Safe Learning-Based Control using Gaussian Processes

Prof. Angela Schoellig

IFAC World Congress 2020 – Learning for Control Tutorial

Institute for Aerospace Studies
UNIVERSITY OF TORONTO

VECTOR INSTITUTE

DYNAMIC SYSTEMS LAB

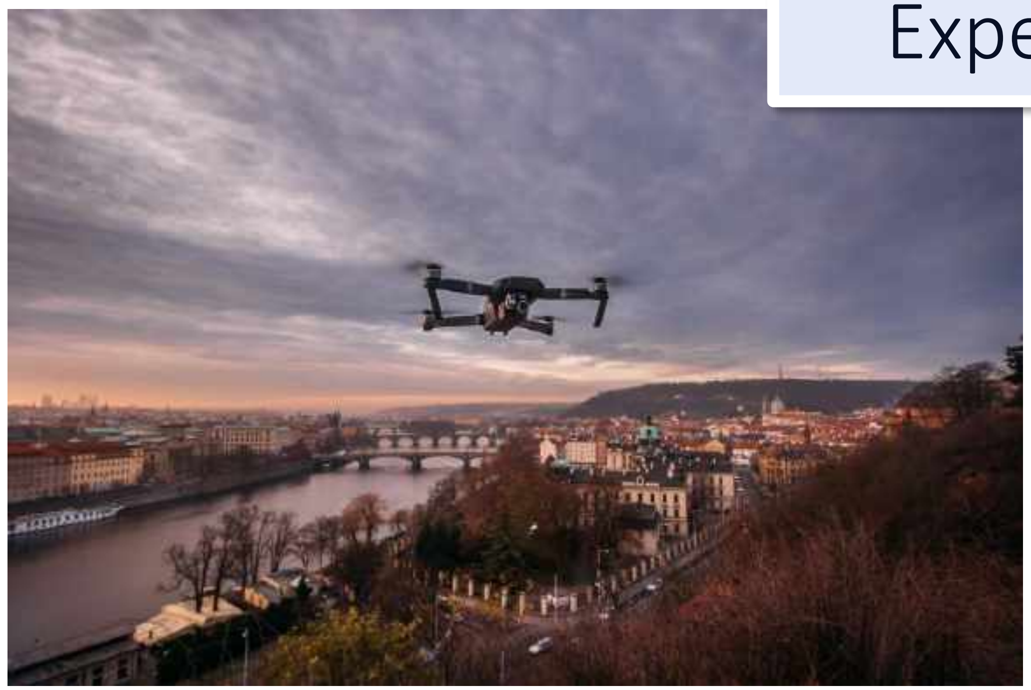
The Future of Automation

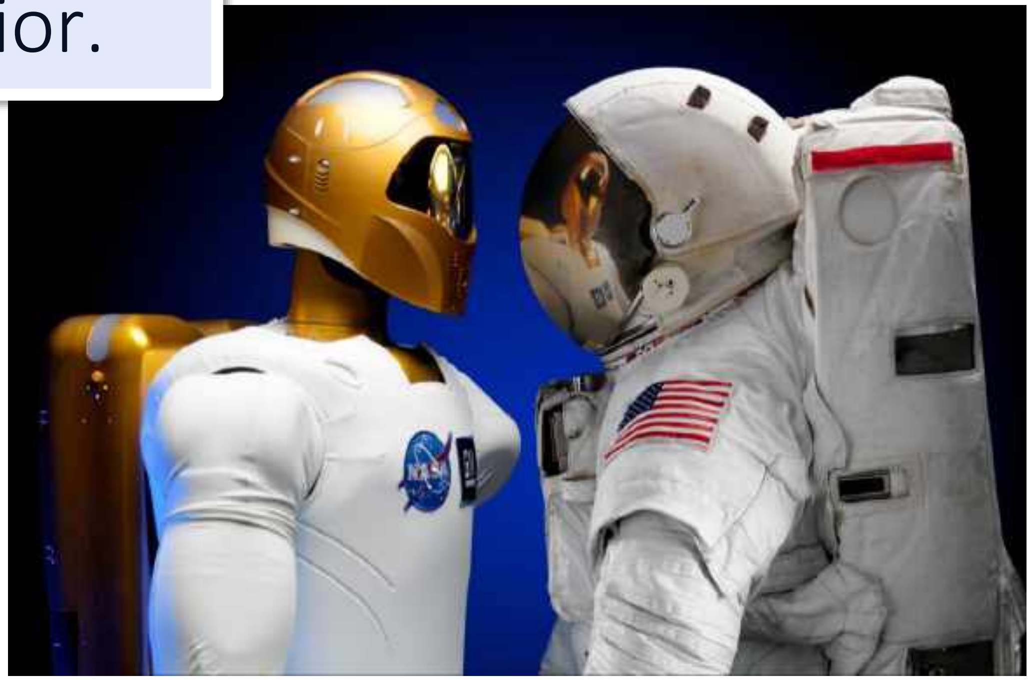
DYNAMIC SYSTEMS LAB

Large prior uncertainties.

Active decision making.

Expect safe and high-performance behavior.

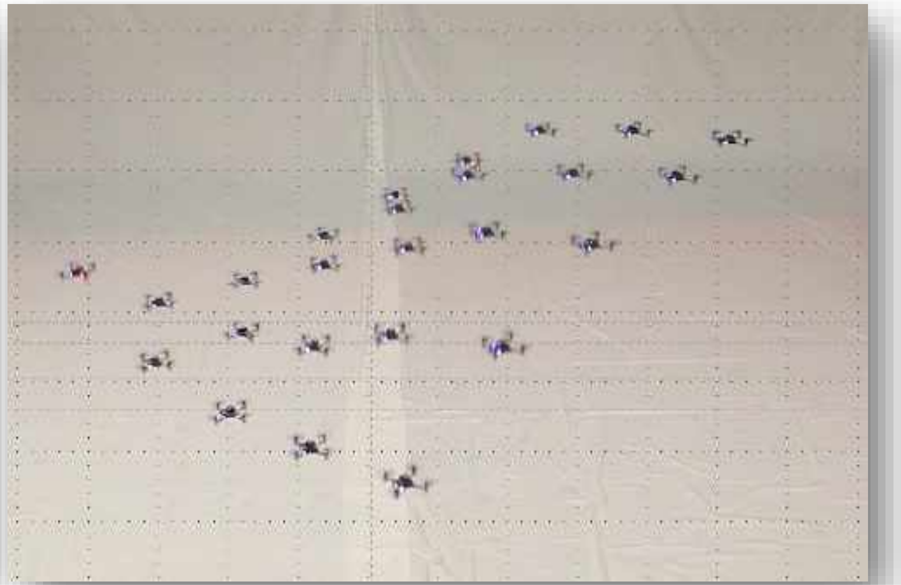




Robots in My Lab

Model uncertainties that limit performance:

