EPFL

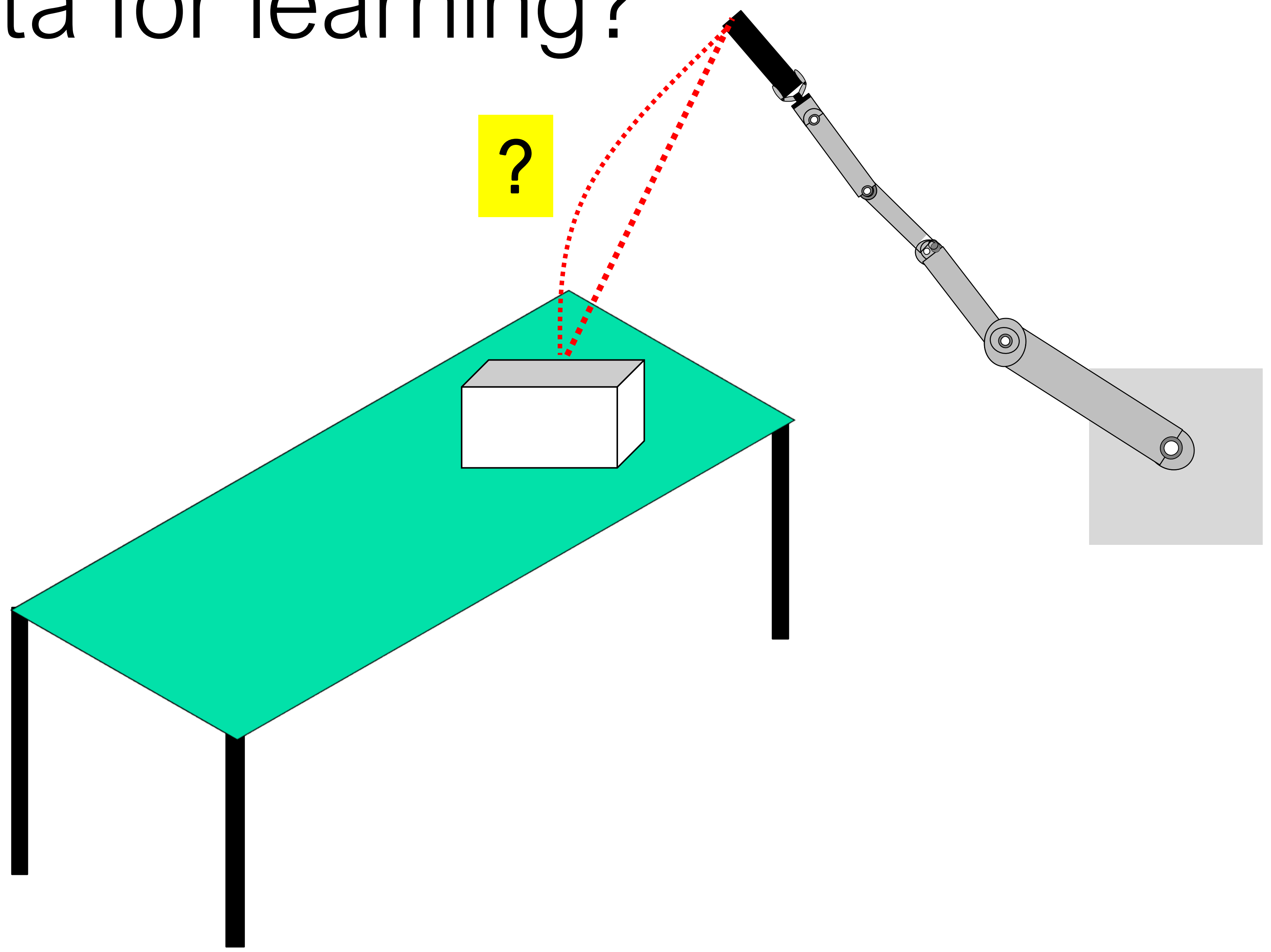
Data-Driven Learning
How to Transmit Skills to Robots

Chapter 2, Book

Motivation

How can we learn optimal controllers to perform a task from data?

- Use data-driven approaches to learn optimal controllers
- How do we gather data for learning?



Main Methods to Train Robots

Reinforcement Learning

Robot learns on its own, by trial and error

Bootstrap the search & reduce search to feasible set of parameters

Simulation

Human

Learning from demonstration

Robot learns by imitating an expert

Expert data

Human

Optimal control

Main Methods to Train Robots

Reinforcement Learning

Robot learns on its own, by trial and error

Bootstrap the search & reduce search to feasible set of parameters

Human

Learning from demonstration

Robot learns by imitating an expert

Expert data

Human

DIFFICULTIES: Humans and robots bodies differ - need proper interfaces;
Requires a knowledgeable human

Learning from Demonstration

Learning from demonstration

Robot learns by imitating an expert

Optimal Control

Solutions found by solving an **optimal control** problem can be used as **expert demonstrations.**

Search for paths that are optimal and **feasible.**

Generate **many solutions** (warm start with different initial conditions, exploit non convexity of the problem).

See matlab exercises today

Programming by Demonstration

Learning from Demonstration

Initial Approaches

- Programming by Demonstration
  - Started in the 1980s
  - Primarily used teleoperation to provide demonstrations to the robot
  - Demonstrations consisted of position and orientation that robot would track

Programming By Demonstration

- Teleoperation
  - Manually move robot through motions for task
  - Generate Motion Primitives to enable task segmentation
- Symbolic Reasoning
  - Generate state-action-state sequences to represent task
  - Use ”if-then” rules to construct symbolic task representation
  - Originally these symbolic representations were defined as prior knowledge to the system, not learned

More Recent Approaches

- Multitude of Algorithms for Learning from Demonstration (LfD) with multitude of applications
- From learning simple trajectories to learning sequences of tasks
- How we gather data from humans revolves around three main themes:
  - **Teleoperation**: user controls the robot through interface
  - **Kinesthetic Teaching**: user physically moves the robot
  - **Observational learning**: robot learns from visual observation of demonstration

Teleoperation – as a mean to teach robots

Learning from Demonstration: Tele-operation

- Users control robots using some interface to perform task
- Demonstrations are used in LfD algorithm
- The quality of learning and performance is sensitive to:
  - Interface design
  - Teacher experience

