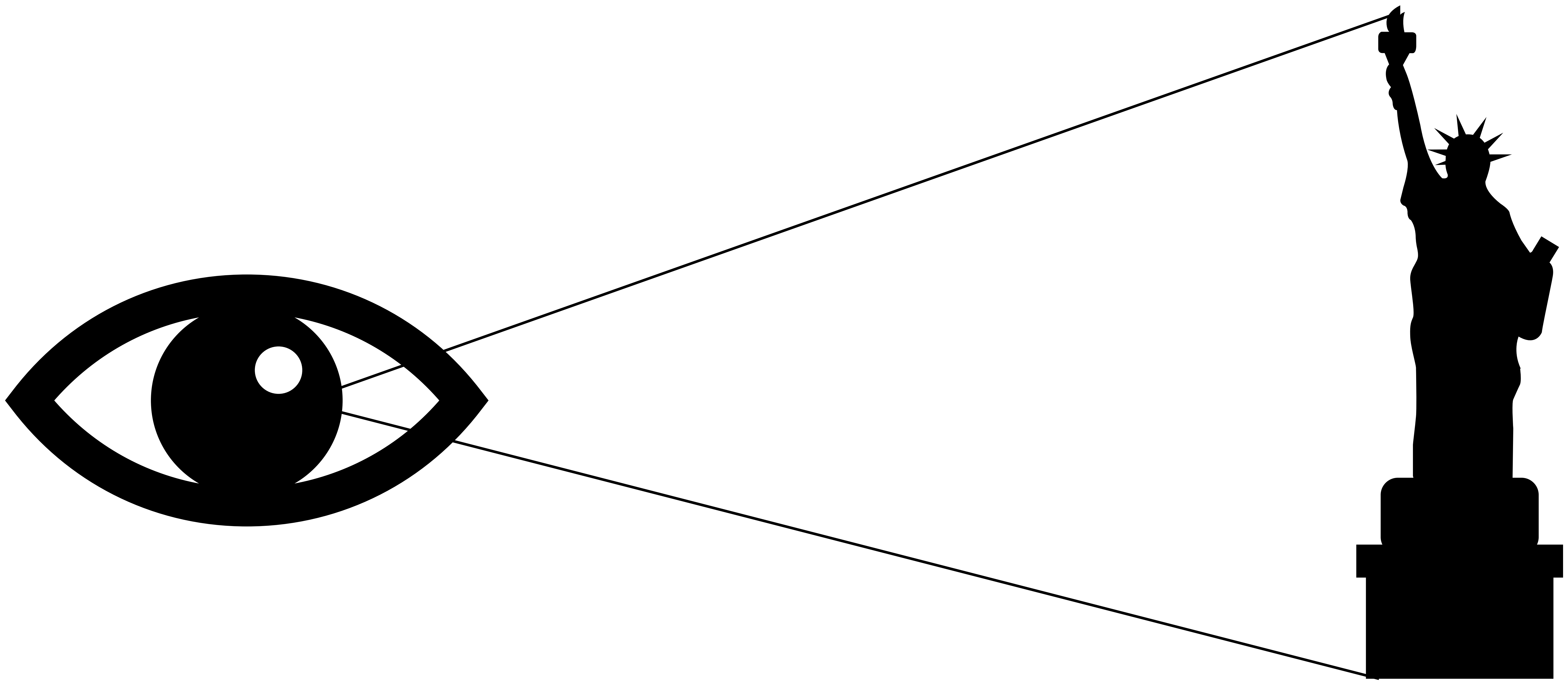
Simulation:
What made us intelligent will make our robots intelligent

Antonio Loquercio

Penn
University of Pennsylvania

What is Perception?



Aristotele (*De Anima*): “Perception is a kind of reception of the form of the perceived object without the matter.”

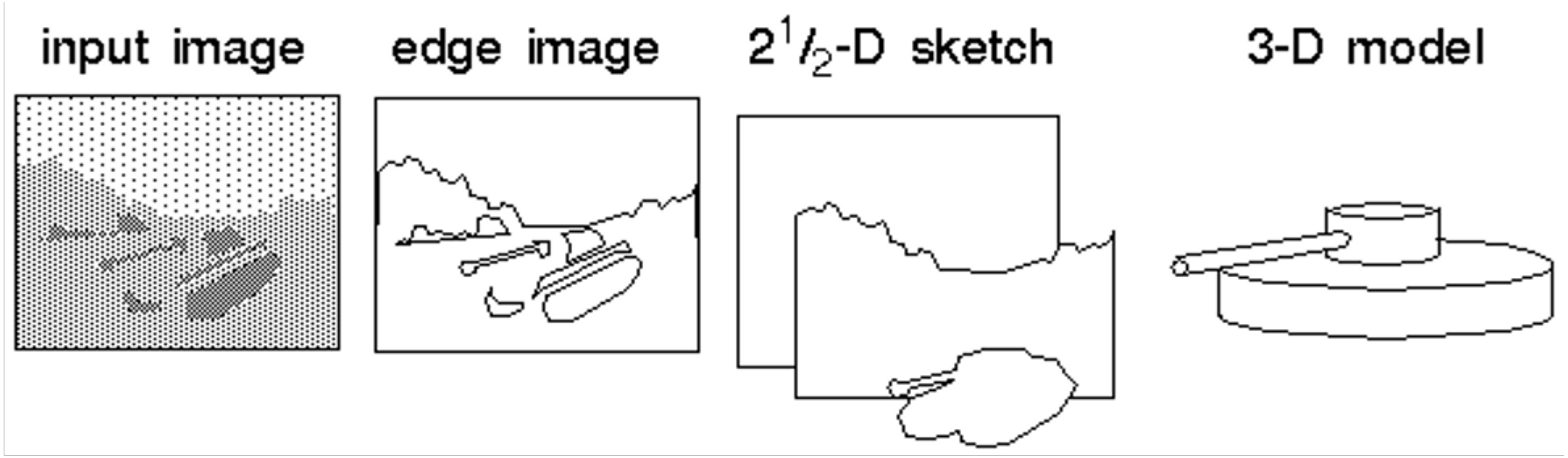
Vision (1982) By David Marr

- PhD in Neuroscience, Cambridge
- Professor of Psychology at MIT (1977-1980)
- Posthumous book: Vision (1982)

> In December 1977, certain events occurred that forced me to write this book a few years earlier than I had planned. Although the book has important gaps, which I hope will soon be filled, a new framework for studying vision is already clear and supported by enough solid results to be worth setting down as a coherent whole.

An *Information Processing* Theory of Vision

Vision as Inverse Graphics or Inverse Optics

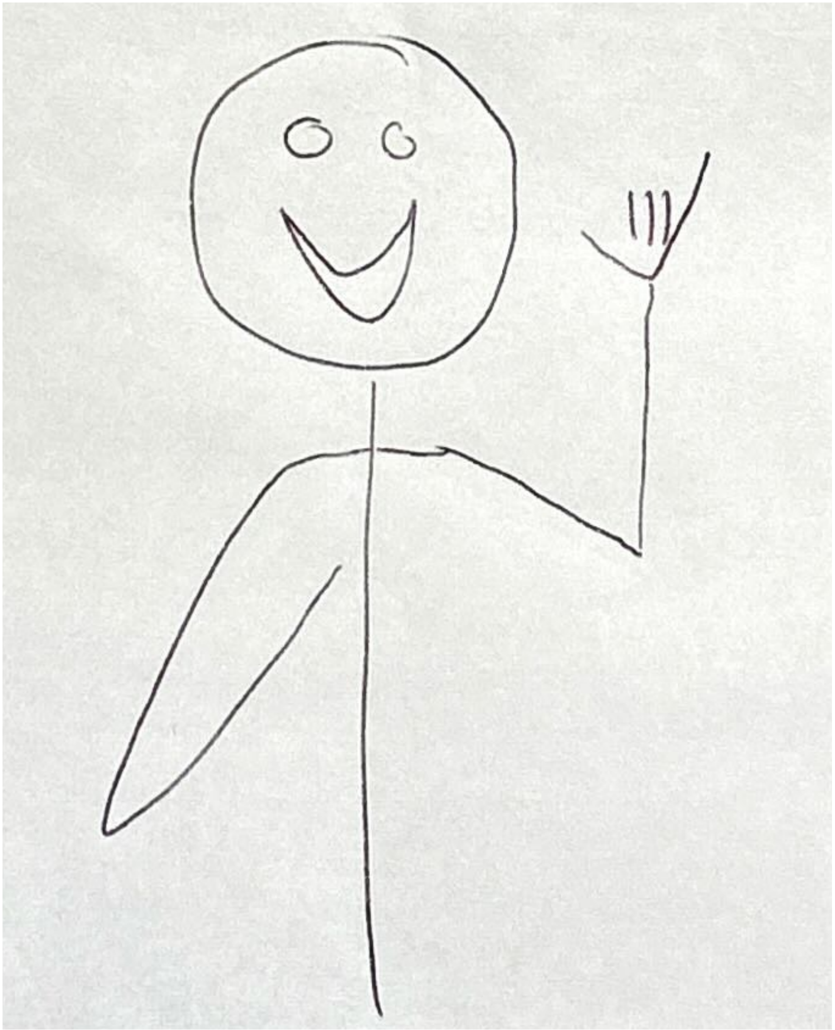


- Very appealing from a “software engineering” perspective.
  - Modularity, feedforward pipeline

The Theory Does not Fit The Data

- Humans do not recover veridical, task-independent 3D representations
- Principles of modularity and feedforward processing don't hold for human vision

What we *think* we see



What we *really* see

W. Warren, [Does This Computational Theory Solve the Right Problem? Marr, Gibson, and the Goal of Vision](), Perception 41(9), 2012

Our Brain 200 Million Years Ago

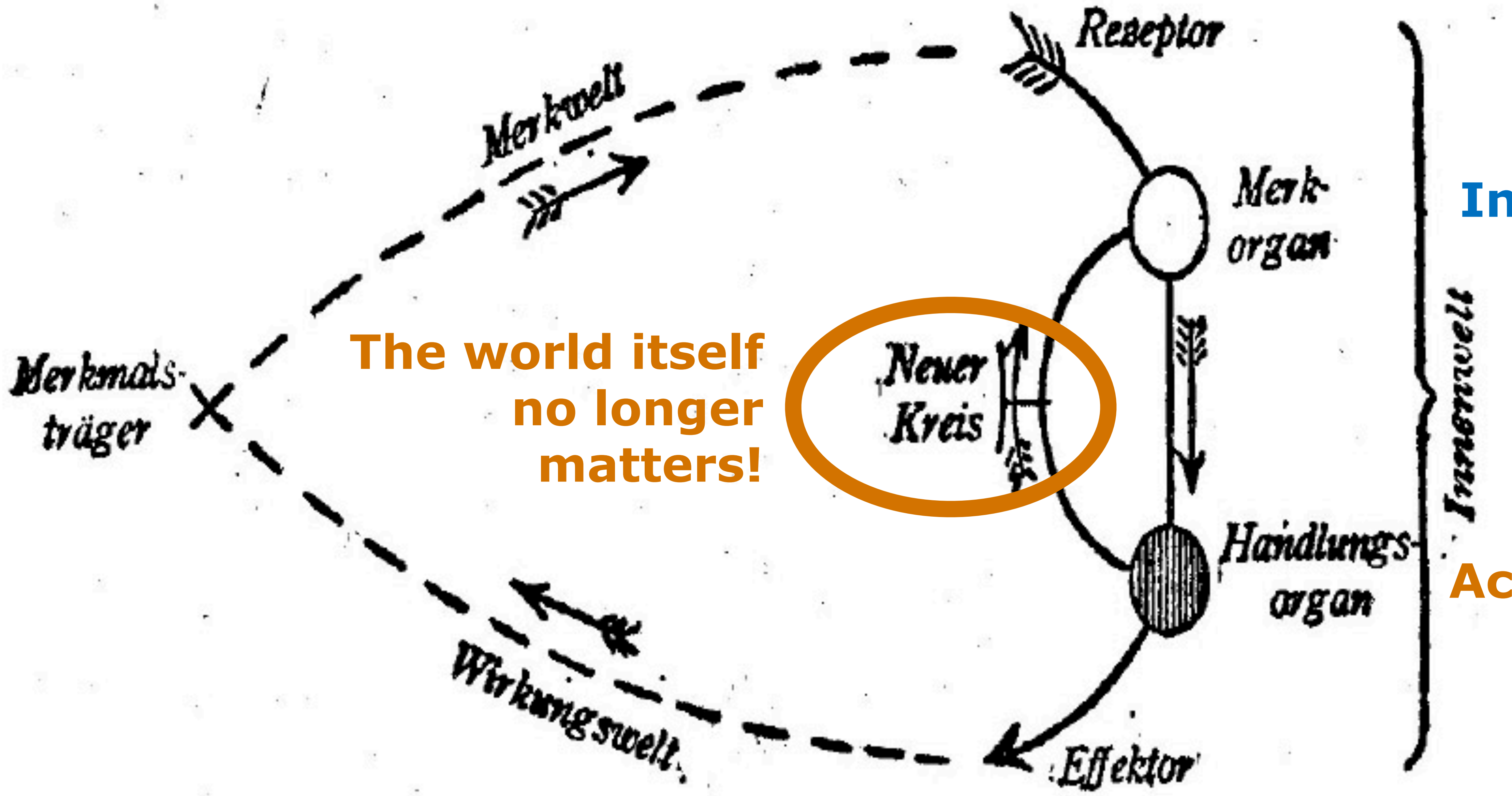
Illustration by Rebecca Gelernter.
From the book "*A brief history of intelligence*"

- Simulate the outcome of actions before they occur.
- Enabled by a neurological breakthrough: the neocortex.
- 76% of our brain consists of the neocortex.

An Alternative Theory of Perception

- Jakob von Uexküll, German biologist (1864-1944)
- Theory of “Umwelt” (Sensory-action surrounding worlds).

Feedback about the action



Input from the world

Action affects the world

Differences between organisms

Reflex agent

- Consider how the world IS
- Choose action based only on current percept
- Do not consider the future consequences of actions

Predictive agent

- Consider how the world WOULD BE
- Decisions based on (hypothesized) consequences of actions
- Must have a model of how the world evolves in response to actions

An Actionable Theory of Perception

- Natural selection optimizes fitness, not veridicality.

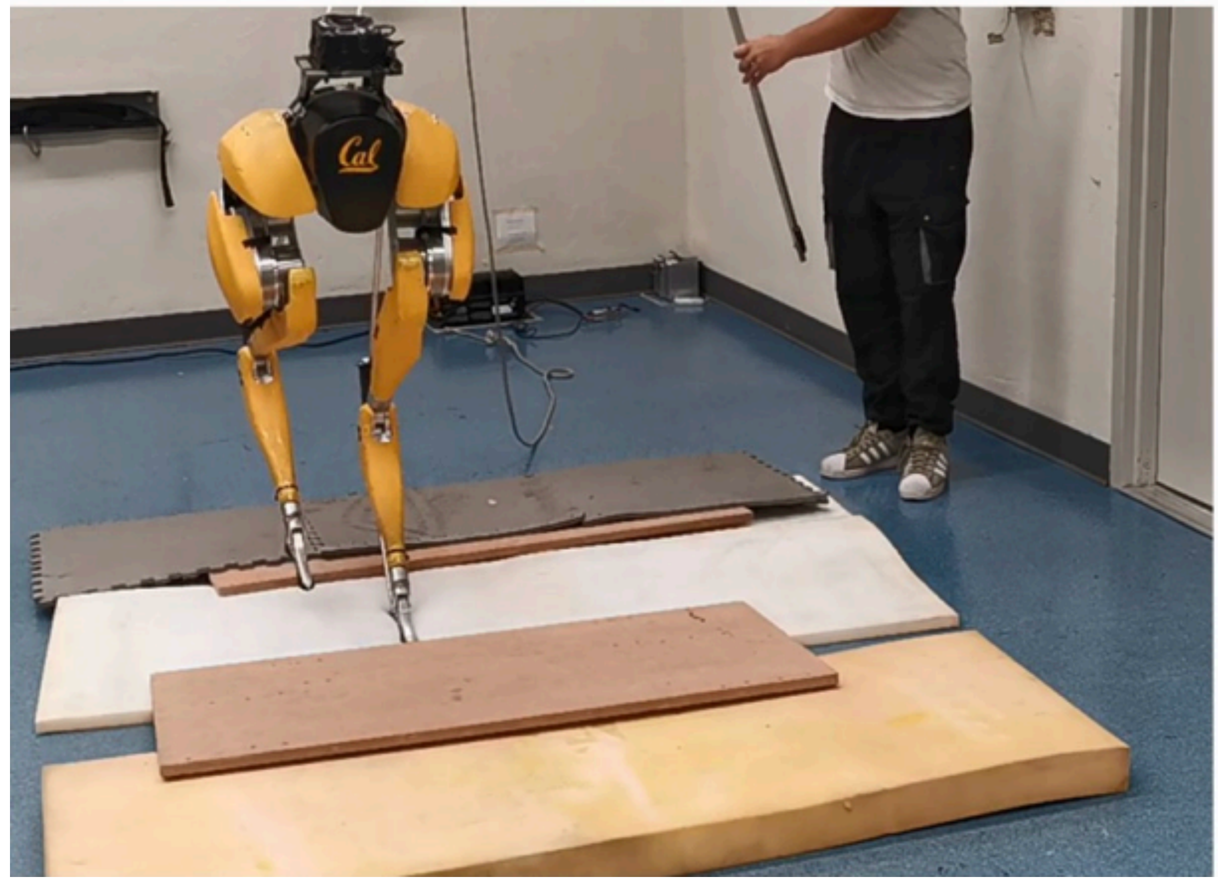
- The model of the world that an organism builds from perception is imperfect in many ways, but useful.

- **Discrepancies** between the predictions and the observed state of the world generate **“sparks of awareness”** (a view held by Erwin Schrödinger)

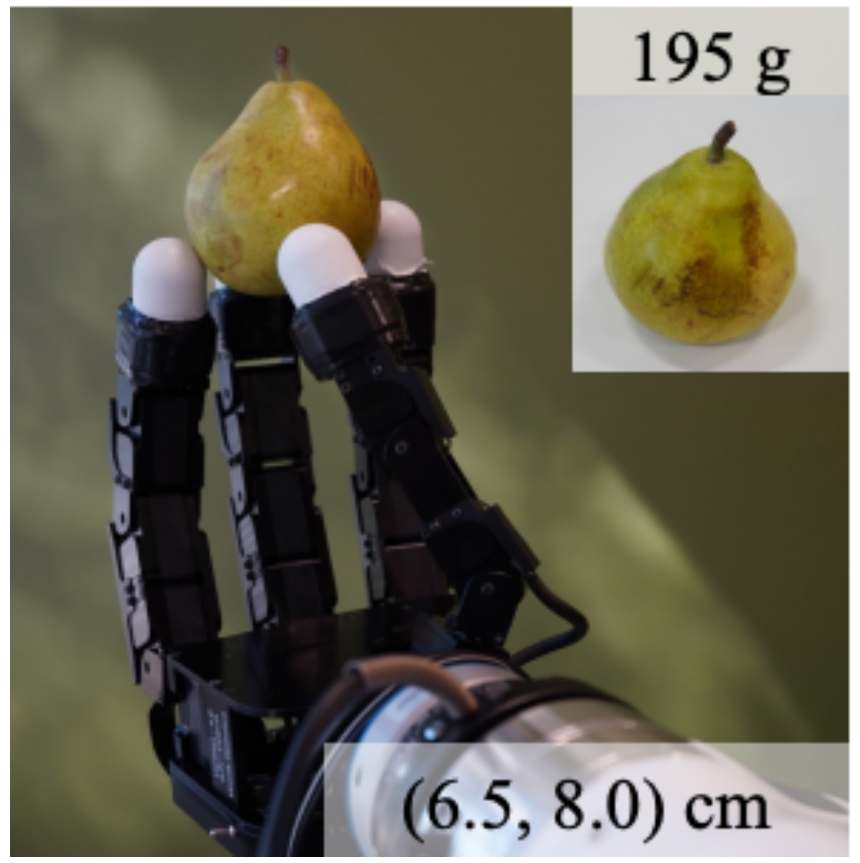
A Quiet Revolution in Robotics (2020-2022)

Loquercio et al, 2021

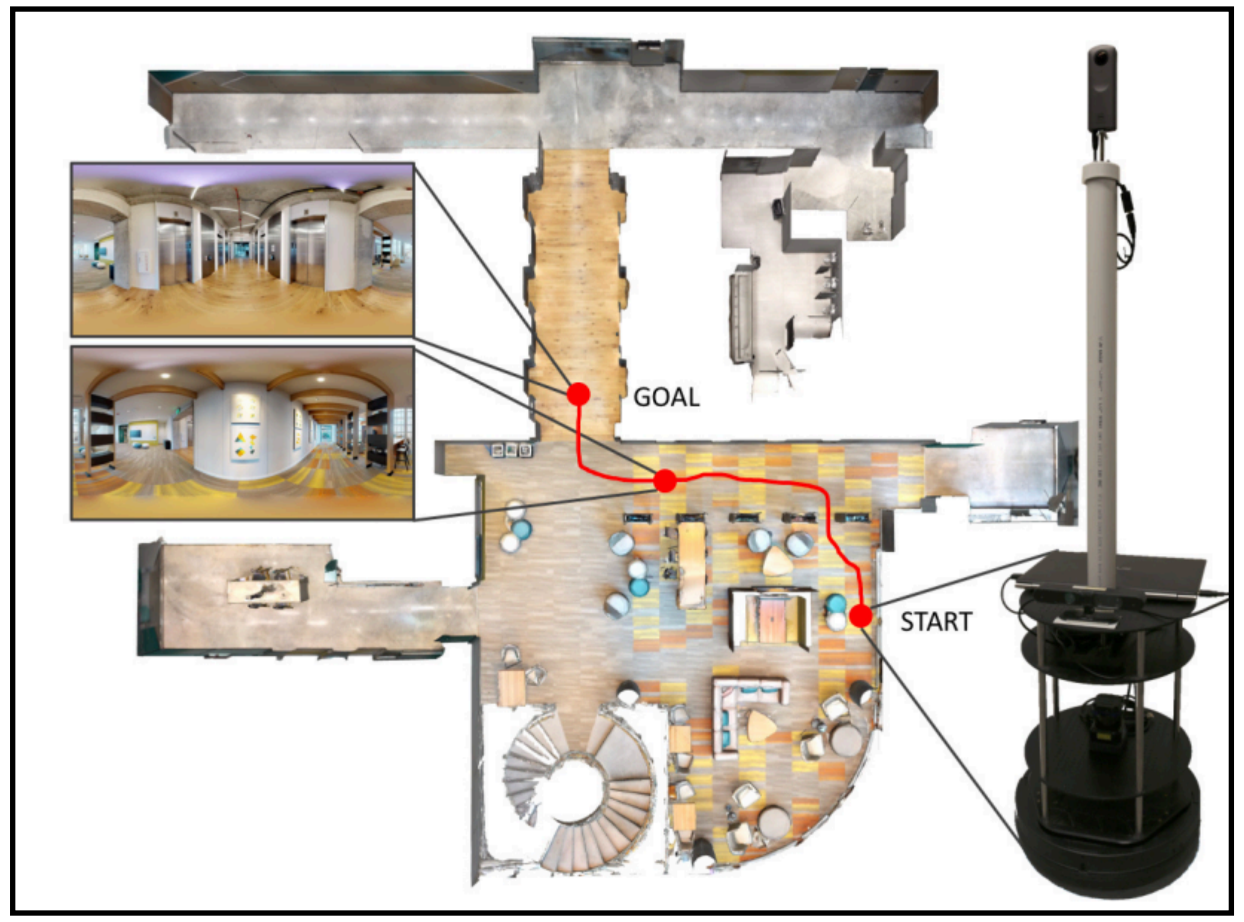
Lee et al, 2020



Kumar et al, 2022



Qi et al, 2022



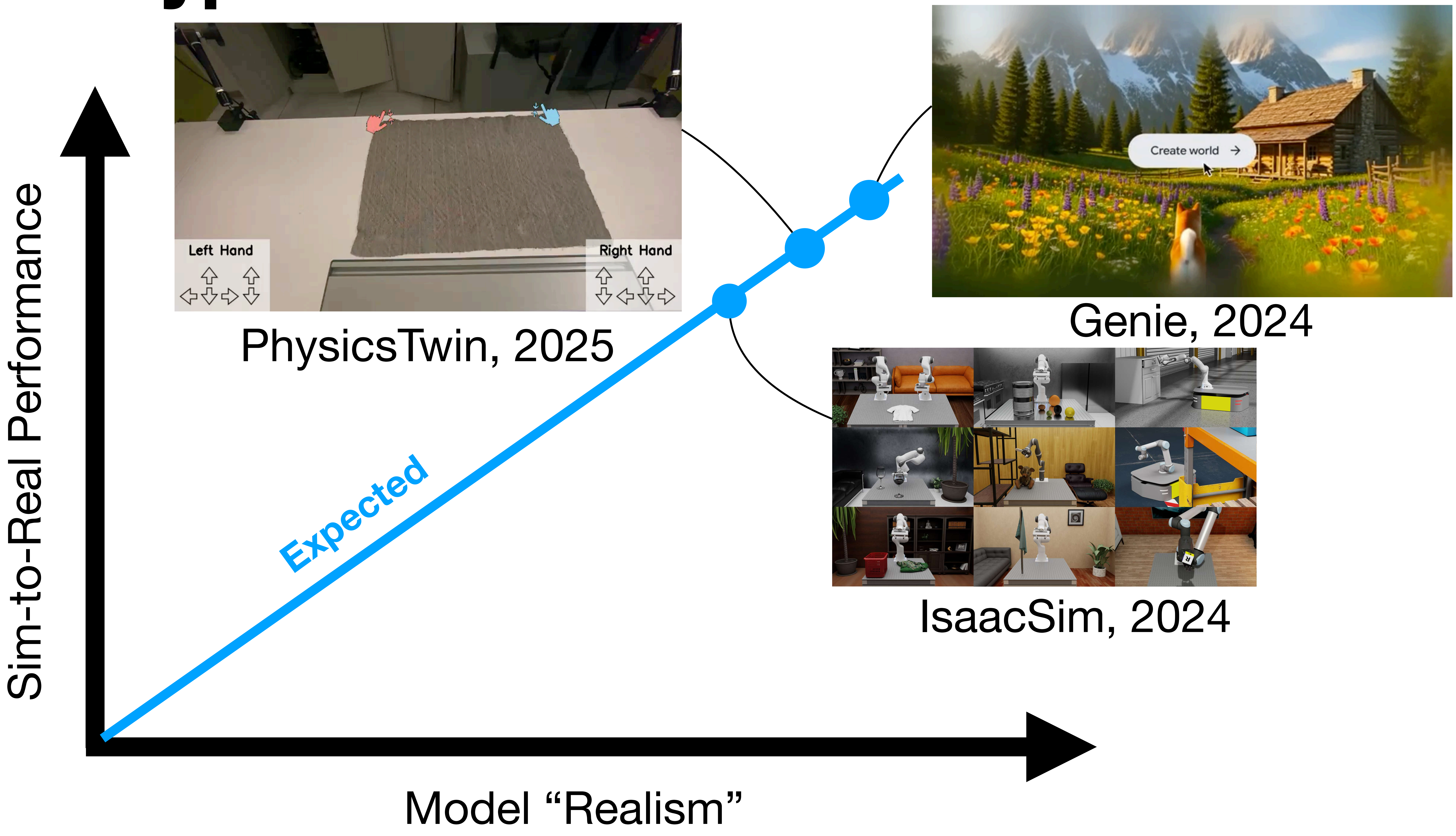
Anderson et al, 2021

Simulation-to-Real-World Transfer

Learning High-Speed Flight in the Wild, Science Robotics 2021

Antonio Loquercio, Elia Kaufmann, Rene Ranflt, Matthias Mueller, Vladlen Koltun, Davide Scaramuzza

Towards Hyper-Realistic Models



A Powerful Argument Against Simulators

The computational power of the physical world scales linearly with the number of particles. Think of pouring sand. Every sand particle computes its interactions with other particles. The more particles, the more computation. Thus the amount of computation scales linearly with the number of particles. **None of our computer devices can provide such computational power.** Assuming that accurate simulation oh physical behavior is important for robotics, this appears to be a fundamental challenge for the simulation approach.

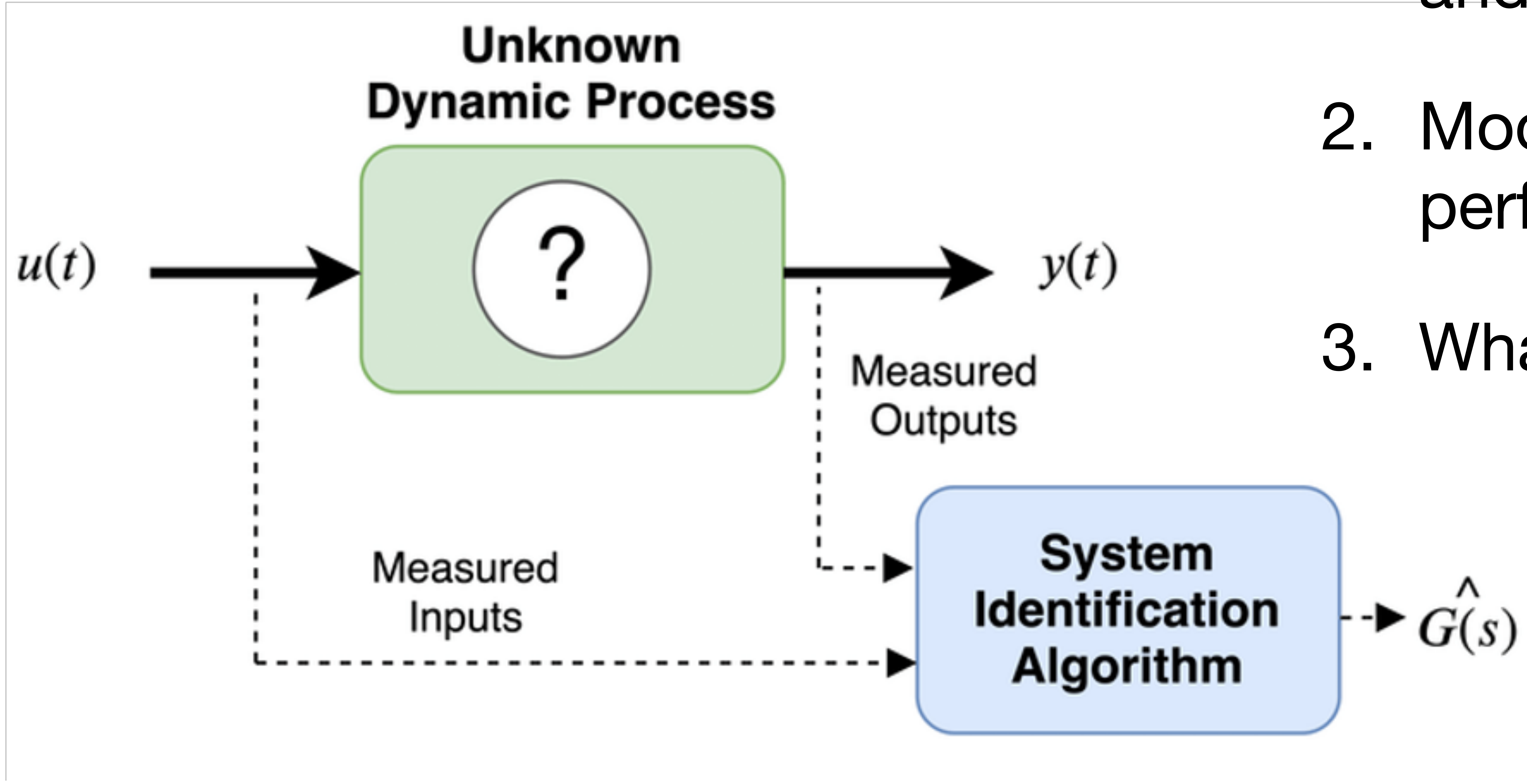
By Rodney Brooks

A Powerful Argument Against Simulators

The computational power of the physical world scales linearly with the number of particles. Think of pouring sand. Every sand particle computes its interactions with other particles. The more particles, the more computation. Thus the amount of computation scales linearly with the number of particles. **None of our computer devices can provide such computational power. Assuming that accurate simulation oh physical behavior is important for robotics**, this appears to be a fundamental challenge for the simulation approach.