Unknown terrain and topography

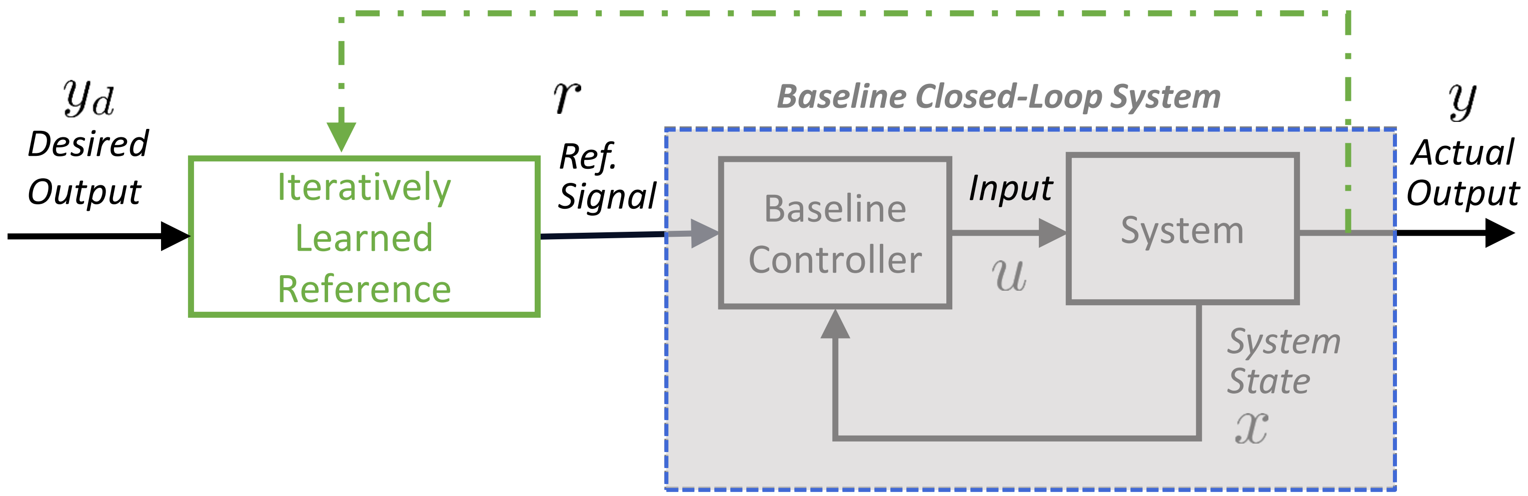


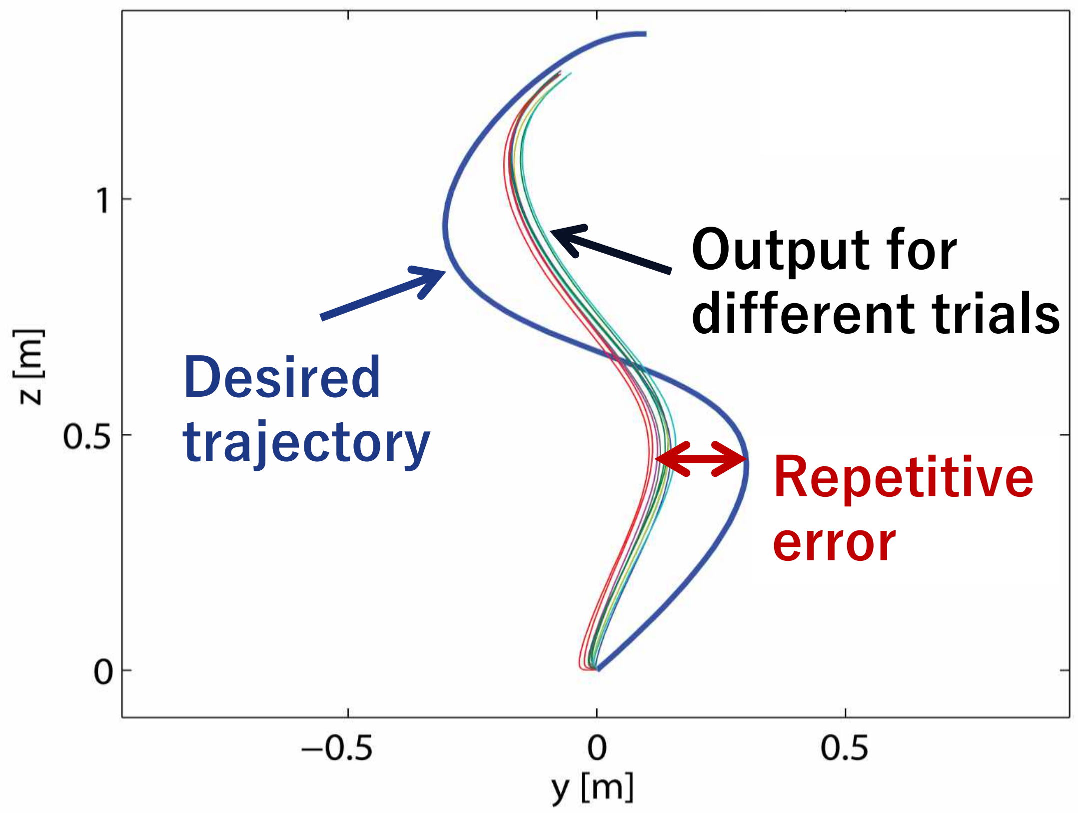
Unknown aerodynamic effects

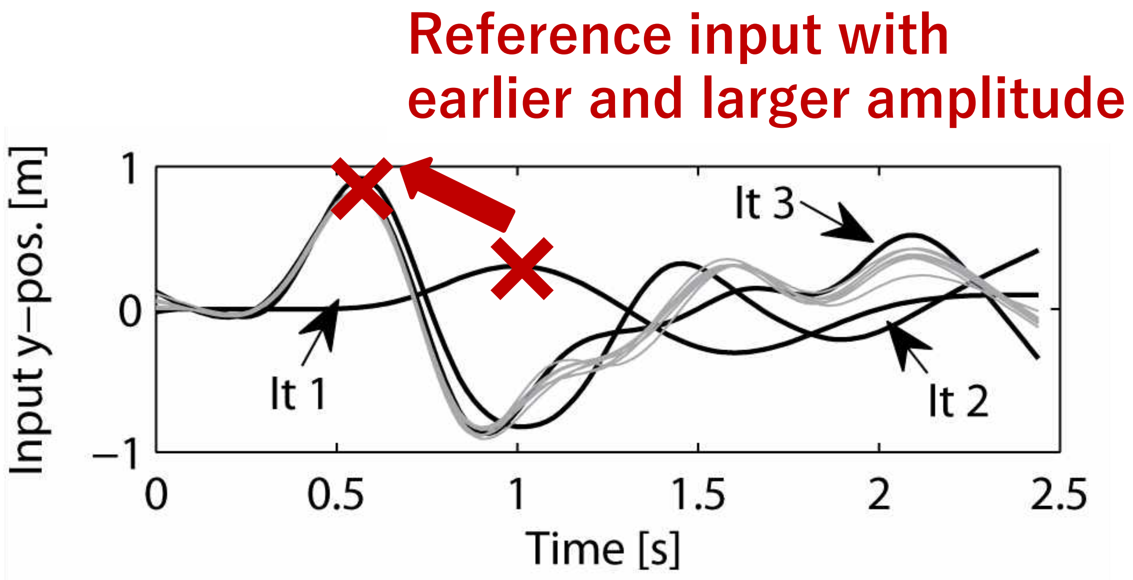
Unknown weather conditions

Interaction with unknown objects

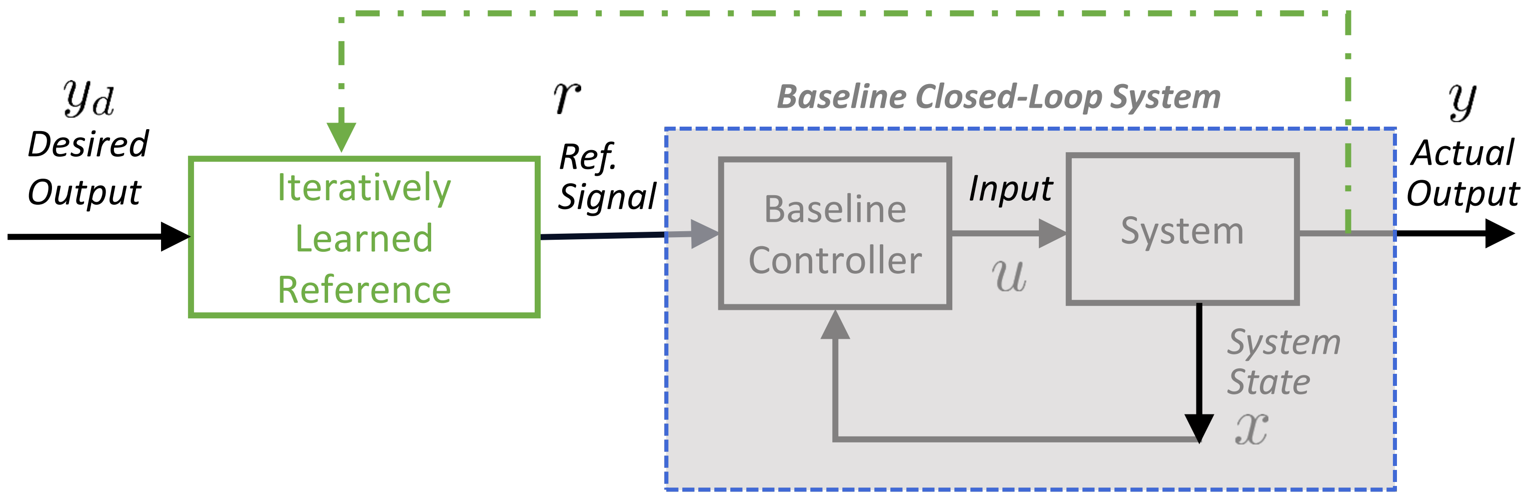
Learning from data can improve performance.