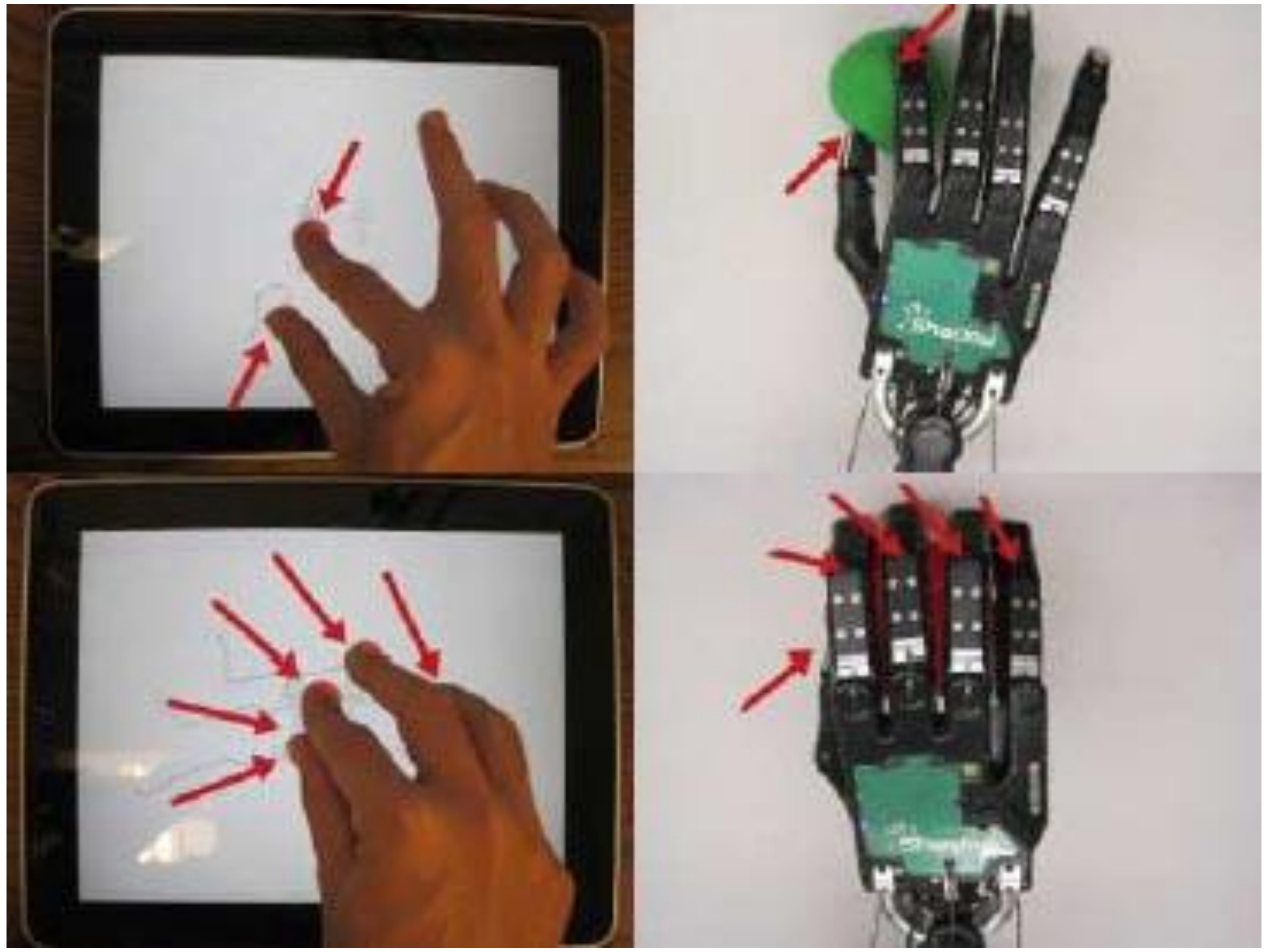
Learning from Demonstration: Tele-operation

Graphical user interface/Tablet

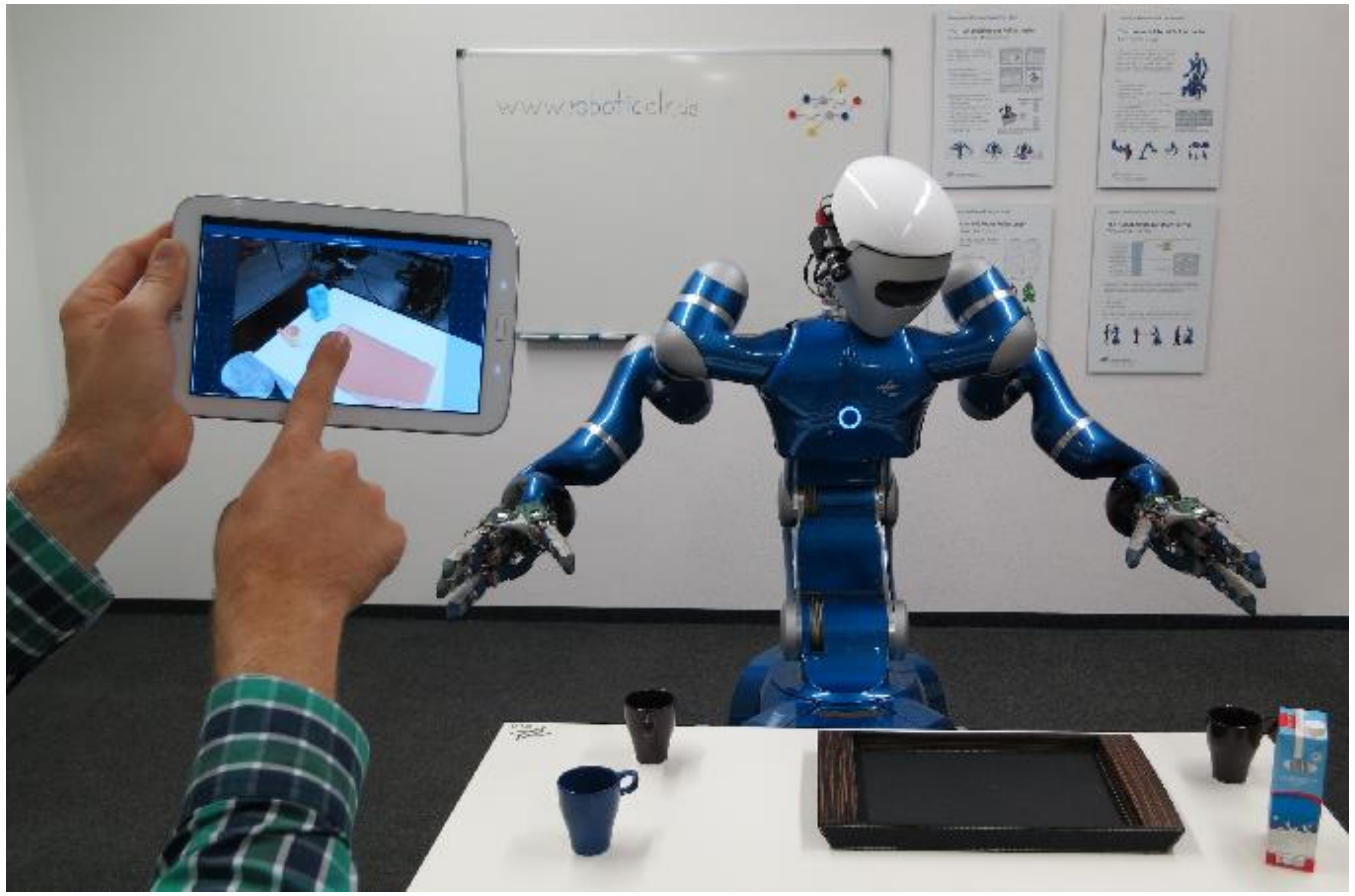
Can communicate the desired motion by mimicking the motion on the tablet, or desired target

+ User-friendly, cheap, easily integrable to existing set-up

- Limited in what can be transferred (displacement in a plane, reduced field of motion)



Dexterous Telemanipulation With a Multi-Touch Interface. Toh et al.
http://graphics.cs.cmu.edu/?p=223



A Knowledge-Driven Shared Autonomy Human-Robot Interface for Tablet Computers. Birkenkampf et al.
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7041352

