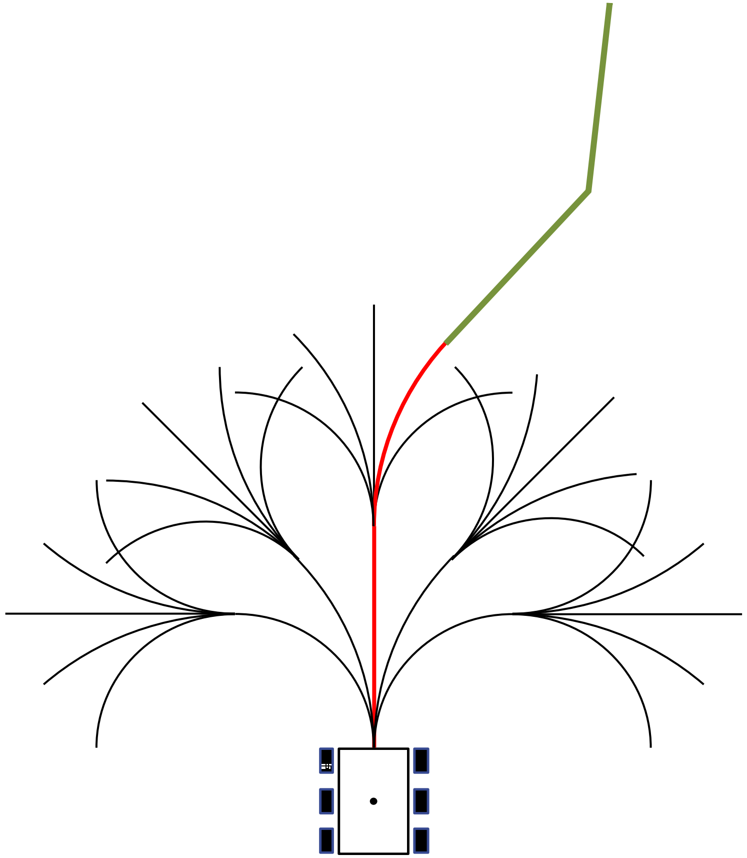
MLNav: Learning-Augmented Rover Navigation


Shreyansh Daftry


Hiro Ono


Neil Abcouwer





MLNav: Learning to Navigate on Martian Terrains, Shreyansh Daftry et al., R-AL 2022

