# Learning Deep Control Policies for Autonomous Aerial Vehicles with MPC-Guided Policy Search

**Tianhao Zhang, Gregory Kahn, Sergey Levine, Pieter Abbeel**

**Berkeley Artificial Intelligence Research Laboratory (BAIR)**

# Motivation

- Enable autonomous aerial vehicles (AAVs) to navigate complex, unstructured environments

# Challenges

- Complex, unstructured environments
  - no explicit state estimation
- Use raw observations from onboard sensors
  - high dimensionality and non-linearity
- Real-time evaluation at test time
  - computationally efficient

Guided Policy Search *

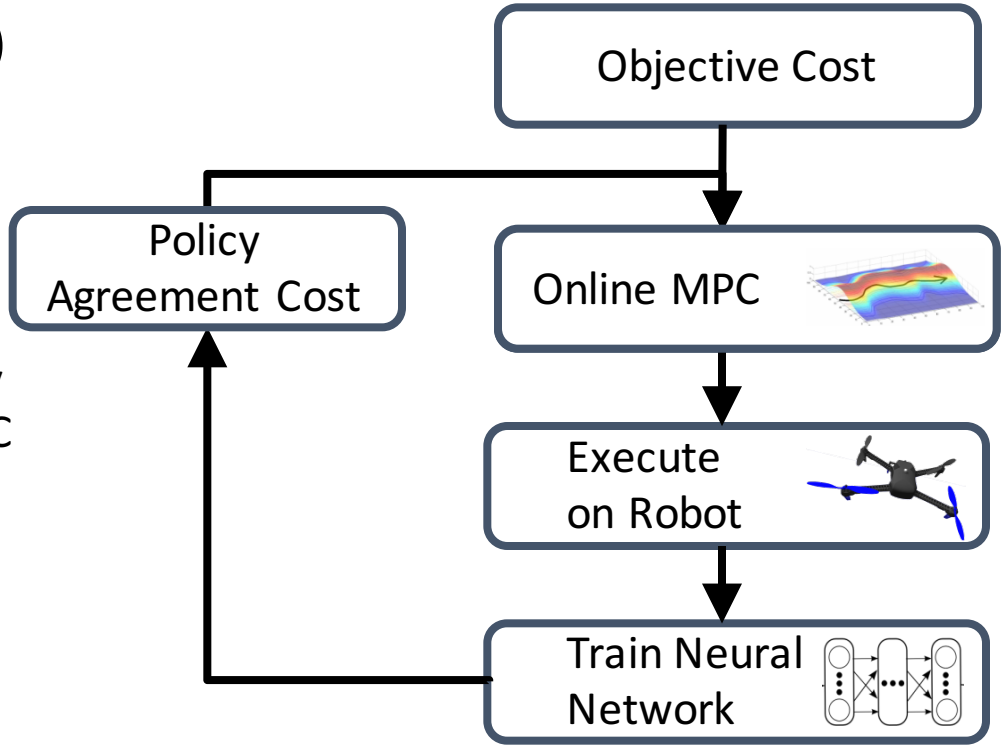
- Robust to model errors and environment disturbances
  - flying systems are prone to catastrophic failures

Model Predictive Control

* S. Levine & P. Abbeel. "Learning neural network policies with guided policy search under unknown dynamics." *NIPS*. 2014.
S. Levine et al. "End-to-end training of deep visuomotor policies." JMLR. 2015.