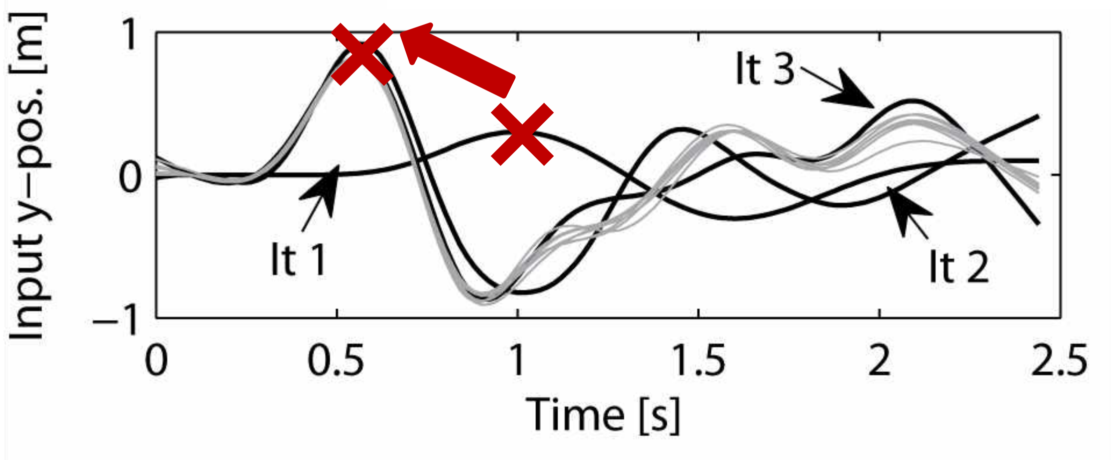




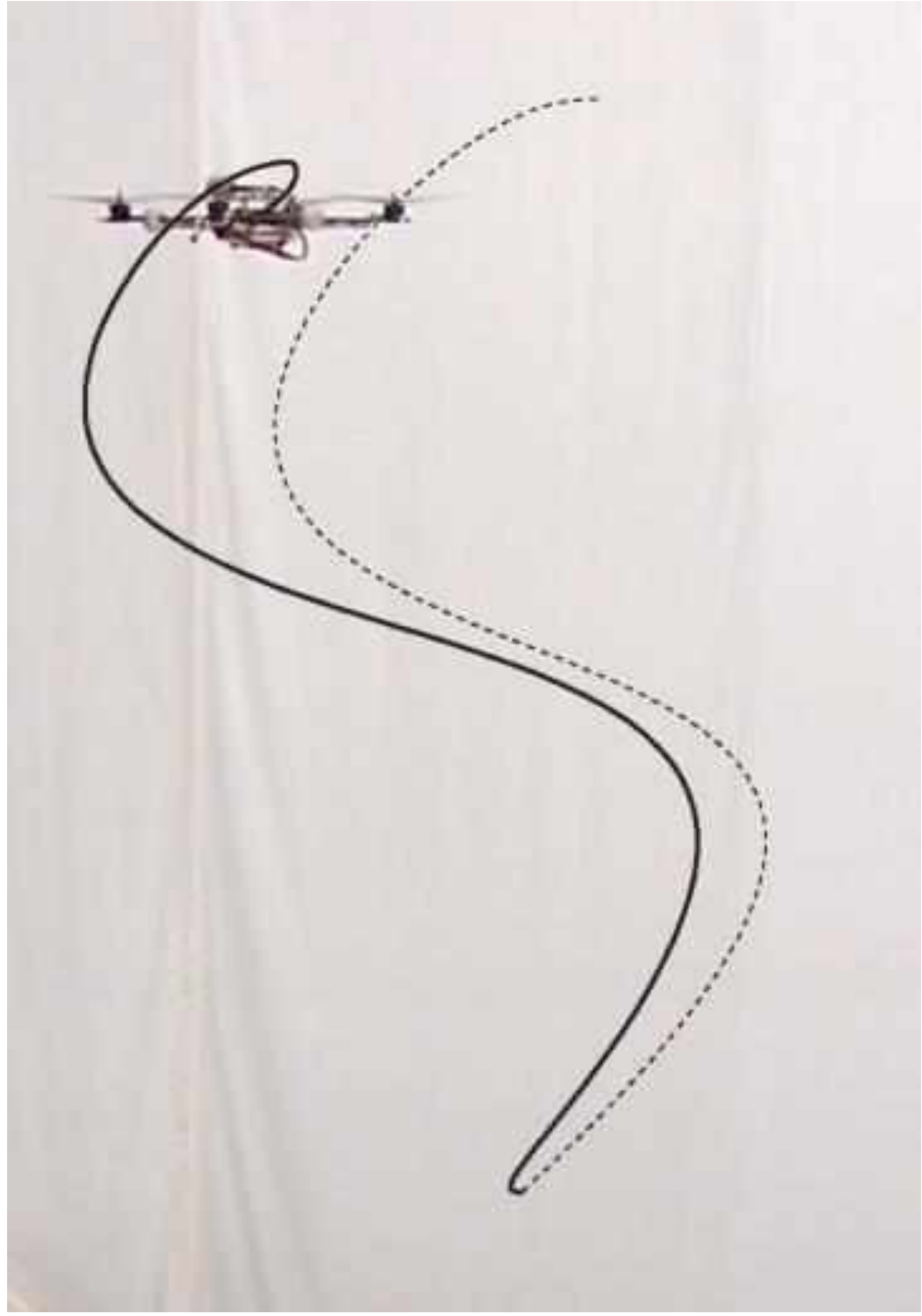
Learning from data can improve performance.



Reference input with earlier and larger amplitude



Video 2x

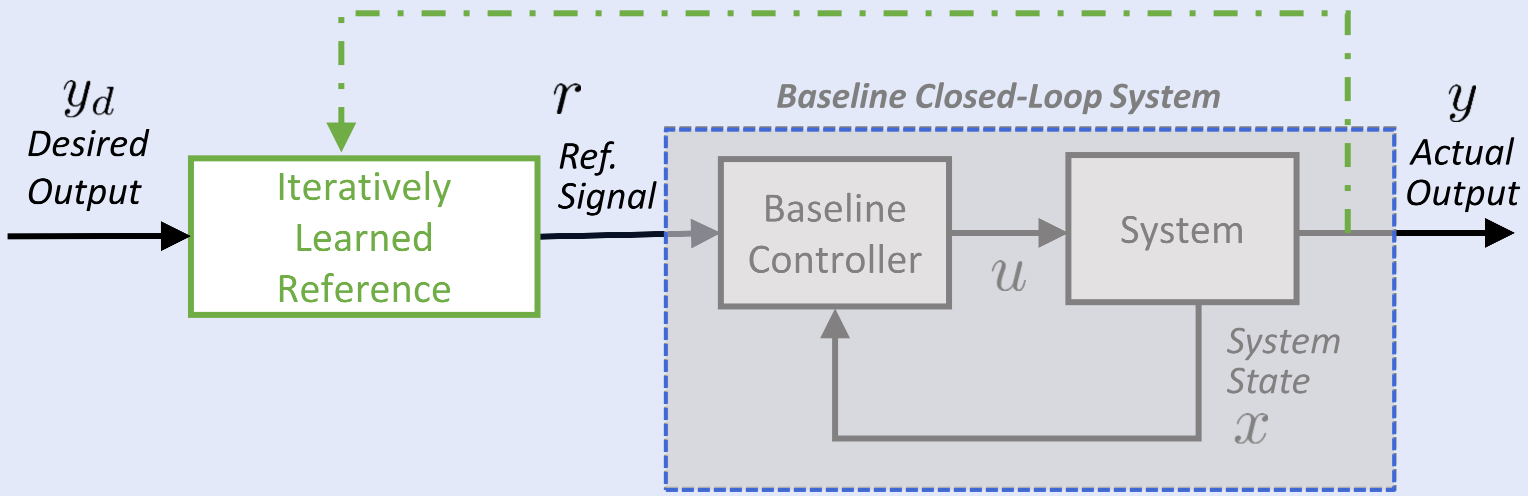


Learned Triple Flip [ICRA10] https://youtu.be/bWExDW9J9sA

The image part with relationship ID rId6 was not found in the file.

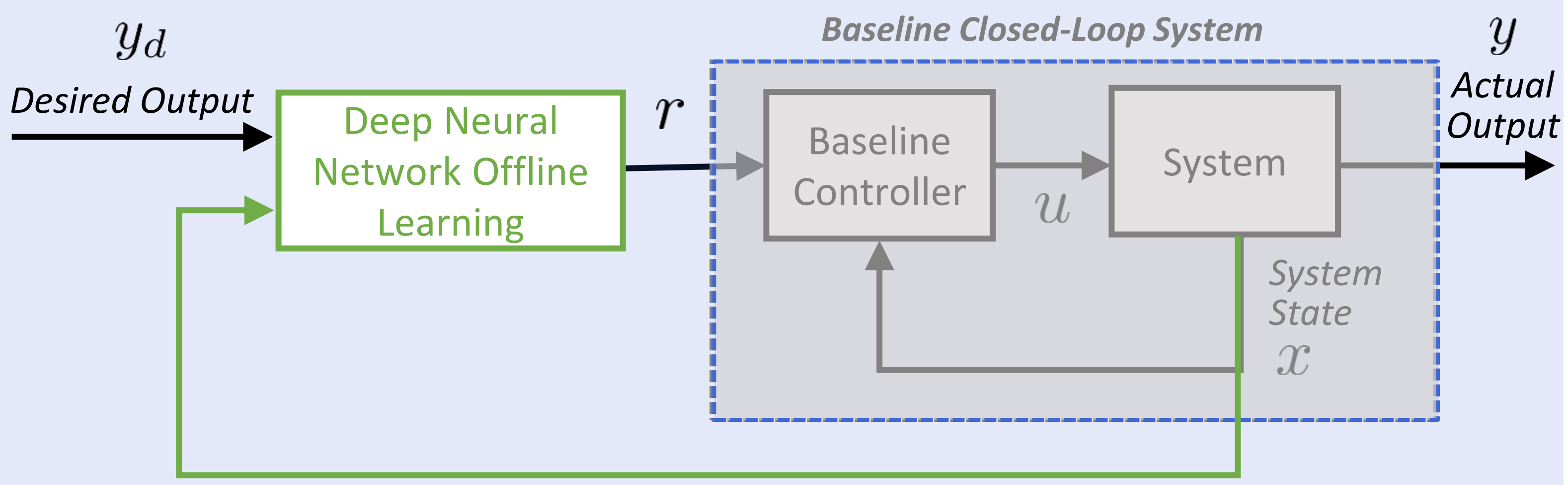
Learning from data can improve performance.

Learning a single task through repetition



[ECC'09, IROS'12, AURO'12]

Offline learning of inverse model



[ICRA'16, CDC'17, RAL'18, ECC'19]

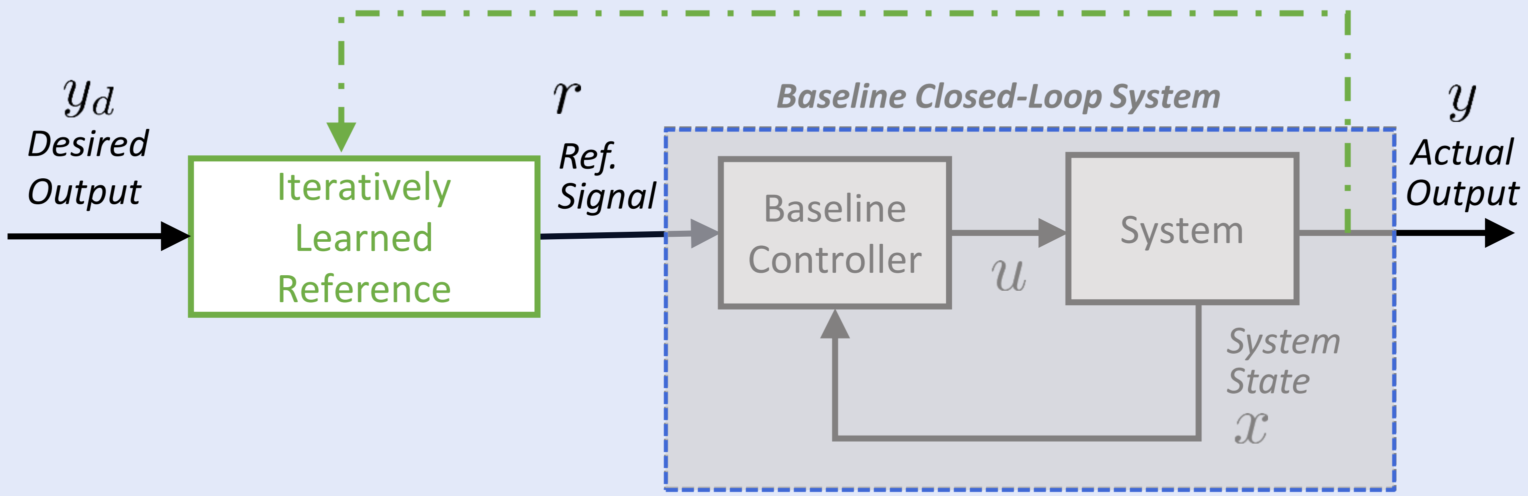
Mobile Manipulator Control [IROS'20] http://tiny.cc/ball_catch

DYNAMIC SYSTEMS LAB

Learning from data can improve performance.