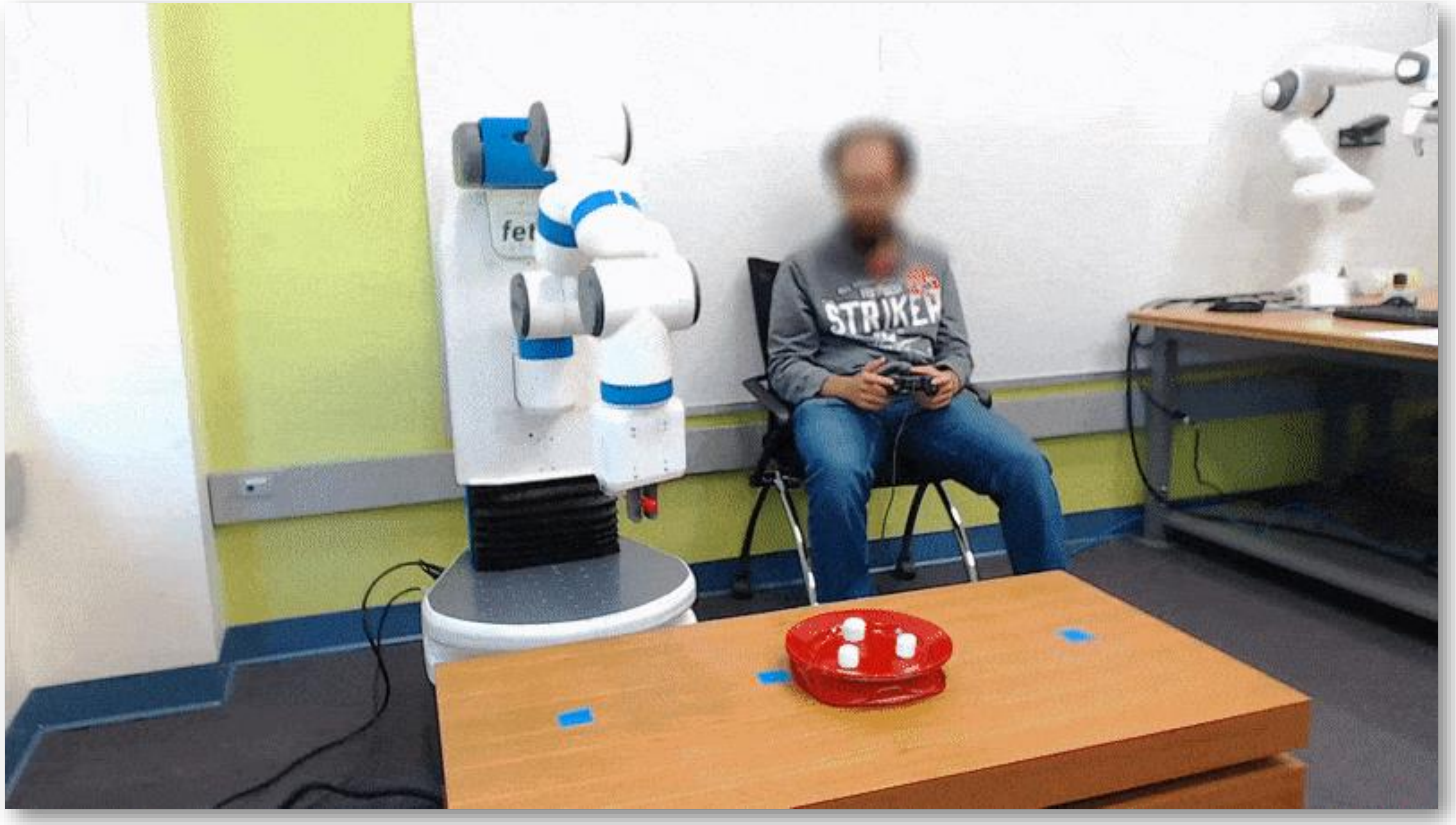
Learning from Demonstration: Tele-operation

Joysticks: control the robot’s end-effector through a 3DOFs or 6DOFs joystick

+ Can communicate the desired motion in 3D or event 6D, easily amenable to control 6DOFs robot arm in position and orientation

+ User-friendly, cheap, easily integrable for a vast range of applications

- Requires often the use of the two hands
- Limited in what can be transferred (displacement and speed, not forces)



Video Source: https://iliad.stanford.edu/research/interactions

Losey, Dylan P., et al. "Controlling assistive robots with learned latent actions." 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020.

Learning from Demonstration: Tele-operation

Joysticks: control the robot's end-effector through a 3DOFs or 6DOFs joystick

+ Can communicate the desired motion in 3D or event 6D, easily amenable to control robots that differ from humans (here drones) - control in position and orientation

- Sensitive to experience of teacher

A. Ng, A. Coates, M. Diel, V. Ganapathi, J. Schulte, B. Tse, E. Berger, E. Liang, Inverted autonomous helicopter flight via reinforcement learning, in: International Symposium on Experimental Robotics, 2004

Learning from Demonstration: Tele-operation

Exoskeletons

+ Conveys directly the dynamics of the motion (embodied transmission)
+ Allows to control all joints as well as movement in Cartesian space

- Heavy, cumbersome
- Does not fit all sizes and strengths
- Does not convey nor render forces at contact

Capio Upper Body Exoskeleton for Teleoperation by the DFKI GmbH Robotics Innovation Center.
https://robotik.dfki-bremen.de/en/research/projects/capio.html

Learning from Demonstration: Tele-operation

Telepresence

+ Enables to perform tasks when remote

+ Can offer a more efficient and more secure mean to intervene (in disastrous or dangerous environments)