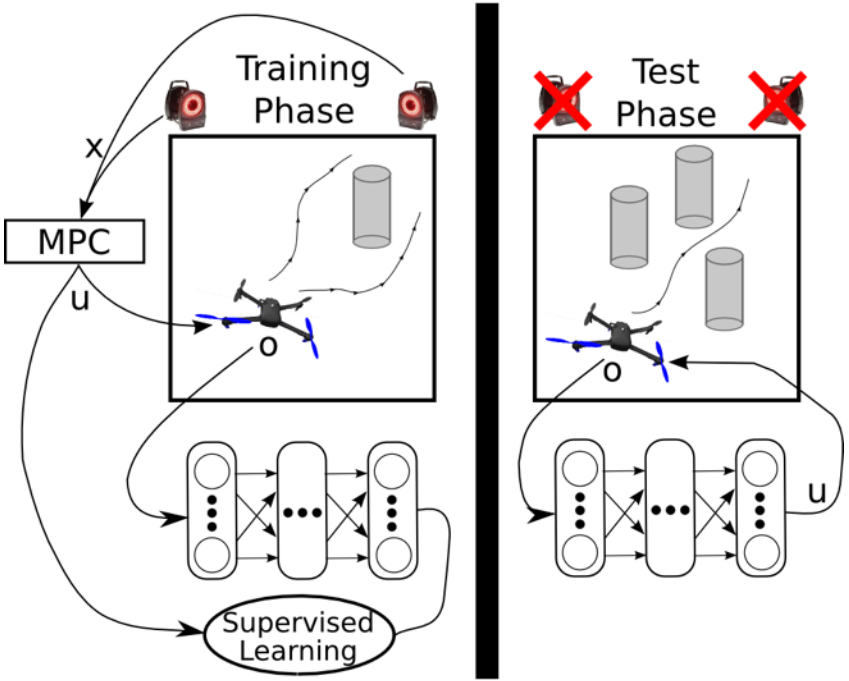
# Approach: MPC-GPS

- Guided Policy Search (GPS)
  - Trajectory optimization
  - Supervised learning
  - Policy agreement
- MPC-GPS
  - Substitute offline trajectory optimization for online MPC

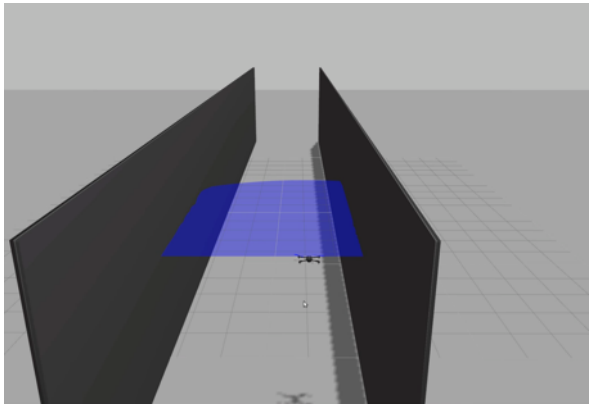
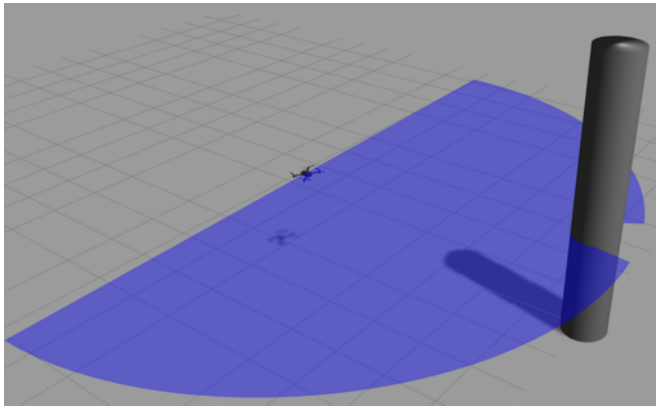


# Approach: MPC-GPS (cont.)

- Training:
  - use instrumented setup to obtain full state information **x**
  - MPC uses **x** to generate trajectories
  - Record observations **o**
  - Policy is trained to map from **o** to **u**
- Test:
  - No need for instrumented setup
  - Policy runs in closed loop