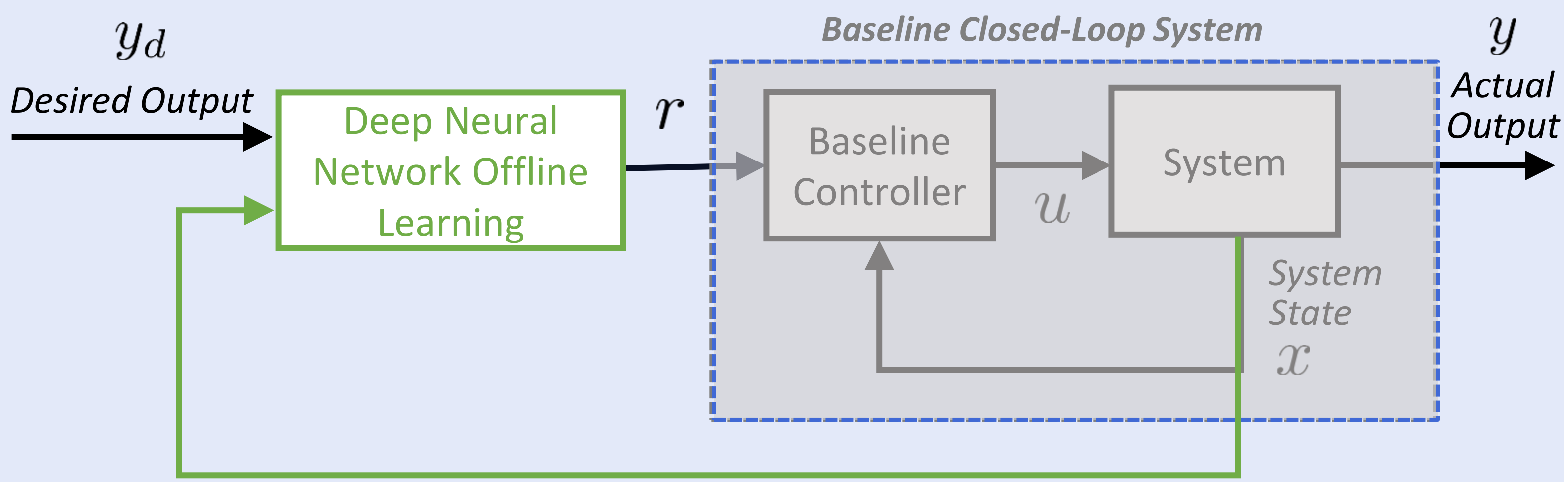
DYNAMIC SYSTEMS LAB

Learning a single task through repetition



[ECC'09, IROS'12, AURO'12]

Offline learning of inverse model



[ICRA'16, CDC'17, RAL'18, ECC'19]

- **+ Input-output stability** if baseline system is stable
- **+ Acausal corrections** possible
- **− Baseline controller required**
- **− Training phase**
- **− State constraints** not considered

Problem Statement

Design a controller for systems with prior uncertainty that learns online and continuously improves performance while satisfying safety constraints.

Considered system dynamics:

$$x_{t+1} = \underbrace{f(x_t, u_t)}_{\text{a priori model}} + \underbrace{g(x_t, u_t)}_{\text{unknown}} + \underbrace{\xi_t}_{\text{noise}}$$

Key features:

- Nonparametric model
- Improved performance with more data

Compare to (simplified view):

$$x_{t+1} = \underbrace{Ax_t + Bu_t}_{\text{a priori model}} + \underbrace{\Delta A\, x_t + \Delta B\, u_t}_{\text{unknown}}$$

with a-priori given sets $\Delta A \in \mathcal{A},\ \Delta B \in \mathcal{B}$

- **Robust control:** finds controller that achieves stability and performance for all possible $\Delta A, \Delta B$
- **Adaptive control:** estimates $\Delta A, \Delta B$ and uses estimate in controller