- Need to provide good visual rendering of the scene

- Delays in transmission can be detrimental if task is too rapid

- Force not always (well) rendered

Davinci Surgical Robot

Learning from Demonstration: Tele-operation

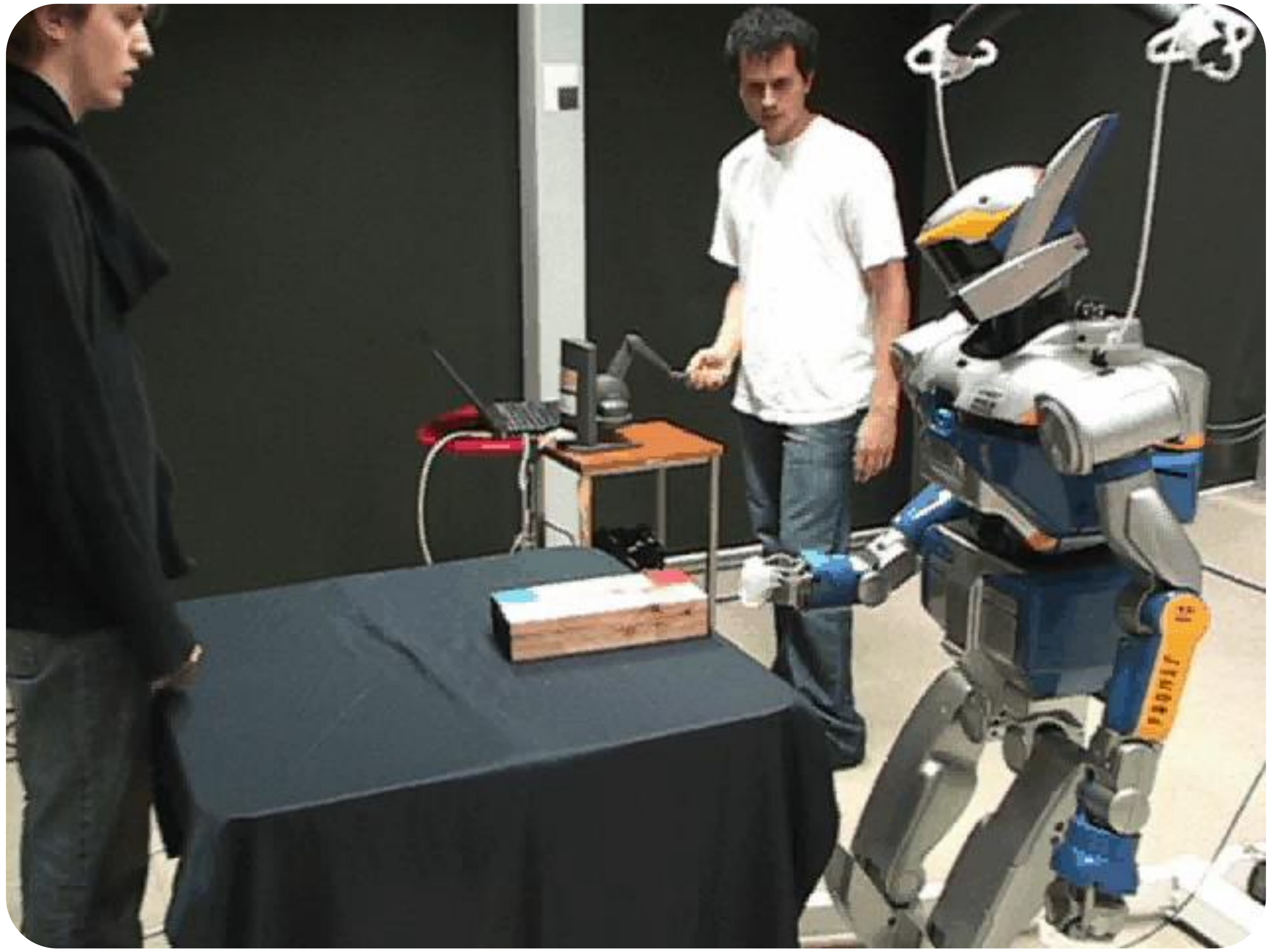
Haptic interfaces

+ Transmit the forces applied by the user

+ Renders the forces perceived by the robot to the user

- Close-loop system induces delays of 100-400ms depending on distance and medium used

- Delays may lead to incorrect response and instable behavior



LASA Lab / EPF in collaboration with JRL / TsukubaL

Learning from Demonstration: Tele-operation

Haptice Devices & Teleoperation

New finger-based haptic devices leverage on sense of touch mounted on robots' fingers to perceive contact and render these through hand exoskeleton

+ offer higher resolution

+ Closer to human touch

- Covers only fingertips, but sense of touch is all along fingers, palm, top of the hand

- Suffers from drifts, calibration can be an issue

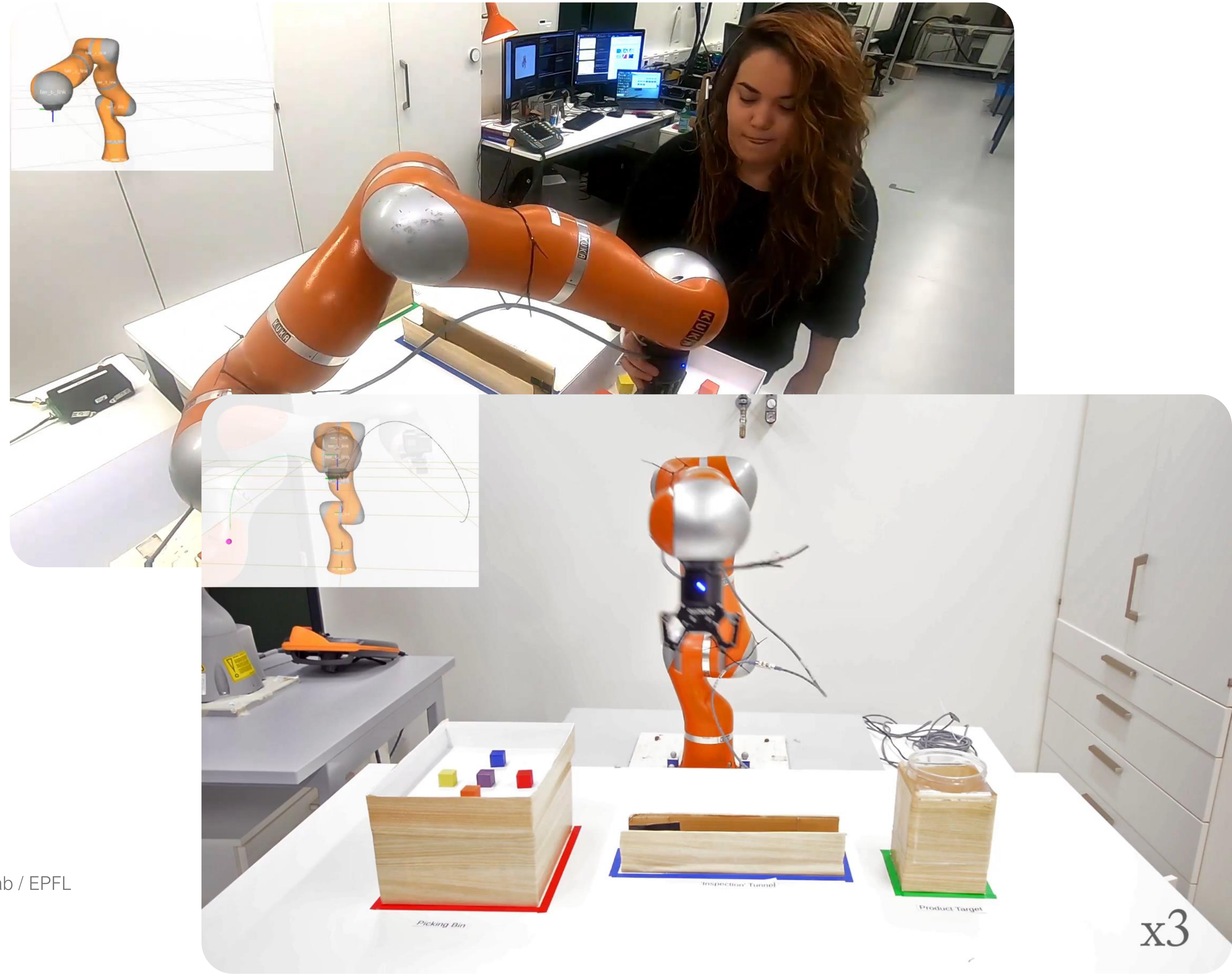
HaptX haptic glove integrated with Shadow Robot hand
Footage from Adam Savage's Tested + ShadowRobot Company and Syntouch
https://www.youtube.com/watch?v=rEoq7DMgaEc&t=24s