By Rodney Brooks

Is Physical Realism Necessary?



1. The key is the dynamics of the observable and controllable subspace.
2. Modeling beyond that may not improve performance.
3. What can be safely ignored?

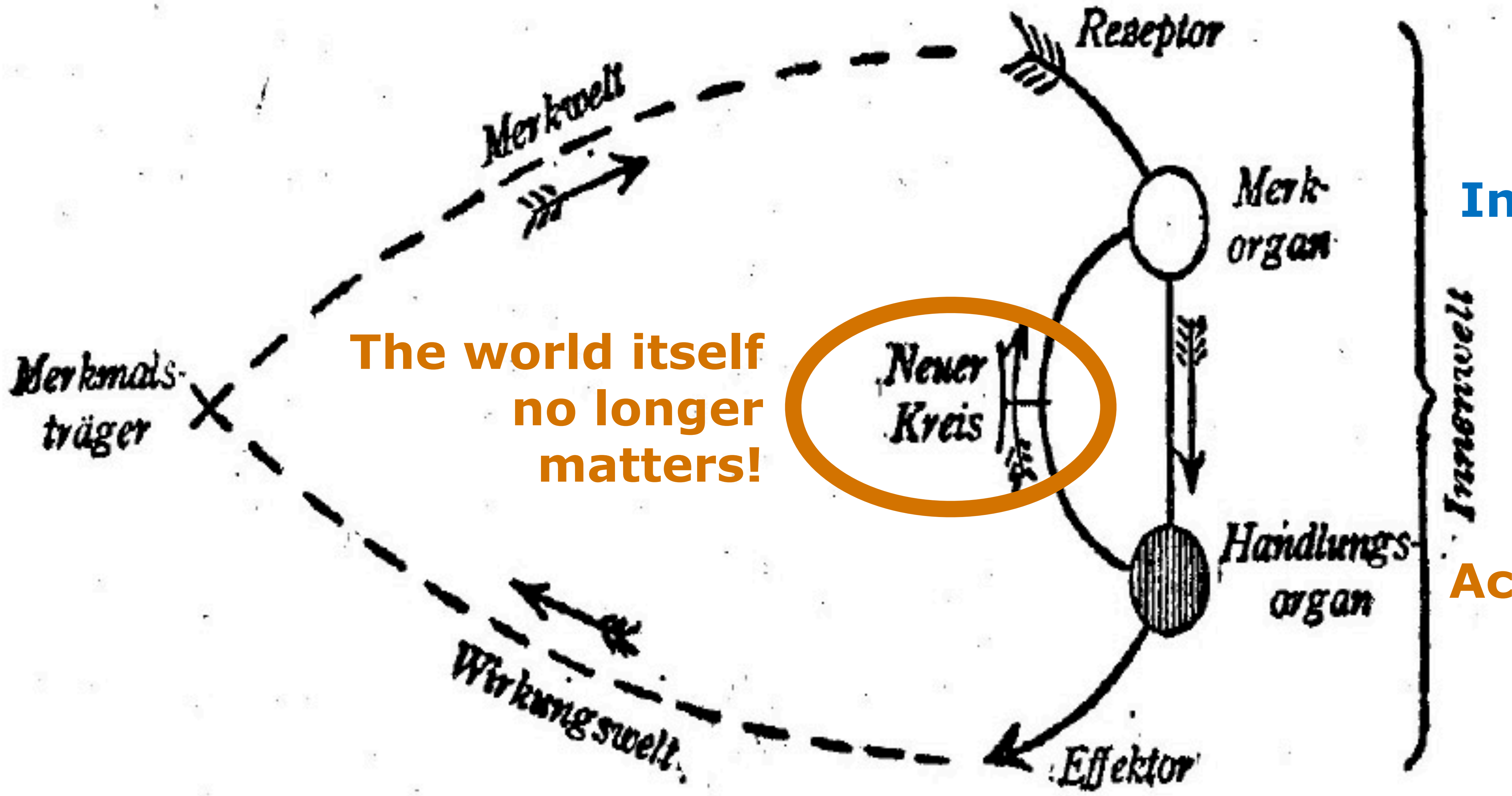
Ho and Kalman (1965), "Effective construction of linear state-variable models from input-output data," Regelungstechnik, Vol. 12, pp. 545-548.
Åström and Bohlin (1965), "Numerical identification of linear dynamic systems from normal operating records," Proc. IFAC Symp on Self Adaptive Systems, Teddington, UK

An Alternative Theory of Perception

- Jakob von Uexküll, German biologist (1864-1944)
- Theory of “Umwelt” (Sensory-action surrounding worlds).

Feedback about the action

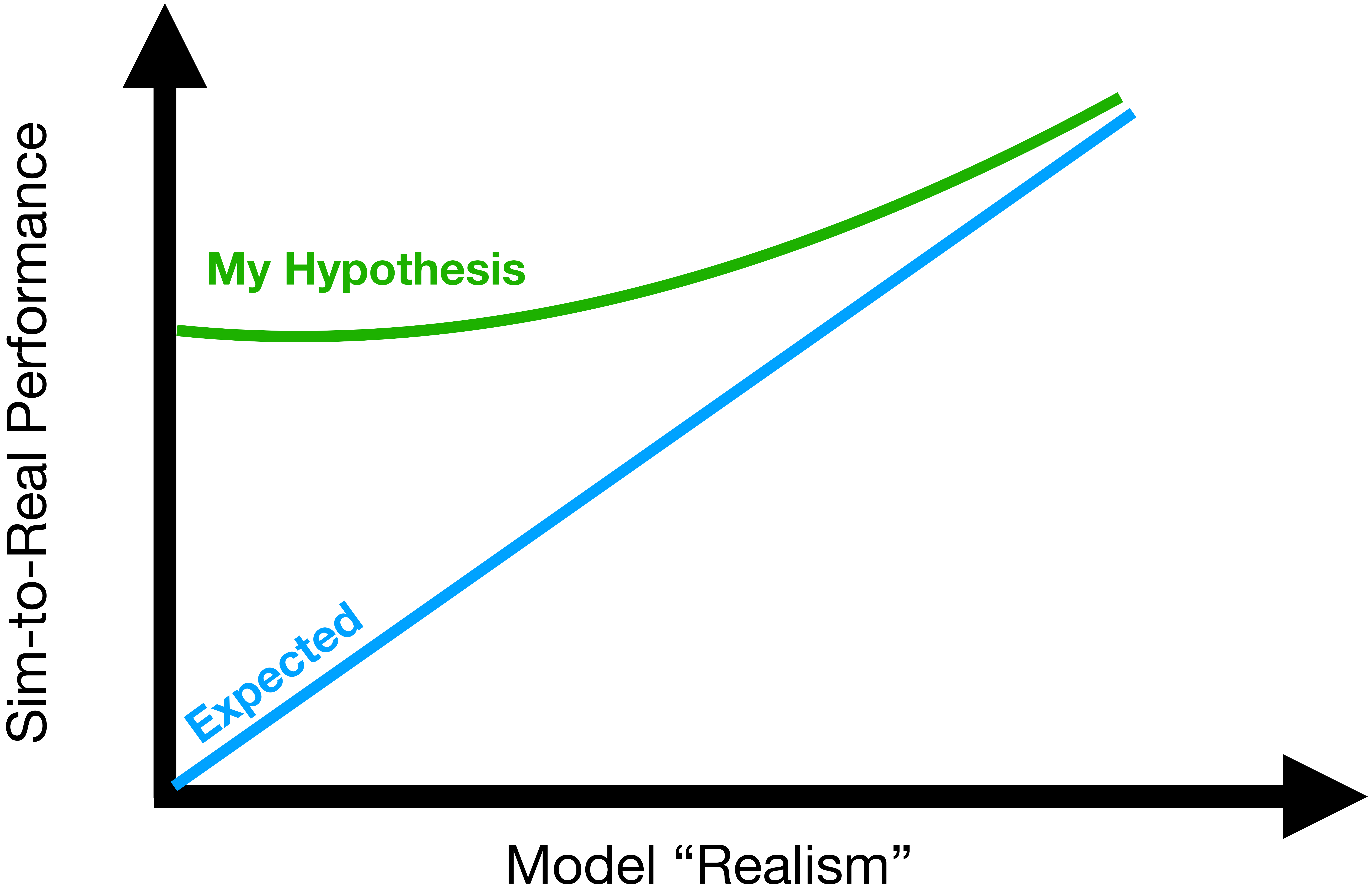
The world itself no longer matters!



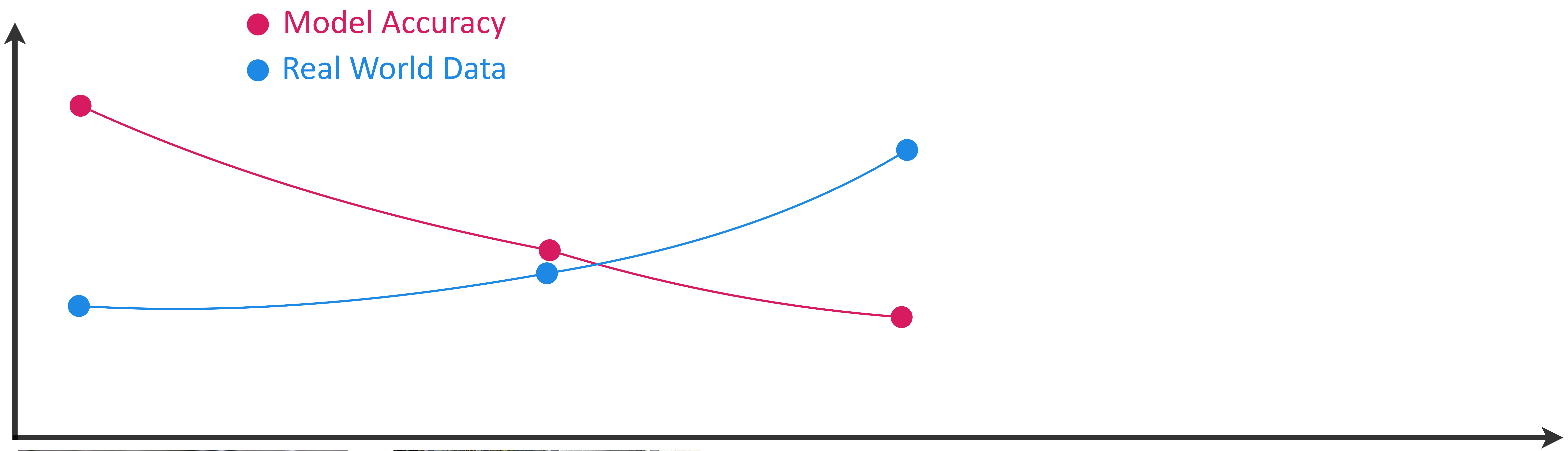
Input from the world

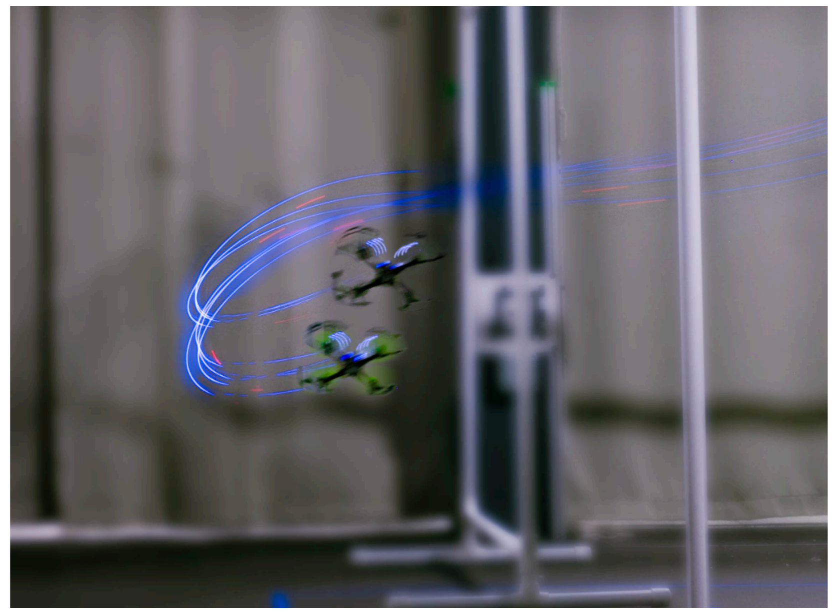
Action affects the world

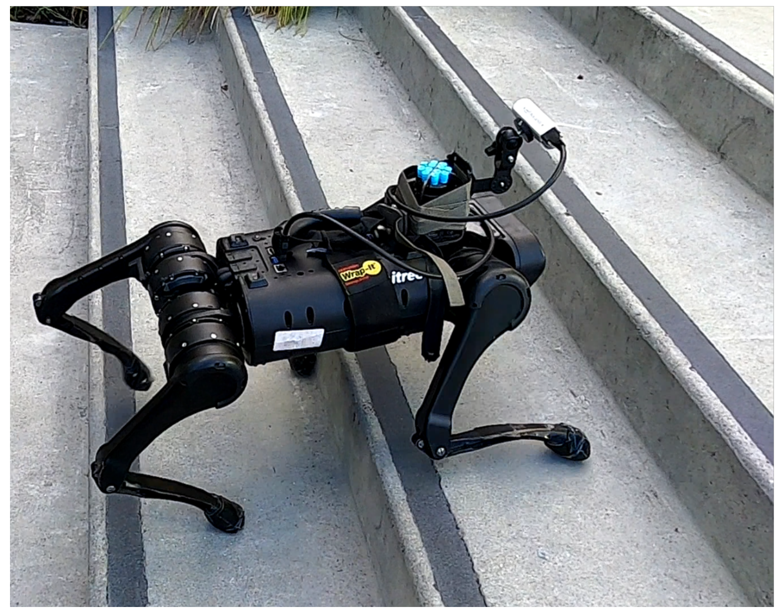
Sensorimotor Reduced-Order Models

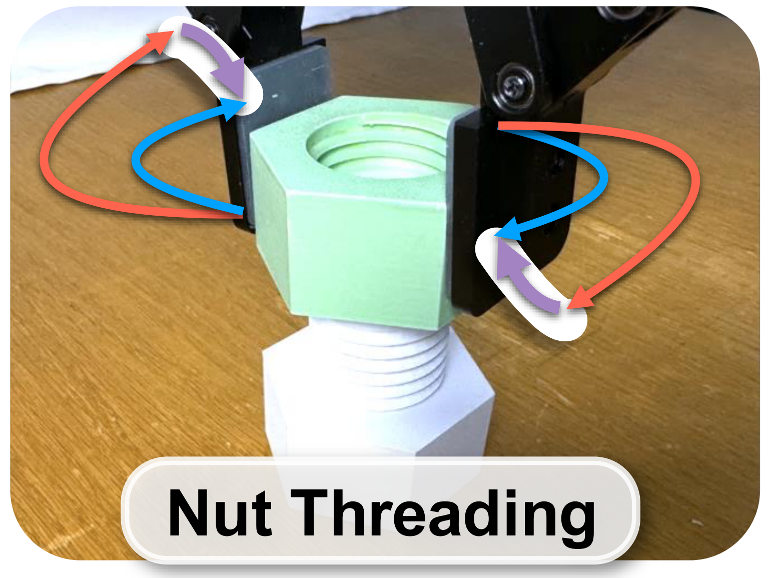


Today's Talk

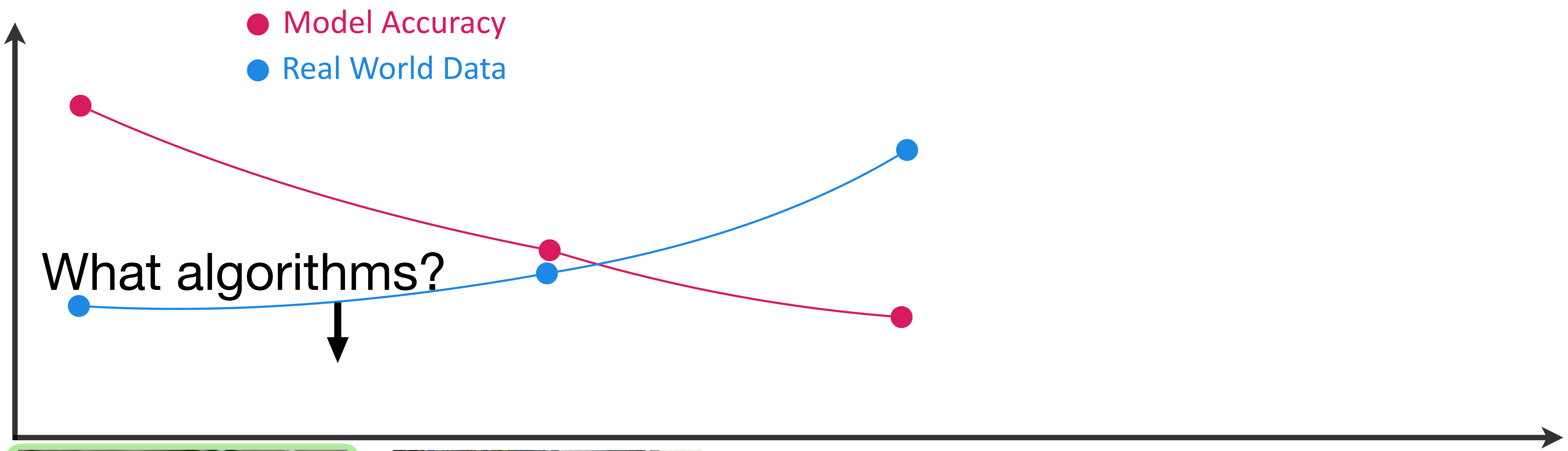




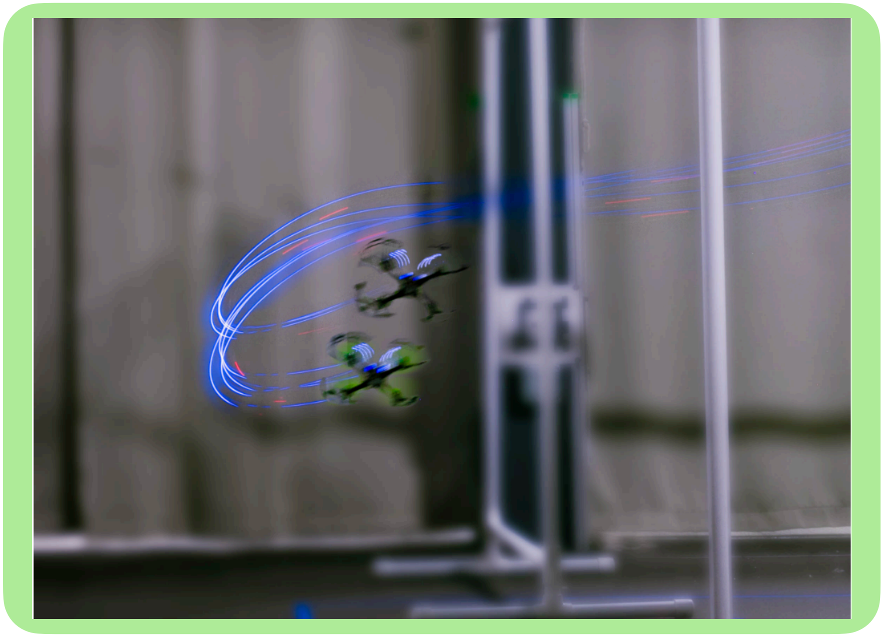


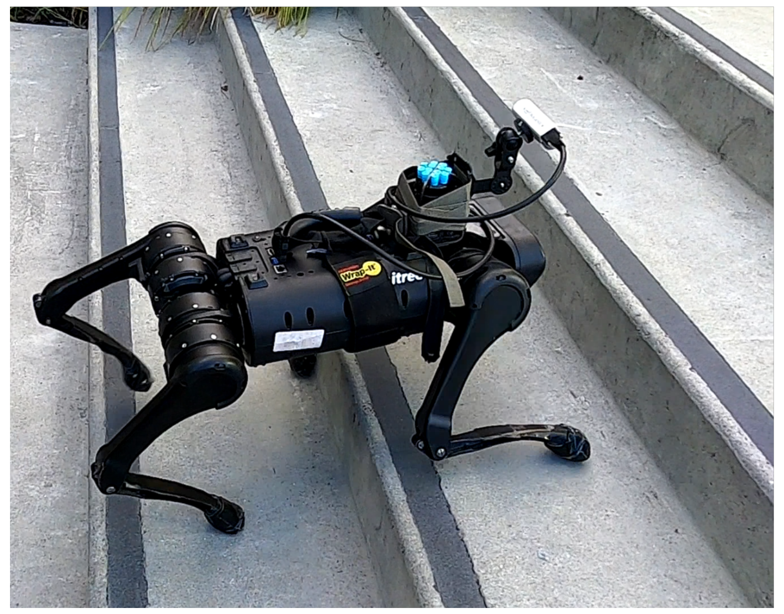


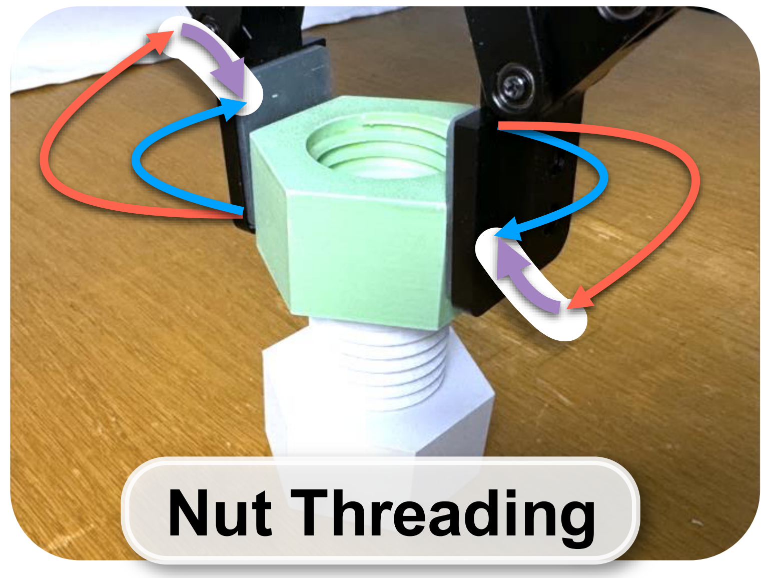
Today's Talk



Model Accuracy
Real World Data
What algorithms?







Nut Threading

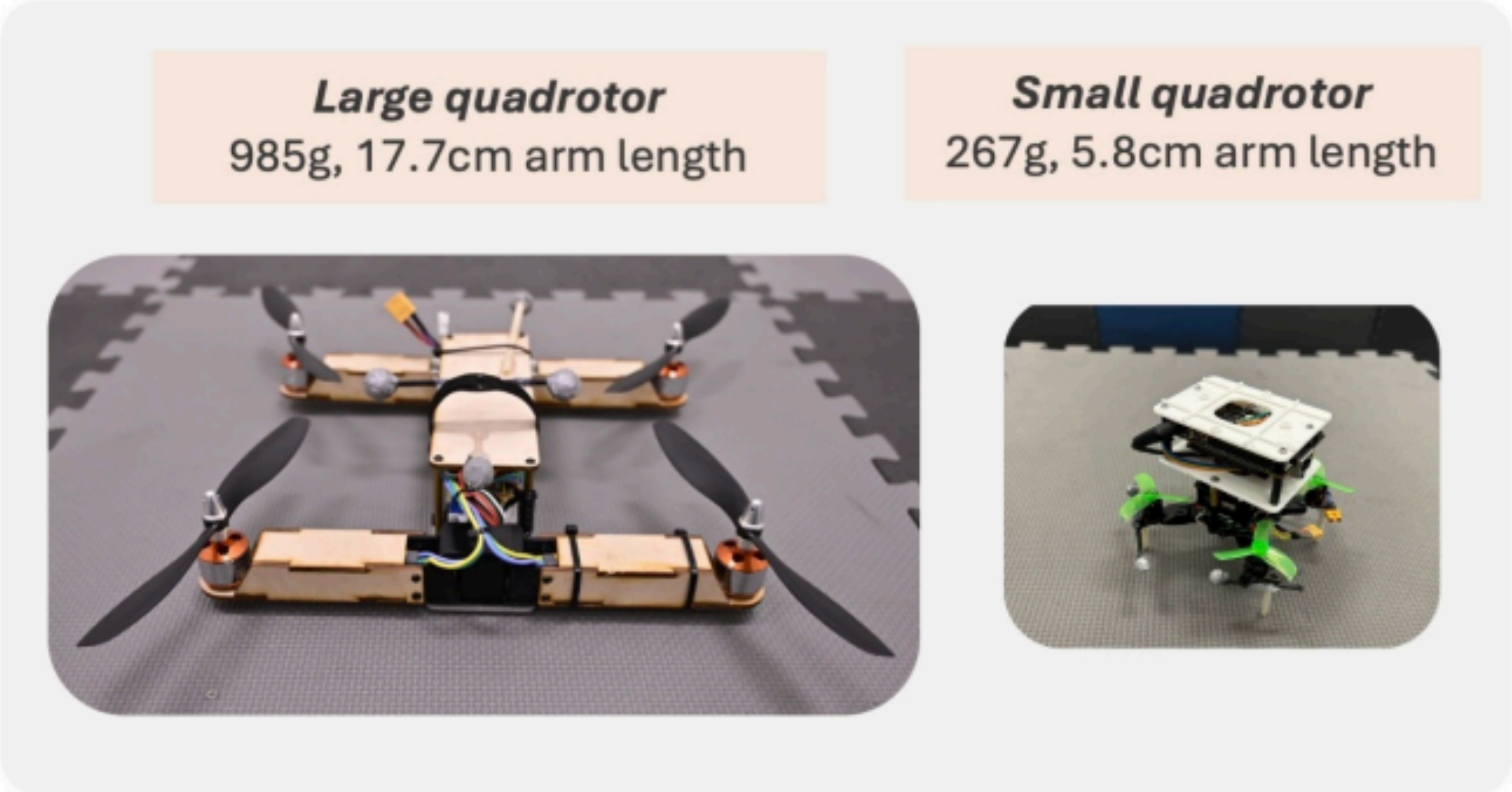
Simulation for Quadcopter Flight

This paper has been accepted for publication in the IEEE Transactions on Robotics (T-RO), 2025. © IEEE

A Learning-based Quadcopter Controller with Extreme Adaptation

Dingqi Zhang[1], Antonio Loquercio[2], Jerry Tang[1], Ting-Hao Wang[1],
Jitendra Malik[3], and Mark W. Mueller[1]

Abstract—**This paper introduces a learning-based low-level controller for quadcopters, which adaptively controls quadcopters with significant variations in mass, size, and actuator capabilities. Our approach leverages a combination of imitation learning and reinforcement learning, creating a fast-adapting and general control framework for quadcopters that eliminates the need for precise model estimation or manual tuning. The controller estimates a latent representation of the vehicle's system parameters from sensor-action history, enabling it to adapt swiftly to diverse dynamics. Extensive evaluations in simulation demonstrate the controller's ability to generalize to unseen quadcopter parameters, with an adaptation range up to 16 times broader than the training set. In real-world tests, the controller is successfully deployed**

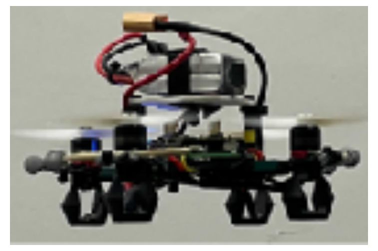


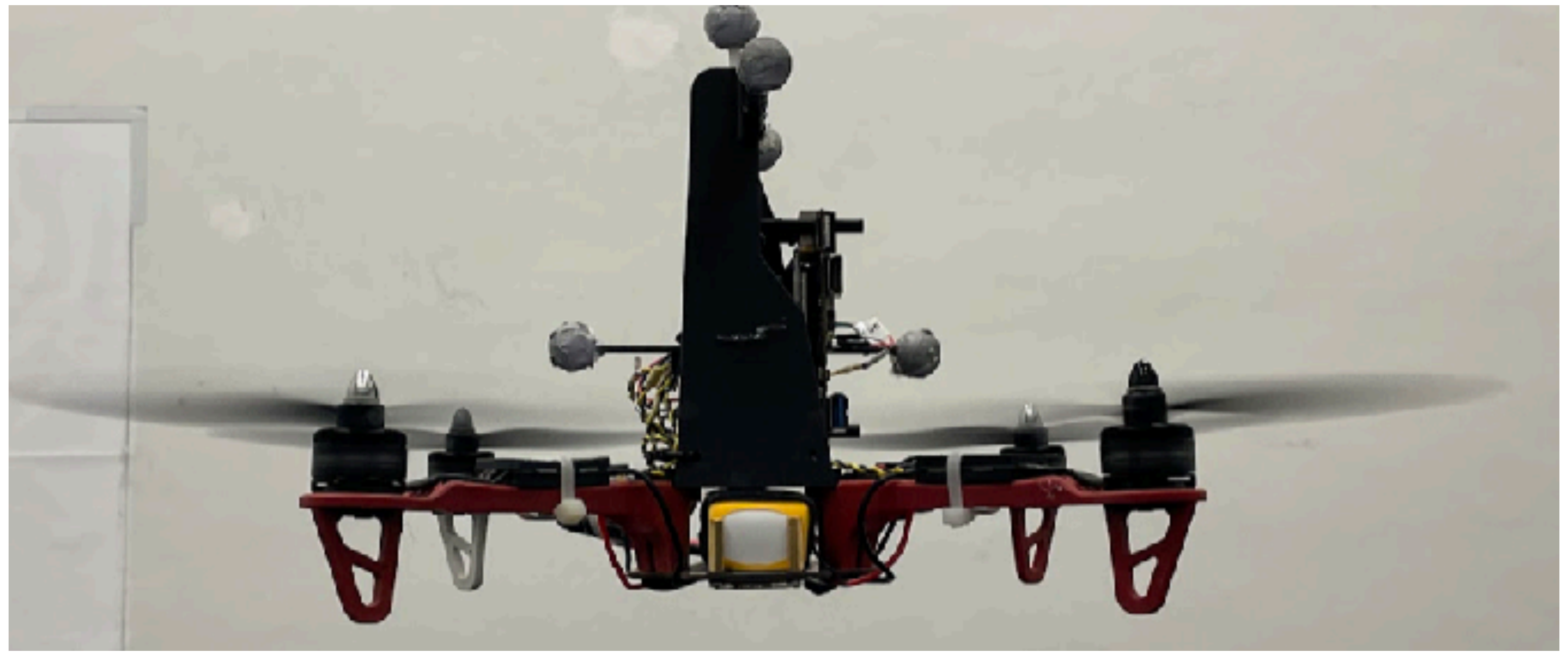
One Adaptive Low-level Controller For All Drones