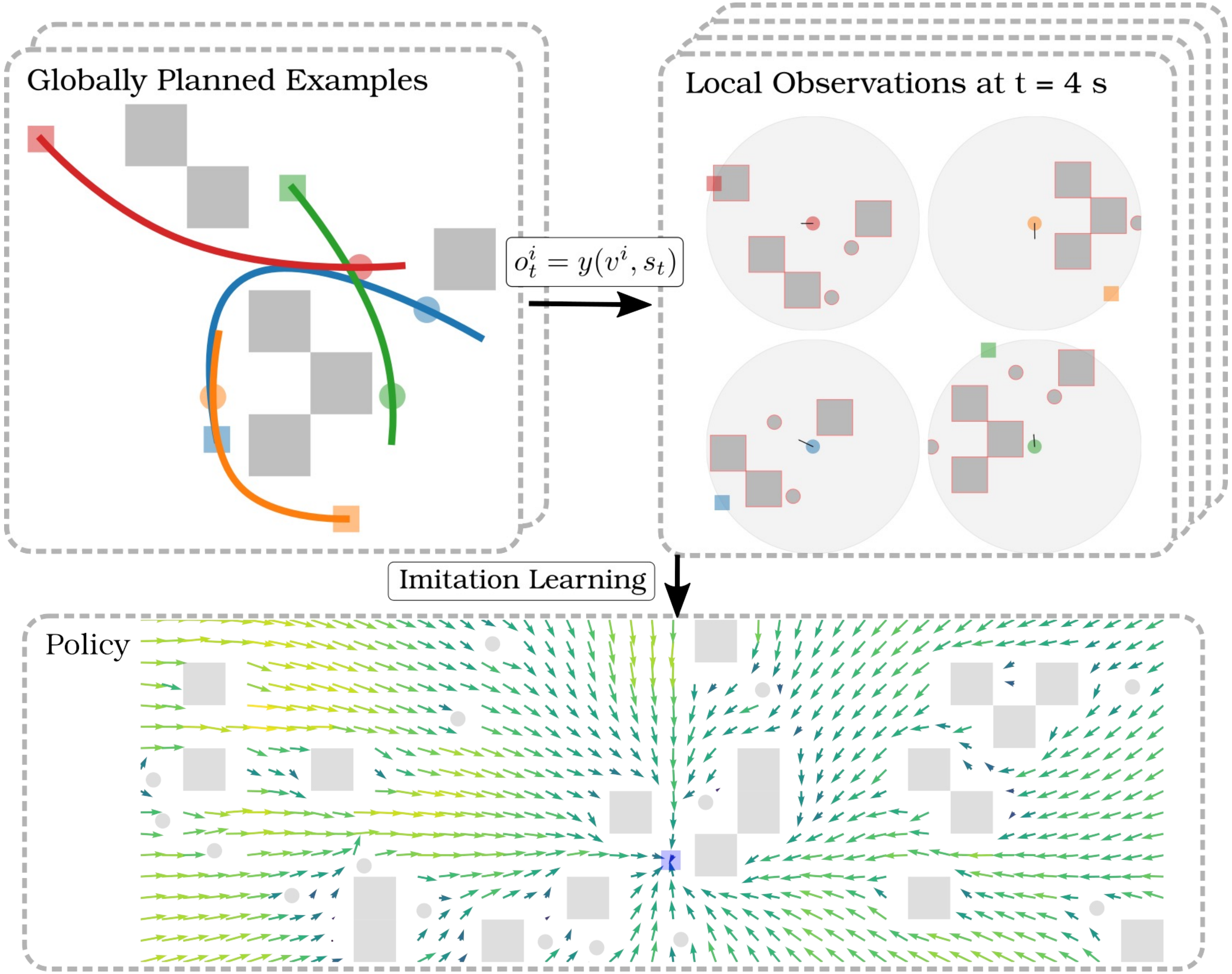

Perseverance Sol 122
Sol = Martian day
Image: NASA/JPL-Caltech