EPFL

Learning from Demonstration: Kinesthetic Teaching

- Teacher physically moves robot

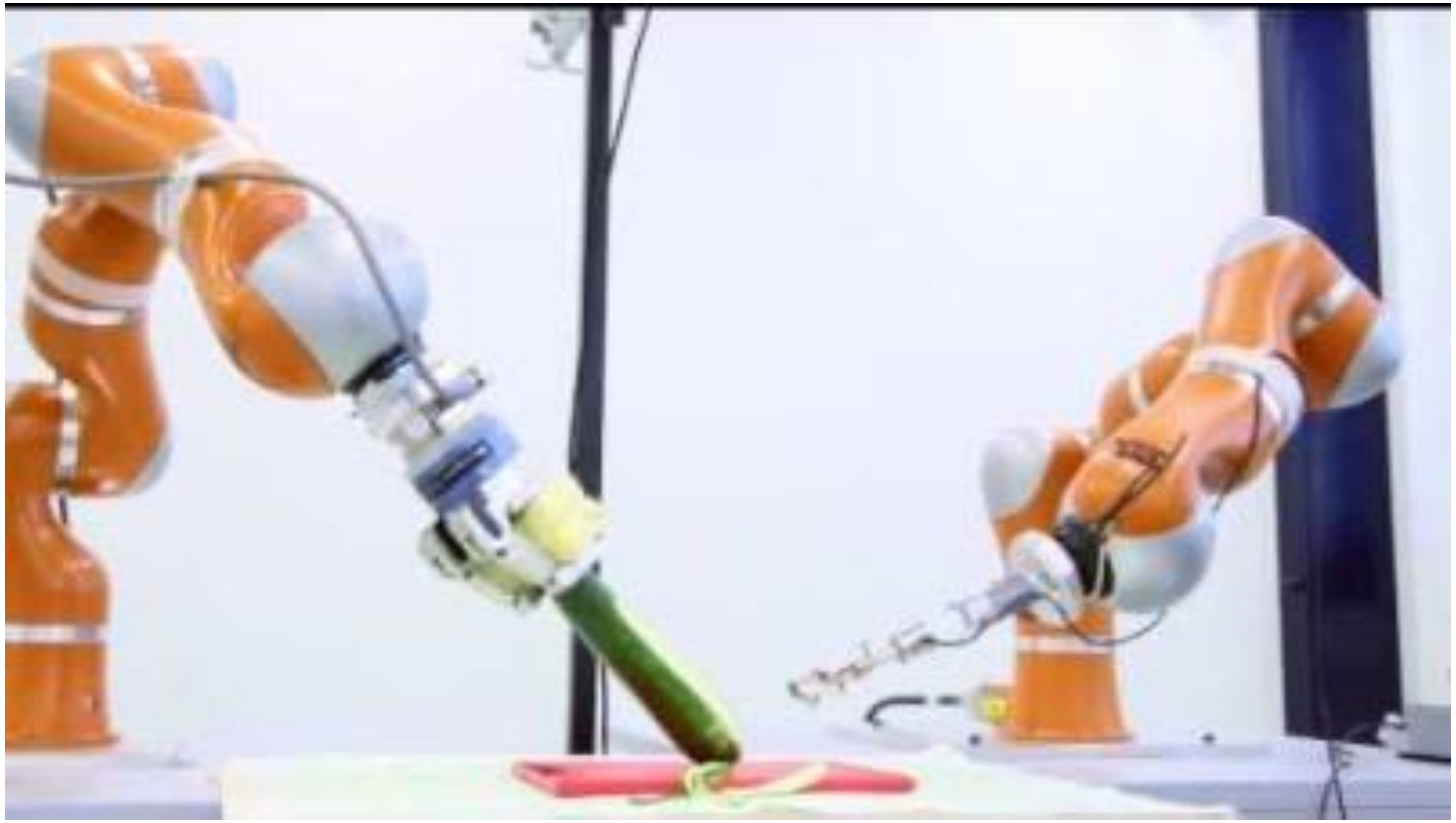
+ Direct control motion of robot
+ Can transmit forces
+ Can perceive forces at contact



LASA Lab / EPFL

Learning from Demonstration: Kinesthetic Teaching

- Cumbersome
- Limited in the number of joints / limbs one can move at once



LASA Lab / EPFL

Video Source: https://www.youtube.com/watch?v=xIK6U52TjRM
Learning by Demonstration, Bimanual Coordinated Task, LASA

Observational Learning: Vision Systems

Use RGB-D camera, automatic reconstruction of body motion

+ Enable users to perform task in a natural manner
+ Can be extended to analysing any videos of human motion, not necessarily videos of human teaching a robot

- Can be slow (live analysis of camera image <50Hz)
- Can be imprecise, especially in face of large occlusion (forces user to face cameras)

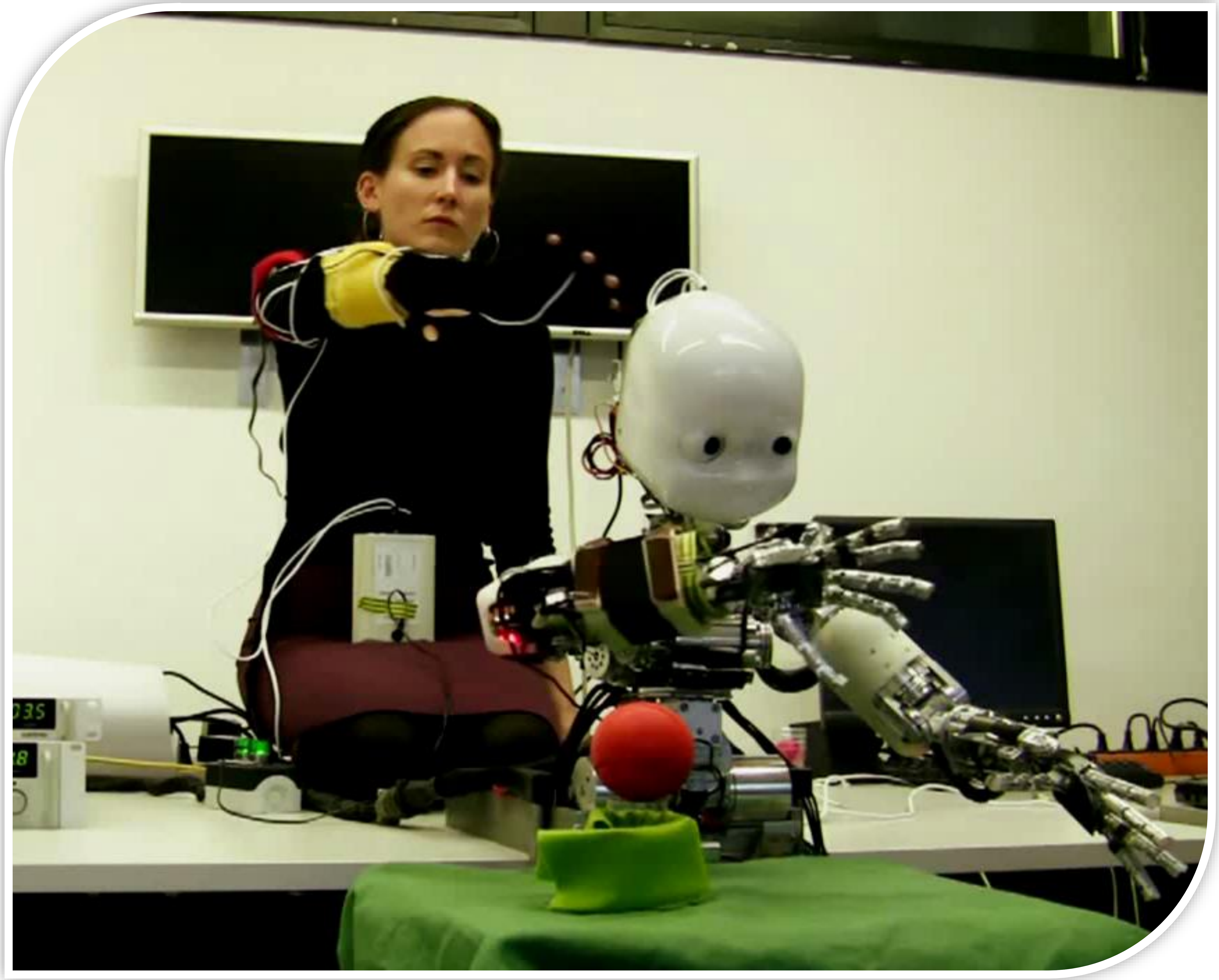
DexPilot: Vision Based Teleoperation of Dexterous Robotic Hand-Arm System
Handa et al. ICRA 2020