No Prior Model Knowledge Required

Disturbance Rejection

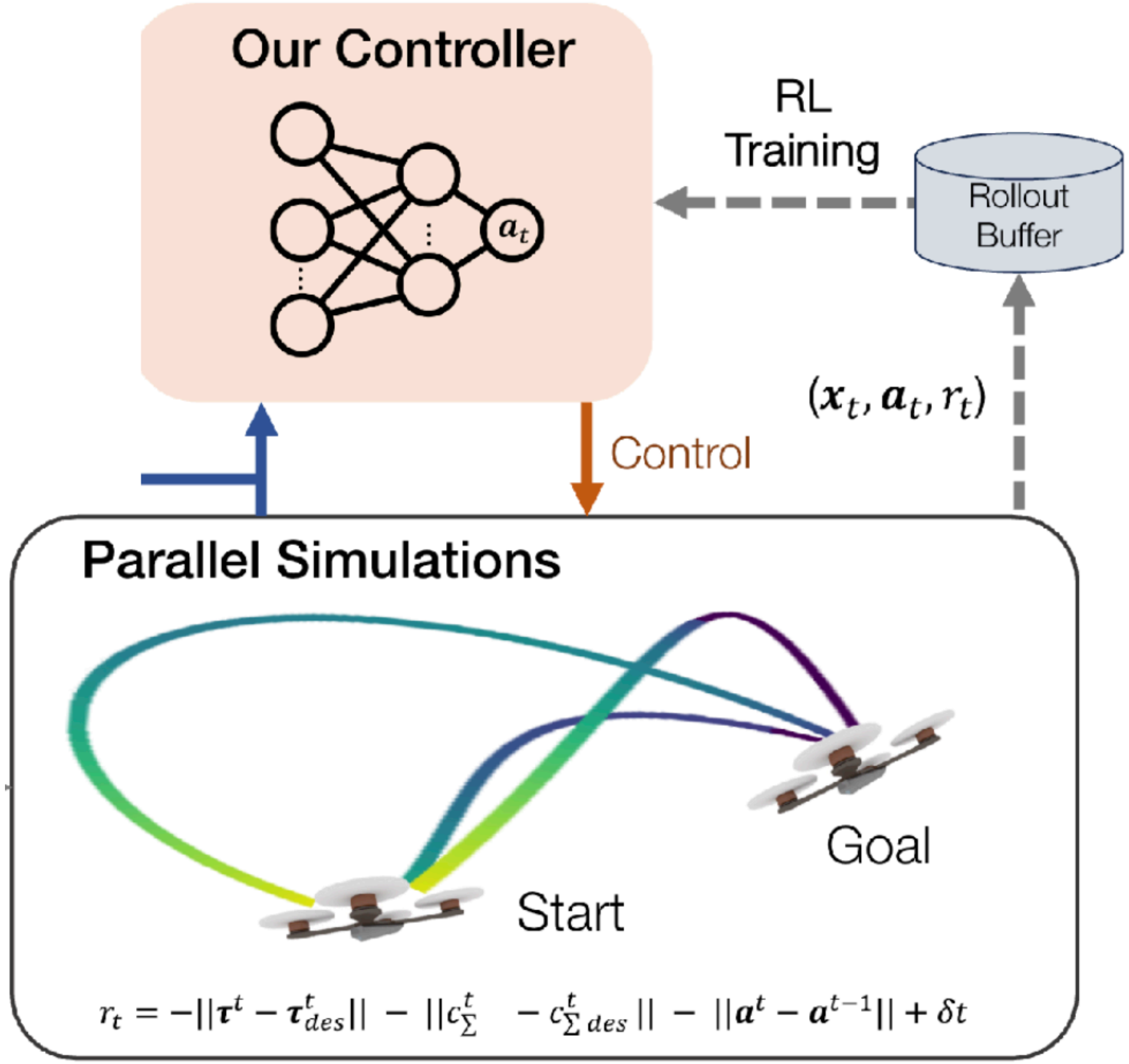
Flying Vastly Different Quadrotors

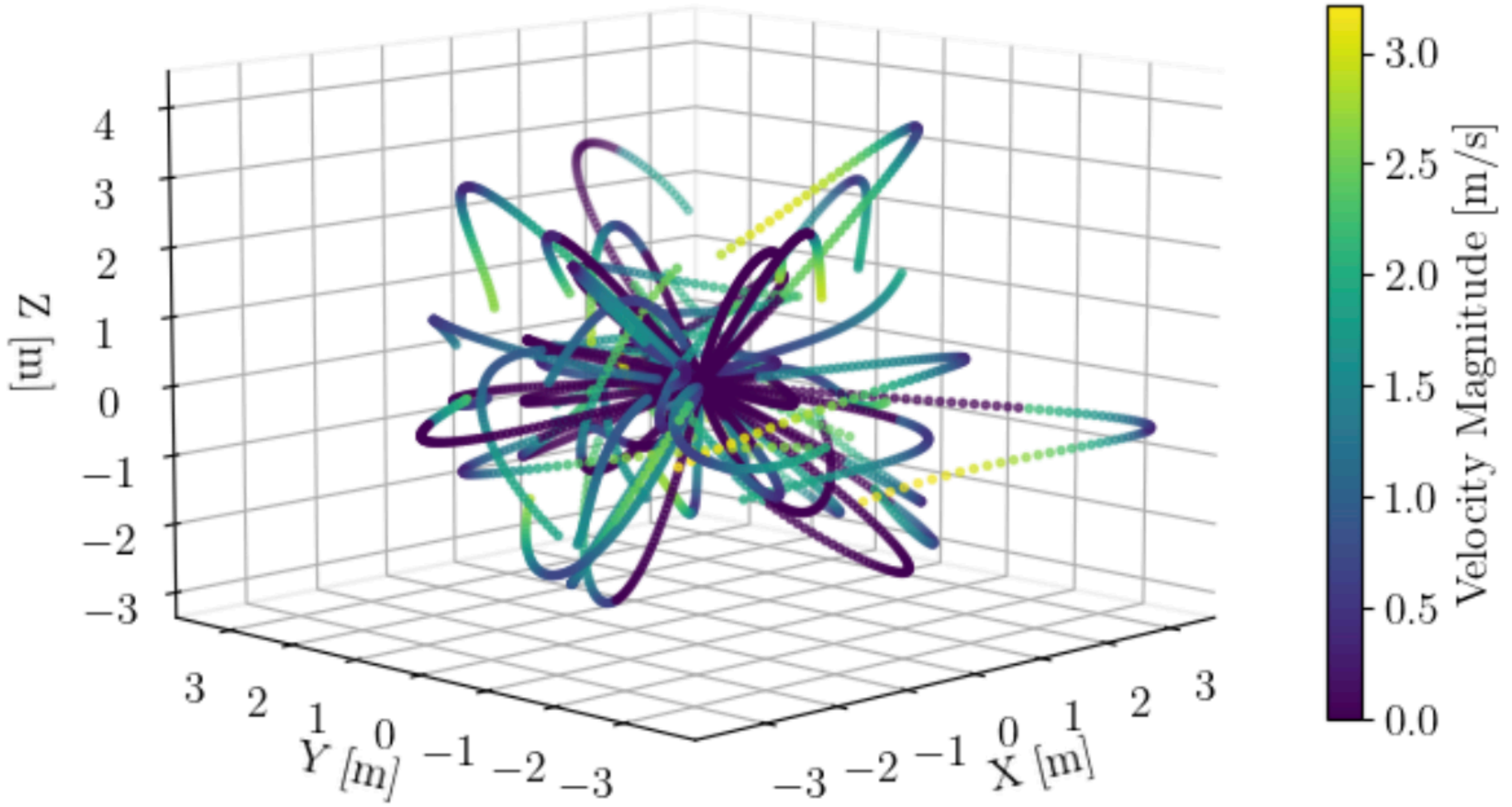




Mass [kg]	[0.142, 0.950]	**669.0%**
Arm length [m]	[0.046, 0.20]	**434.8%**
Mass moment of inertia around x, y [kg m^2]	[7.42e-5, 5.6e-3]	**7547.2%**
Mass moment of inertia around z [kg m^2]	[1.2e-4, 8.8e-3]	**7333.3%**
Propeller constant /kappa [m]	[0.0041, 0.0168]	**409.8%**
Motor constant [N/rad^2]	[1.145e-7, 7.64e-6]	**6672.5%**
Max. motor speed [rad]	[1108, 8034]	**725%**

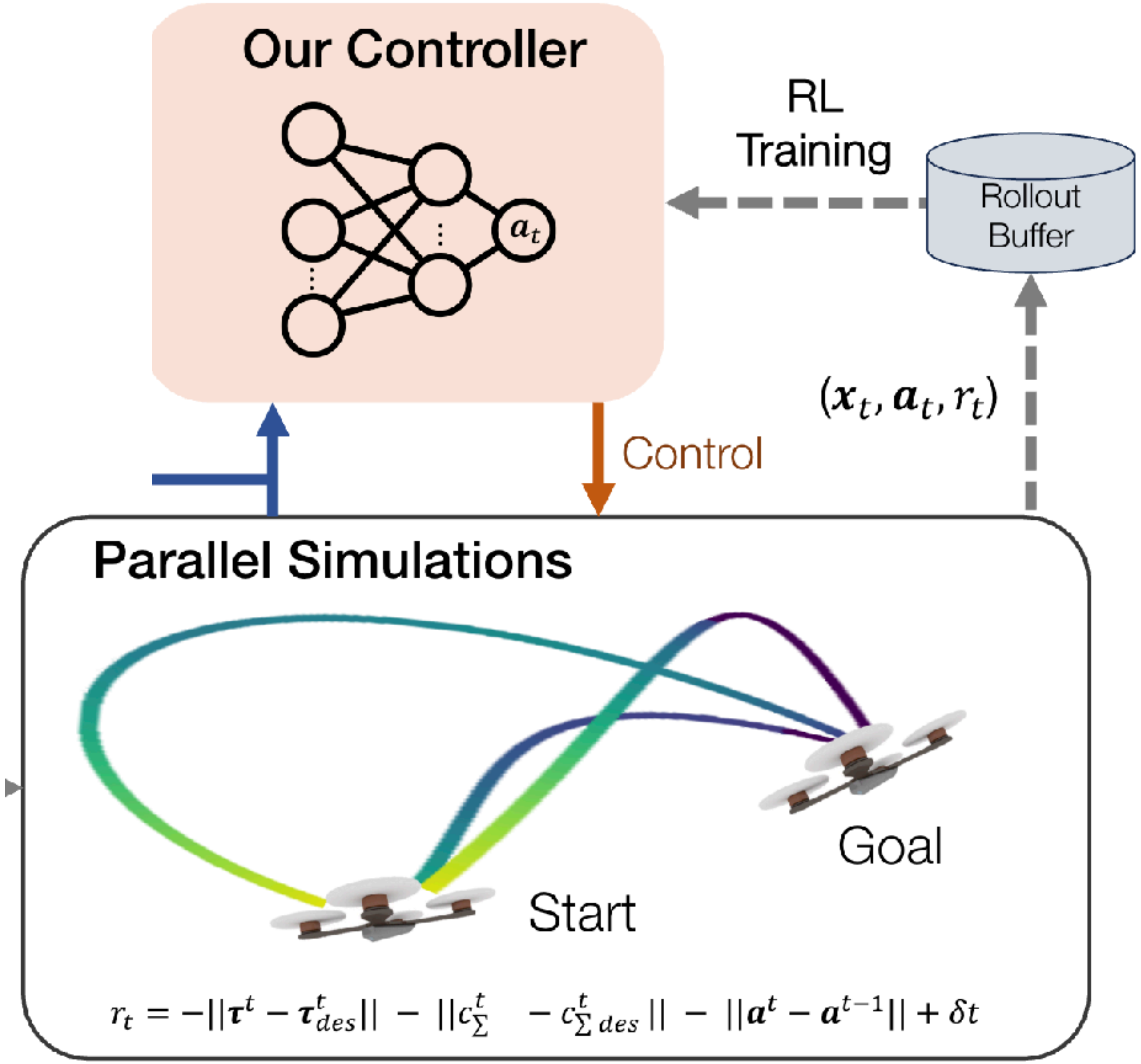
Training in Simulation with Domain Randomization



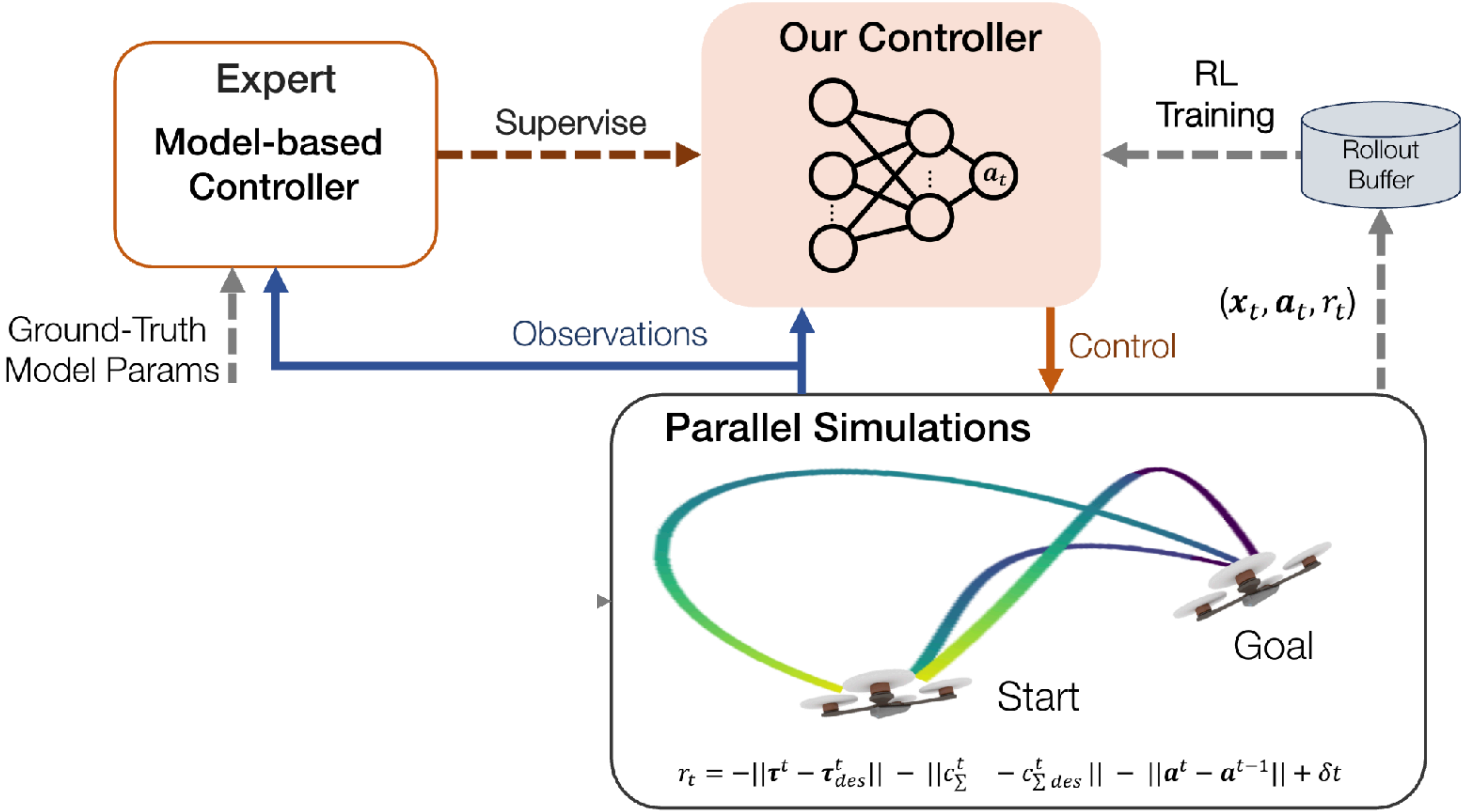


This will fail if done naively!

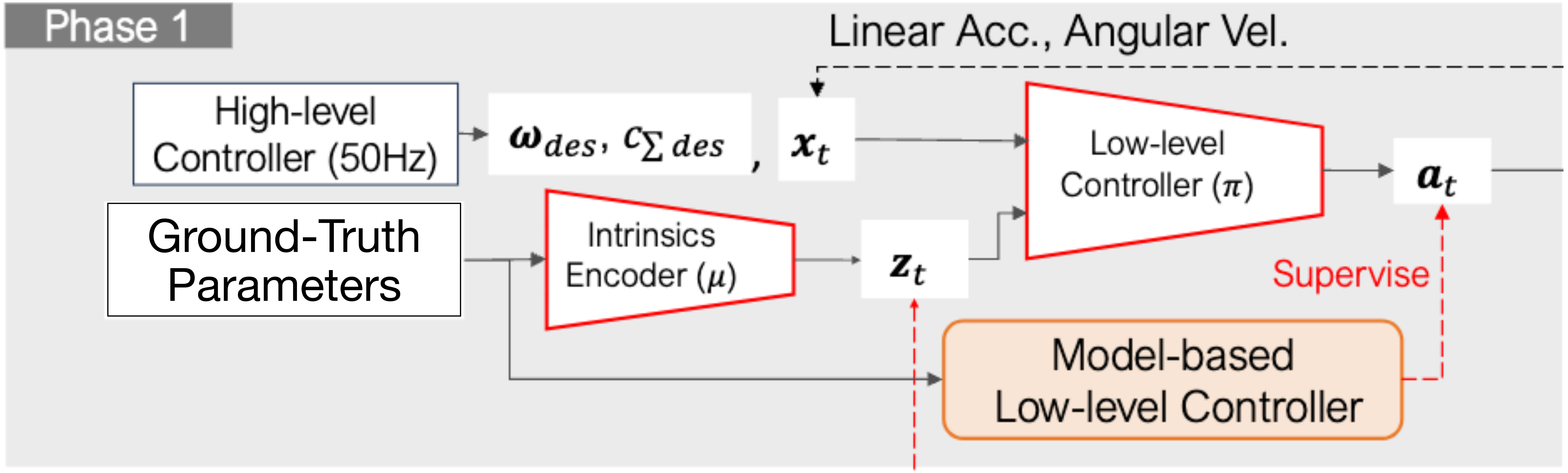
1. Model-Based Guidance



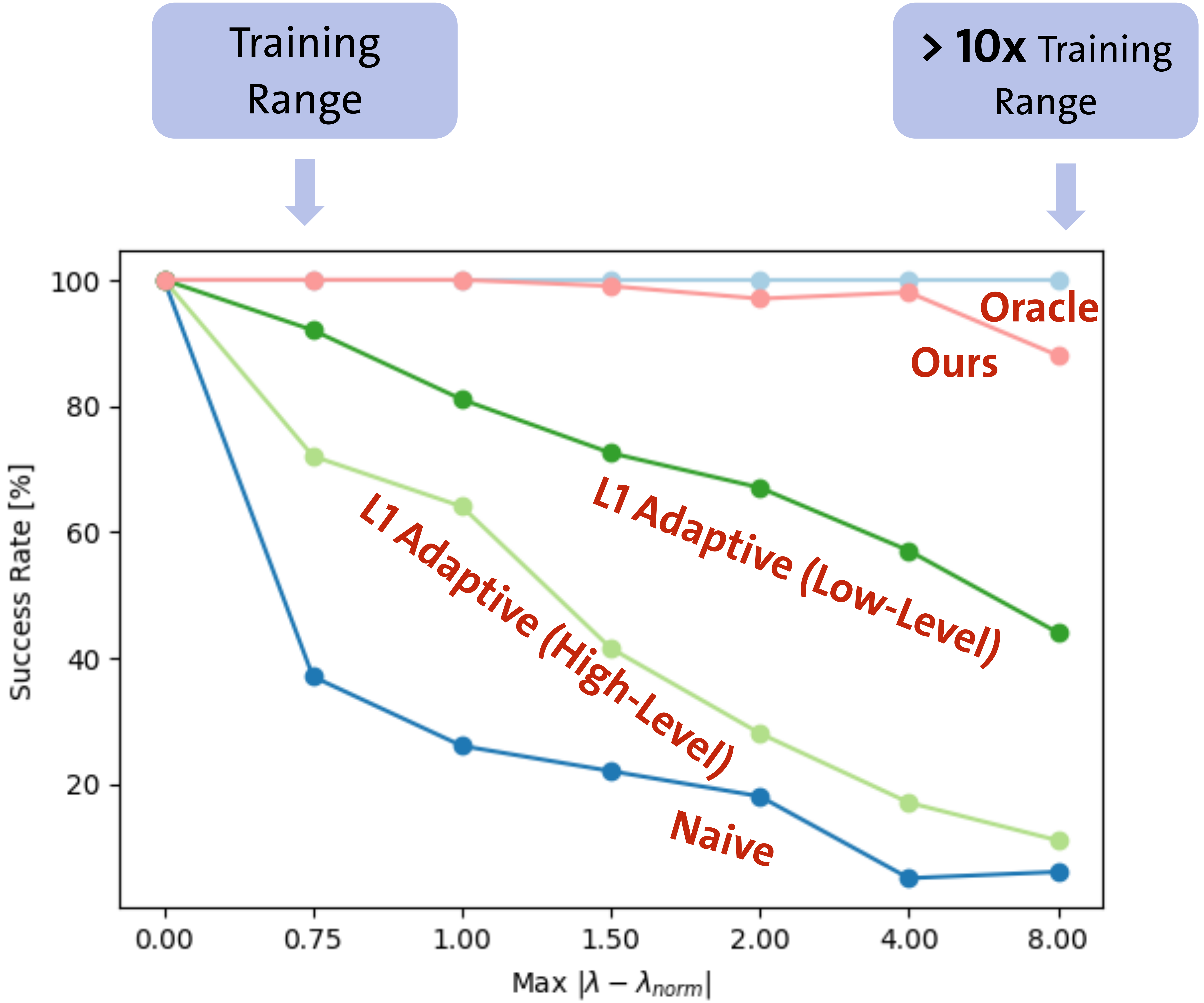
2. Physics-Informed Domain Randomization



3. Latent-Space System Idenfication

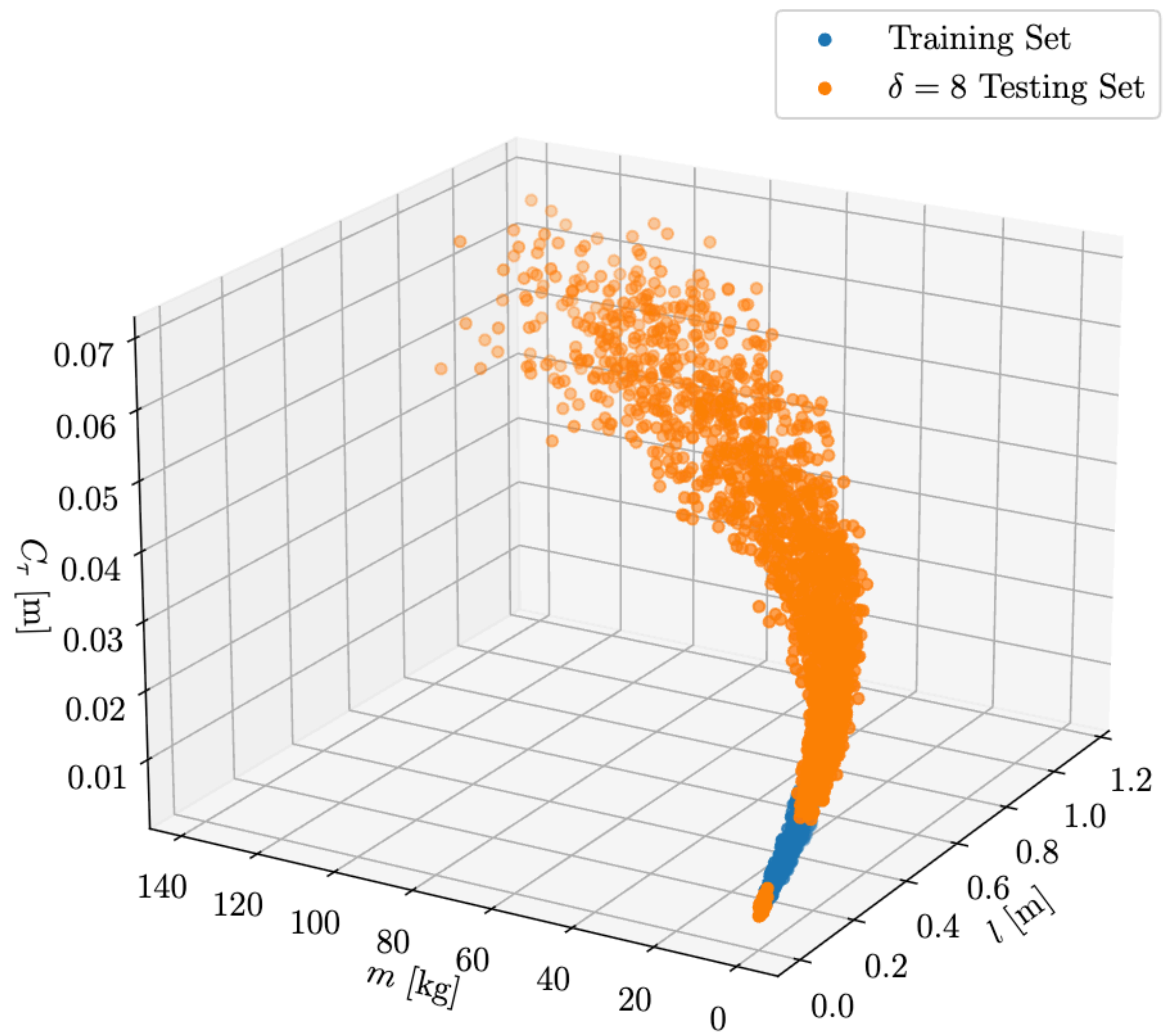


Impressive Generalization

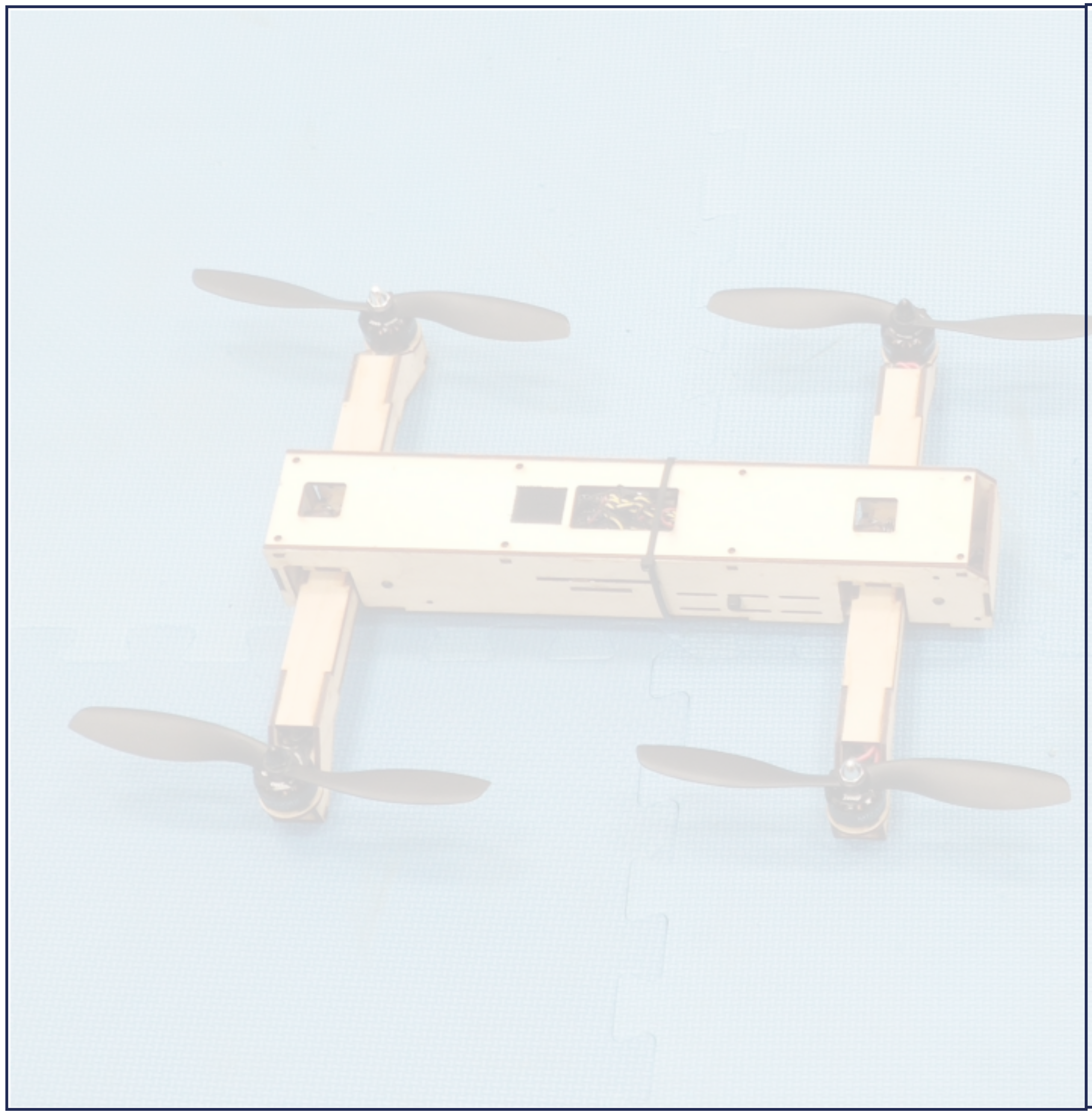


Training up to 10Kg Drones

Generalizes up to 120Kg Drones



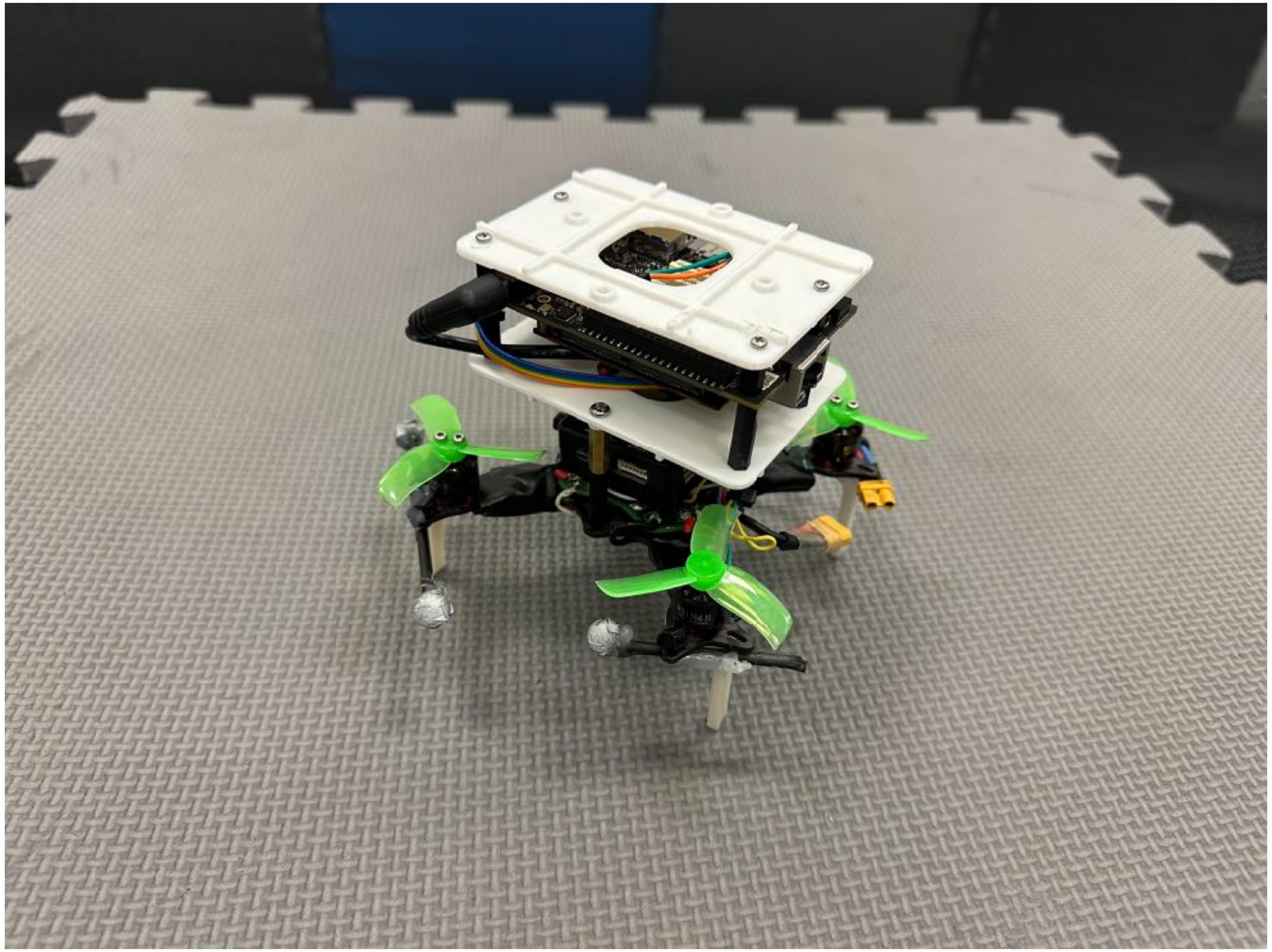
Flying Morphologies Unseen at Training Time



QUaRTM: A Quadcopter with Unactuated Rotor Tilting Mechanism Capable of Faster, More Agile, and More Efficient Flight,
Jerry Tang, Karan P. Jain, and Mark W. Mueller

Further Robustness Test

Flying Morphologies Unseen at Training Time



- ~30% mass above the propellers
- 3X larger x-z inertia than a normal drone.