Approach

DYNAMIC SYSTEMS LAB

Nonparametric model for unknown model error

Gaussian processes

reliable confidence intervals

Algorithm to safely acquire data and optimize task

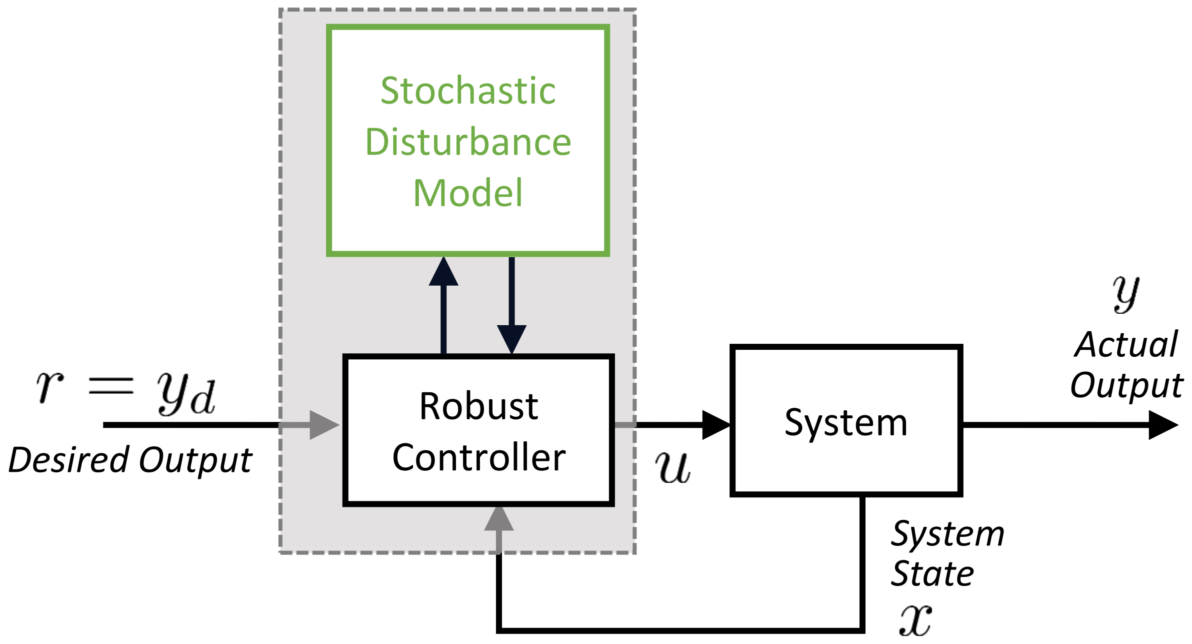
Robust control

stability & performance under uncertainty

Defining and analyzing closed-loop safety

Lyapunov analysis

stability of learned models



= safe model-based reinforcement learning

Approach

Nonparametric model for unknown model error

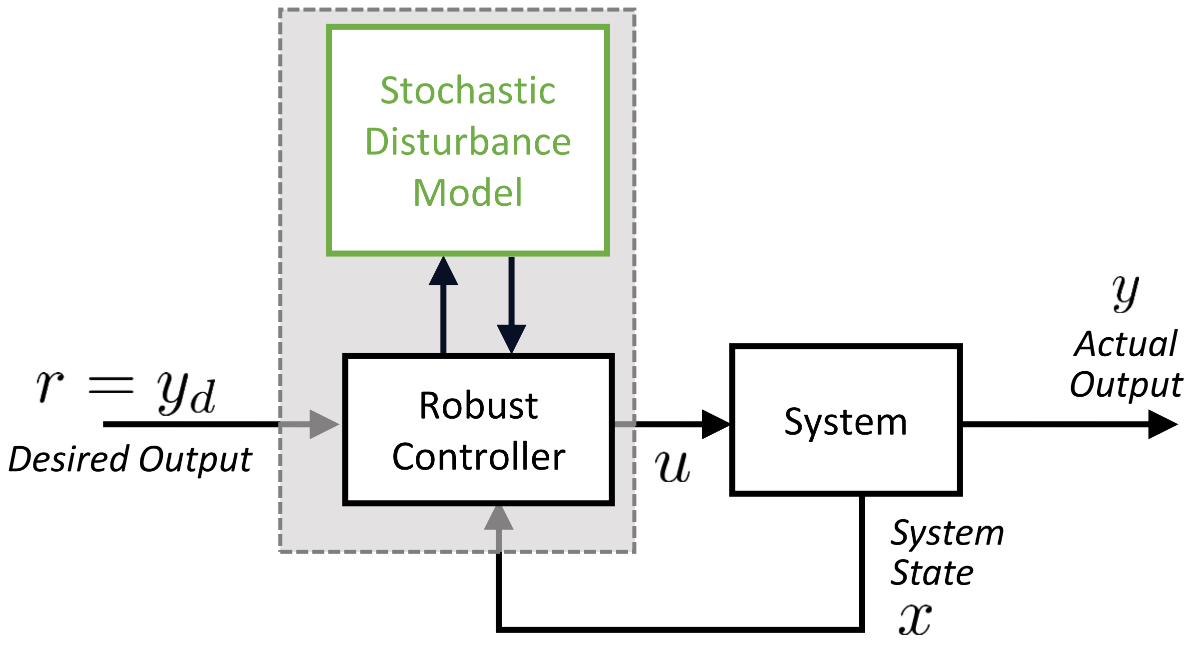
Gaussian processes

reliable confidence intervals

Algorithm to safely acquire data and optimize task

Robust control

stability & performance under uncertainty



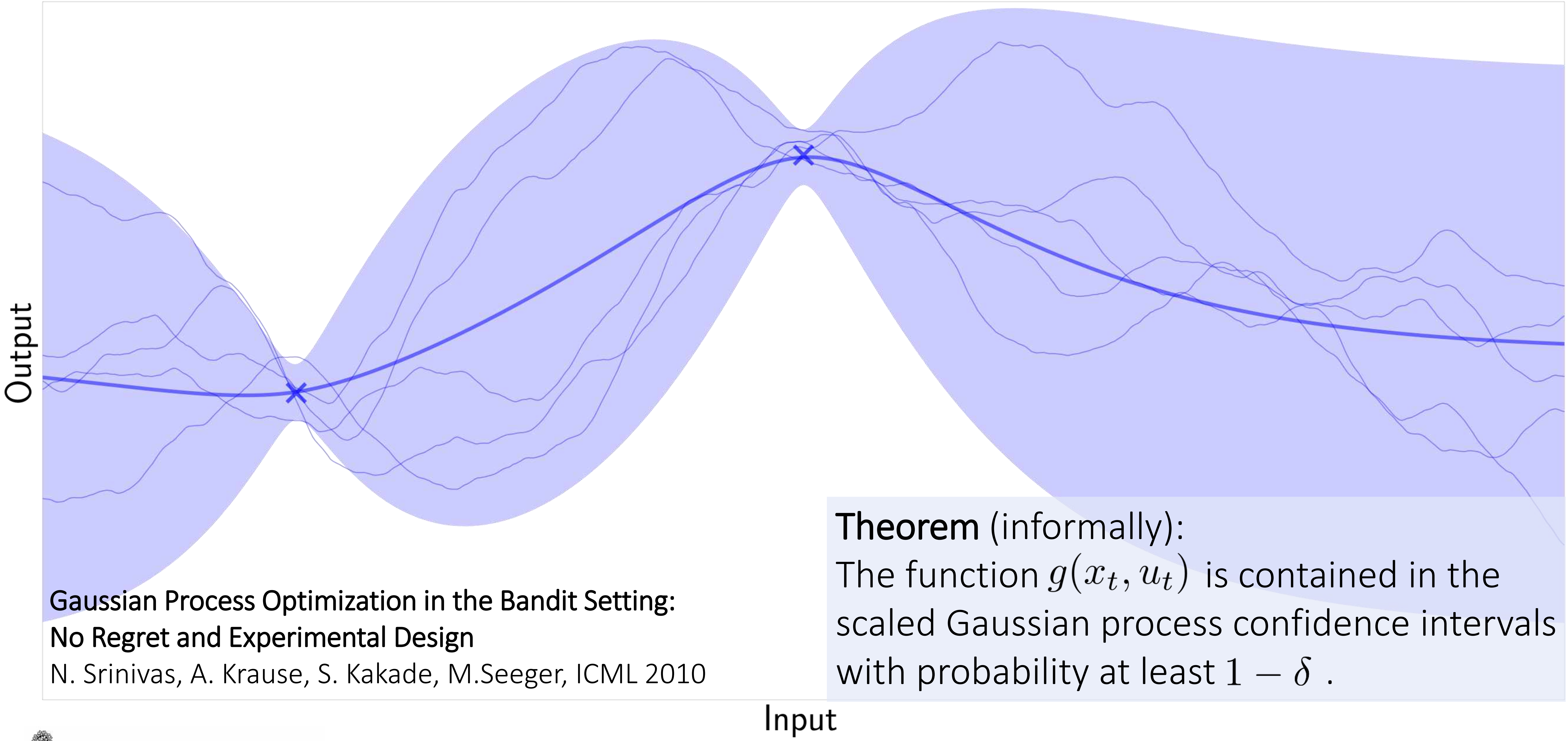
Defining and analyzing closed-loop safety

Lyapunov analysis

stability of learned models

= safe model-based reinforcement learning

Gaussian Process



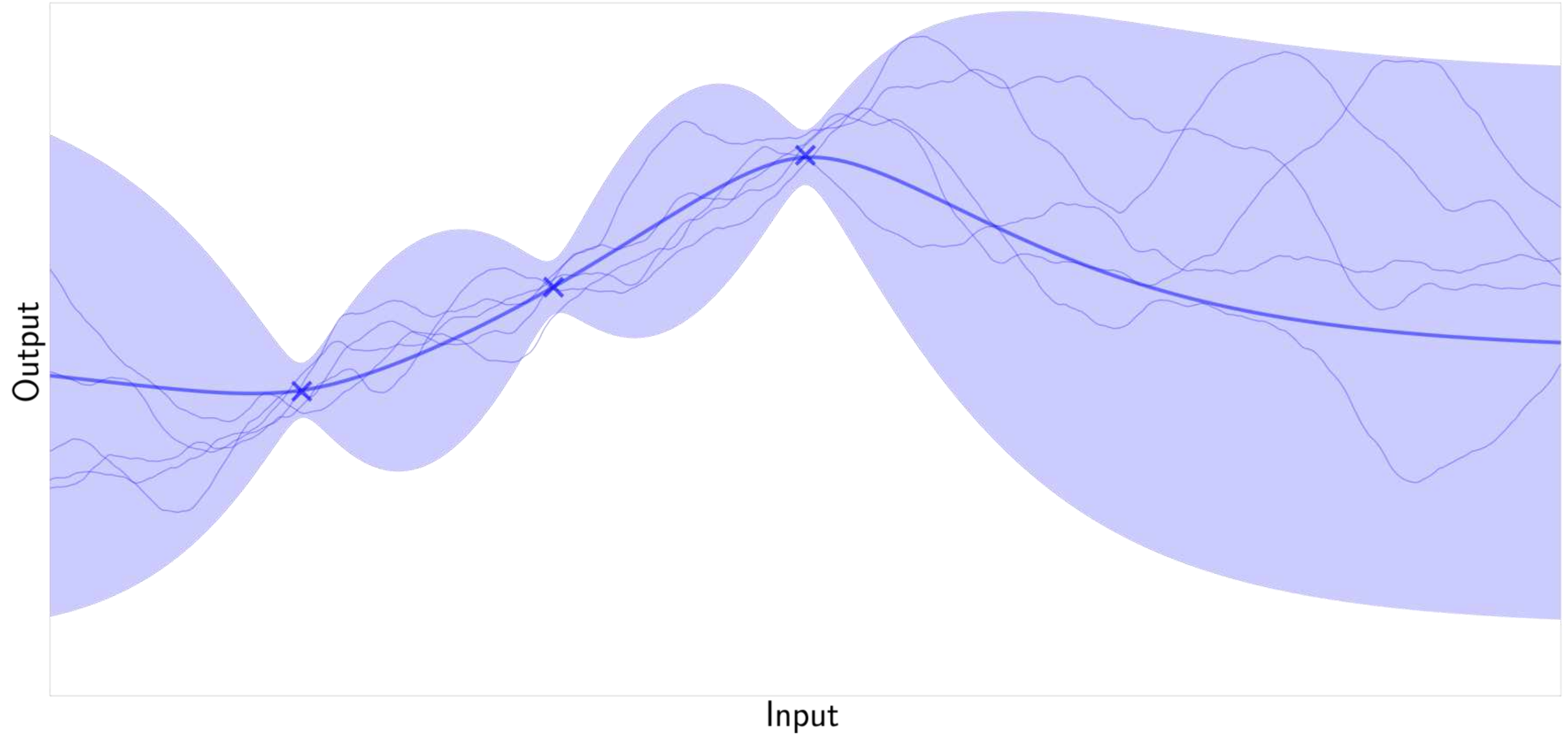
Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design
N. Srinivas, A. Krause, S. Kakade, M.Seeger, ICML 2010

Theorem (informally):
The function $g(x_t, u_t)$ is contained in the scaled Gaussian process confidence intervals with probability at least $1 - \delta$.

Gaussian Process

- Can model arbitrary smooth functions.
- For a given input, it provides an interval in which the function value lies with high probability.
- As more data is gathered, the uncertainty is reduced.

Our model framework for developing reinforcement learning algorithms with safety guarantees.



Approach

Nonparametric model for unknown model error

Algorithm to safely acquire data and optimize task

Defining and analyzing closed-loop safety

Gaussian processes
reliable confidence intervals

Robust control
stability & performance under uncertainty

Lyapunov analysis
stability of learned models

1. Linear
2. Nonlinear
3. Nonlinear, predictive

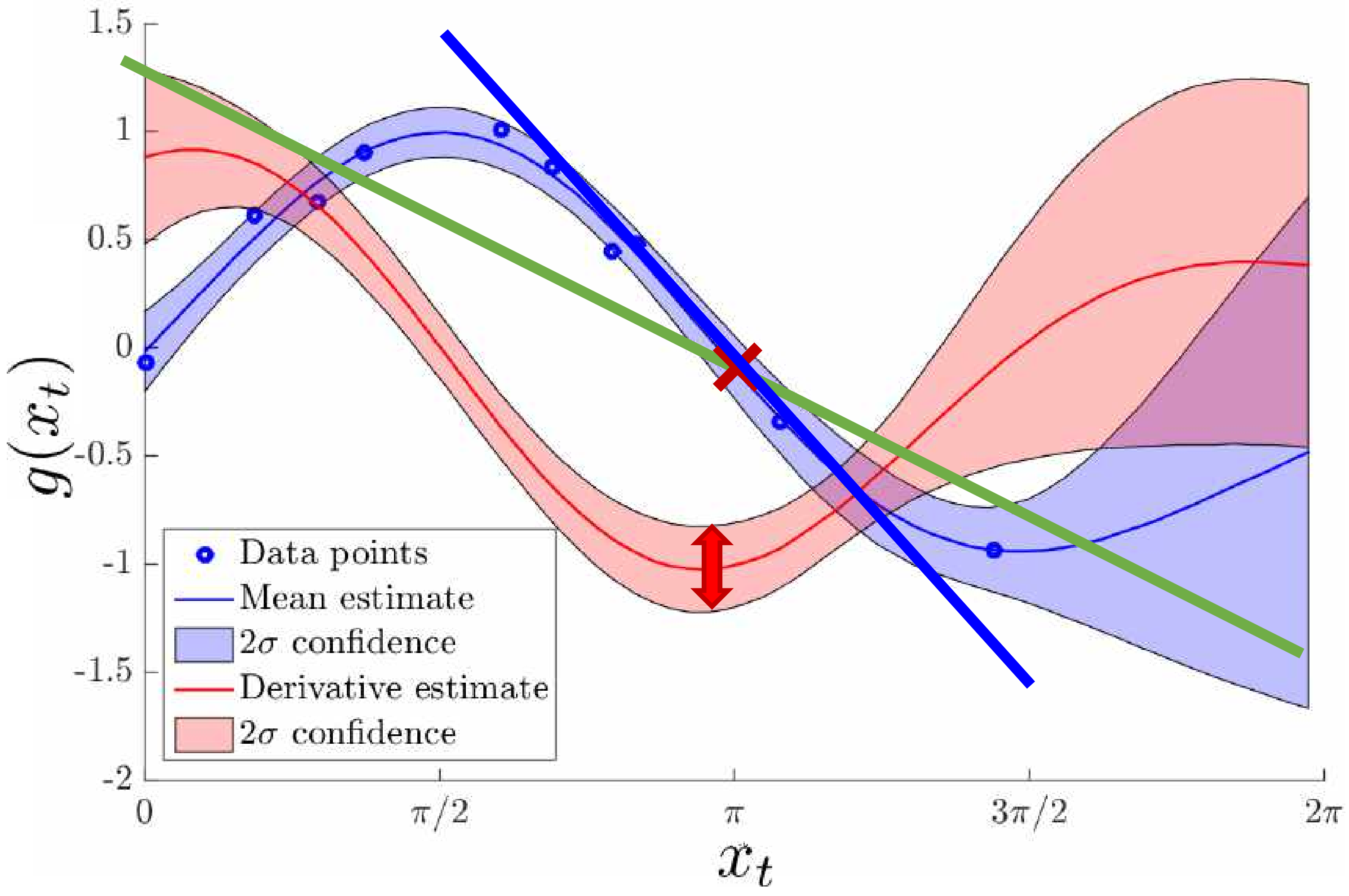
= safe model-based reinforcement learning