Observational Learning: Motion Capture System

Markers attached to body parts are tracked by infra-red cameras at high resolution

+ Conveys directly the dynamics of the motion (embodied transmission)
+ Allows to control all joints including hand movements

- Requires a set-up with several high resolution cameras
- Suits does not fit all sizes and strengths
- Does not convey nor render forces at contact



LASA Lab / EPFL

Observational Learning: Motion Capture System

Video Source: https://youtu.be/ggLge1Rw2z4?t=77

C. Stanton, A. Bogdanovych, E. Ratanasena: Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning. In proceedings of Australasian Conference on Robotics and Automation (ACRA 2012), Wellington, New Zealand, 3-5 December 2012.

Observational Learning: Motion Capture System

Video Source: https://youtu.be/LM4rDfW8-TU
HAL Robotics

Reinforcement Learning

Reinforcement Learning & Learning from Demonstration

- Teach the robot how to do it through demonstrations or let the robot learn on its own

Reinforcement Learning

Let a robot explore on its own and learn an optimal controller through trial and error

Challenges:

- Large amounts of data
- Time and energy to collect experimental data
- Safety during learning
- Typically use simulation: sim2real reality gap
- **Defining reward function**

Google Robotics Arm Farm

4 robotic arms, 800,000 grasp attempts

Reinforcement Learning

Learning in simulation first

Challenges:

- Large amounts of data
- Time and energy to collect experimental data (even in simulation)
- Requires very accurate simulation of the physics of the world

OpenAI

44 CPU cores and 8 GPUs to train our policy, collecting about one hundred years o

Reinforcement Learning & Learning from Demonstration

Let a robot explore on its own and learn an optimal controller through trial and error