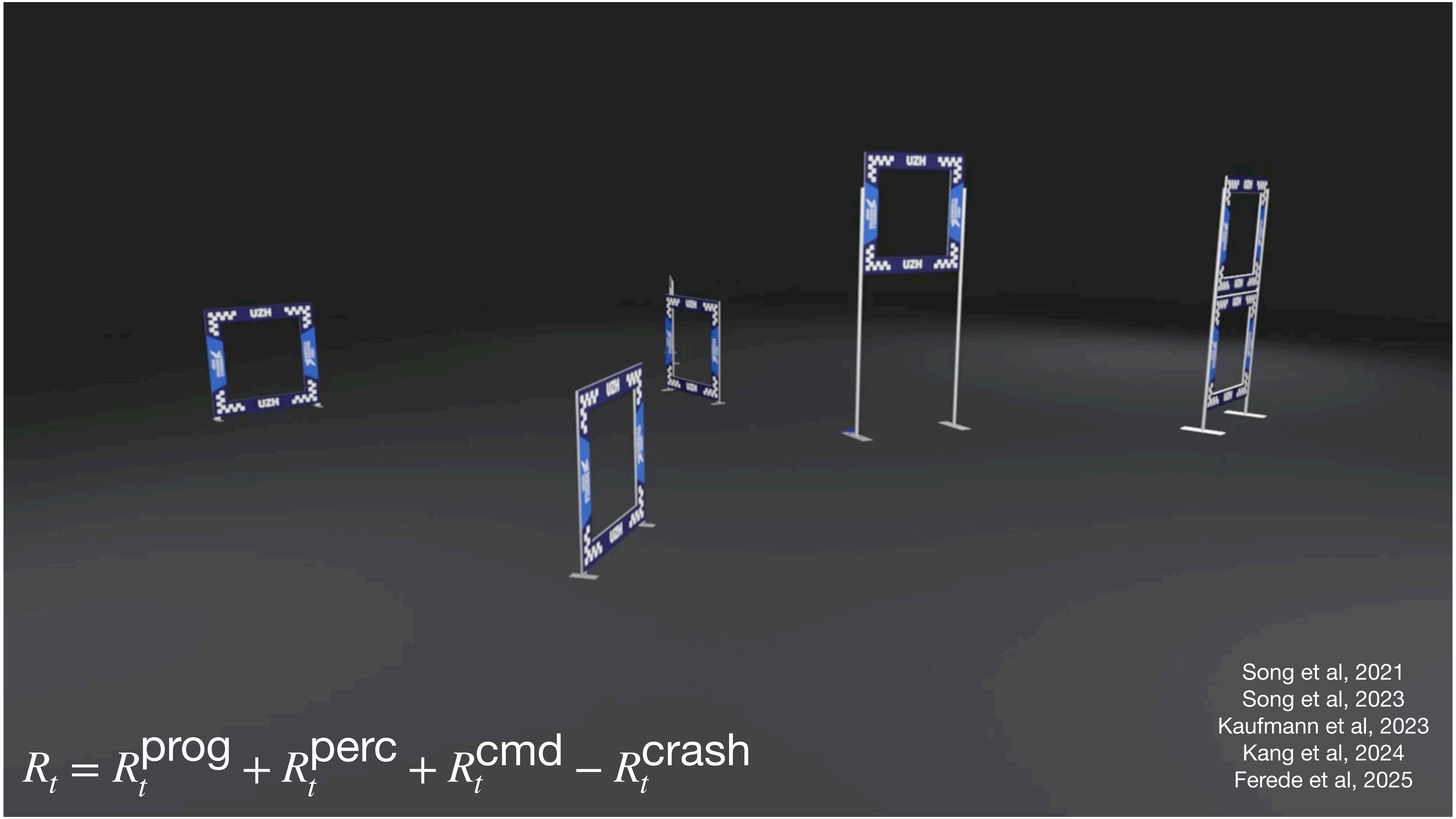

Ours
Mini Quadcopter
HiPeR Lab

Summary up to now

- **Imposing structure in the randomization space, together with real-time adaptation, leads to generalization beyond the training environment.**

Kaufmann, Bauersfeld, **Loquercio**, Koltun, Scaramuzza. *Champion-Level Drone Racing using Deep Reinforcement Learning*, Nature 2023

Training with Reinforcement Learning in Simulation



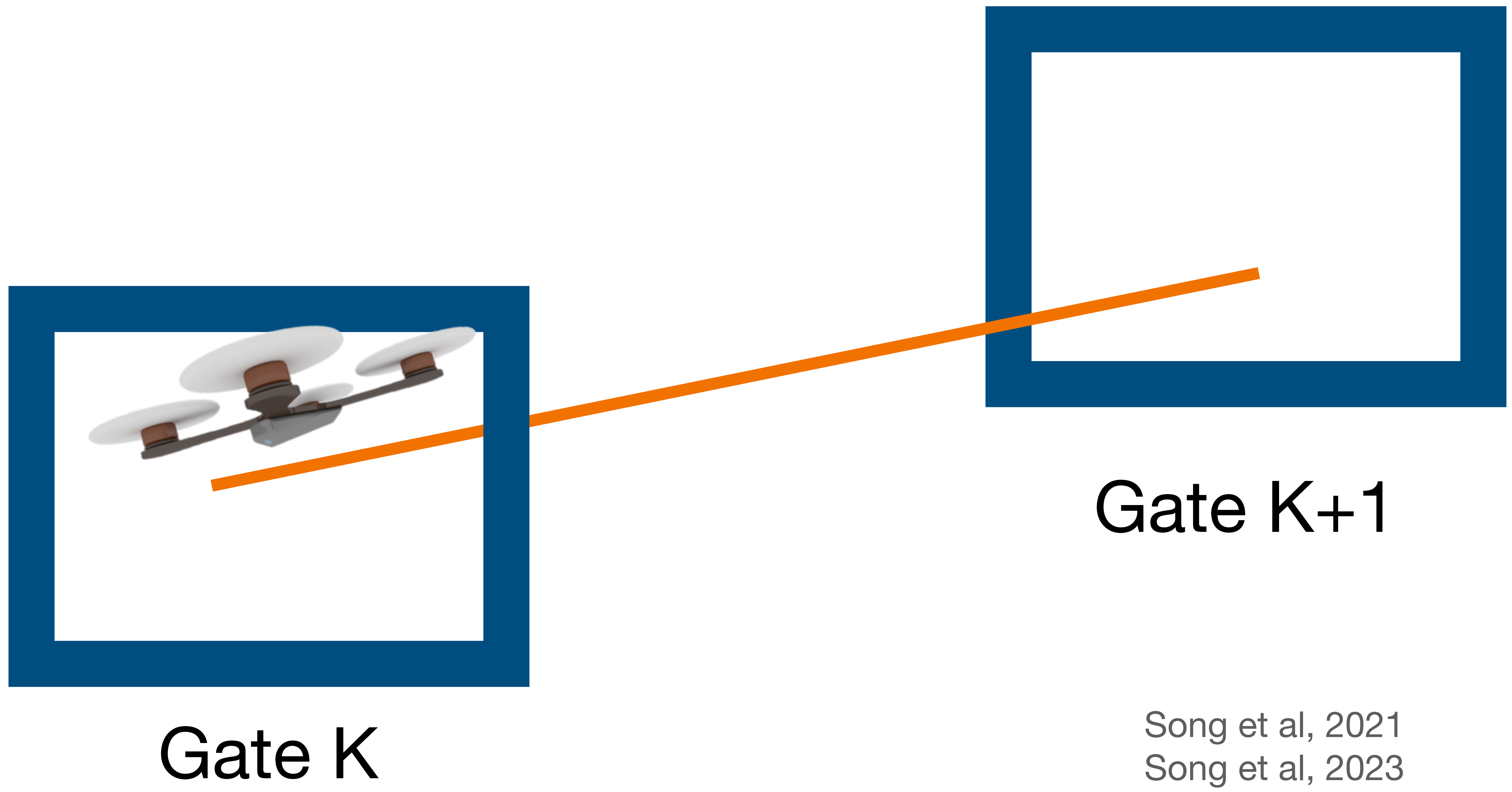
$$R_t = R_t^{\text{prog}} + R_t^{\text{perc}} + R_t^{\text{cmd}} - R_t^{\text{crash}}$$

Song et al, 2021
Song et al, 2023
Kaufmann et al, 2023
Kang et al, 2024
Ferede et al, 2025

Sim-to-Real Reinforcement Learning for Drone Racing

$$R_t = R_t^{\text{prog}} + R_t^{\text{perc}} + R_t^{\text{cmd}} - R_t^{\text{crash}}$$

$$R_t^{\text{prog}} = \lambda_1 [d_{t-1}^{\text{Gate}} - d_t^{\text{Gate}}]$$



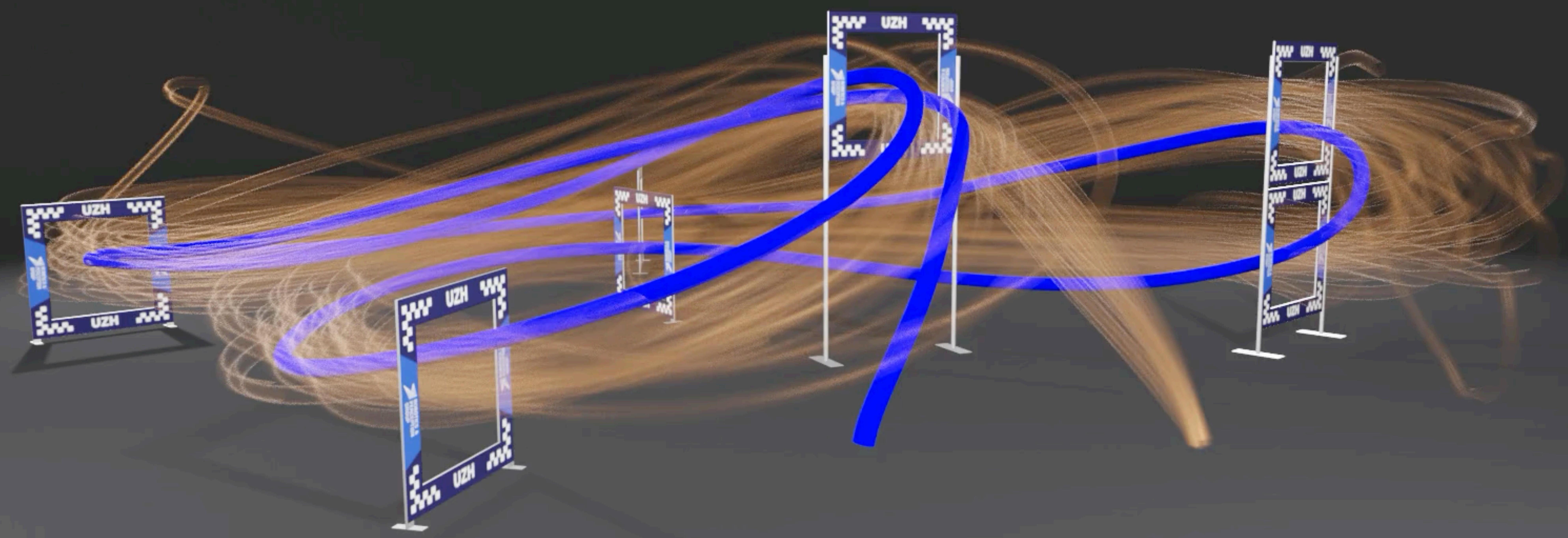
Song et al, 2021
Song et al, 2023
Kaufmann et al, 2023
Kang et al, 2024
Ferede et al, 2025

Prior Work on Drone Racing with RL

- Essentially, trajectory tracking with loose constraints. Can either be solved offline (e.g., with RL) or online (e.g., with MPPI or MPCC).

- Two main problems with this approach:
  - **Policies are trained to complete a lap as soon as possible, not to win a race**. These are two different things.
  - Explores a small subset of the state space. **Easy to exploit model errors.**

Human Pilots Don’t Fly a Fixed Path



Human Pilots

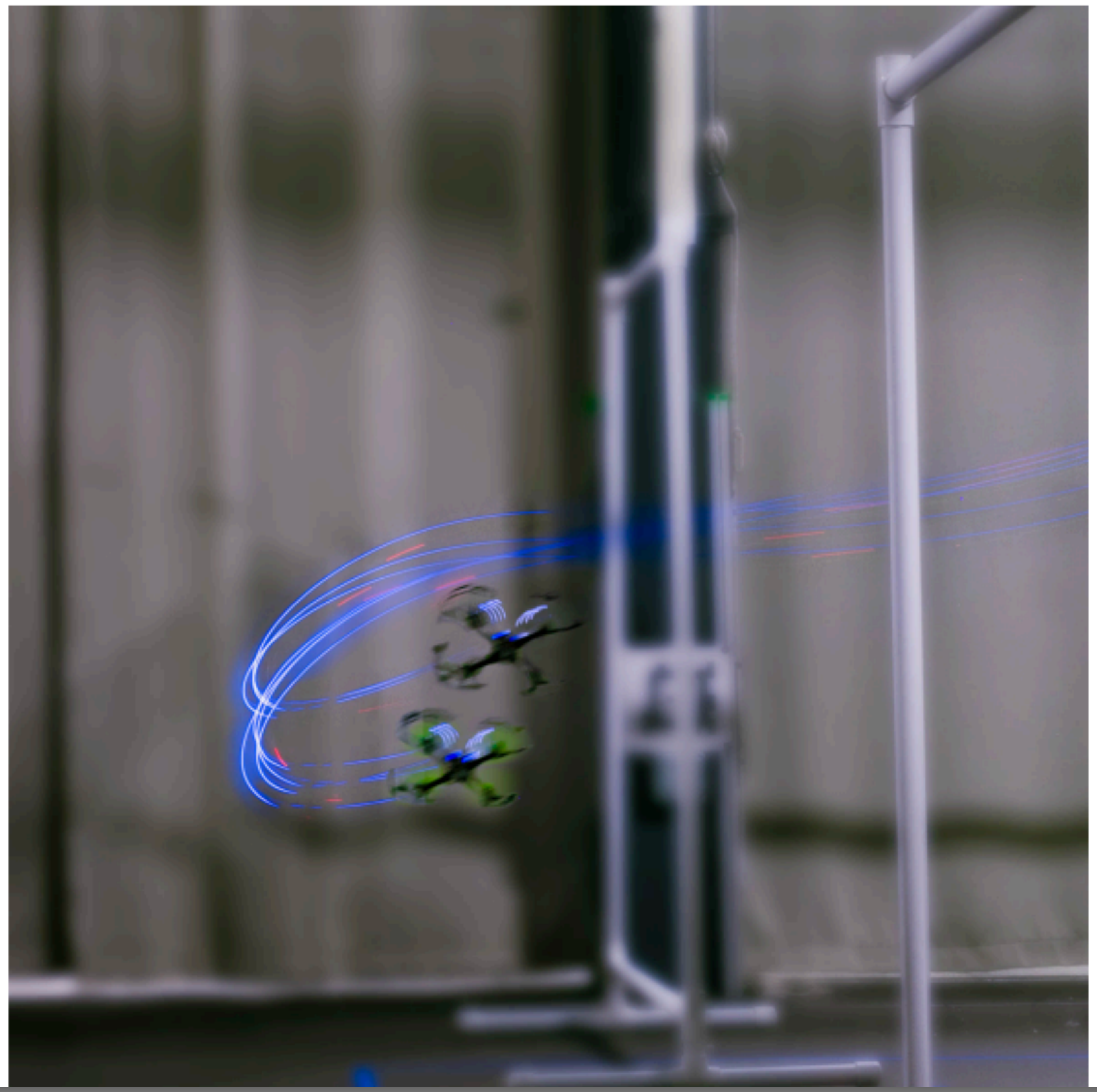
Autonomous Drone

Enlarging the Solution Set in Drone Racing Problems

Agile Flight Emerges from Multi-Agent Competitive Racing

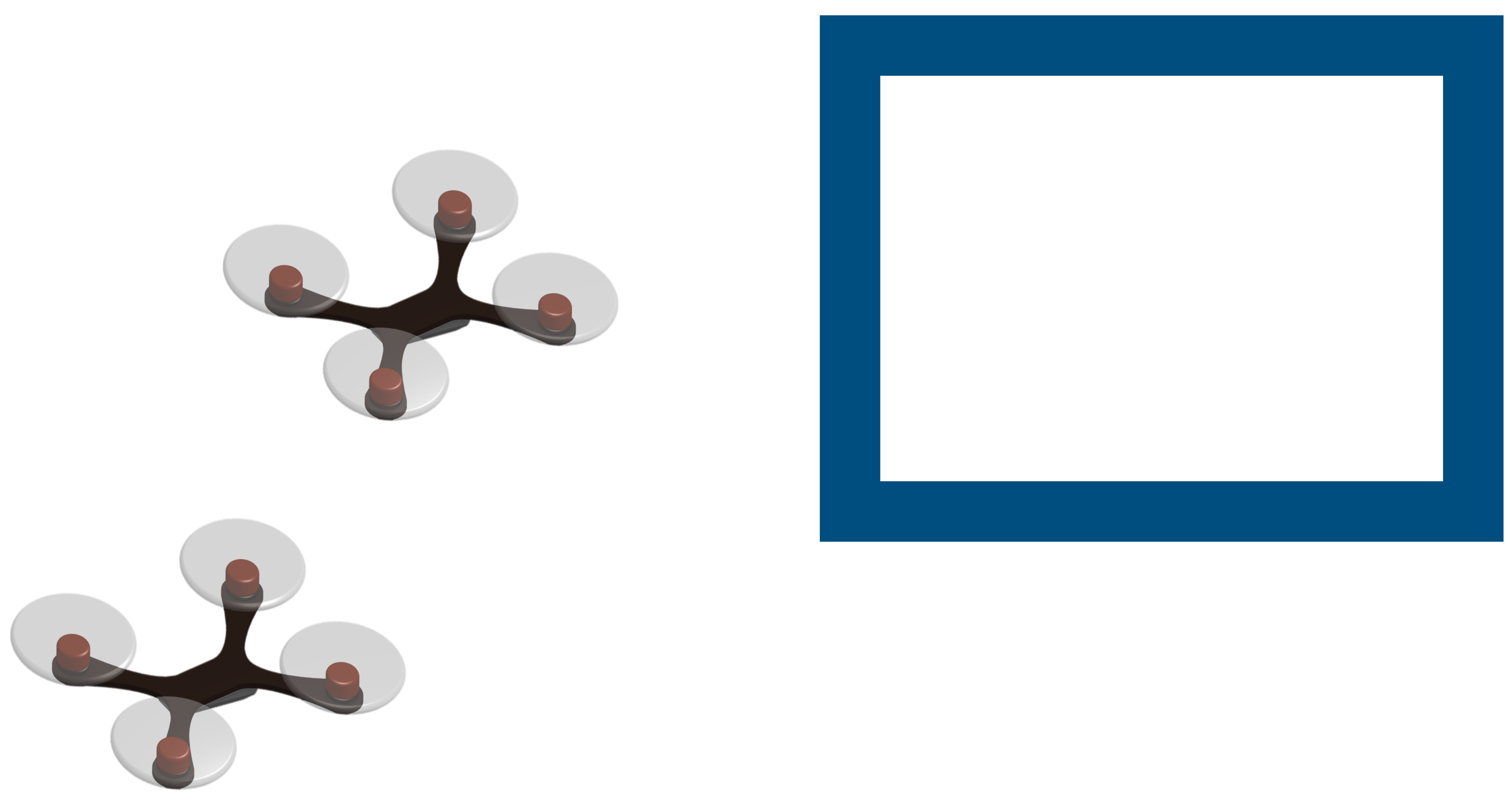
Vineet Pasumarti* Lorenzo Bianchi* Antonio Loquercio

Abstract—**Through multi-agent competition and the sparse high-level objective of winning a race, we find that both agile flight (e.g., high-speed motion pushing the platform to its physical limits) and strategy (e.g., overtaking or blocking) emerge from agents trained with reinforcement learning. We provide evidence in both simulation and the real world that this approach outperforms the common paradigm of training agents in isolation with rewards that prescribe behavior, e.g., progress on the raceline, in particular when the complexity of the environment increases, e.g., in the presence of obstacles. Moreover, we find that multi-agent competition yields policies that transfer more reliably to the real world than policies trained with a single-agent progress-based reward, despite the two methods using the same simulation environment, randomization strategy, and hardware. In addition to improved sim-to-real transfer, the multi-agent policies also exhibit some degree of generalization to opponents unseen at training time. Overall, our work, following in the tradition of multi-agent competitive game-play in digital domains, shows that sparse task-level rewards are sufficient for training agents capable of advanced low-level control in the physical world.** Code Video



Train Agents with Sparse Competition Rewards

Pass every gate before your opponent



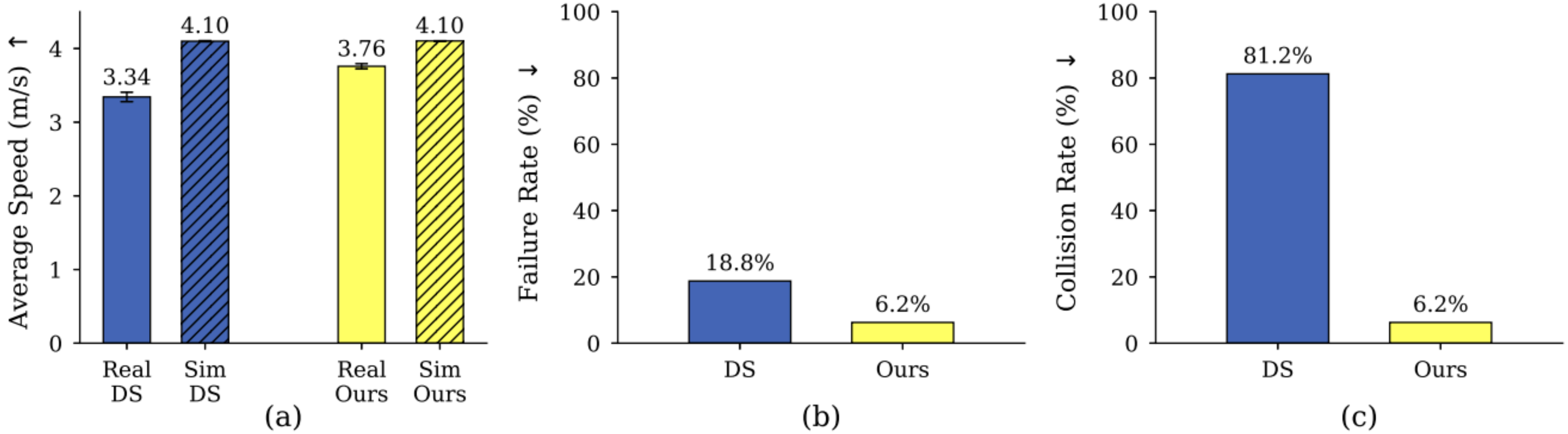
Bonus if you finish a lap first

Blocking (Ours vs Ours)

Blocking (Ours vs Ours)

Physical Contact (Ours vs Ours)

Improved Transfer without Improved Model



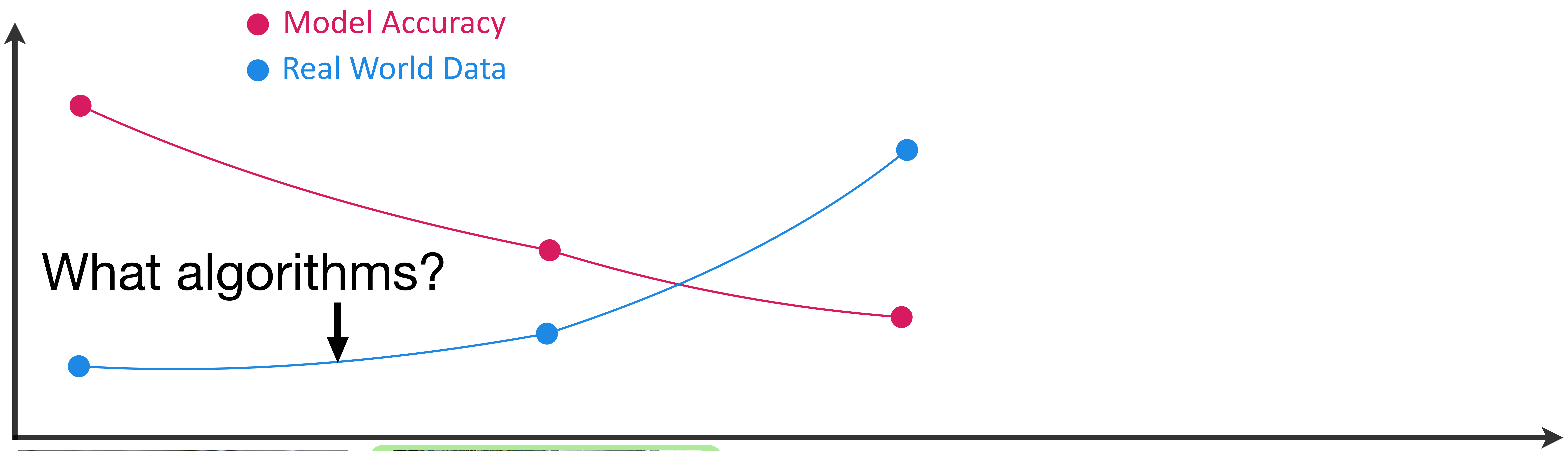
- Smaller sim-to-real gap than prior work.
- Related to the concept of adversarial domain randomization.

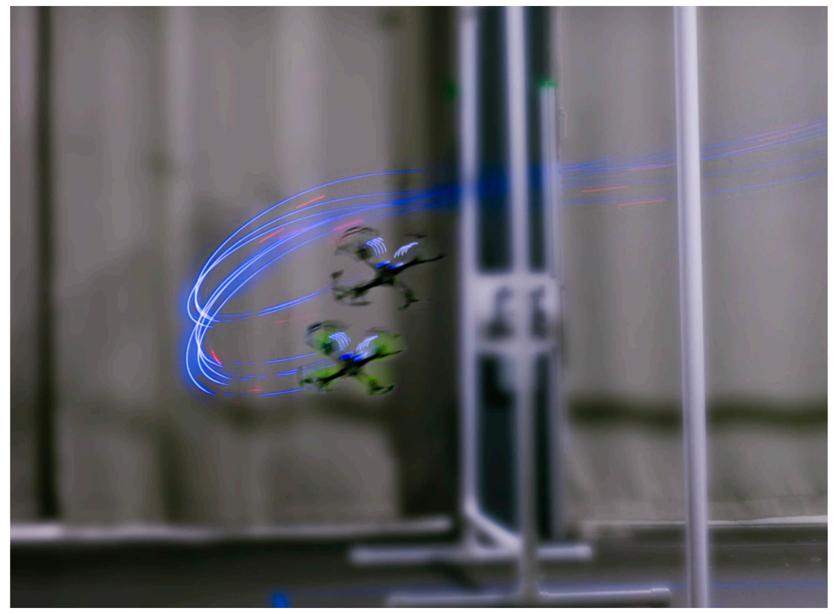
Summary up to now

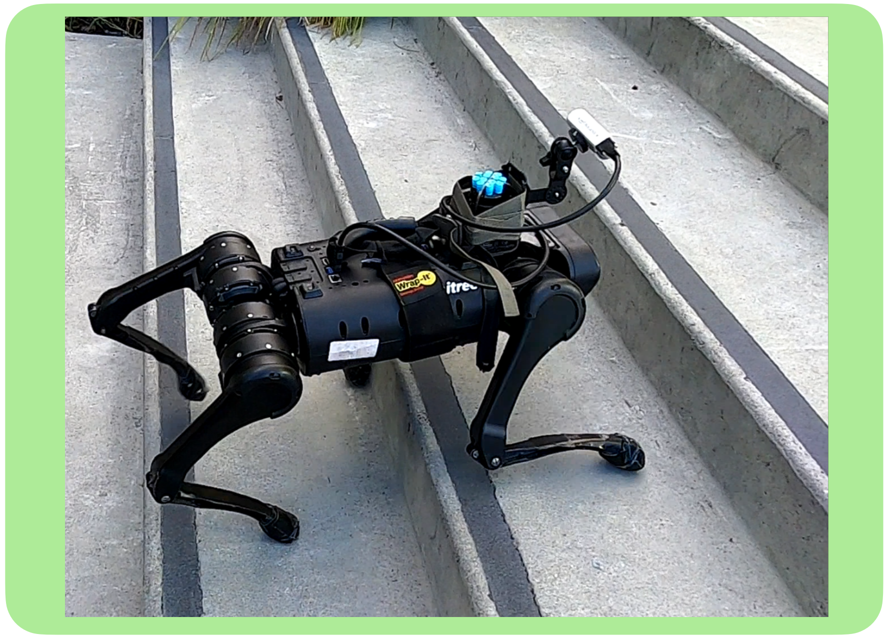
- Imposing structure in the randomization space, together with real-time adaptation, leads to generalization beyond the training environment.

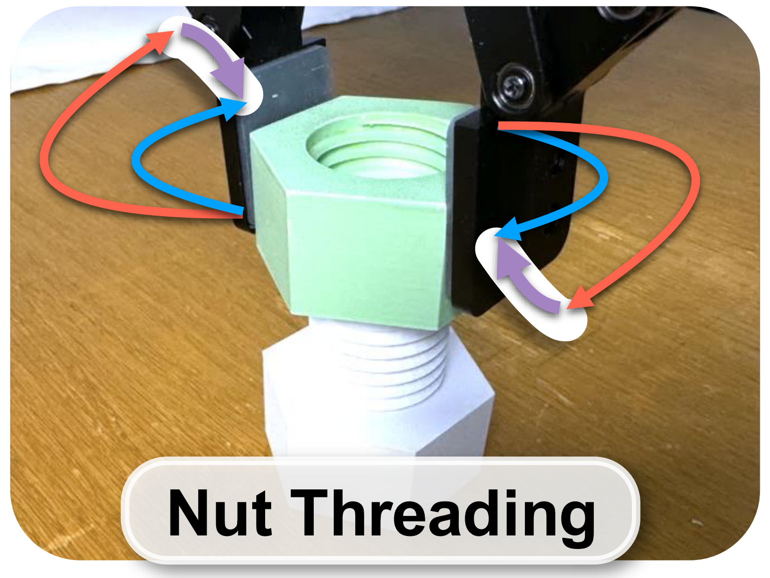
- **Adversarial learning empirically decreases reliance on model accuracy.**

Today's Talk


Model Accuracy
Real World Data
What algorithms?