Learned Decentralized Planner

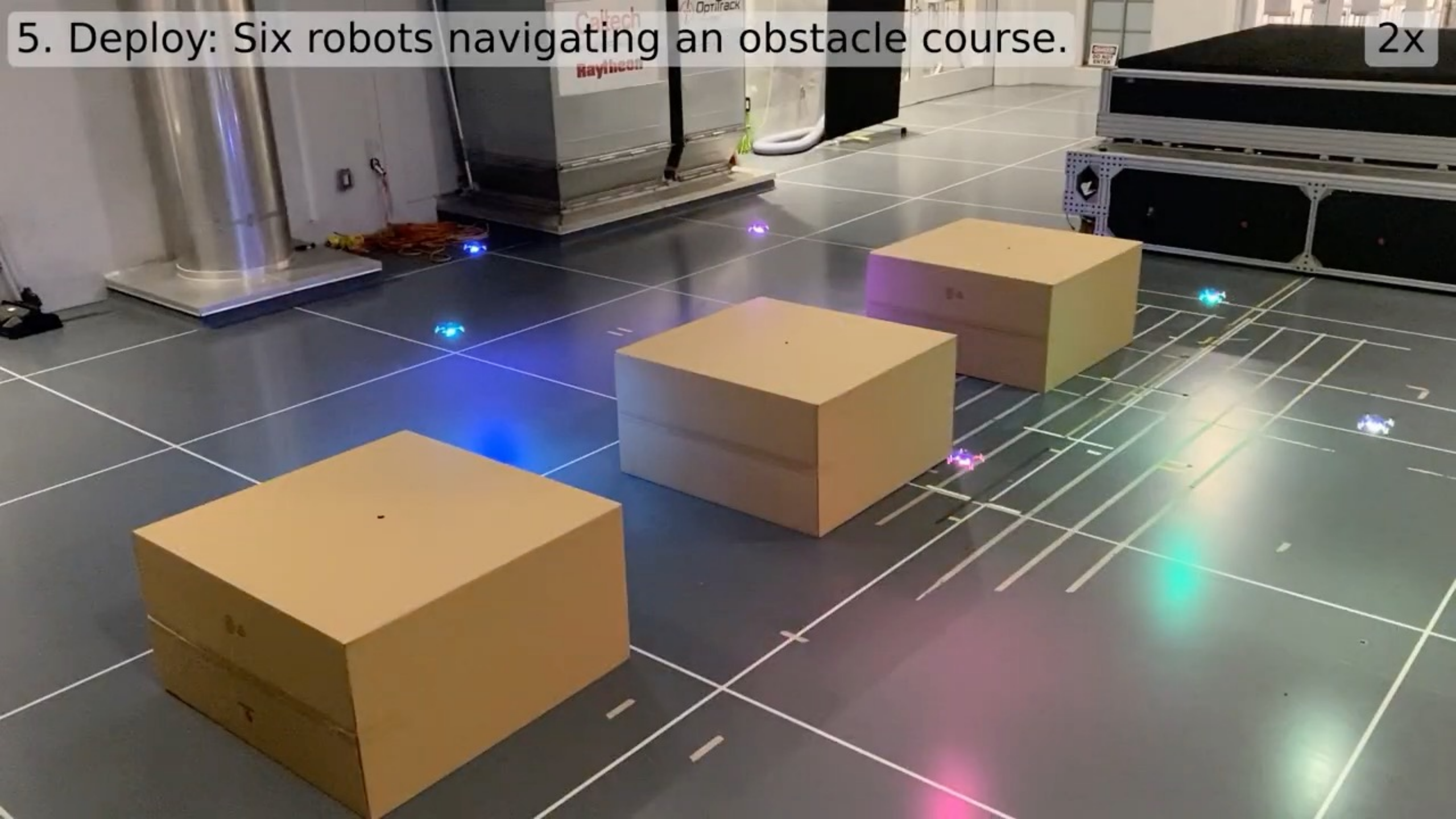
(enforcing safety)