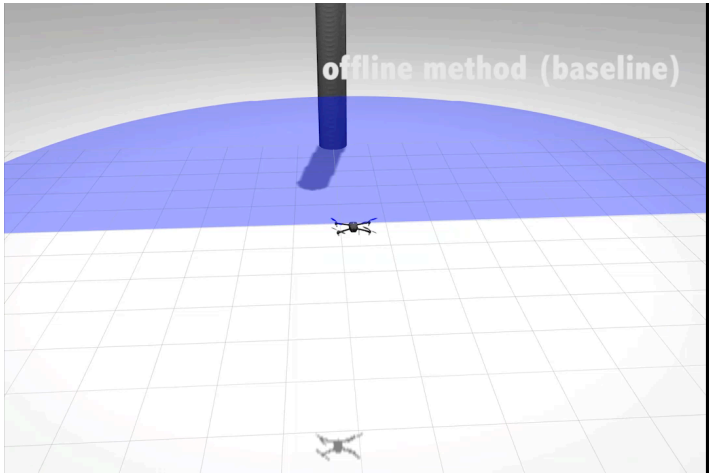
# Experiment Overview

| | Hallway | Cylinder |
|---|---|---|
| Training Environments |  |  |
| Model Errors | • no model error<br>• 8% rotor bias | • 0.05kg mass error<br>• perturbed model params |