Nut Threading

Learning Vision-Based Locomotion


Learning Visual Locomotion with Cross-Modal Supervision

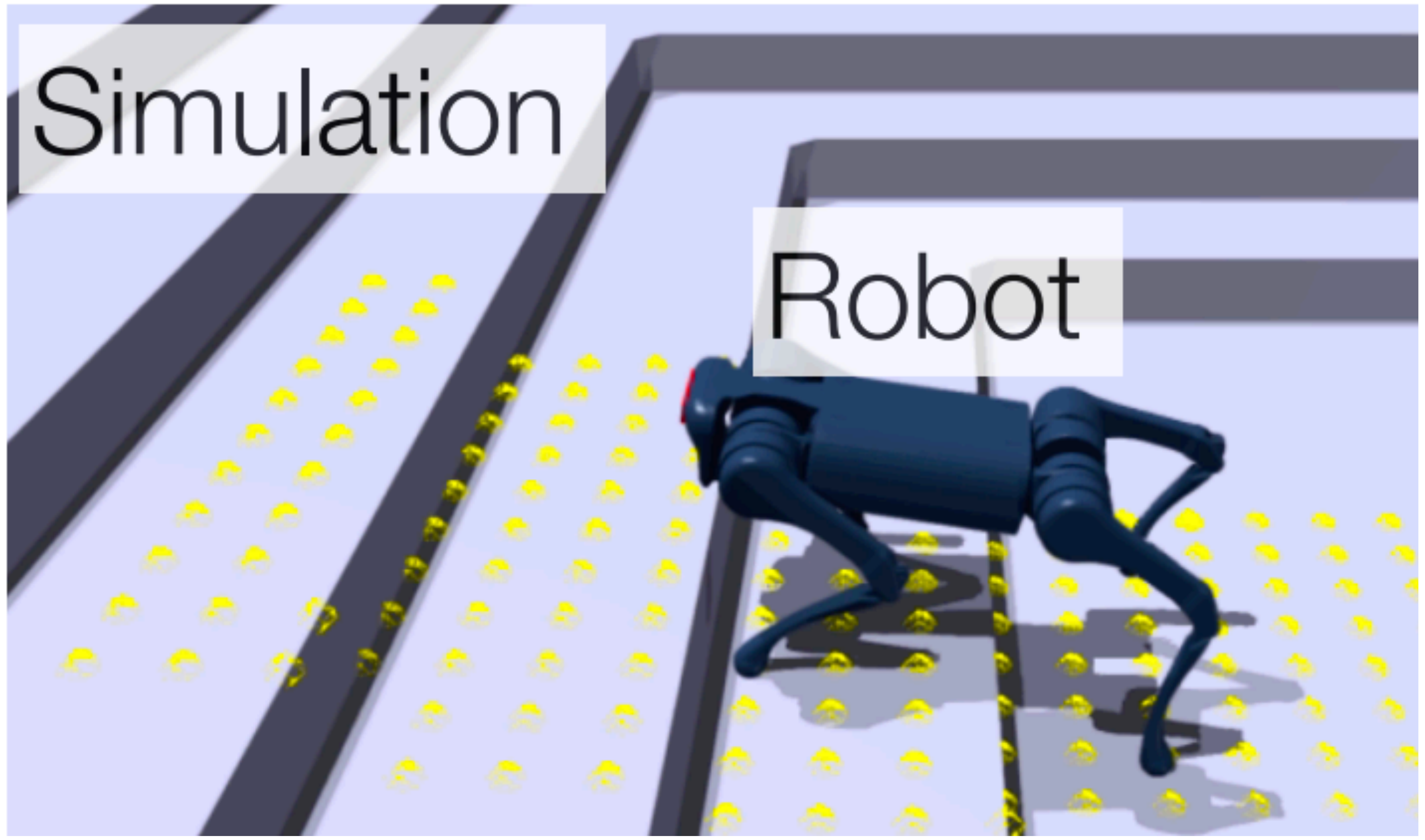
Antonio Loquercio*, Ashish Kumar*, Jitendra Malik

Learning Vision-Based Locomotion

Decision-Making Module

Actions

Learning Vision-based Locomotion in Simulation



Too hard to simulate realistic RGB at high frame rate!

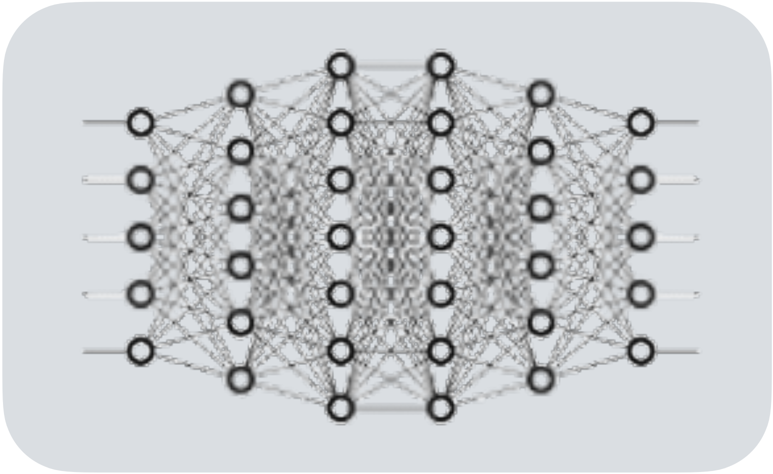
Using Depth Instead of RGB

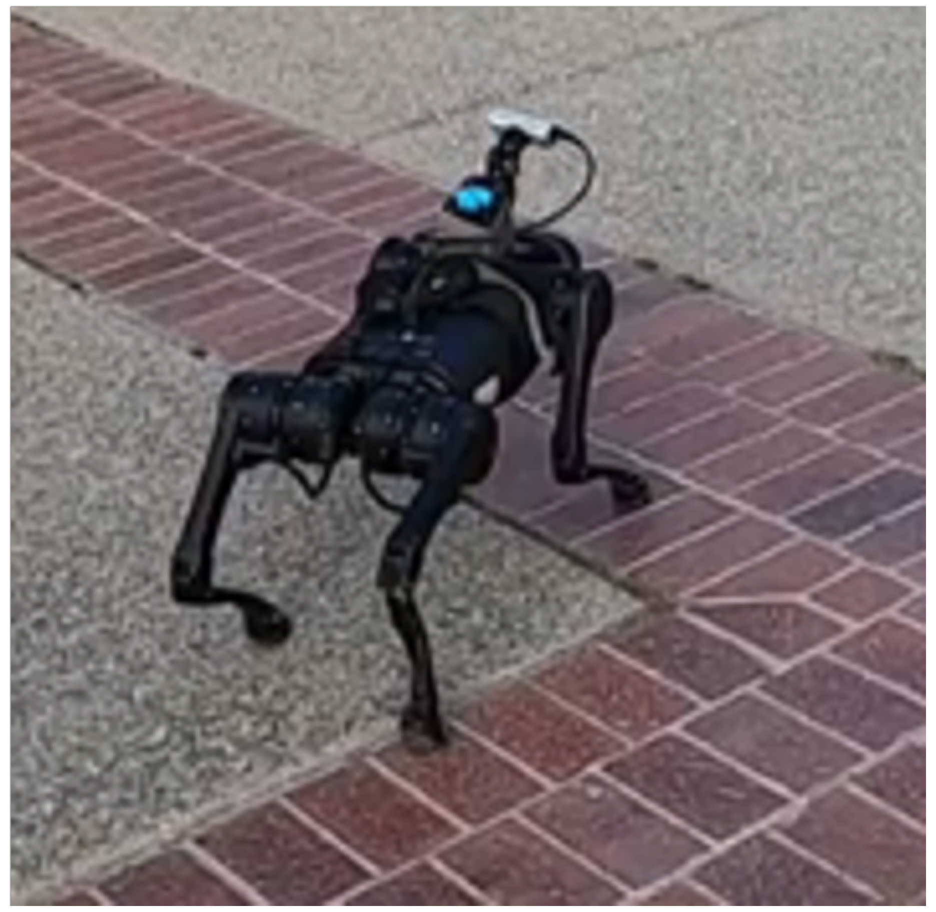
Decision-Making Module

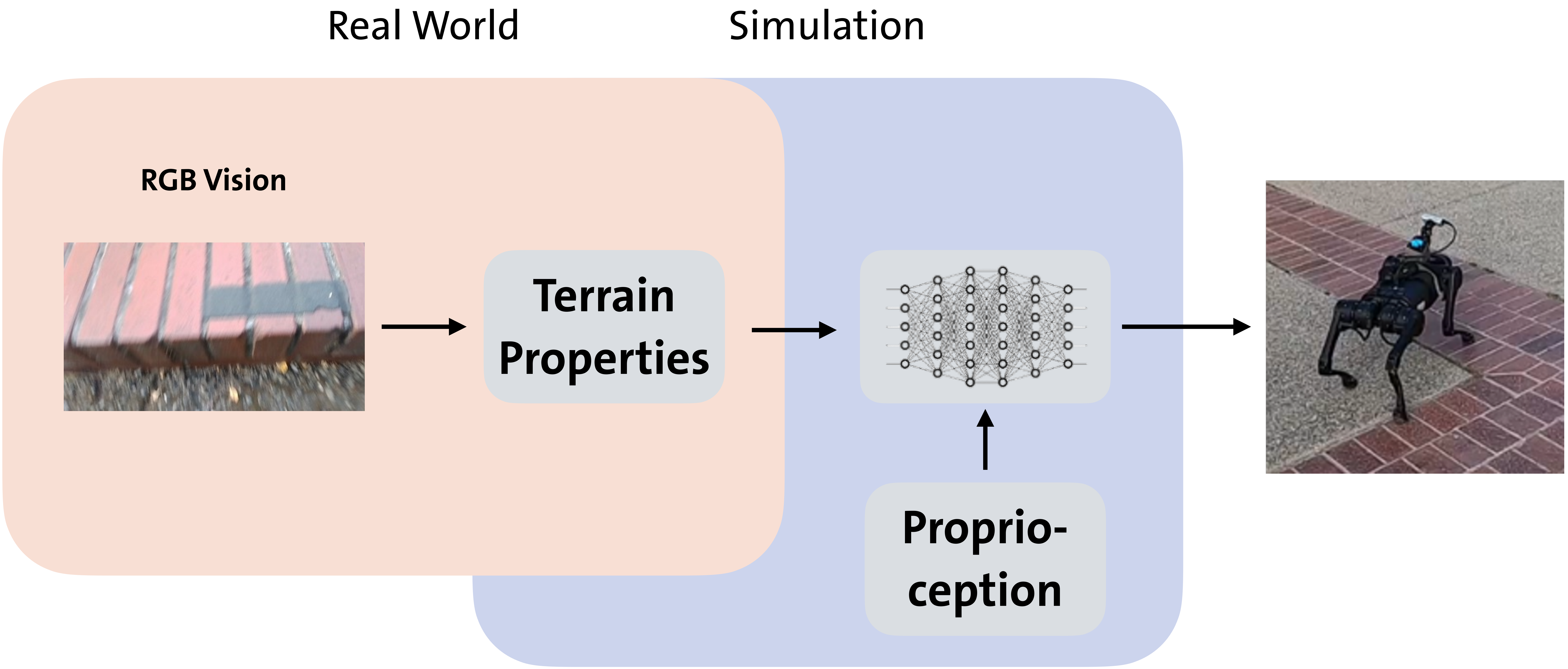
Actions

Agarwal et al., 2022
Miki et al., 2022
...

RGB Vision







Hwangbo et al., 2019
Lee et al., 2020
Kumar et al., 2020

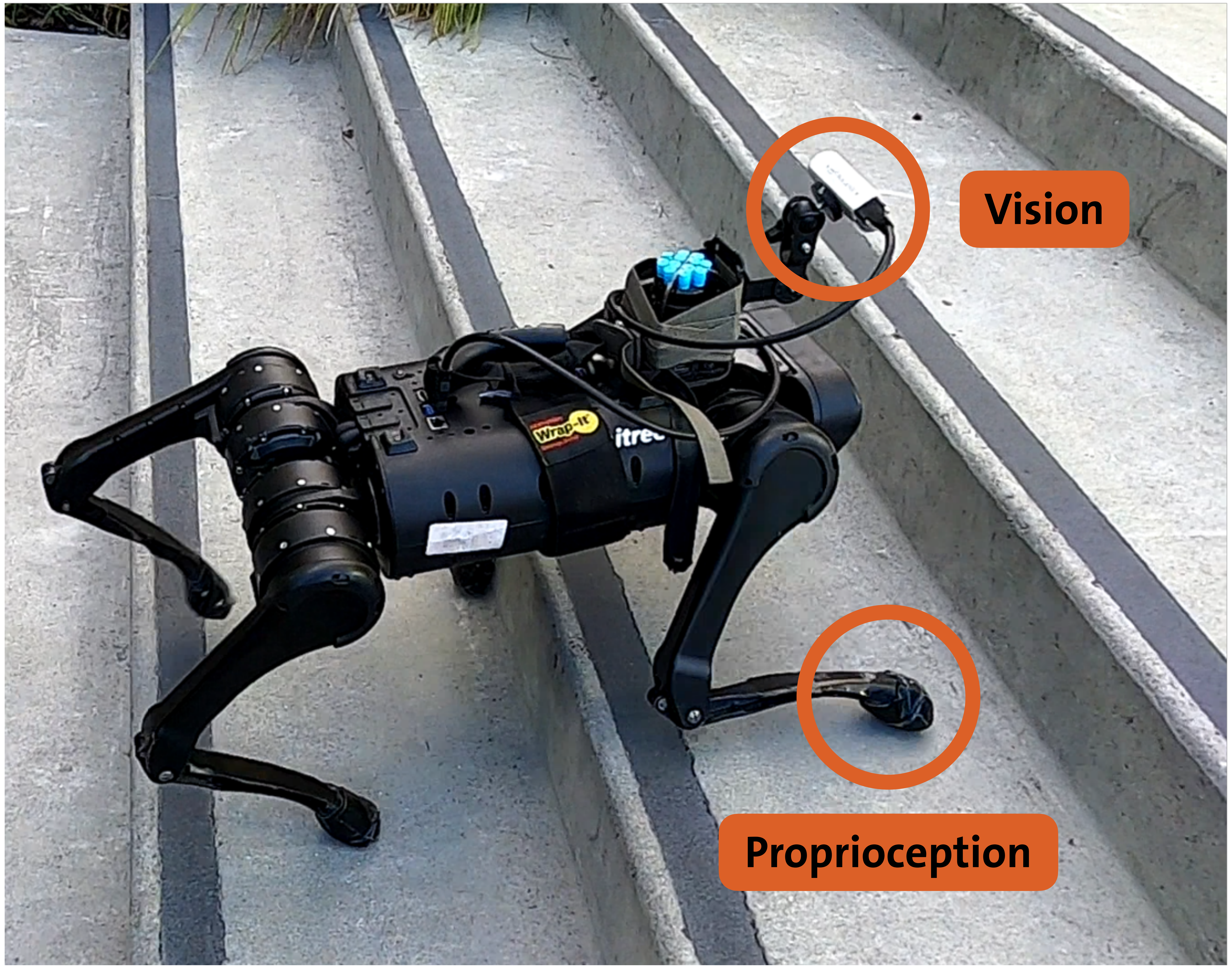
RGB Vision

Terrain Properties

How do we train this abstraction?

1. We can’t use existing datasets
2. Humans can’t provide annotations
3. Can’t use foundation models

Learning Vision from Proprioception



Key Hypothesis:

If proprioception can predict a terrain parameter, then the parameter is important for the task!

Cross-Modal Supervision

Core Concept: Vision will not predict an accurate representation of the environment, but only what proprioception deems important!

Blind

Loquercio et. al,
ICRA, 2023

Vision

Loquercio et. al,
ICRA, 2023

Construction Zone

Loquercio et. al,
ICRA, 2023

Loquercio et. al,
ICRA, 2023

Visual Plasticity

Before Adaptation

After 1min of data

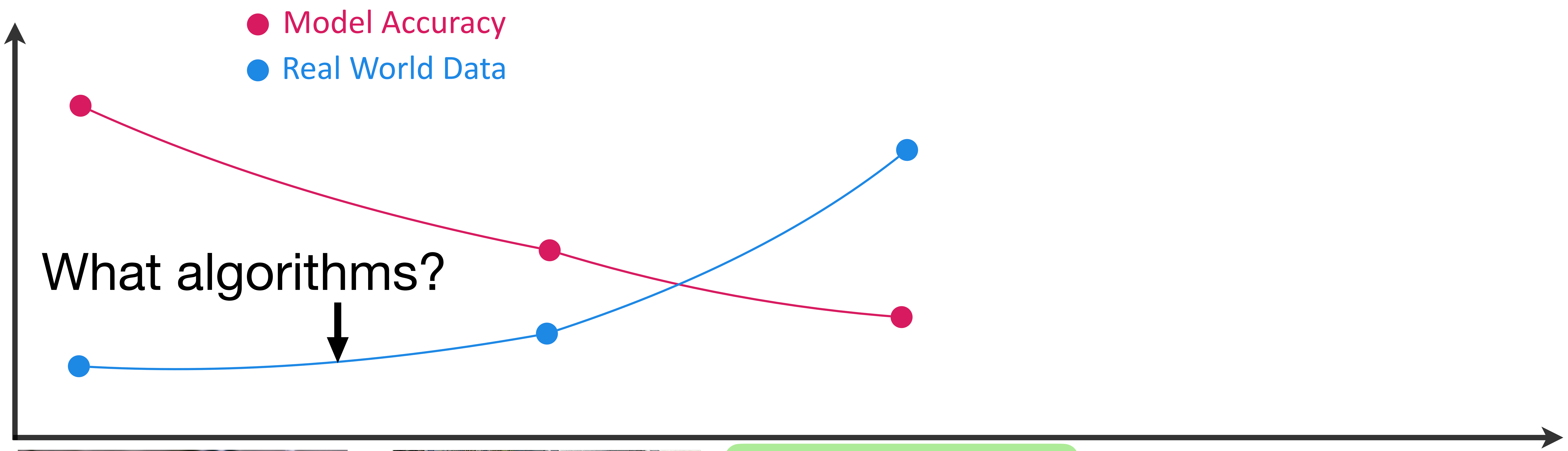
Summary up to now

- Imposing structure in the randomization space, together with real-time adaptation, leads to generalization beyond the training environment.

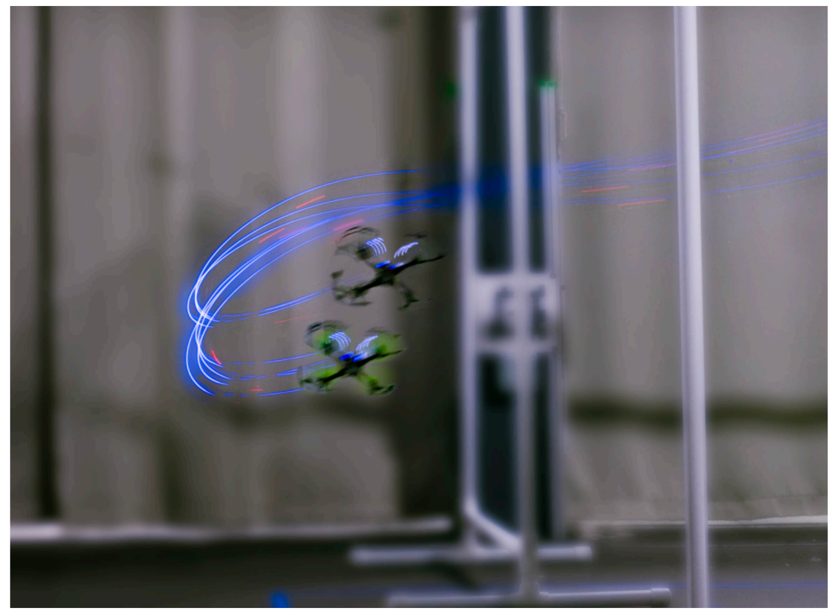
- Adversarial learning empirically decreases reliance on model accuracy.

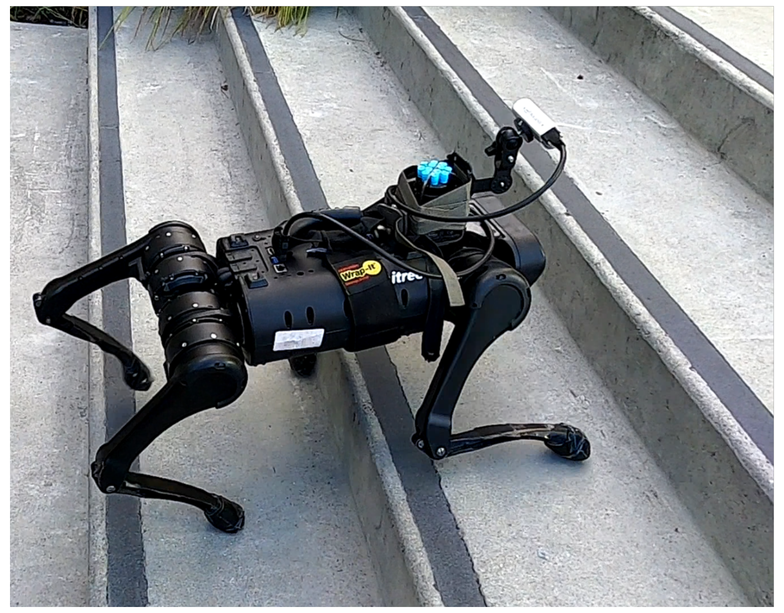
- **There is a coupling between what you can predict and what you need to accomplish a task.**

Today's Talk



Model Accuracy
Real World Data
What algorithms?

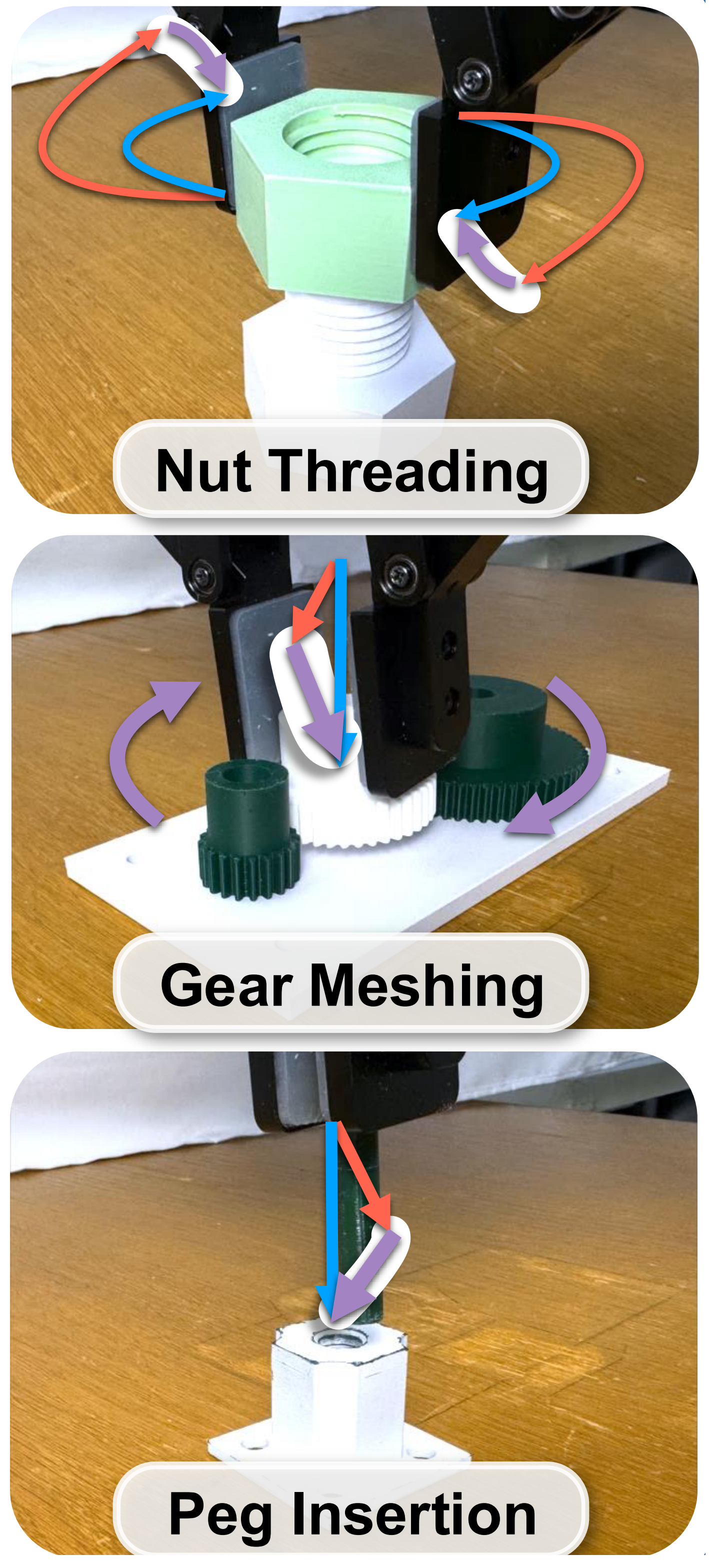


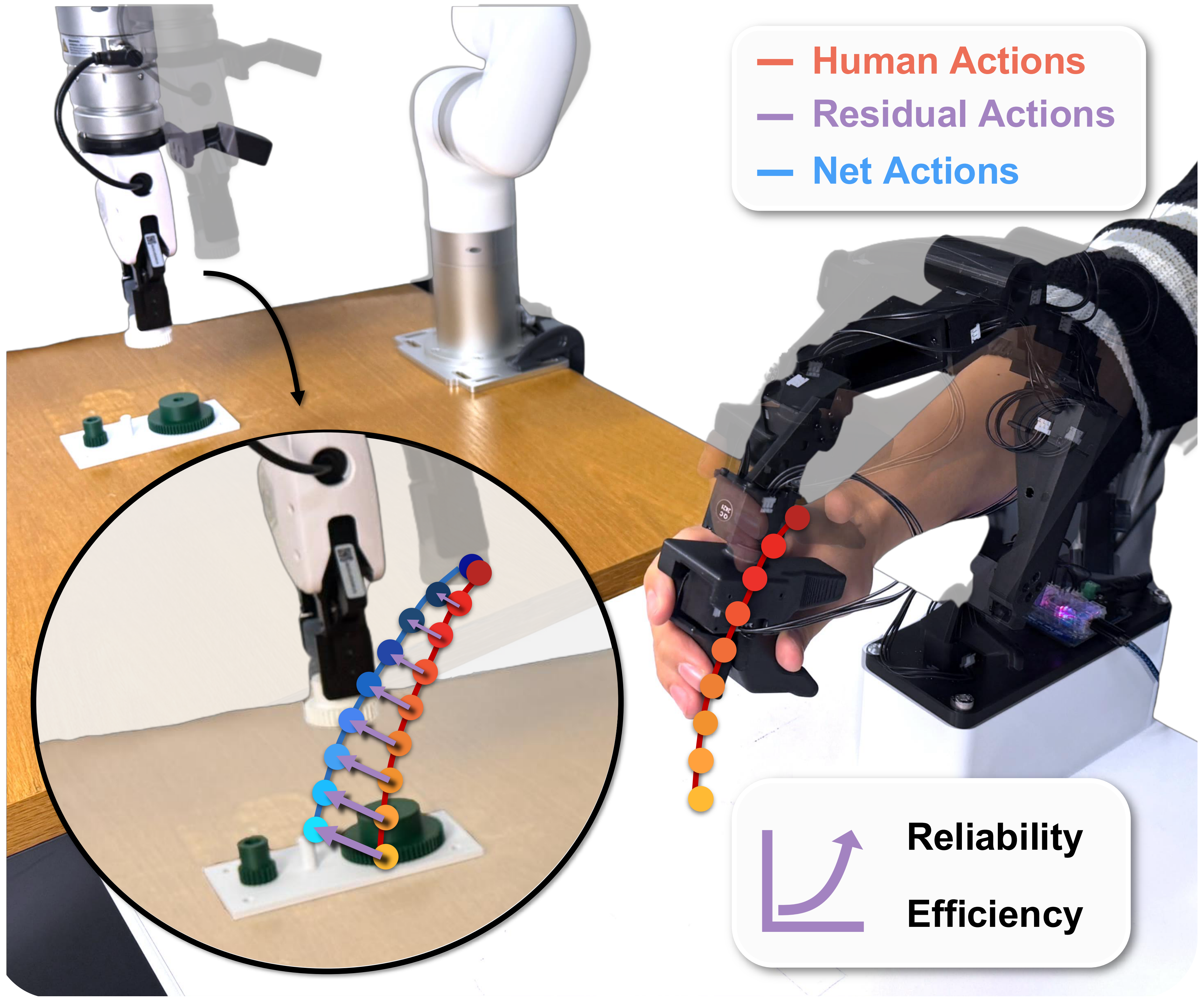


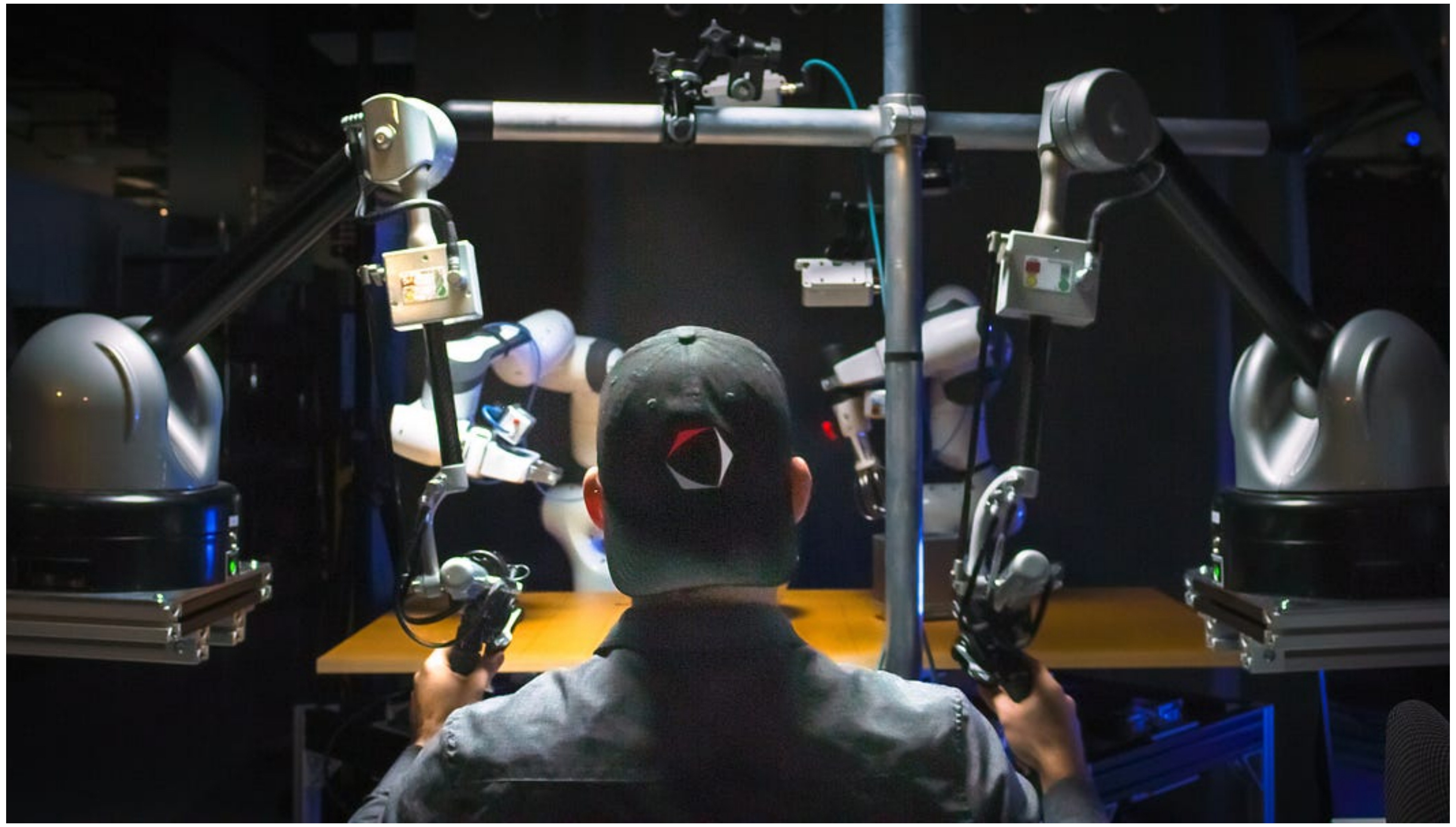
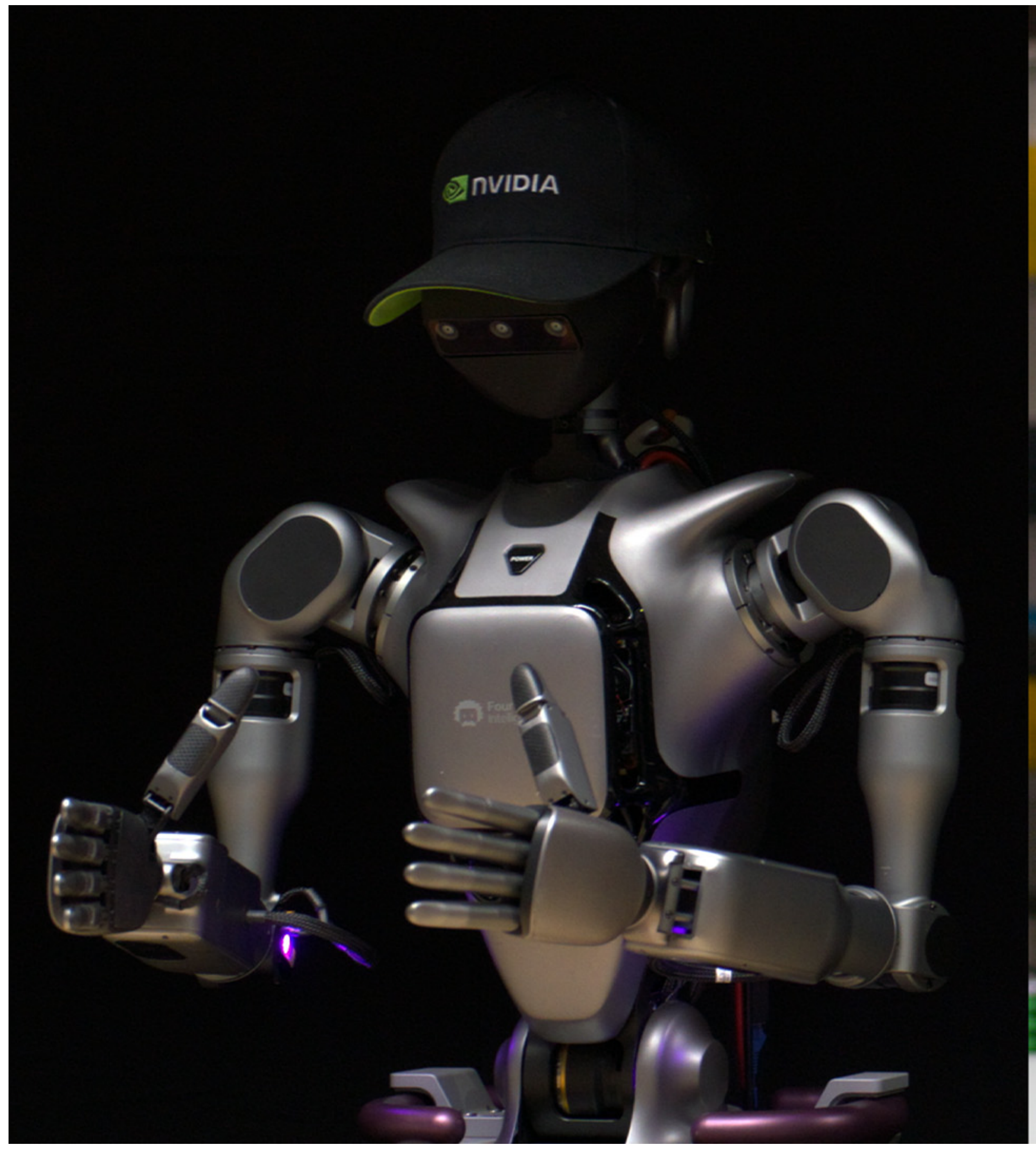
Nut Threading