Linear Robust Control [ECC'15]

- **Gaussian Process Model**

$$x_{t+1} = \underbrace{f(x_t, u_t)}_{\text{a priori model}} + \underbrace{g(x_t, u_t)}_{\text{unknown}} + \underbrace{\xi_t}_{\text{noise}}$$

- **Linear Robust Control**
  - Task: stabilization of an operating point
  - **Linear robust control:**
    - *linearization about operating point*

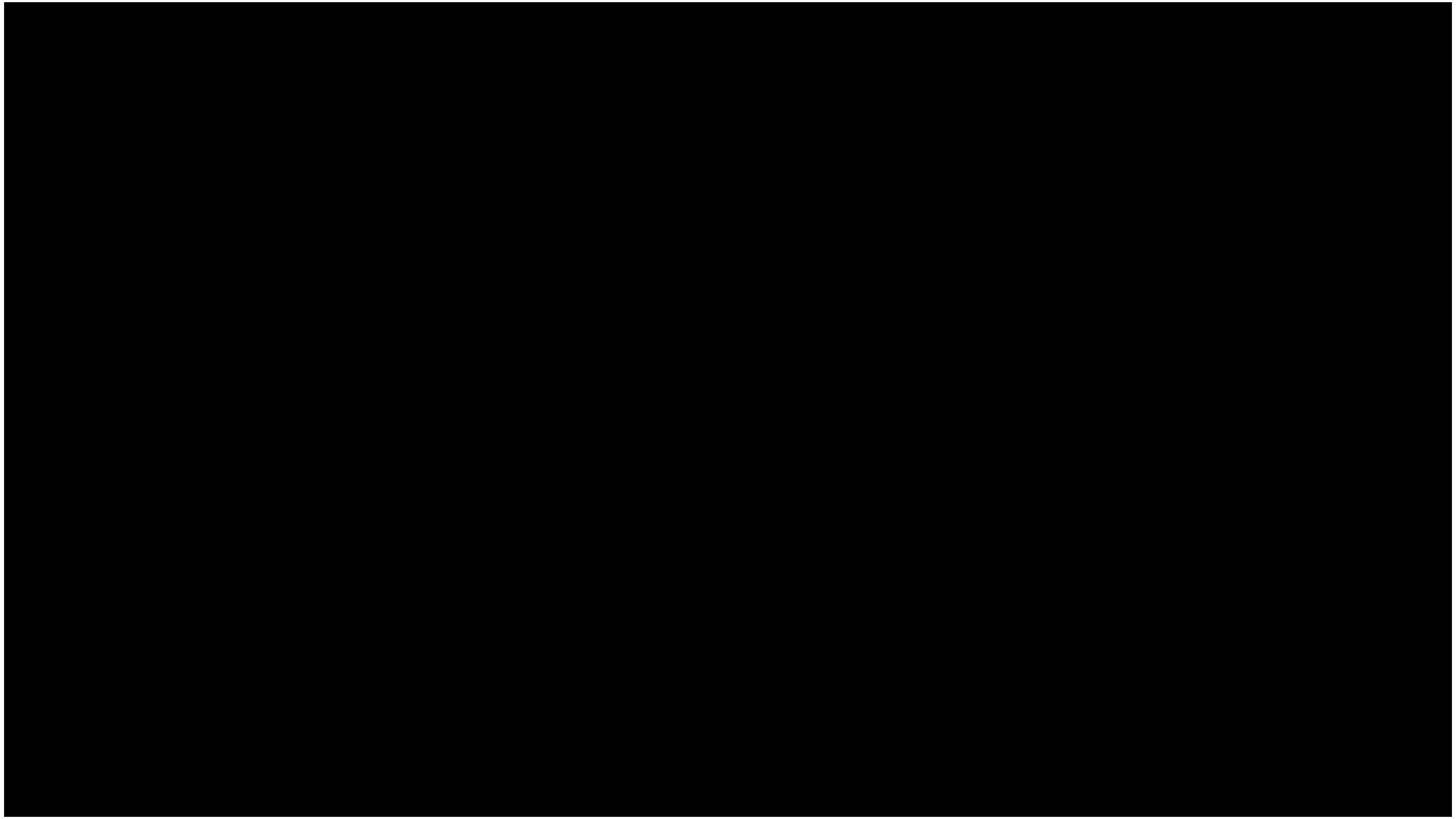


- **<u>Local</u> Stability Guarantees**
  - Local asymptotic **stability around true operating point** with high probability

Linear Robust Control [ECC'15] https://youtu.be/YqhLnCm0KXY

Linear Robust Control [ECC’15] https://youtu.be/YqhLnCm0KXY

DYNAMIC
SYSTEMS LAB



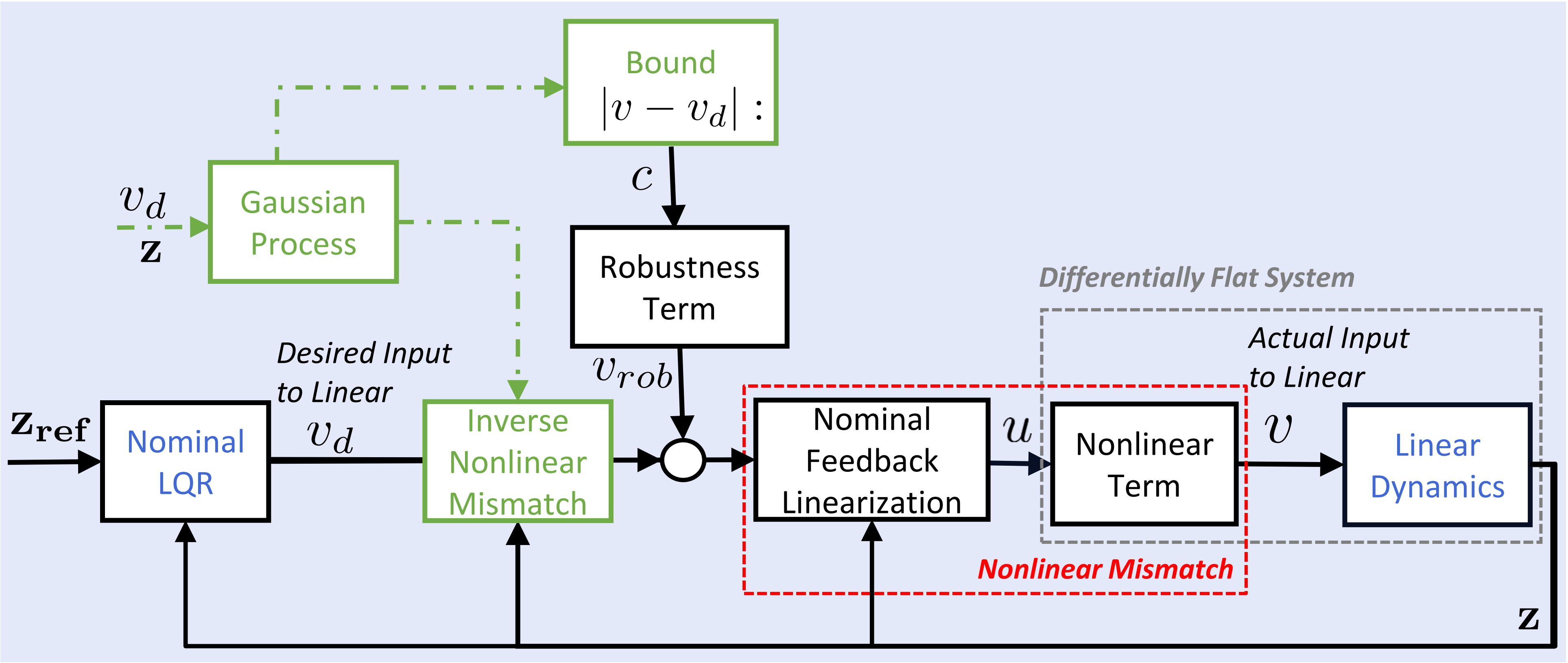
Nonlinear Robust Control for Differentially Flat Systems [L-CSS'20]

- **Model / Assumptions**
  - Differentially flat, control-affine real dynamics and prior model
  - Gaussian Process models inverse nonlinear mismatch

$$x_{t+1} = \underbrace{f_x(x_t) + f_u(x_t)u_t}_{\text{a priori model}} + \underbrace{g_x(x_t) + g_u(x_t)u_t}_{\text{unknown}} + \underbrace{\xi_t}_{\text{noise}}$$

Nonlinear Robust Control for Differentially Flat Systems [L-CSS'20]

- **Model / Assumptions**
  - Differentially flat, control-affine real dynamics and prior model
  - Gaussian Process models inverse nonlinear mismatch
- **Linear Robust Control**
  - Task: high-performance tracking
  - Linear robust control for feedback-linearized system
- **<u>Global</u> Tracking Guarantees**
  - Tracking error is uniformly ultimately bounded with high probability