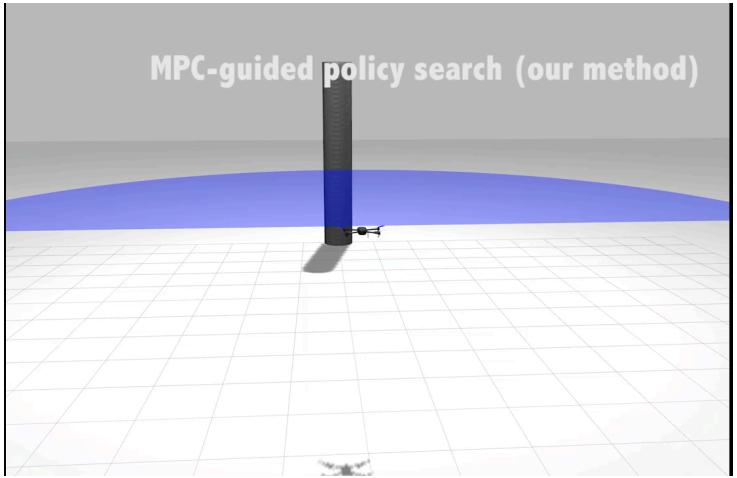
# Experimental Evaluation

Cylinder - 0.05kg mass error

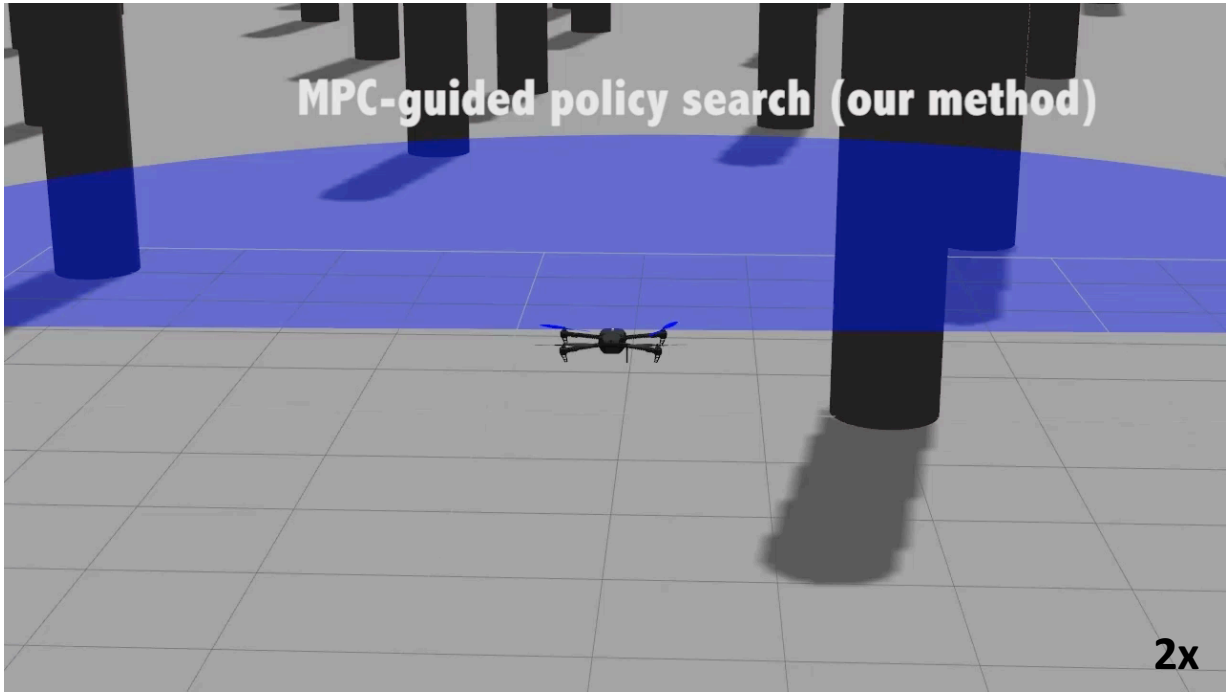
**Baseline**



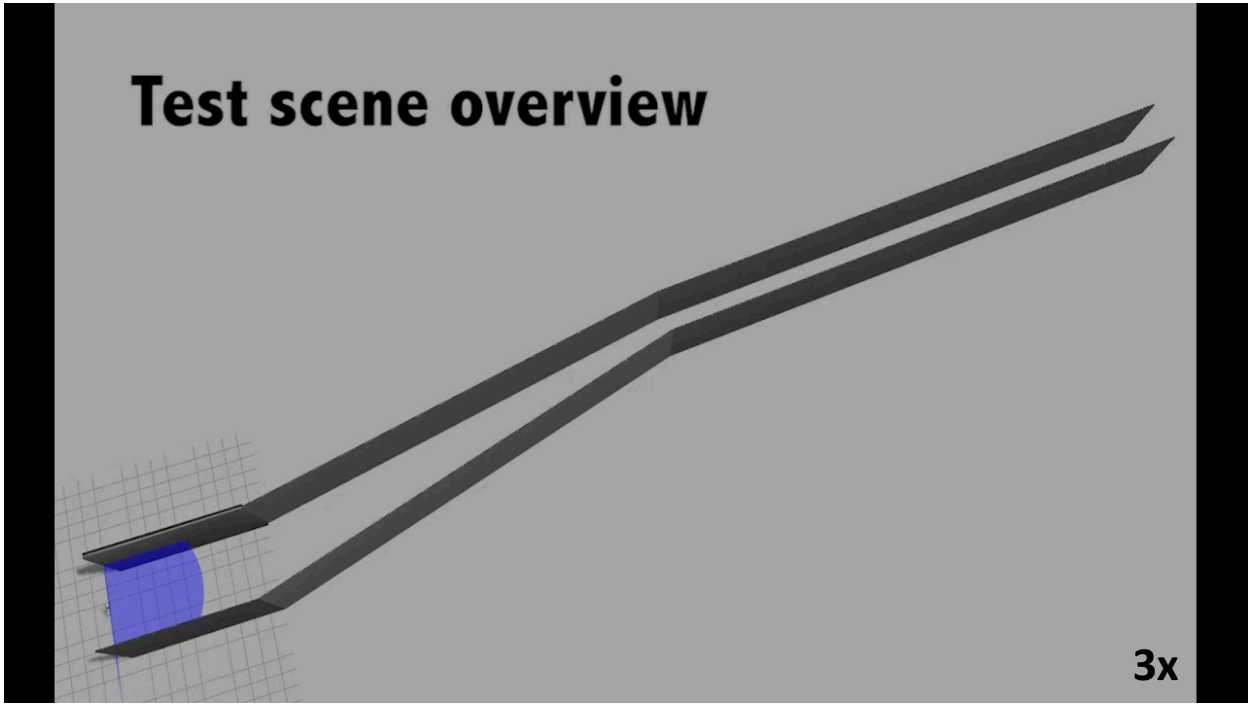

**MPC-GPS**



# Experimental Evaluation (cont.)



# Experimental Evaluation (cont.)



Thank you!