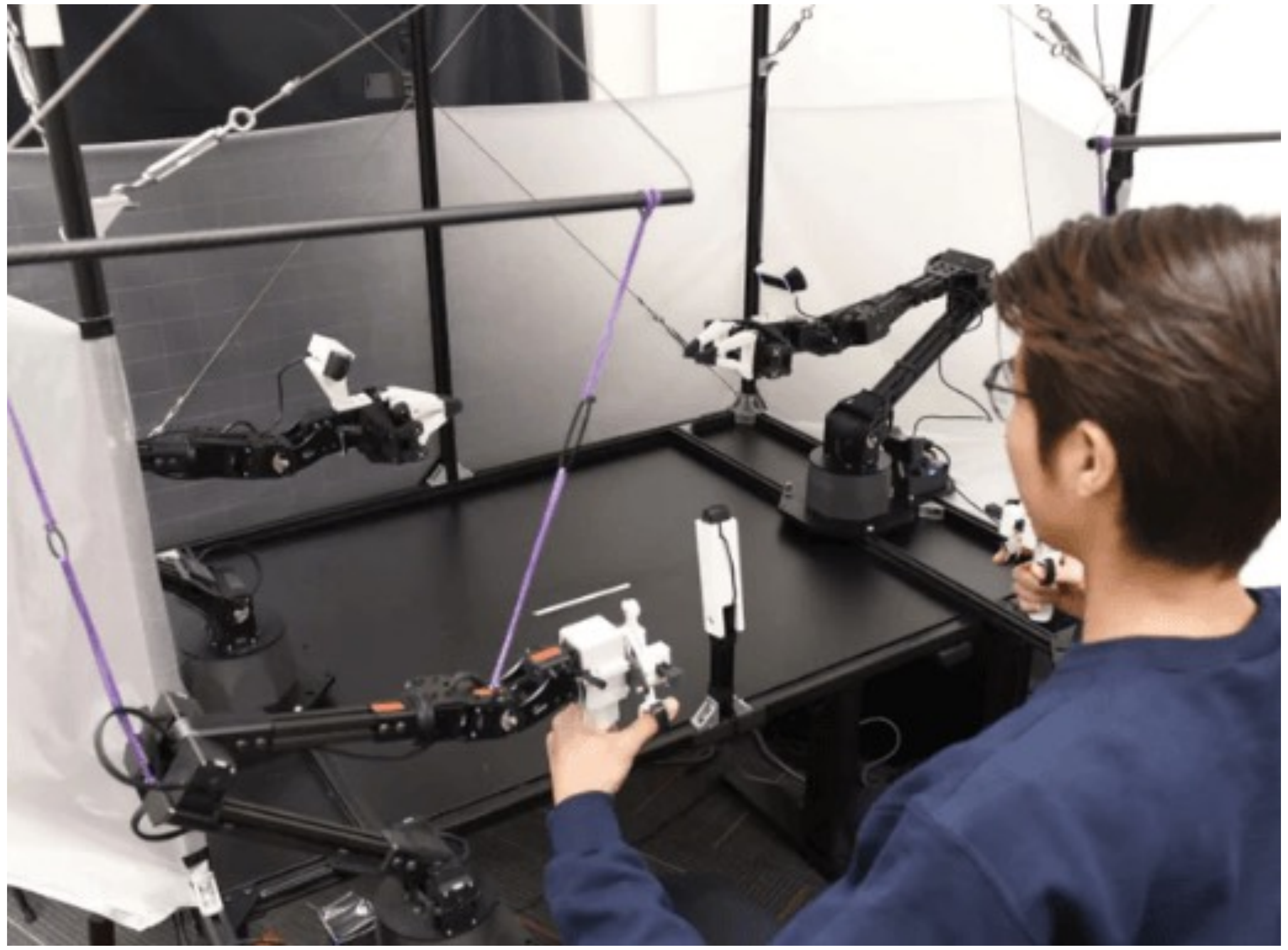
Efficient and Reliable Teleoperation through Real-to-Sim-to-Real and Shared Autonomy

Shuo Sha, Yixuan Wang, Binghao Huang, Antonio Loquercio, Yunzhu Li





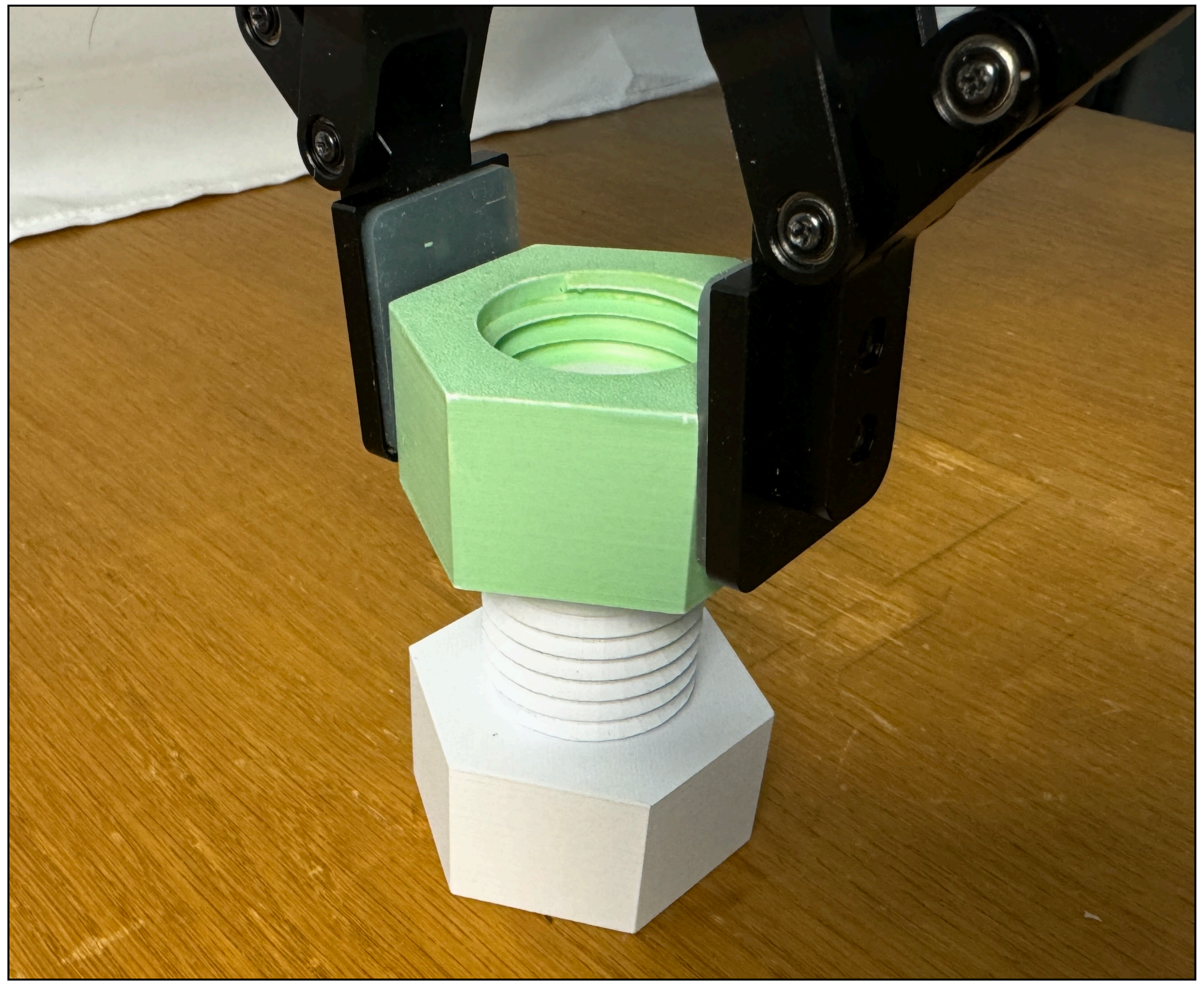


NVIDIA



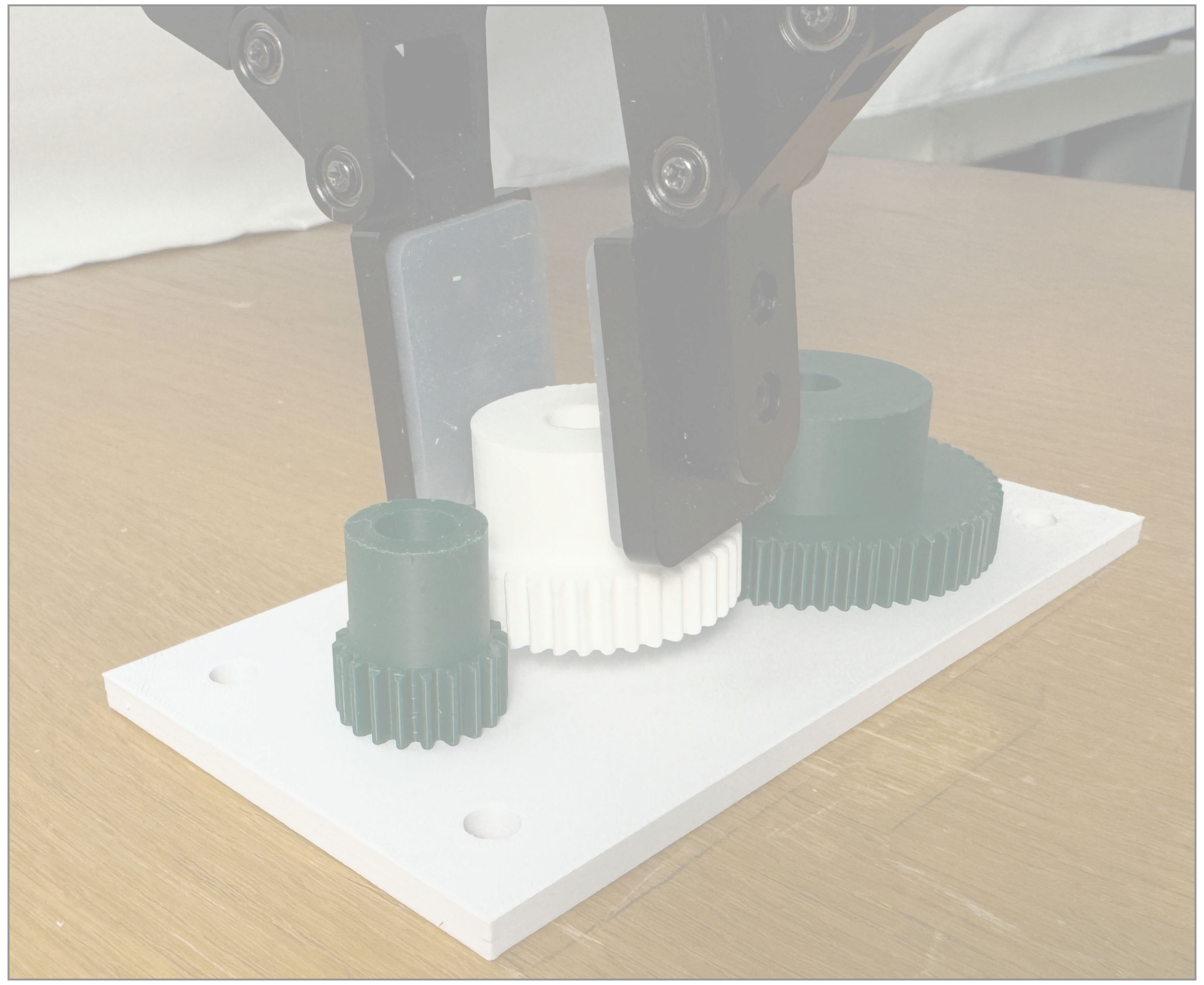
National
University
Hospital

The Limits of Direct Teleoperation

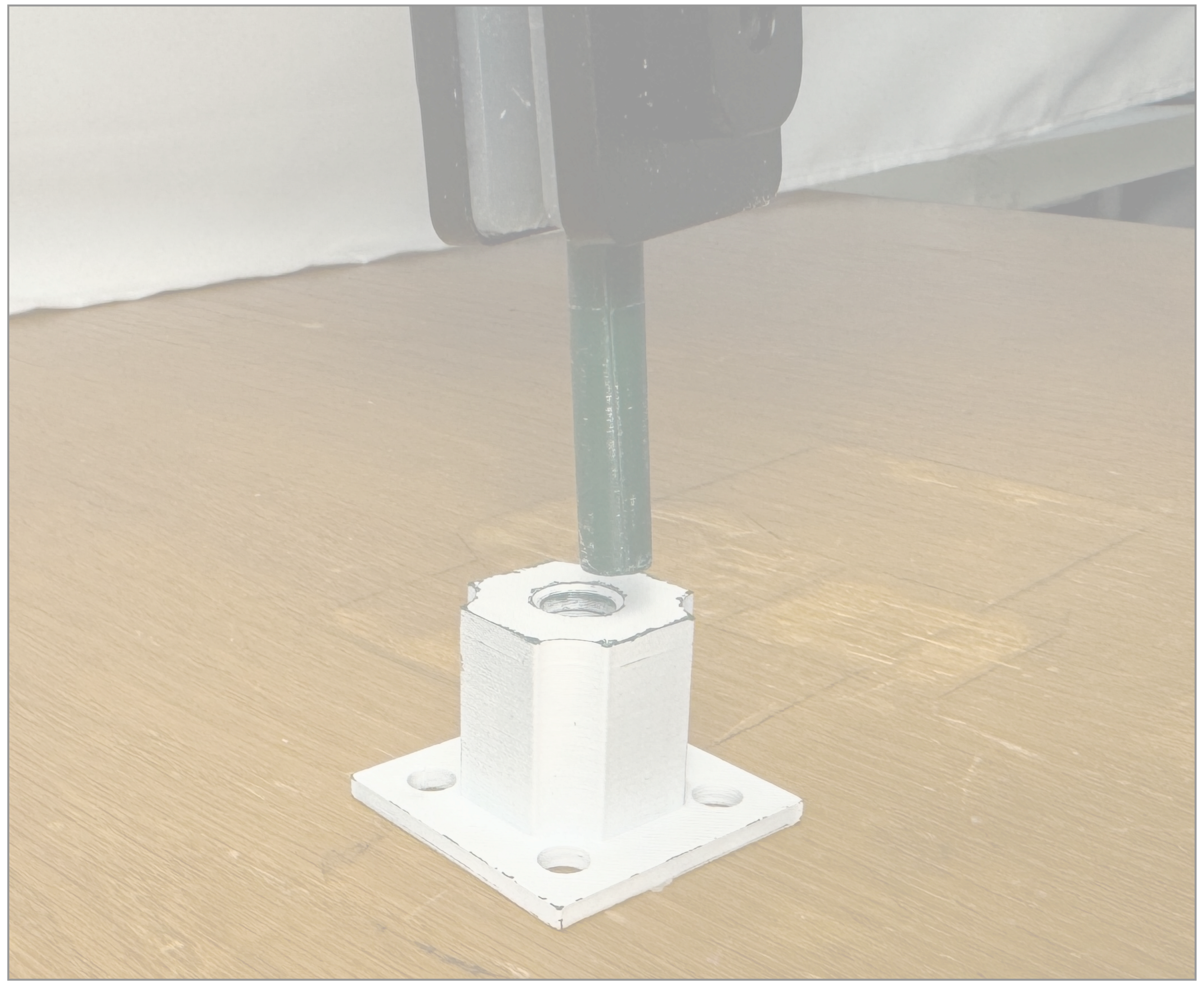
Teleoperation is often inefficient and unreliable for **fine-grained**, **contact-rich** manipulation tasks



Nut Threading



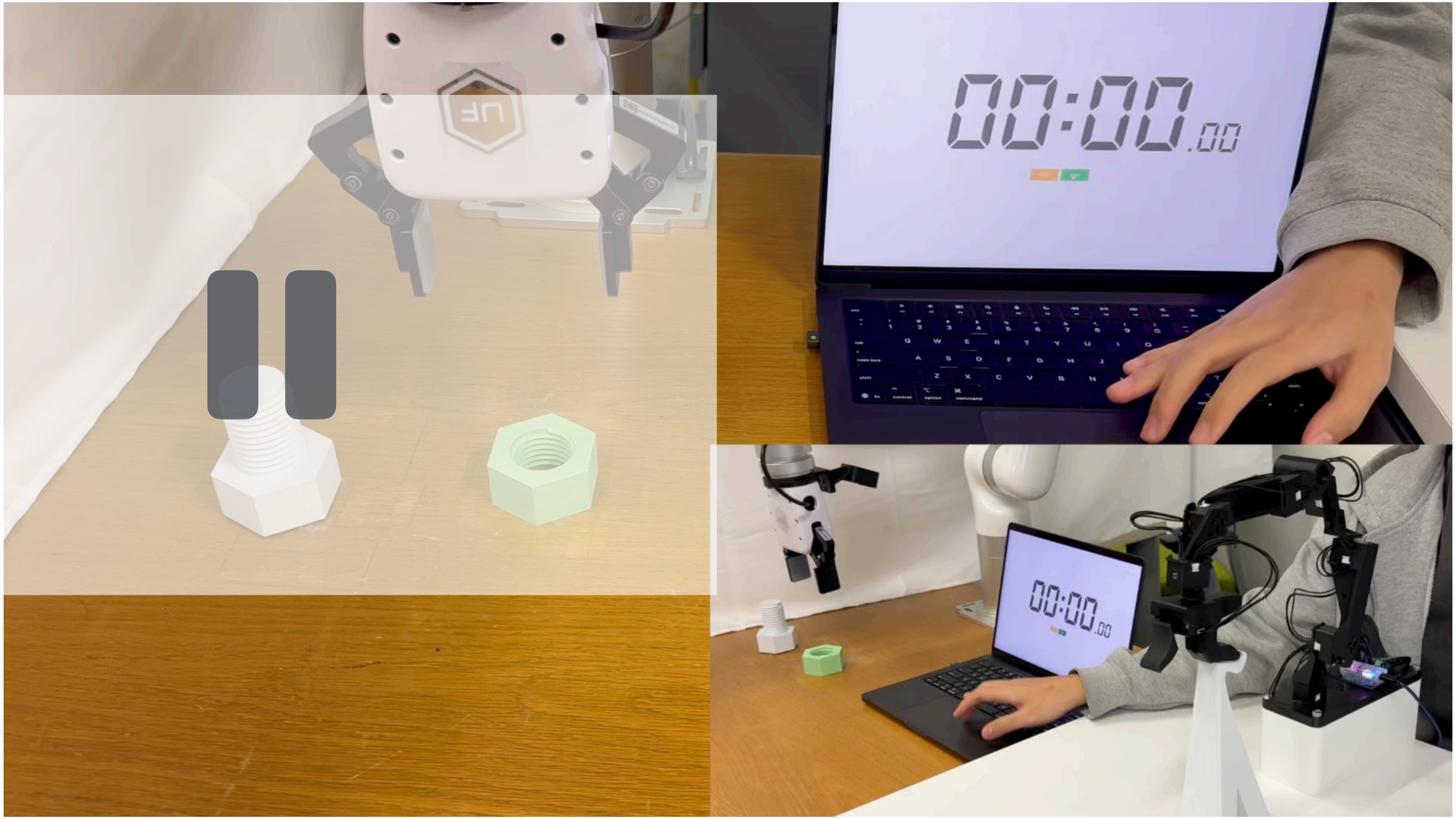
Gear Meshing



Peg Insertion

Copilot-assisted Teleoperation Performance

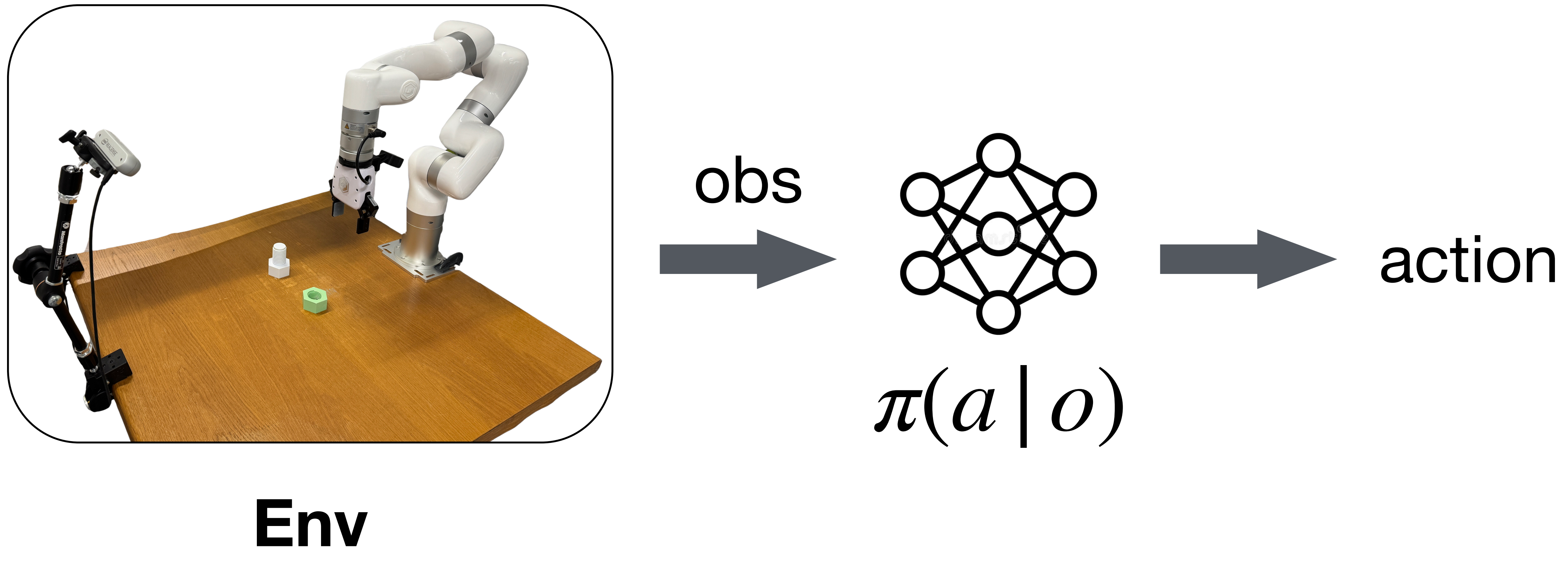
Direct Teleoperation Performance



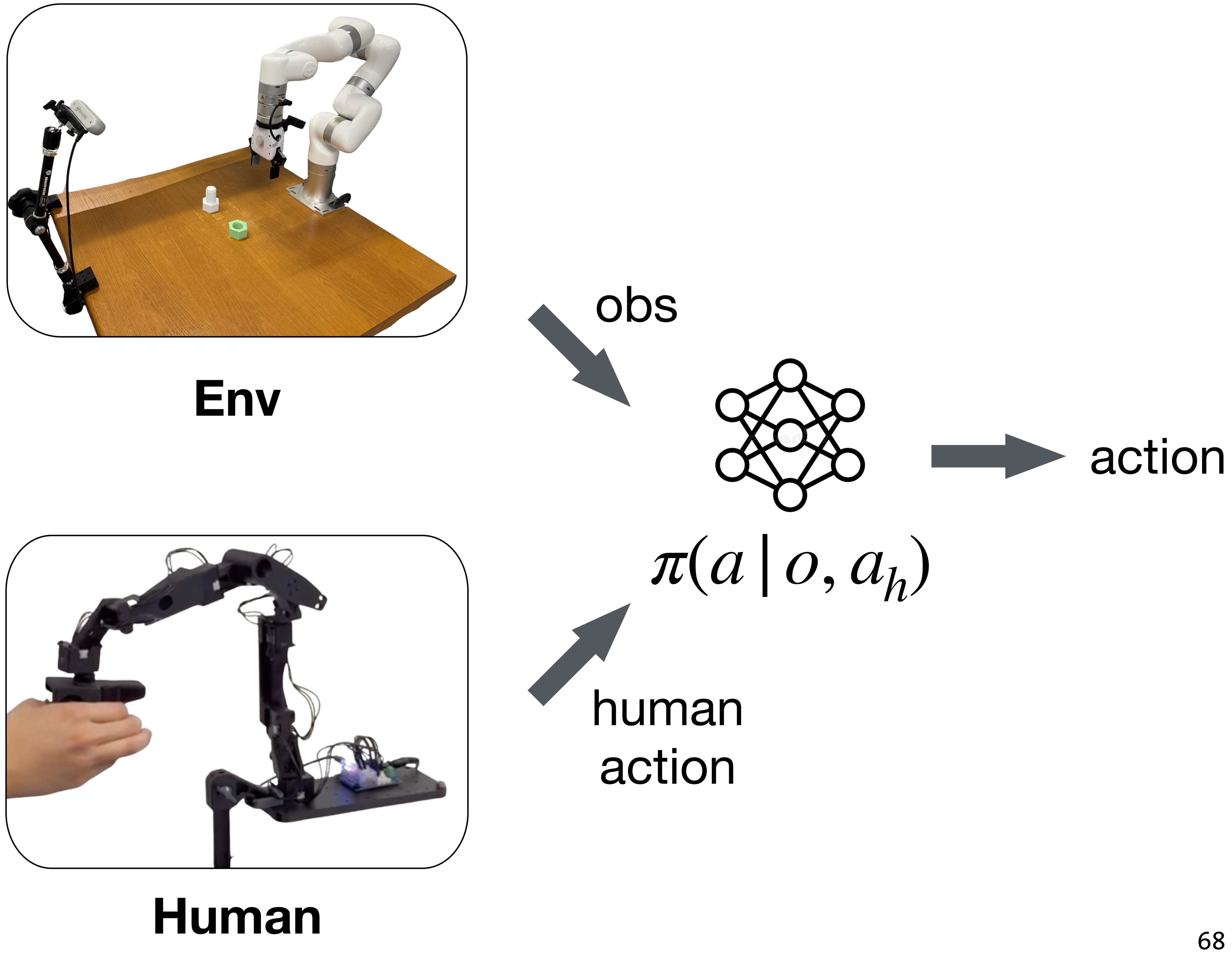
(All videos 1x speed)