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

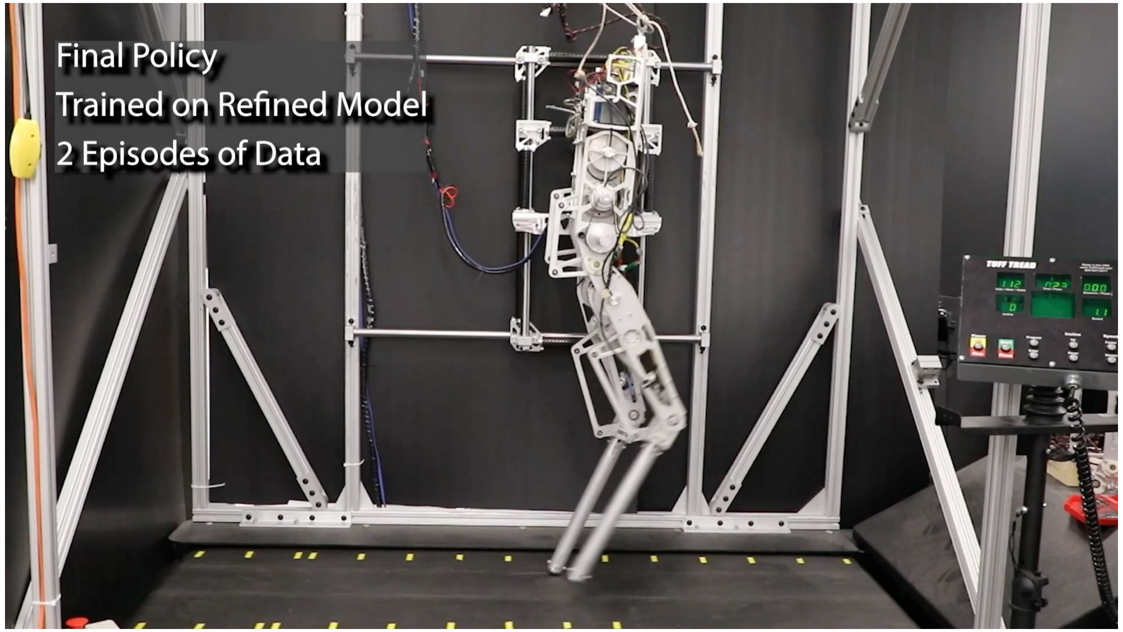
Learned ODE-based Policy

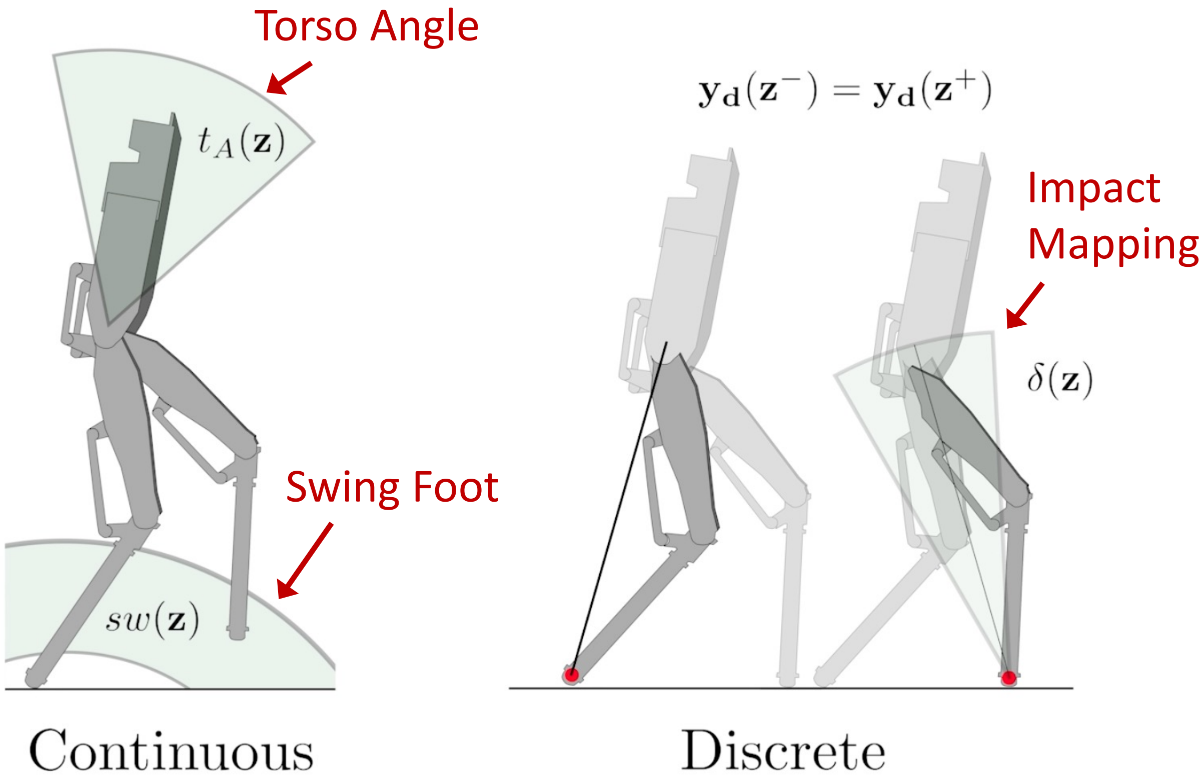
Neural Gaits: Learning Bipedal Locomotion via
Control Barrier Functions and Zero Dynamics Policoes