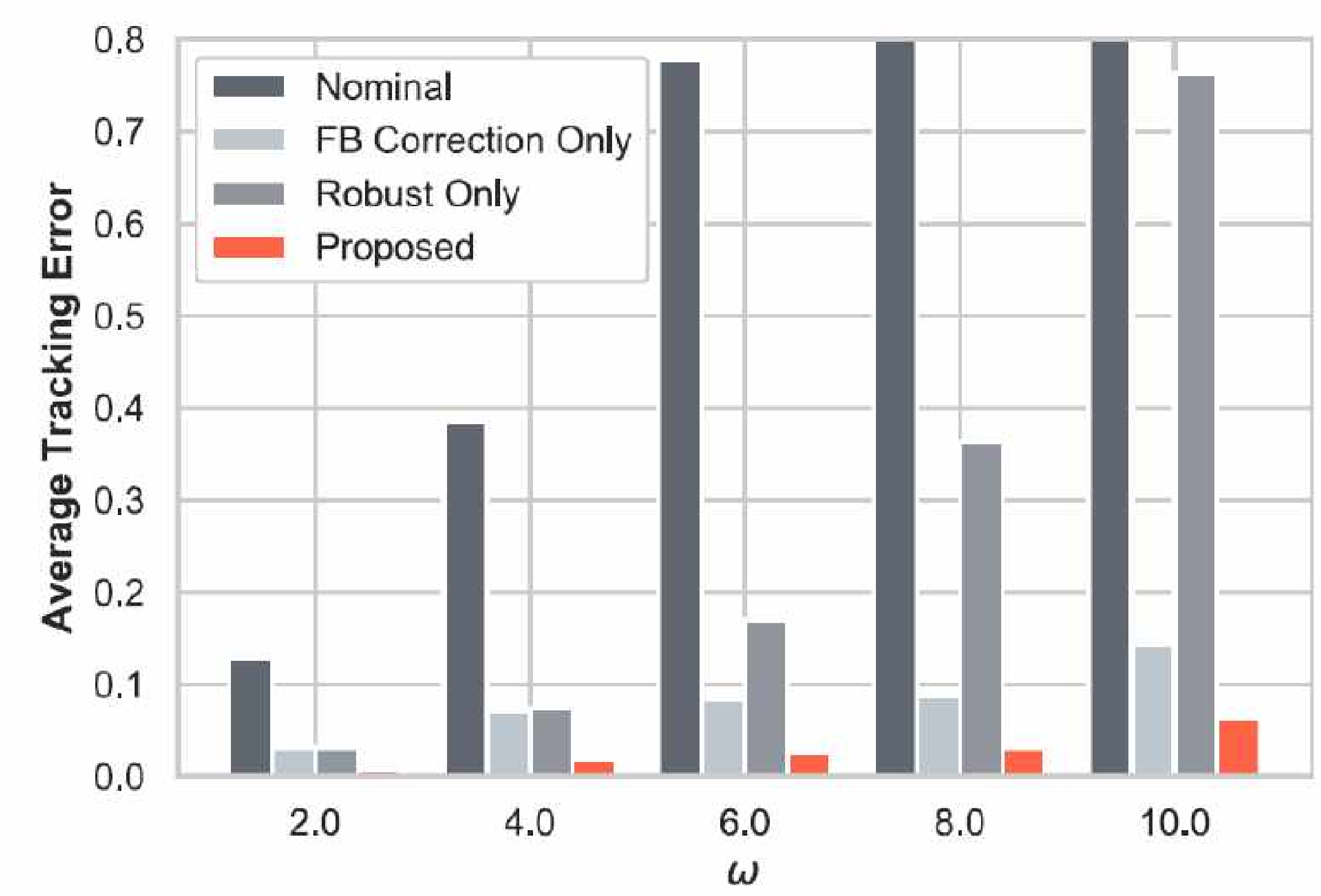
$$\mathbf{z} = [y, \dot{y}, \ldots, y^{(n-1)}]^T$$

Nonlinear Robust Control for Differentially Flat Systems [L-CSS'20]

DYNAMIC SYSTEMS LAB

Cart-pendulum example with model parameter uncertainties:

$$y_d = 0.3t \sin(\omega t)$$



Robust, online learning control with global guarantees on tracking error.

- Predictive capabilities
- State constraints

Robust Predictive Control [IJRR'16, JFR'16]

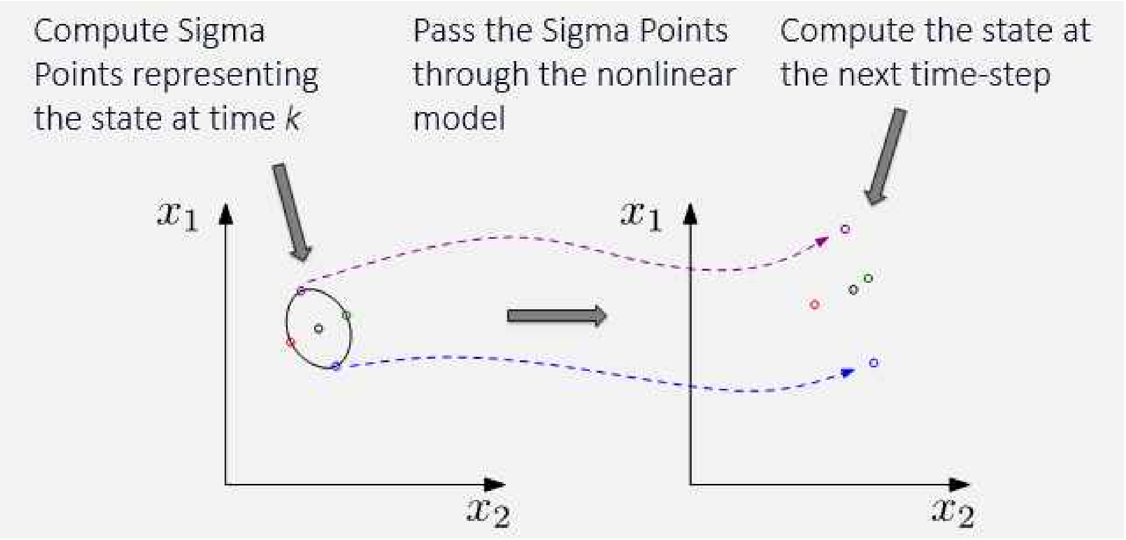
- **Gaussian Process Model**
- **Nonlinear, Robust Model Predictive Control**
  - Task: high-performance tracking
  - Approximations in prediction and nonlinear optimization step
- **Guarantees** [e.g., Tomlin'13, Krause'18, Zeilinger'18]
  - Robustly asymptotically stable
  - Robust constraint satisfaction
  - Recursively guaranteeing the existence of safe control actions

$$\min_{\{u_{0:T-1}\}} E\left[\sum_{k=0}^{T-1} J(x_t, u_t) + J_T(x_t)\right]$$

$$\text{s.t.} \quad x_{t+1} = f(x_t, u_t) + g(x_t, u_t), \quad g \sim GP$$

$$x_0 \sim \mathcal{N}(\bar{x}_0, \Sigma_{x_0})$$

$$\text{Prob}(x_t \in \mathcal{X}_t) \geq 1 - \delta, \quad u_t \in \mathcal{U}_t$$

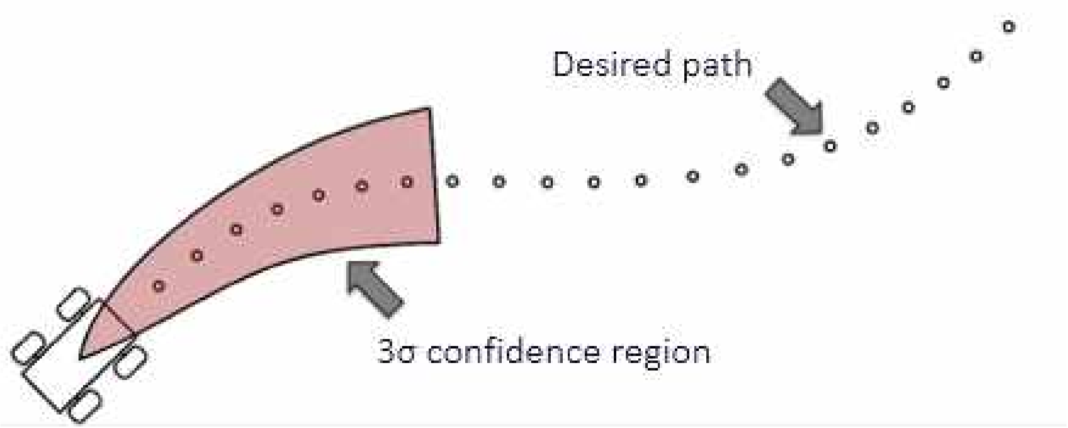


Unscented Transform for prediction

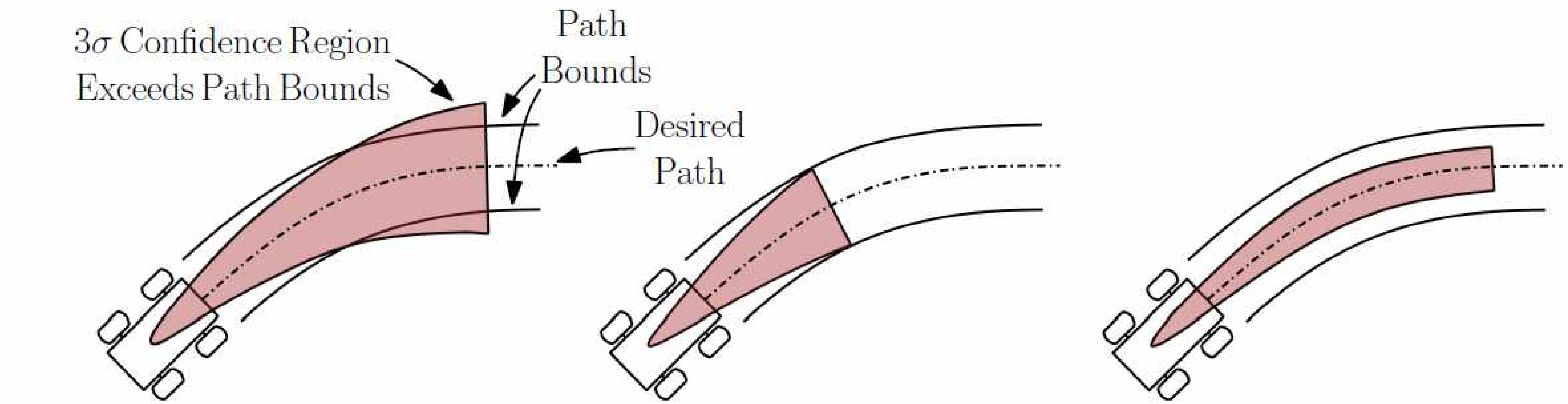
DYNAMIC SYSTEMS LAB

Example: Mobile robot path following

- Problem setup:

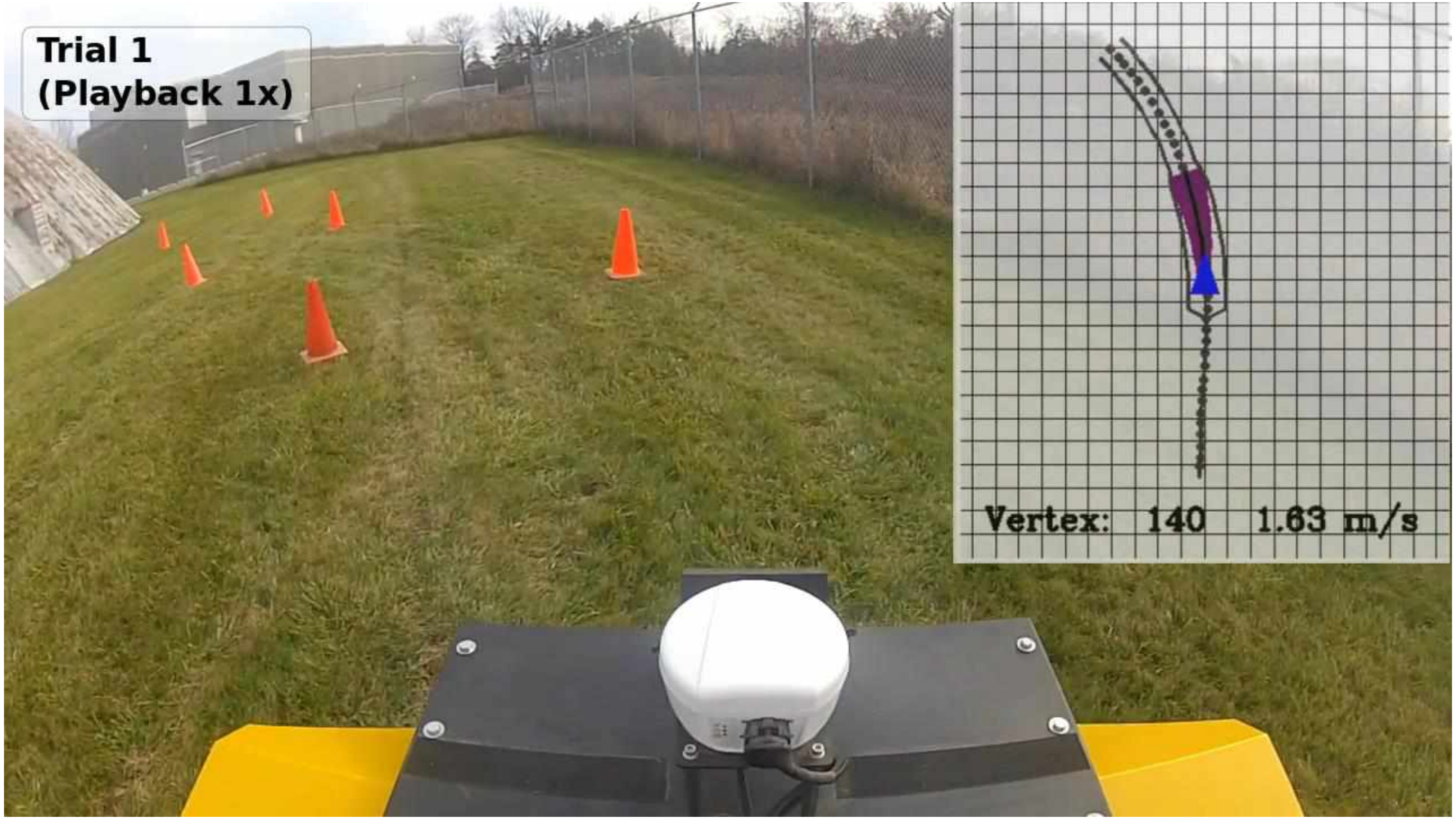


- Learning:



Driving too fast | Slow down for safety | Faster driving after learning

Robust Predictive Control [IJRR’16, JFR’16] https://youtu.be/3xRNmNv5Efk



Summary

Design a controller for systems with prior uncertainty that learns online and continuously improves performance while satisfying safety constraints.

Nonparametric model for unknown model error

Gaussian processes
reliable confidence intervals

Algorithm to safely acquire data and optimize task

Robust control
stability & performance under uncertainty

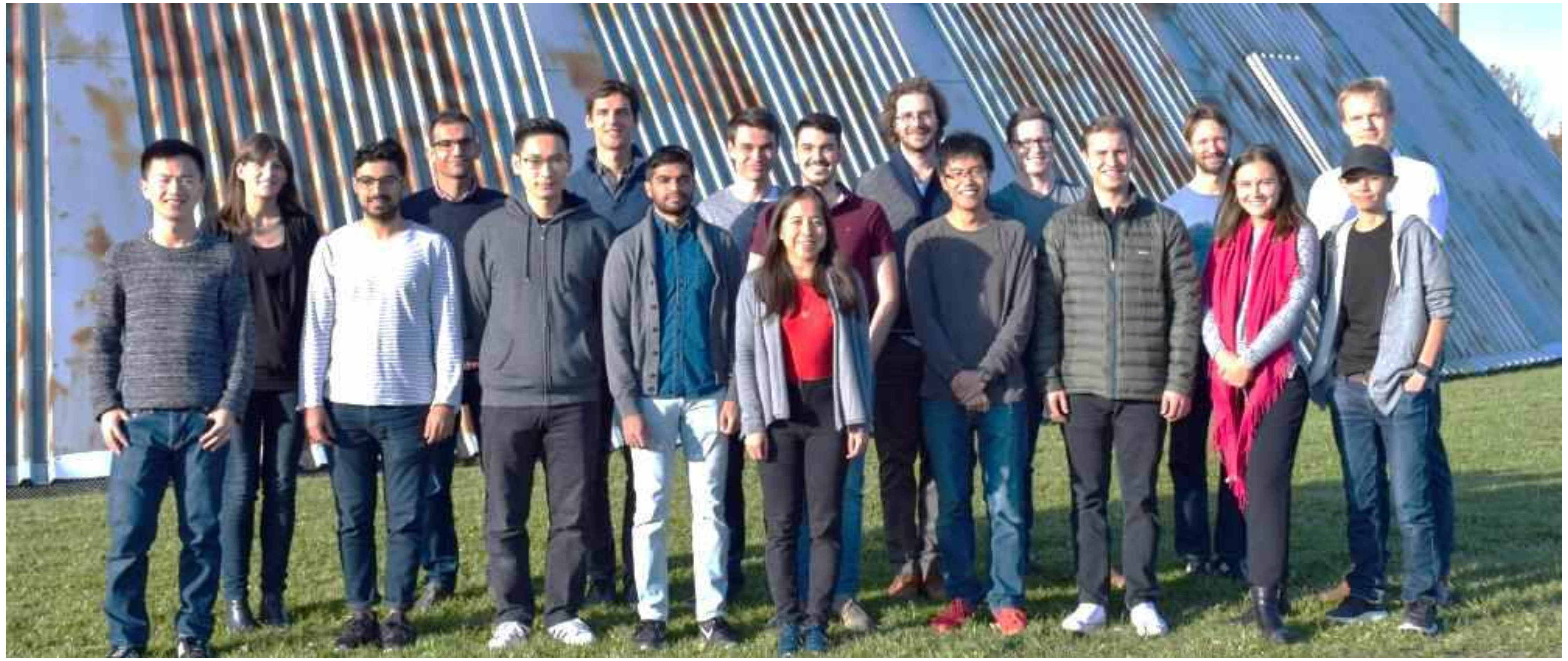
Defining and analyzing closed-loop safety

Lyapunov stability
stability of learned models

1. Linear
   - **Local stability guarantees**
2. Nonlinear
   - **Global tracking error guarantees**
3. Nonlinear, predictive
   - **Probabilistic constraint satisfaction and stability**

Acknowledgements

DYNAMIC SYSTEMS LAB



Follow us!

www.dynsyslab.org

Senior collaborators: Andreas Krause, Tim Barfoot, Raffaello D'Andrea

Funding:

Alfred P. Sloan Foundation

CIFAR Canadian Institute for Advanced Research

Canada Research Chairs www.chairs-chaires.gc.ca

Ontario Research Fund

DRONE DELIVERY CANADA.COM

NSERC CRSNG

INNOVATION.CA Canada Foundation for Innovation | Fondation canadienne pour l'innovation

Ontario Centres of Excellence

Ontario Power Generation

Other Learning Control Results from My Lab

DYNAMIC SYSTEMS LAB

- **Systems with changing dynamics**
  [ICRA’17, IROS’18, RAL’18, JACSP’19, RAL’19]
- **Transfer learning between similar systems** (similarity metric from robust control)
  [IROS’17, ICRA’17, RAL’18, ACSP’18]
- **Collaborative learning of interconnected systems**
  [AURO’19]
- **Active learning**
  [ICRA’16, NeurIPS’17, CDC’19]

M. Paton, “Expanding the Limits of Vision-Based Autonomous Path Following,”, 2017.