The Challenges of Shared Autonomy

Traditional Autonomy



Shared Autonomy



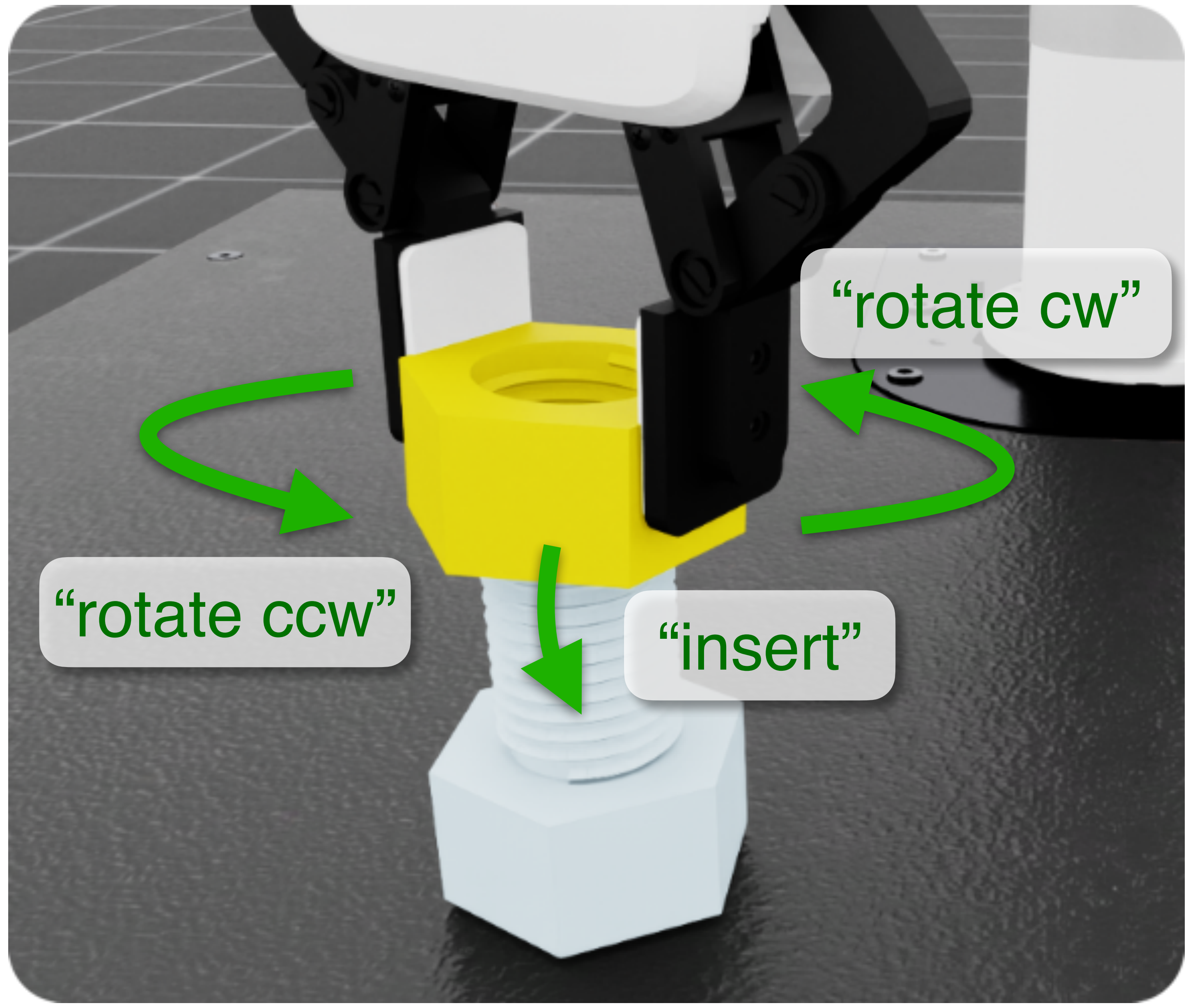
Sim-to-Real for Shared Autonomy

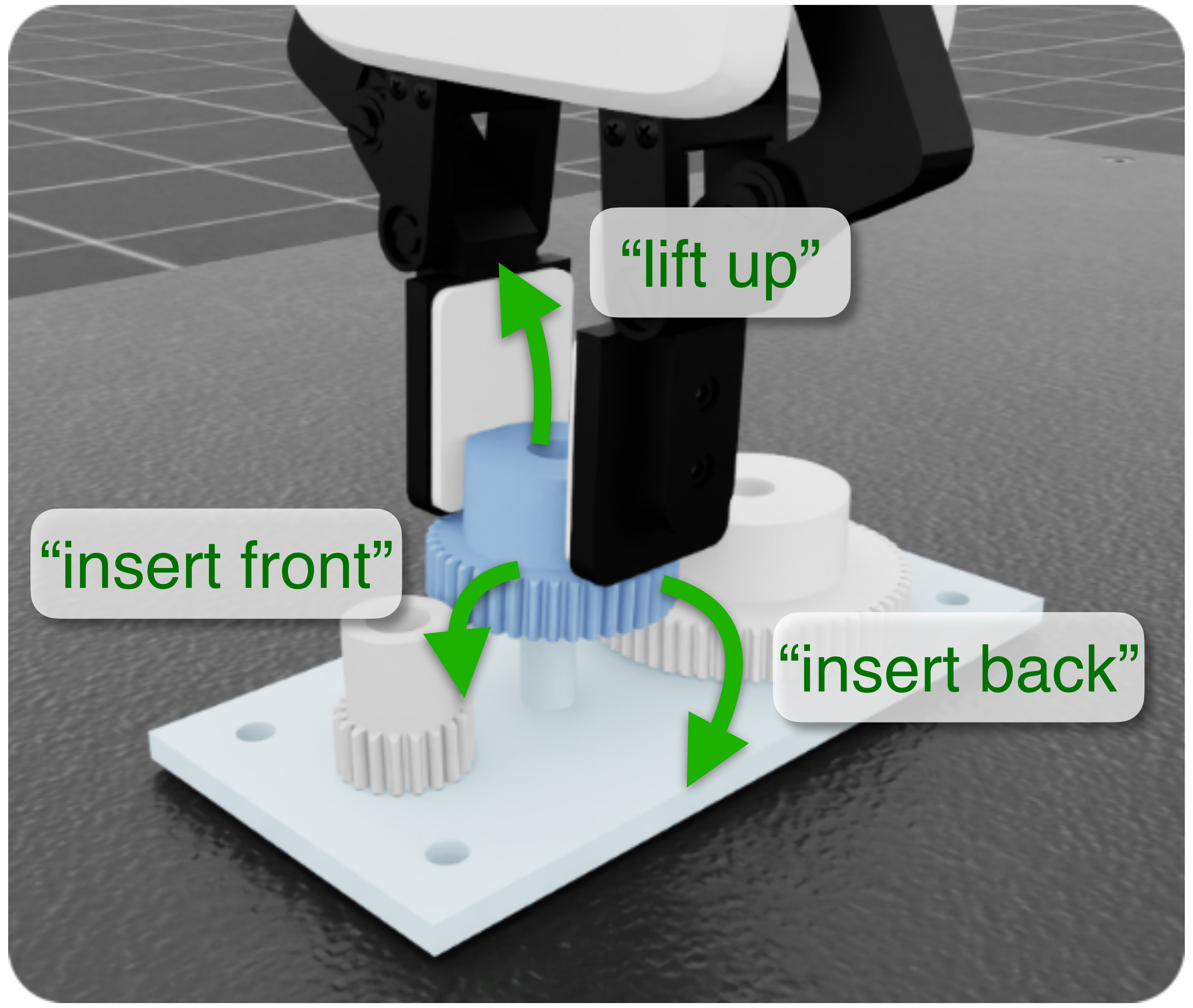
$\pi(a \mid o, a_{human})$

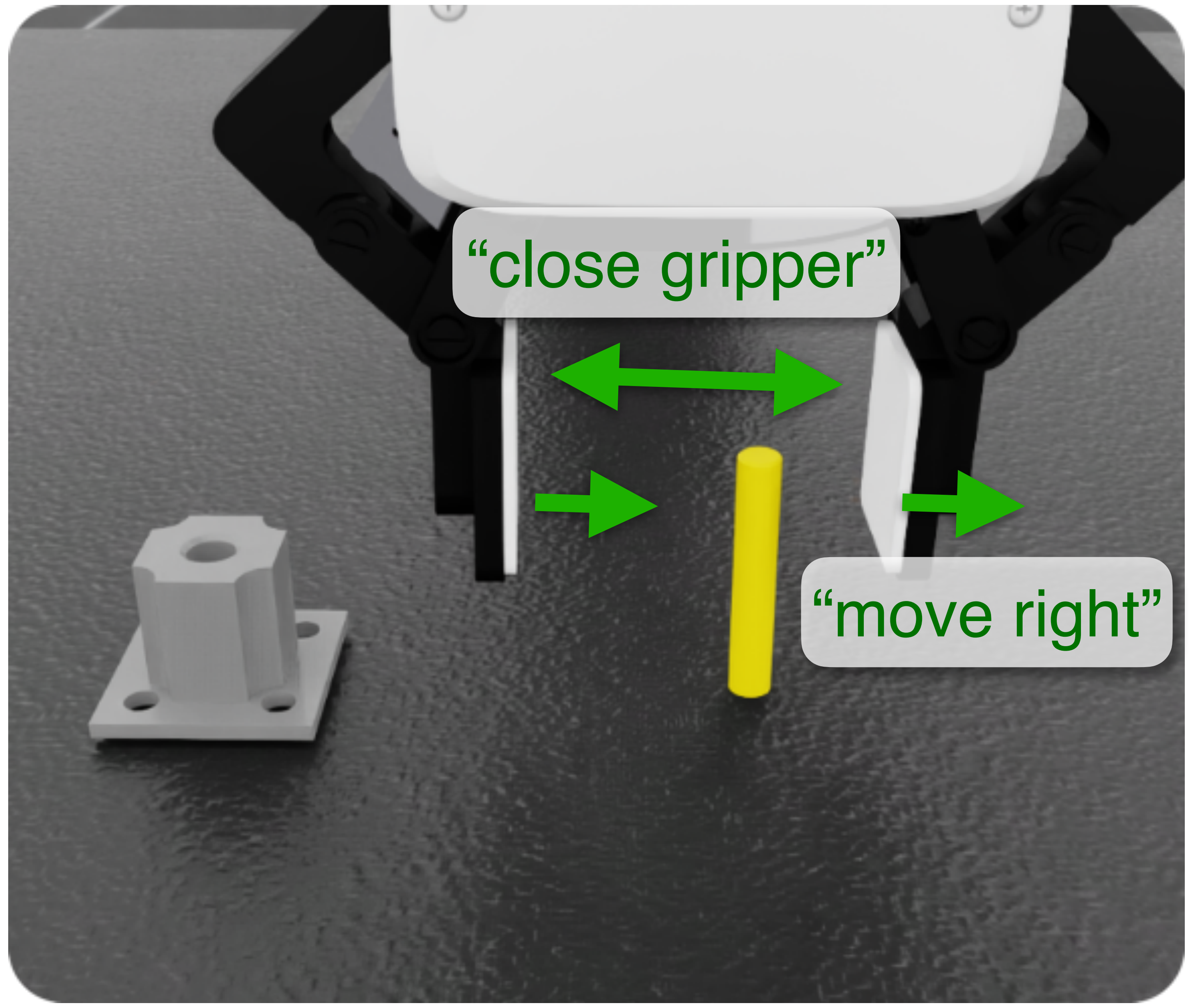
simulating physics
is possible

simulating a human
is hard

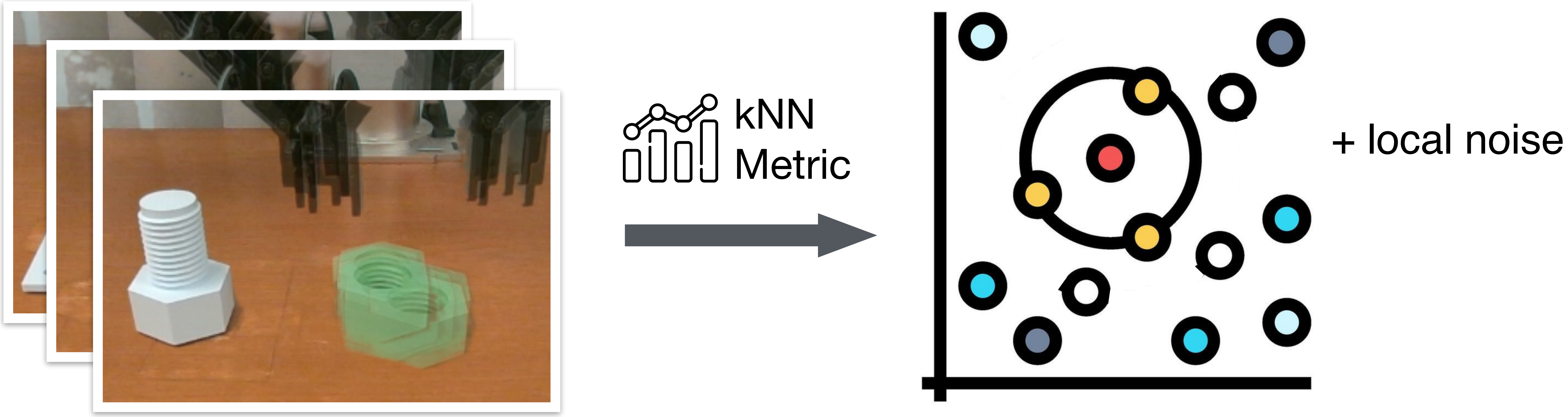
What is a sufficient human surrogate model for training π in simulation?







A Simple and Effective Human Surrogate: kNN



Teleop Data
(< 5 min)

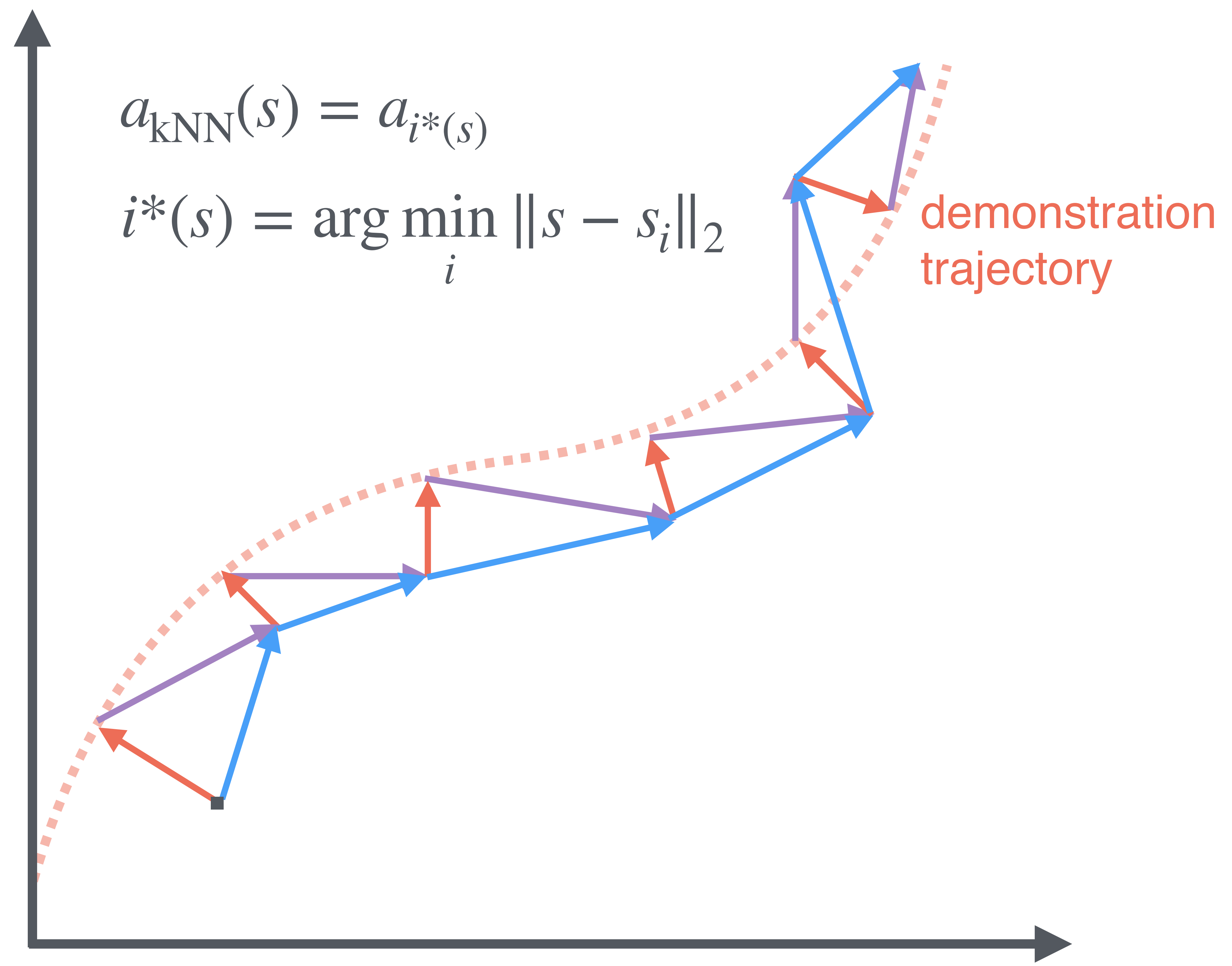
kNN Human
Surrogate

- Data efficient
- Anchors exploration to a manifold of the demonstrations

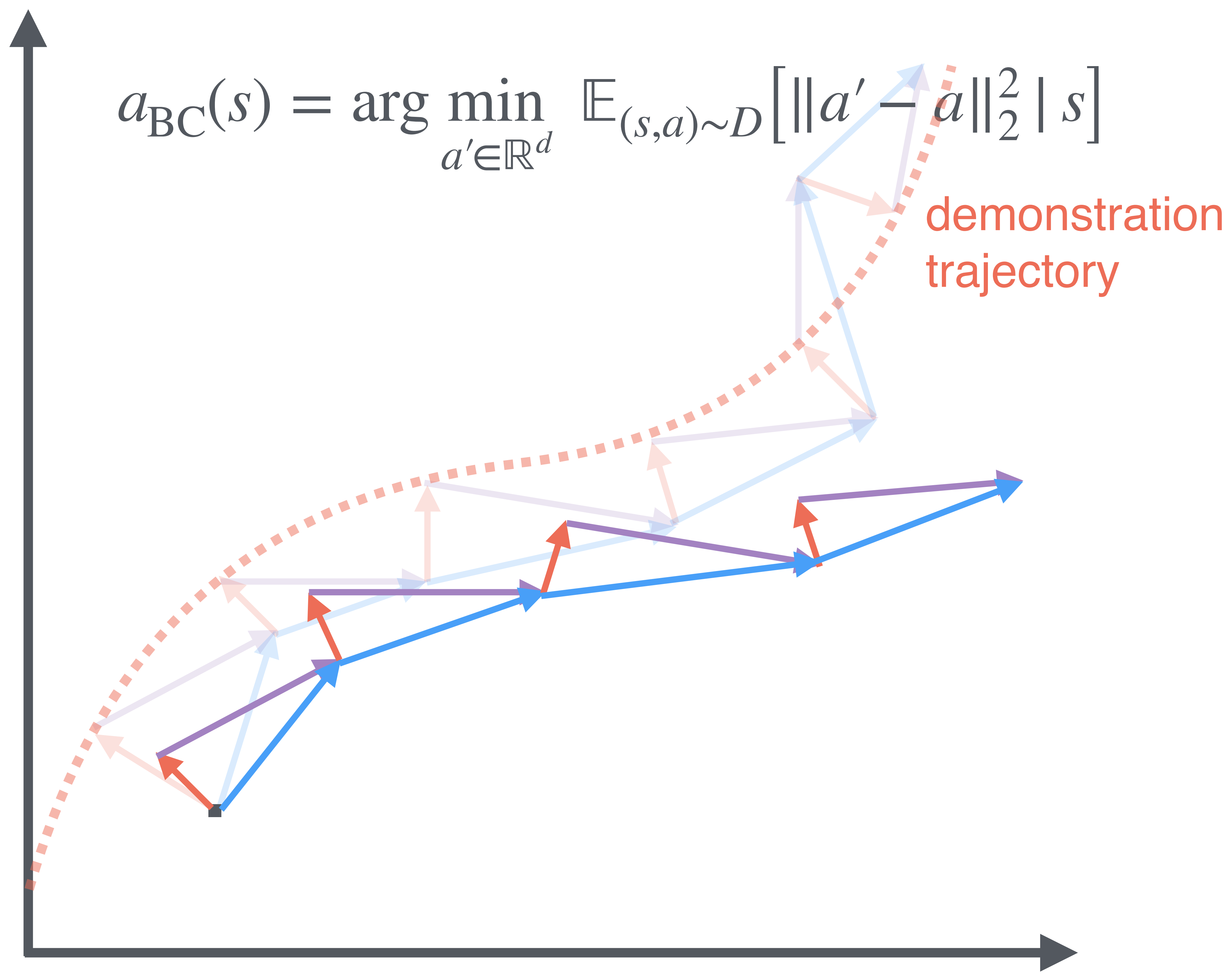
Why a kNN Model as a Human Surrogate?

A kNN model anchors exploration around a local tube of the demonstrations

human actions SA policy π net actions



kNN human model



BC human model

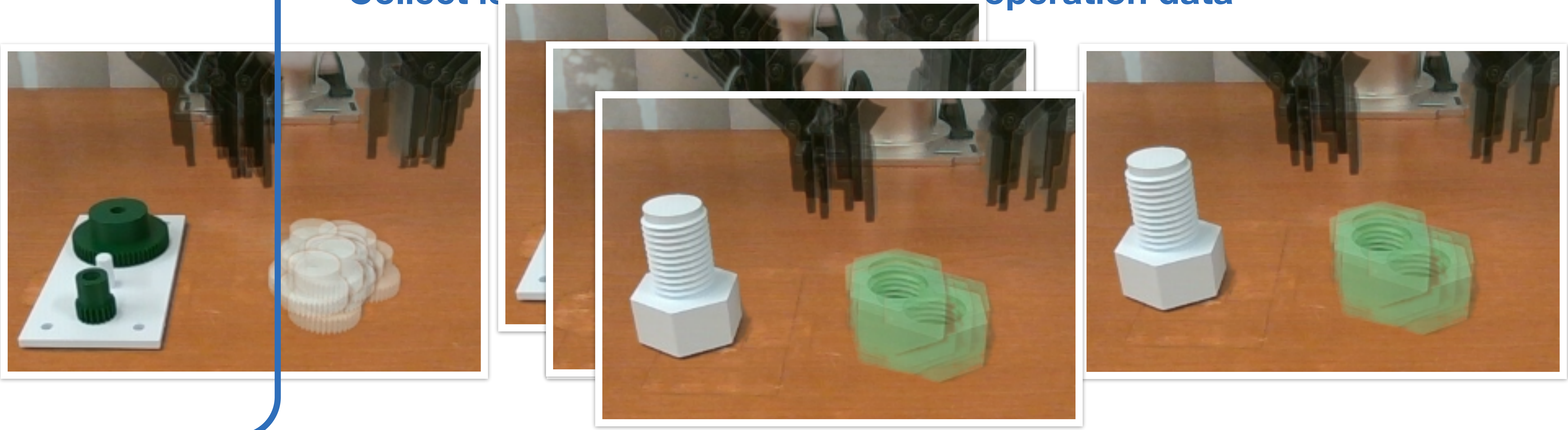
Method Overview

Real World → Simulation → Real World Shared Autonomy

Step 1: Creating a Human Surrogate

Real World

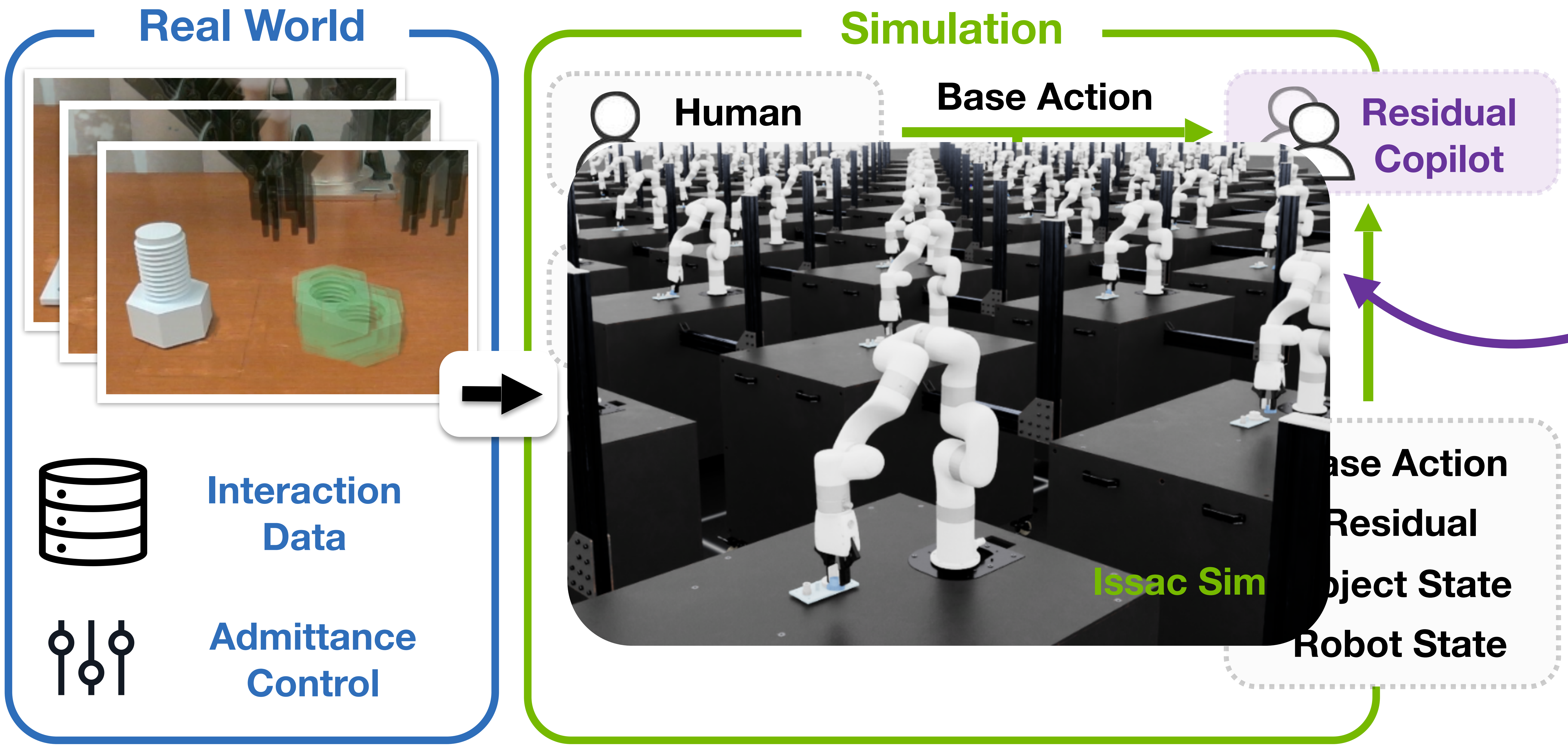
Collect less than 5min real-world teleoperation data



Interaction Data

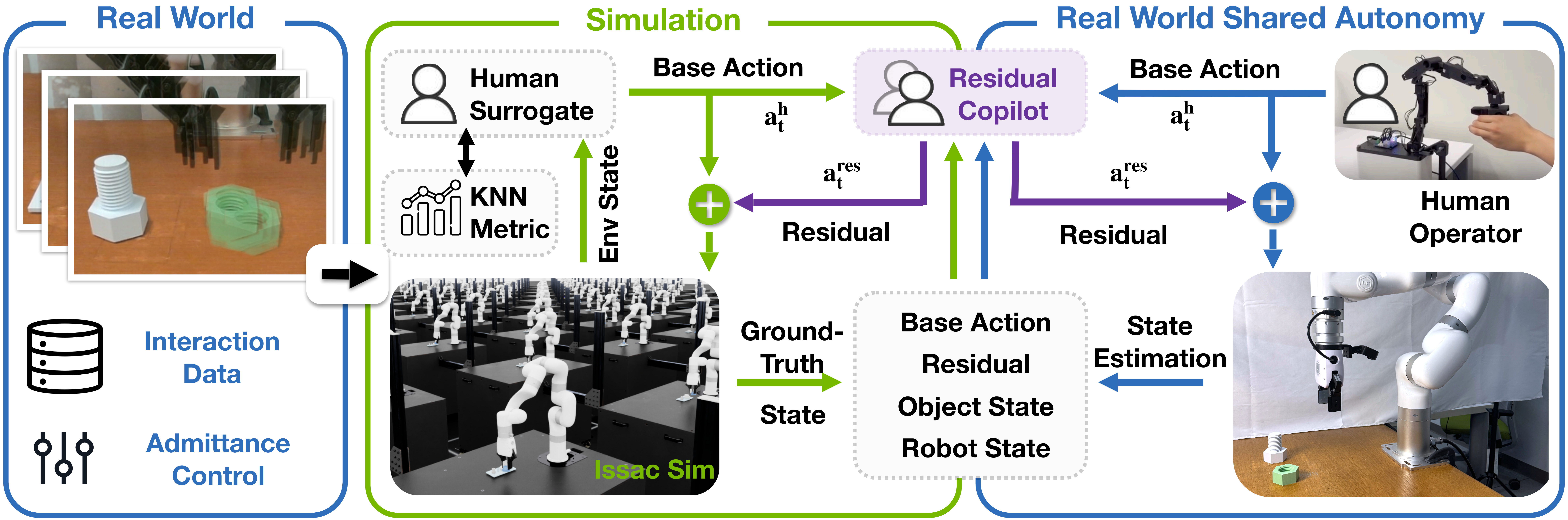
Admittance Control

Step 2: Training a SA policy in Simulation with RL

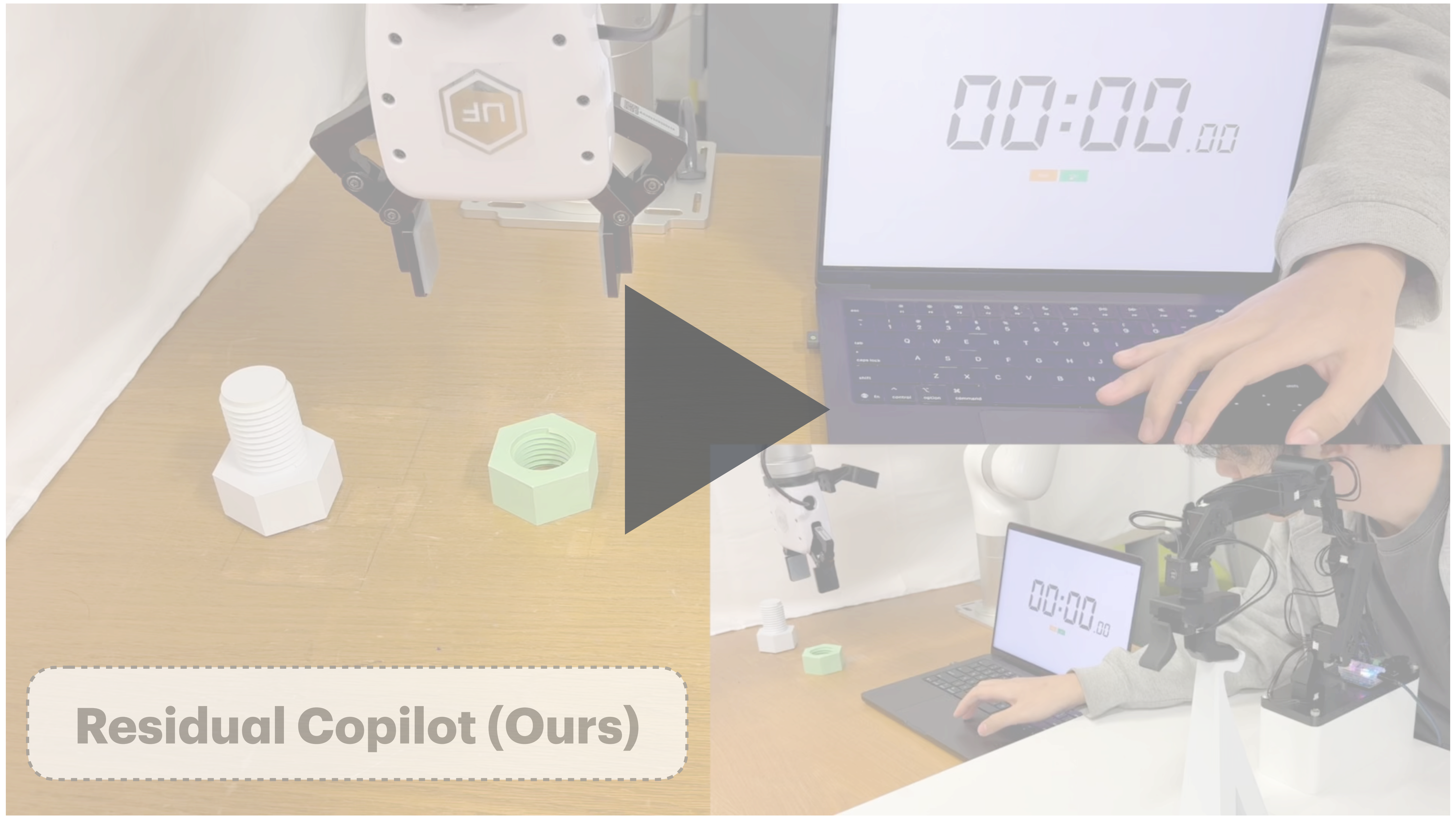


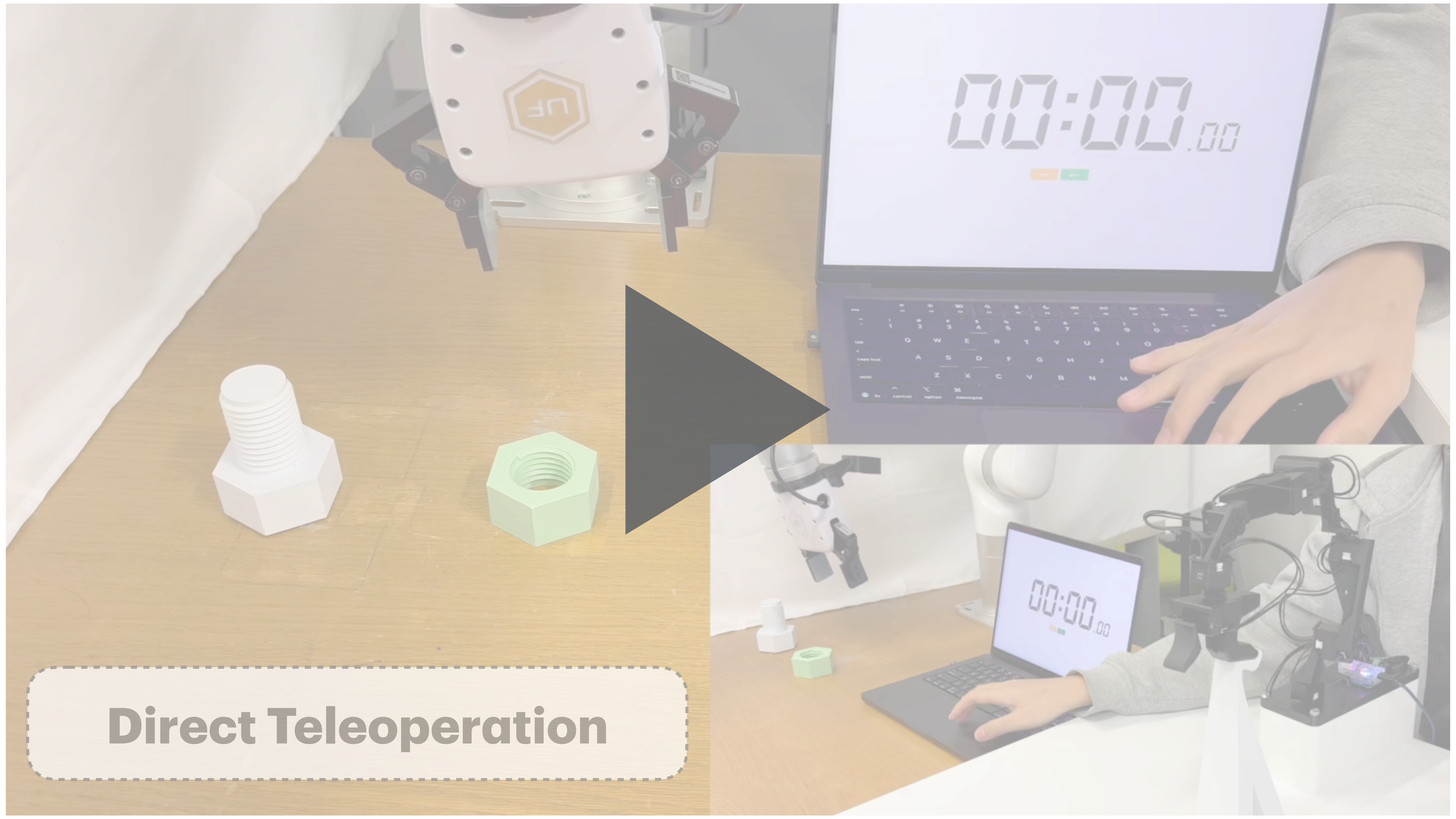
Train residual policy with sparse task rewards and regularization costs.