Kober and Peters, 2008

Reinforcement Learning

Robot learns on its own, by trial and error

Learning from demonstration

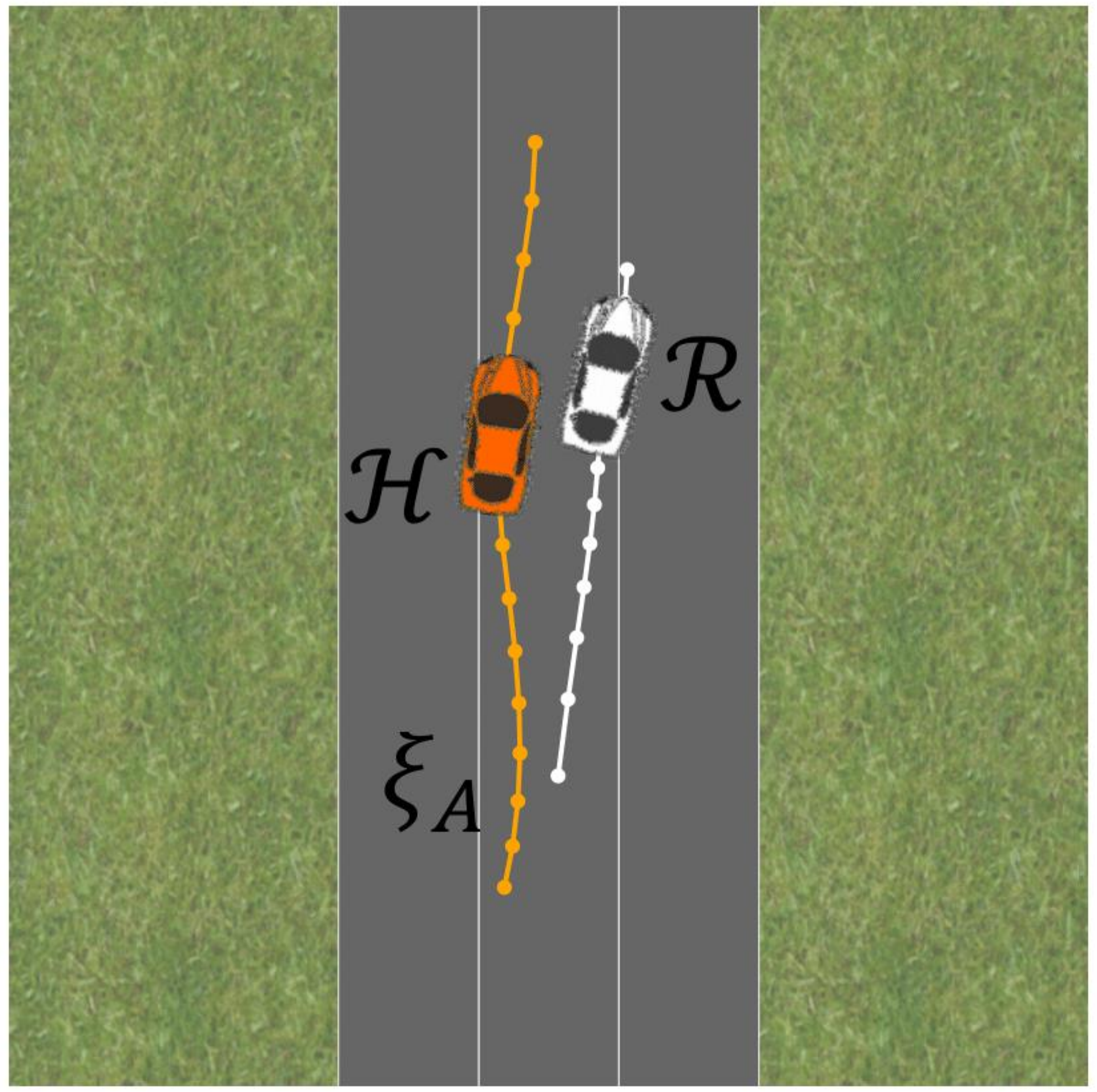
Robot learns by imitating an expert

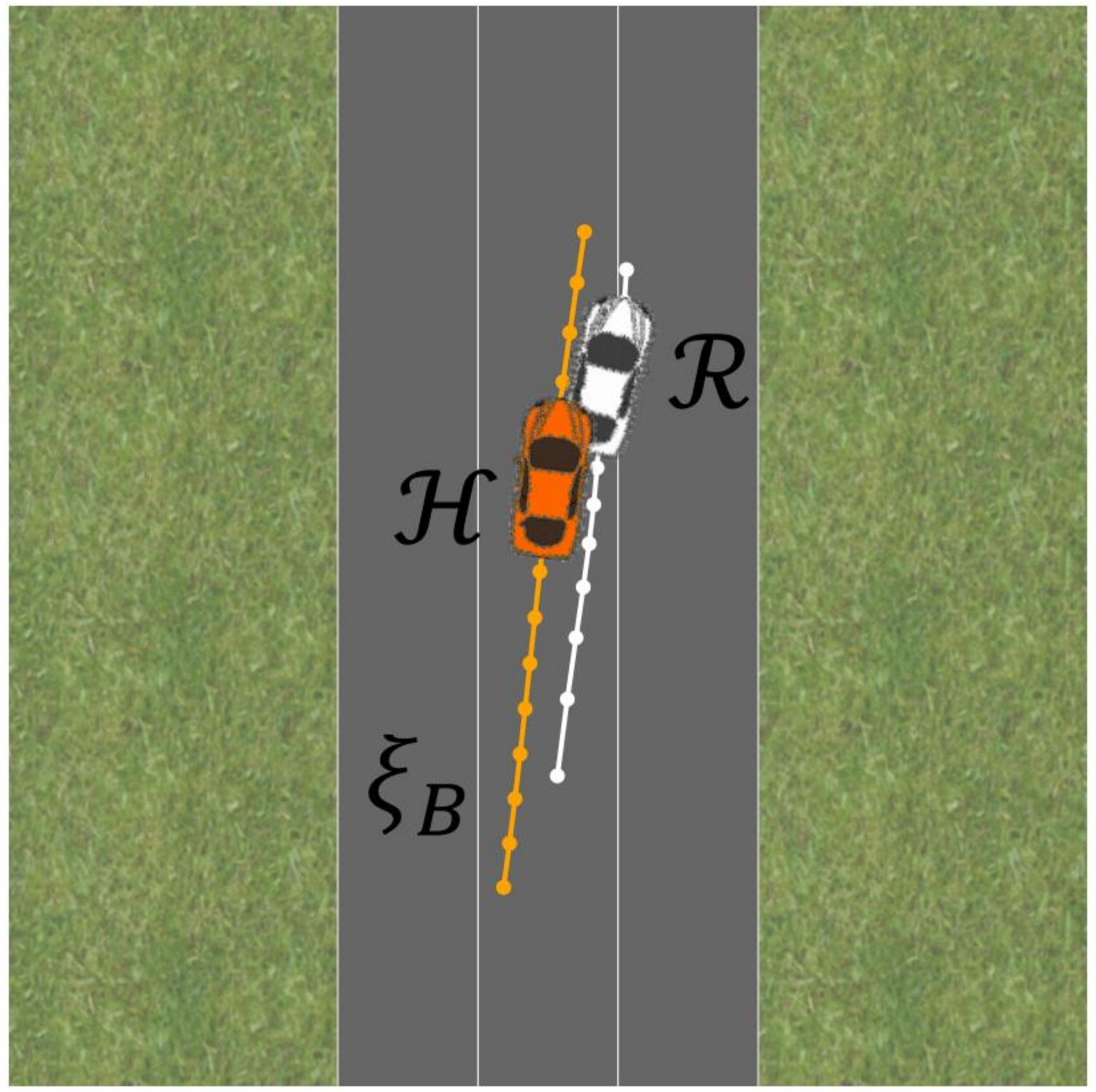
Inverse reinforcement learning

No need to design the reward
Uses human demonstration to guide the search for the reward and the optimal control policy

IRL example : Learning to drive a car

- Gather demonstrations from an expert
- Infer goal/reward function of the environment from the demonstrations





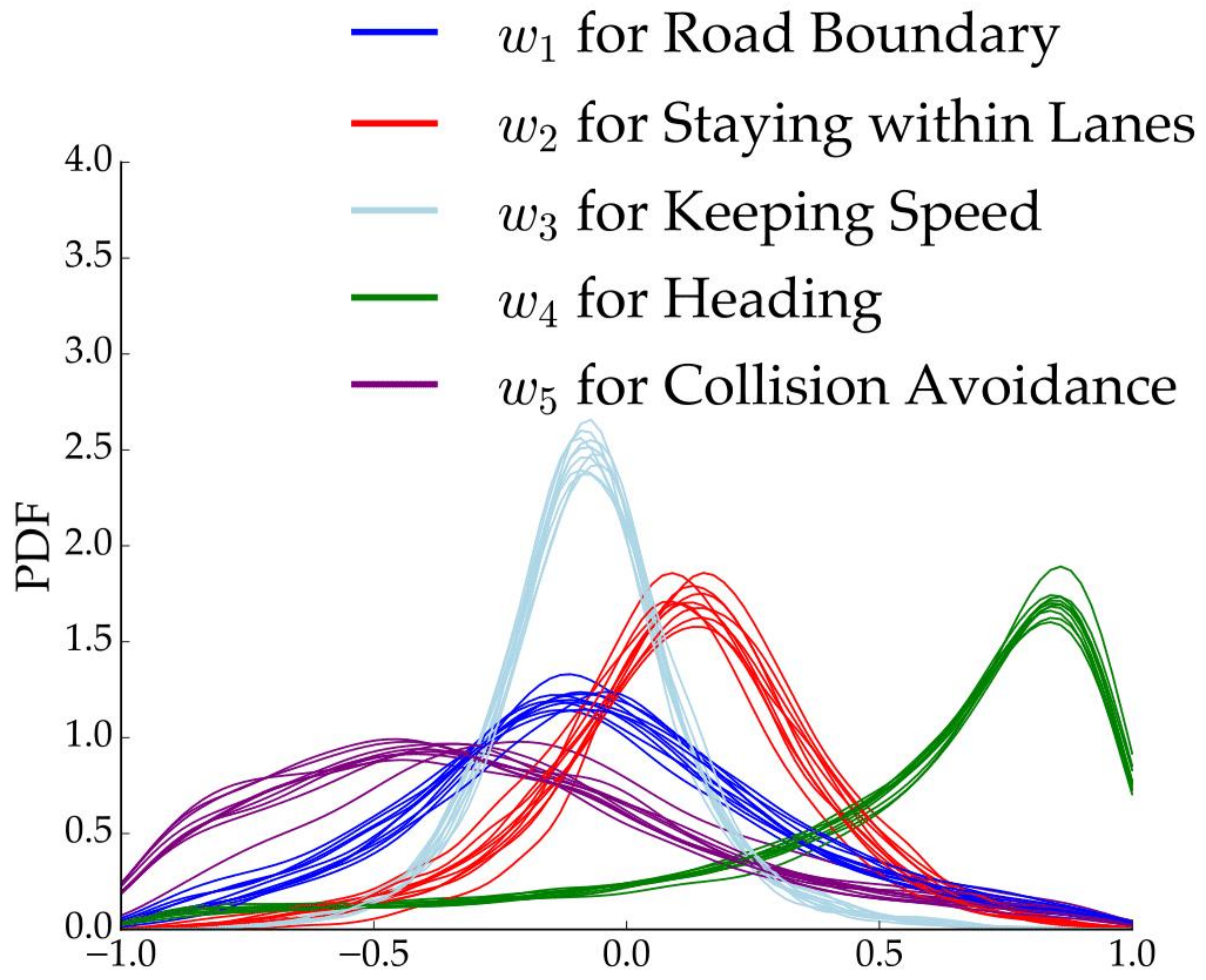
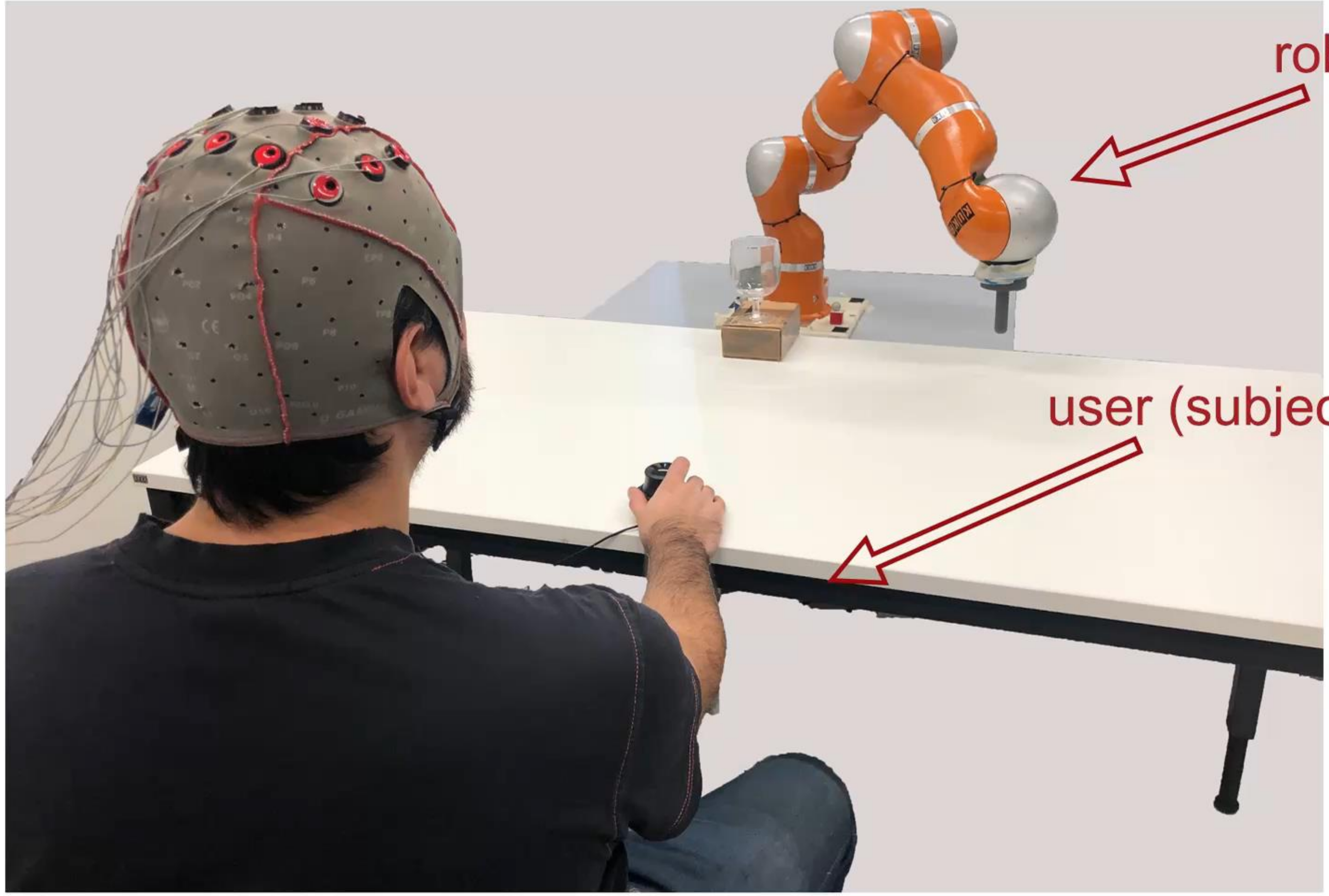