Jimenez Rodriguez, Csomay-Shanklin, et al., L4DC 2022

Ivan
Jimenez
Rodriguez

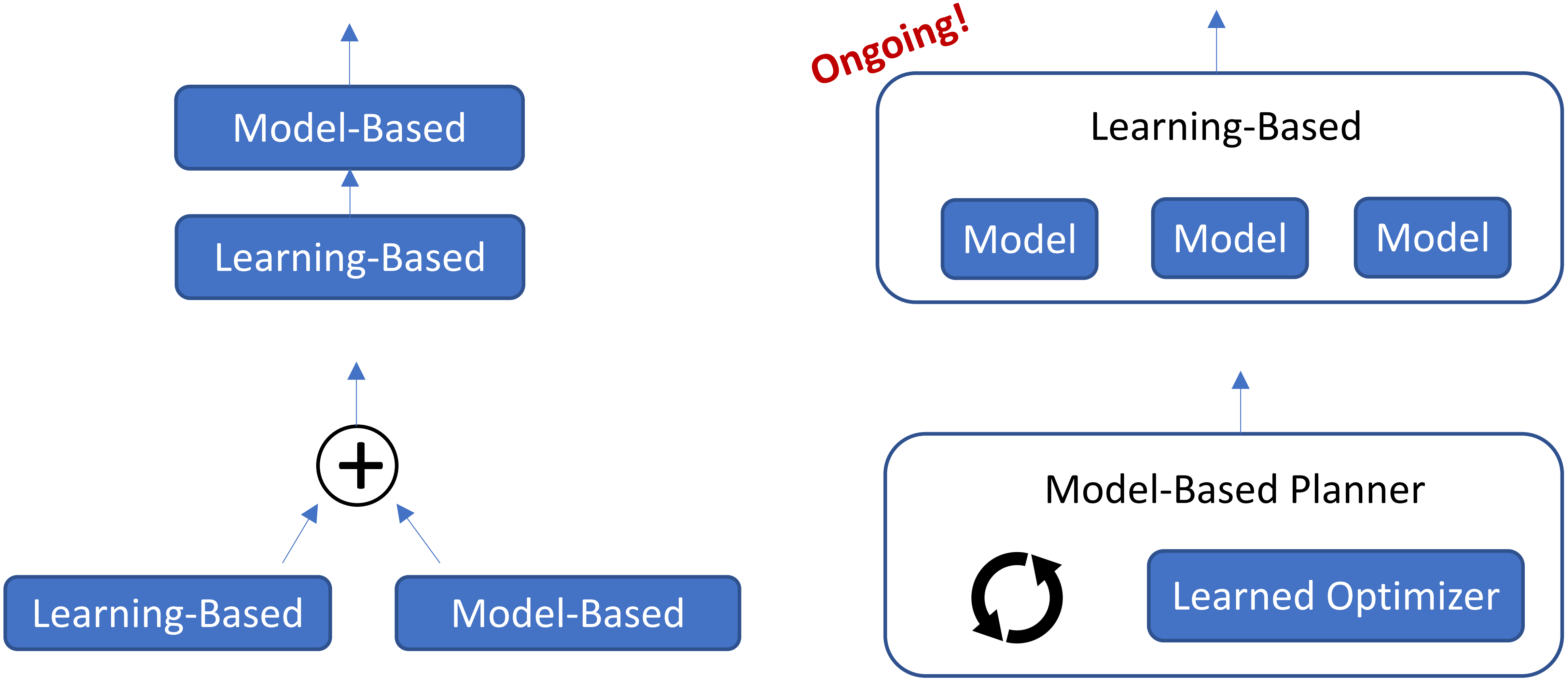
Noel
Csomay-Shanklin



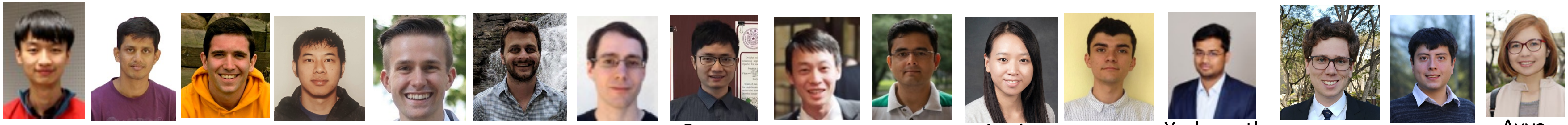


- **Barriers** induce control-theoretic safety conditions
- Conventional: **complicated ODE-based optimization**
- Learn policy as **Neural ODE**
  - **Fast run-time gait generation**
  - **Satisfies safety guarantees**