Step 3: Deploying the SA policy with Real Operators

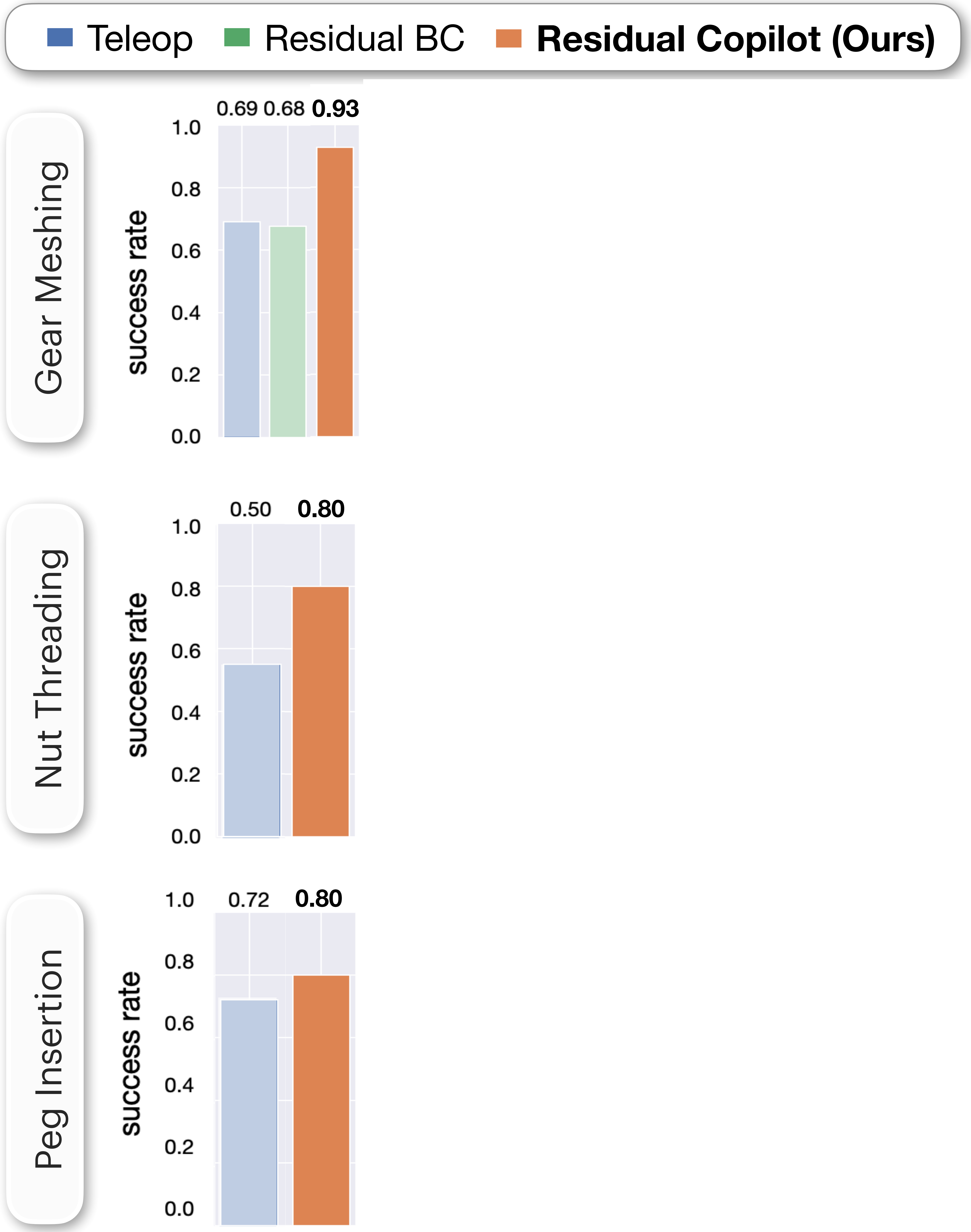


Residual Copilot (Ours) vs Direct Teleoperation



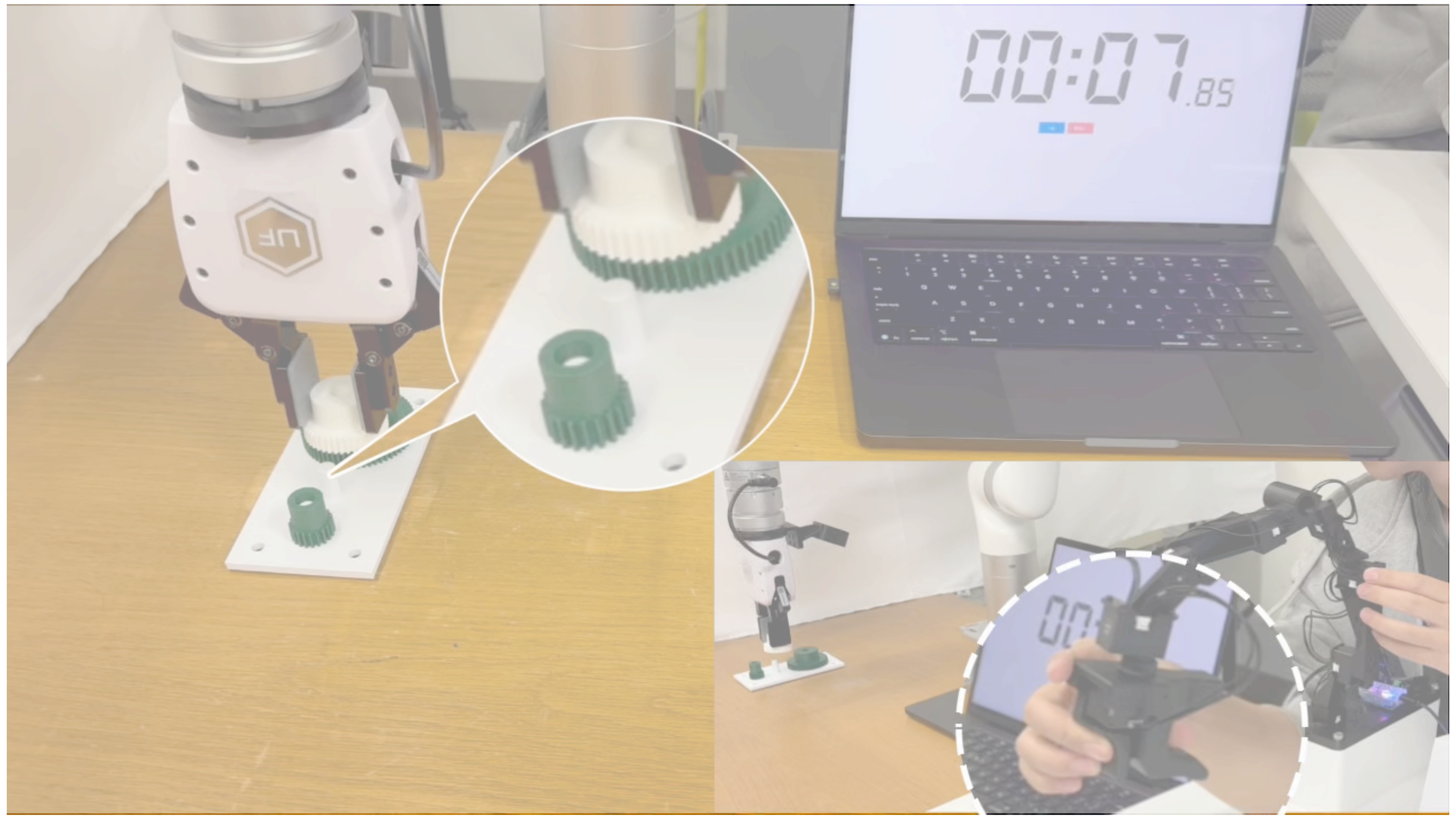


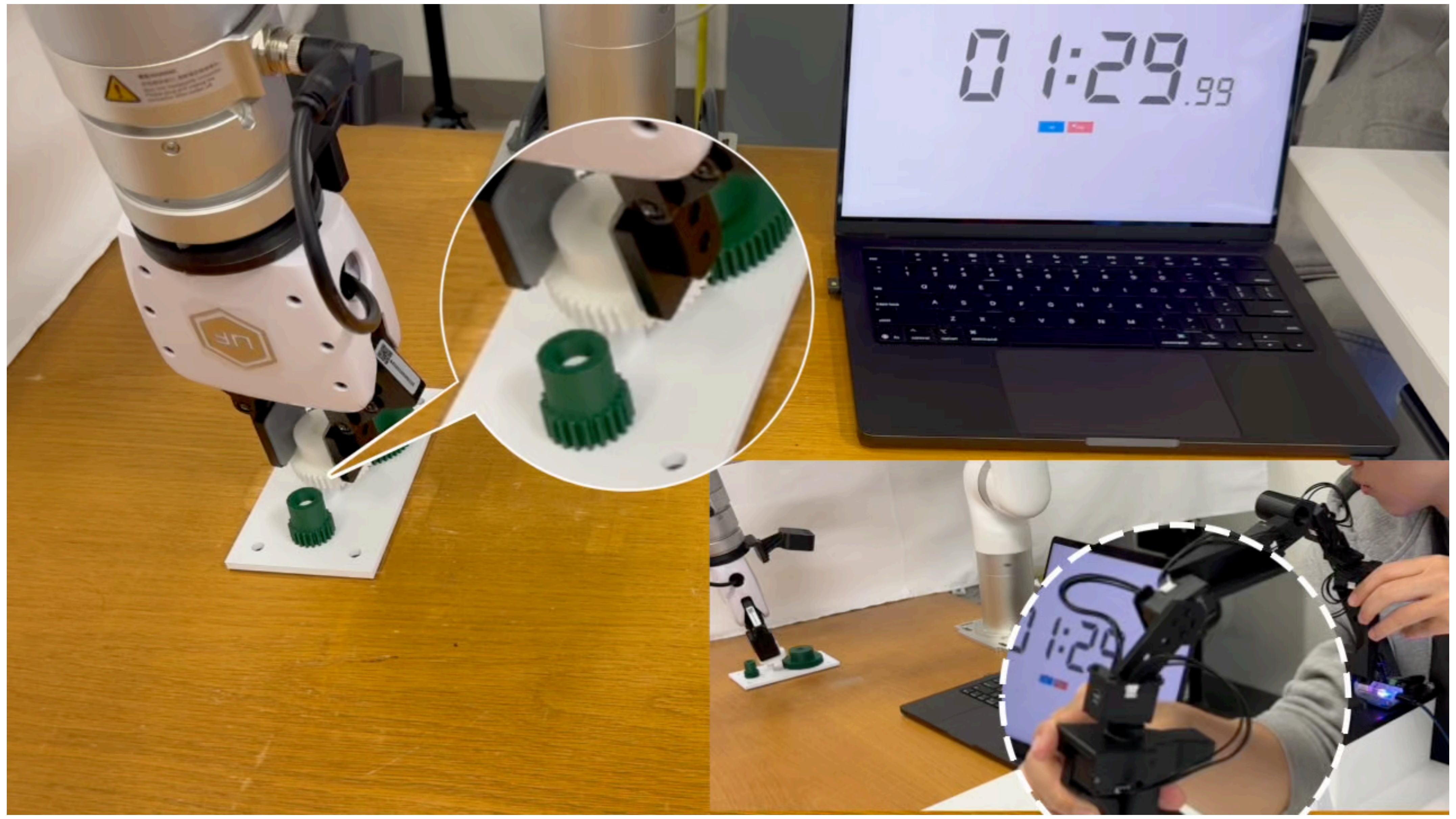
(All videos 20x speed)


Teleop
Residual BC
Residual Copilot (Ours)
Gear Meshing
success rate
0.69 0.68 0.93
Nut Threading
success rate
0.50 0.80
Peg Insertion
success rate
0.72 0.80

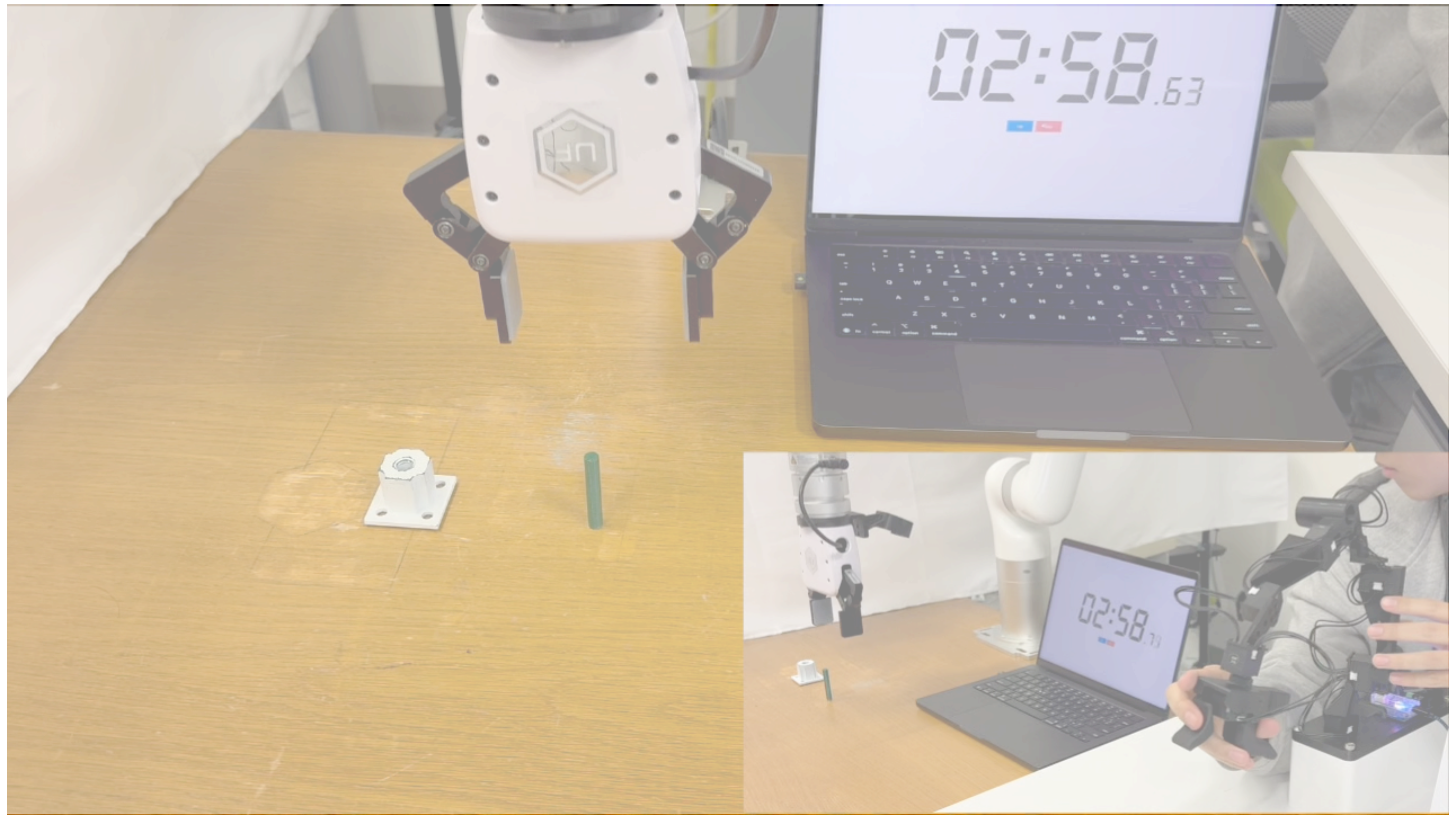
Residual Copilot (Ours) | **Direct Teleoperation** | *(All videos 0.5x speed)*

Gear Meshing

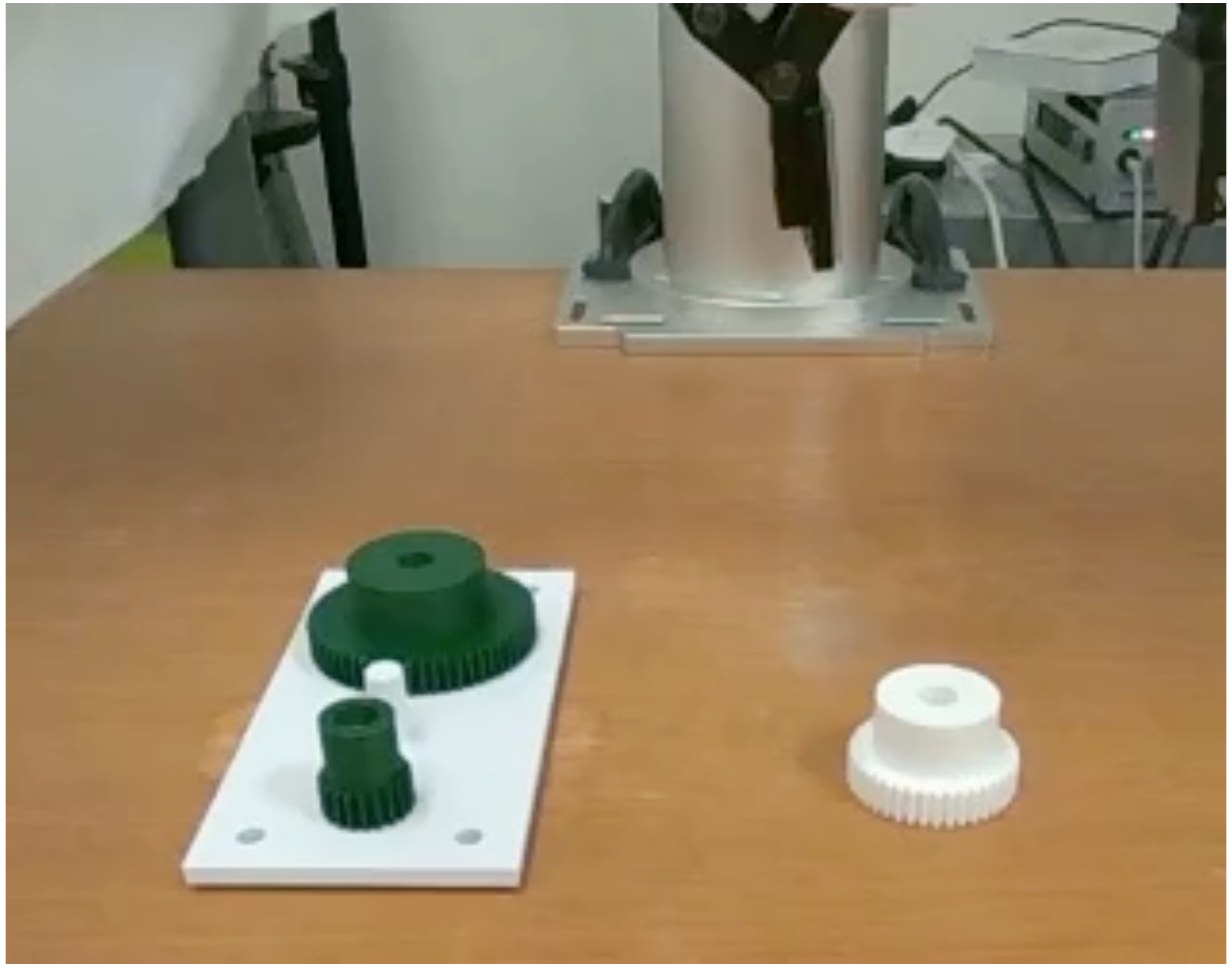




Peg Insertion

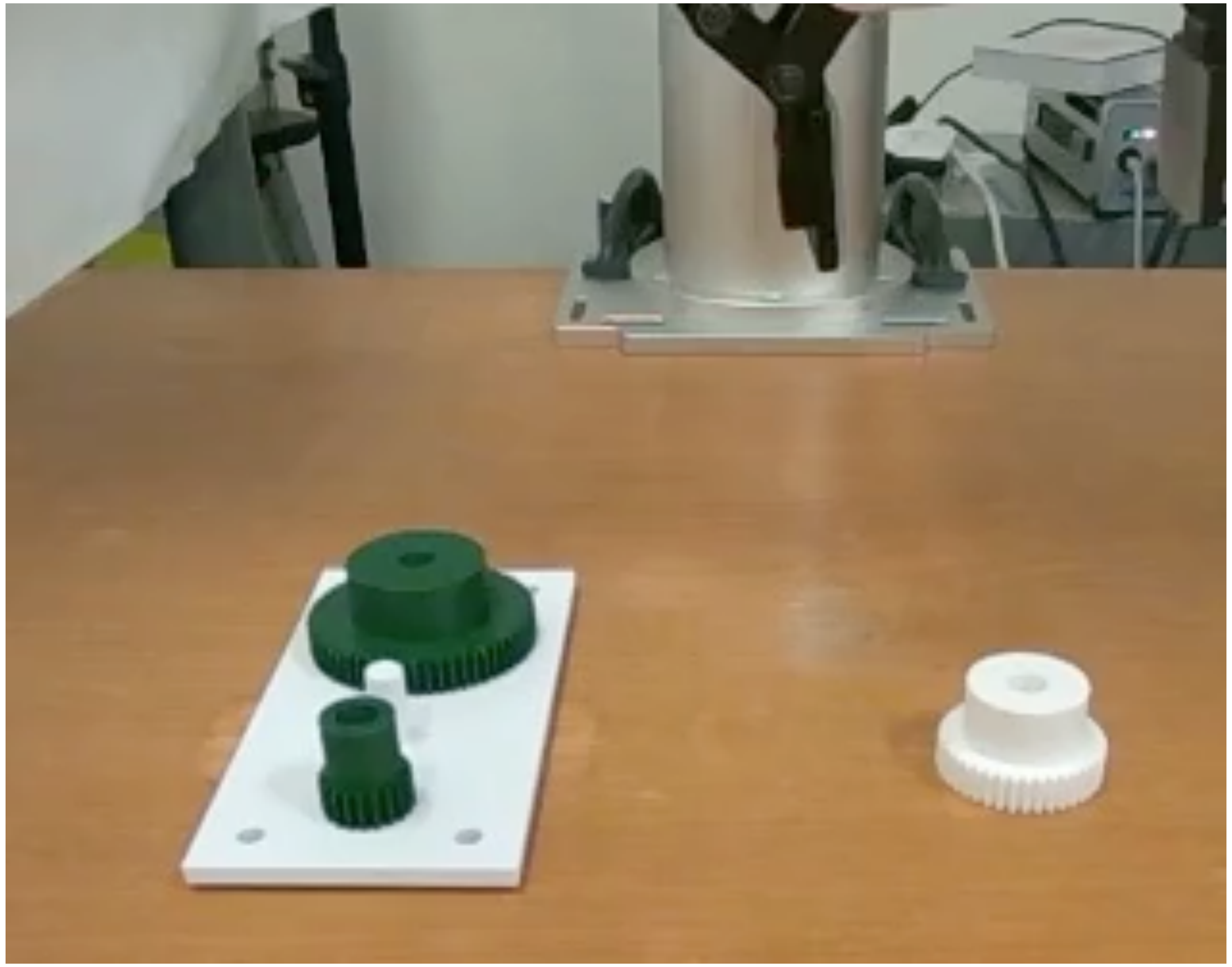


DP trained with **copilot-assisted** teleoperation data



Overall performance:
Grasp: **19/20** Insert: **9/20**

DP trained with **direct** teleoperation data



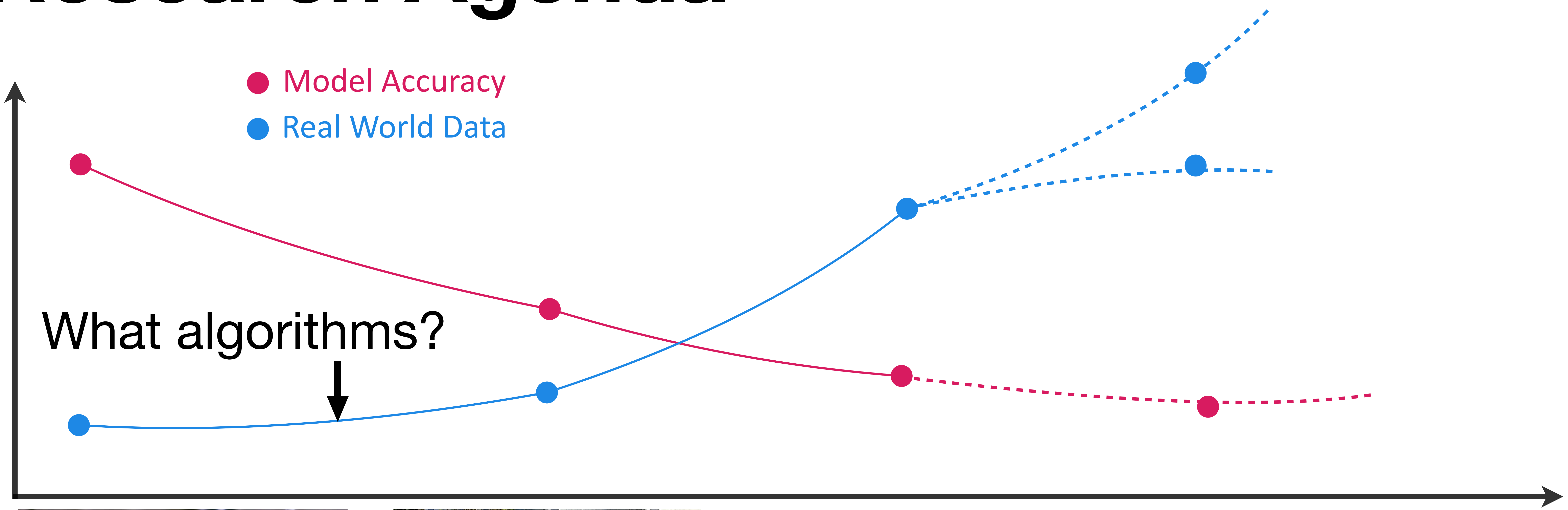
Overall performance:
Grasp: 6/20 Insert: 0/20

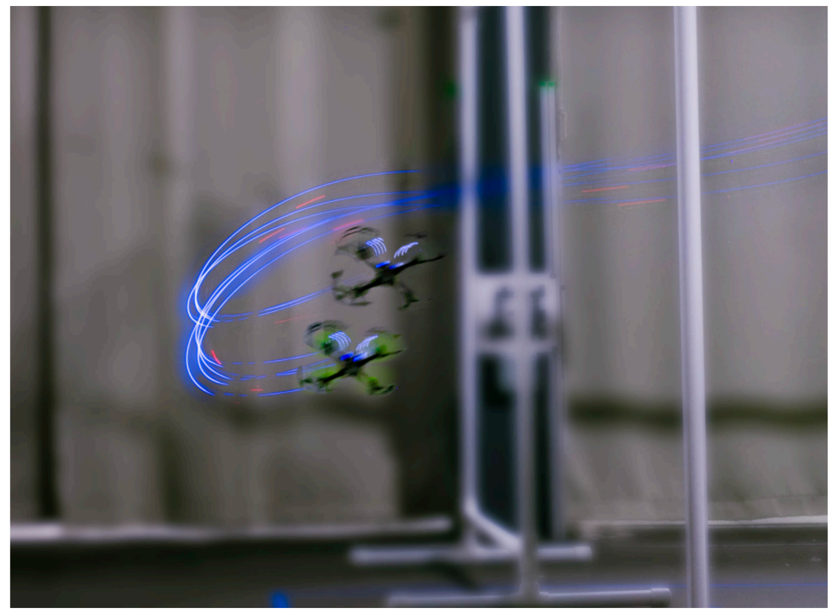
(All videos 2x speed)

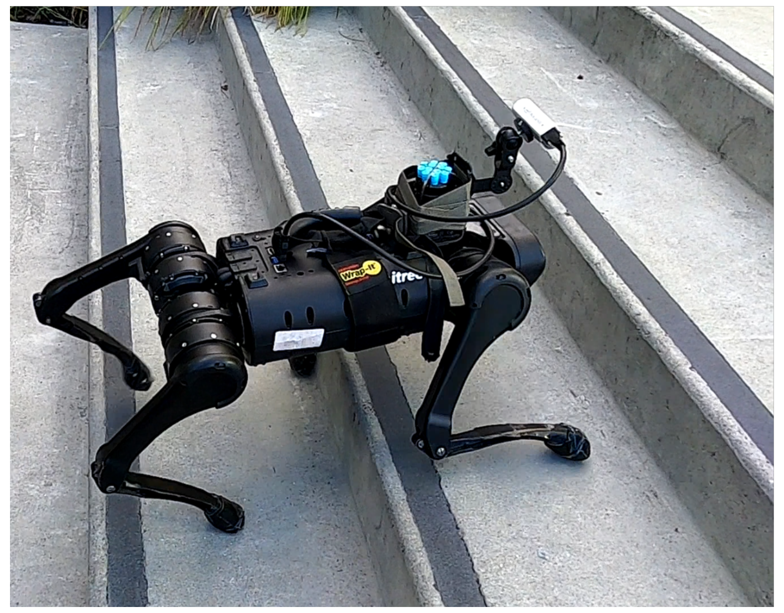
Summary up to now

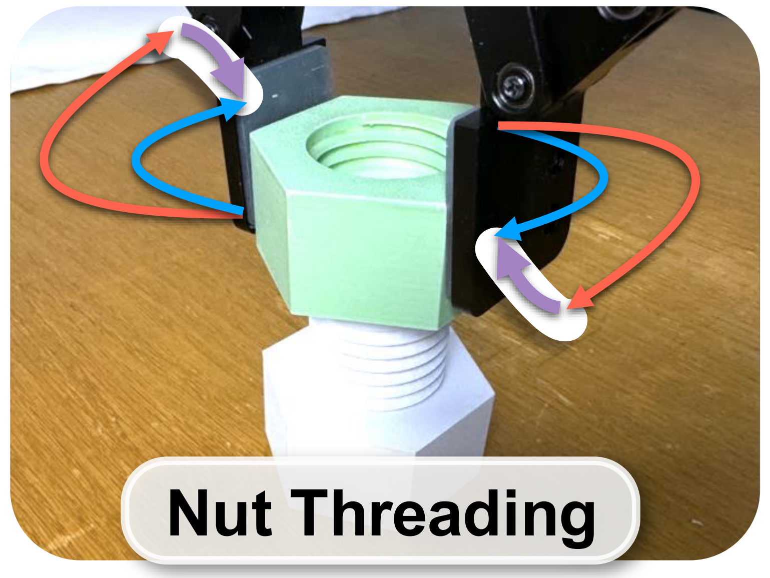
- Imposing structure in the randomization space, together with real-time adaptation, leads to generalization beyond the training environment.
- Adversarial learning empirically decreases reliance on model accuracy.
- There is a coupling between what you can predict and what you need to accomplish a task.
- **Models are designed jointly with the algorithm.**

Research Agenda









Is there a unified paradigm?

What happens at the limit?

Thank you!