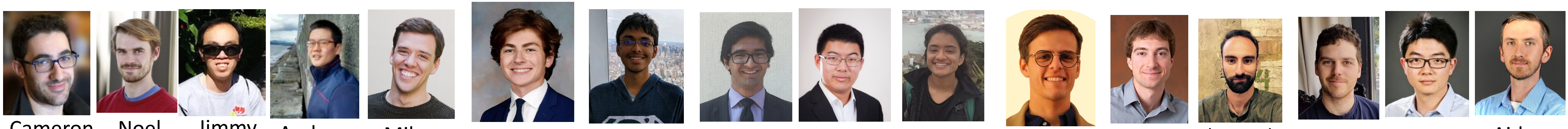
Blending Models/Rules & Black-Box Learning



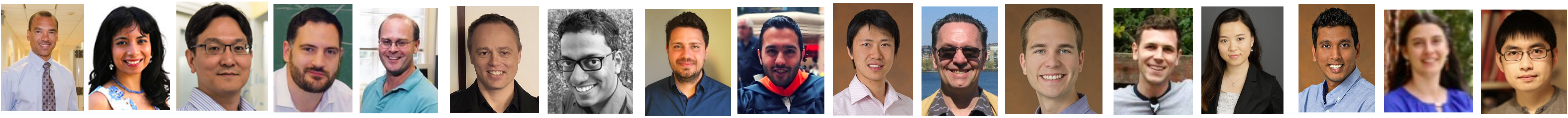
Collaborators



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
Michael O'Connell
Ivan Jimenez Rodriguez
Ayya Alieva



Cameron Voloshin
Noel Csomay -Shanklin
Jimmy Chen
Andrew Kang
Milan Cvitkovic
Lucas Igel
Siddarth Venkatraman
Aadyot Bhatnagar
Albert Zhao
Meera Krishnamoorthy
Ugo Rosolia
Tyler del Sesto
Alessandro Ialongo
Alex Piche
Xichen Shi
Aiden Aceves



Stephen Mayo
Anima Anandkumar
Soon-Jo Chung
Aaron Ames
Joel Burdick
Gabor Orosz
Swarat Chaudhuri
Stephan Mandt
Kamyar Azizzadenesheli
Hiro Ono
Jim Little
Neil Abcouwer
Peter Carr
Rose Yu
Shreyansh Daftry
Bistra Dilkina
Yuxin Chen

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 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

Neural-Swarm2: Planning and Control of Heterogeneous Multirotor Swarms using Learned Interactions, Guanya Shi et al., T-RO 2021.

Neural-Fly Enables Rapid Learning for Agile Flight in Strong Winds, Michael O'Connell, Guanya Shi, et al., Science Robotics 2022

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. R-AL 2021

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., NeurIPS 2020

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 to Make Decisions via Submodular Regularization, Ayya Alieva, Aiden Aceves, et al., ICLR 2021

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

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

Iterative Amortized Policy Optimization, Joe Marino et al., NeurIPS 2021

MLNav: Learning to Safely Navigate on Martian Terrains, Shreyansh Daftry et al., R-AL 2022

Neural Gaits: Learning Bipedal Locomotion via Control Barrier Functions and Zero Dynamics Policies, Jimenez Rodriguez, Csomay-Shanklin, et al., L4DC 2022

Learning to Control an Unstable System with One Minute of Data: Leveraging Gaussian Process Differentiation in Predictive Control, Jimenez Rodriguez, et al., IROS 2021

Thanks!

@yisongyue

http://www.yisongyue.com