Image Source: https://iliad.stanford.edu/research/humans

IRL example : Inferring User's Preference



Batzianoulis, I., Iwane, F., Wei, S., Correia, C.G.P.R., Chavarriaga, R., Millán, J.D.R. and Billard, A., 2021. Customizing skills for assistive robotic manipulators, an inverse reinforcement learning approach with error-related potentials. *Nature Communications biology*, *4*(1), pp.1-14.

EPFL

How do we gather data for learning?

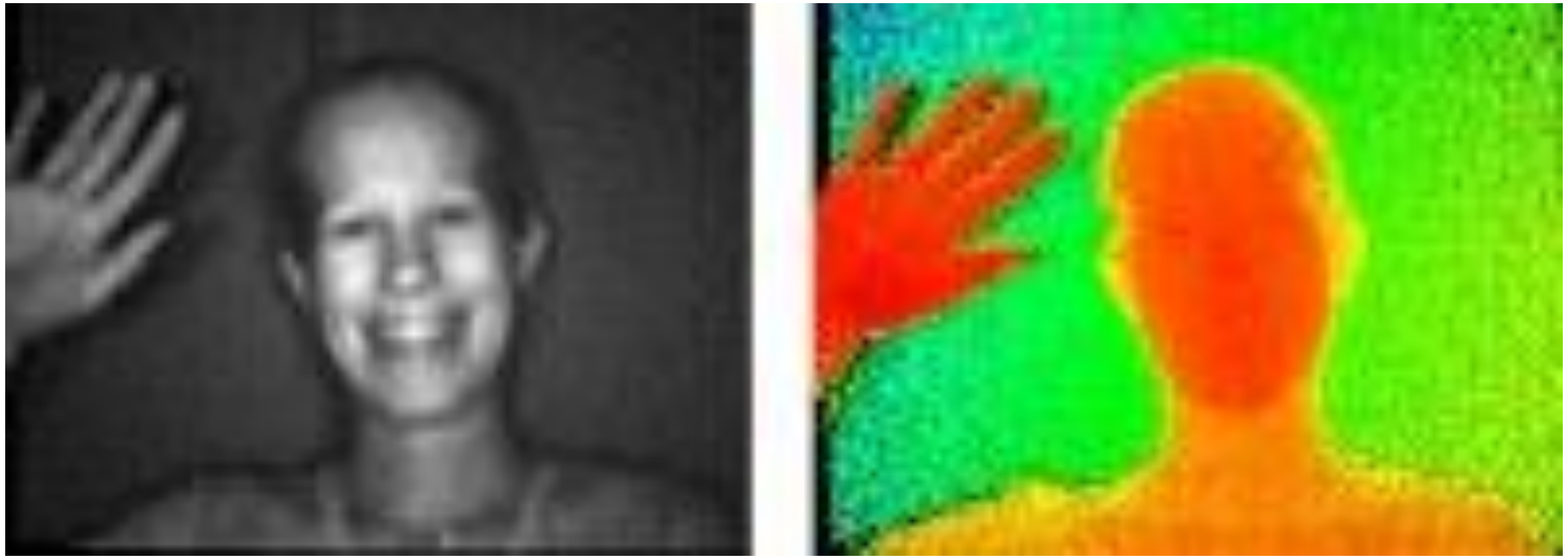
Method to generate the data	Online mode	Need model of robot or world	Trainer	Number of training examples
Learning from human demonstrations	YES	NO	Anyone	<20
Optimal control	NO	YES	Skilled programmer	>100
RL (live)	NO	YES (model-based RL) NO (model-free RL)	Anyone (reward)	>100
RL (simulation)	YES	YES	Skilled programmer	>1,000

Collecting Data for Training Robots: Other challenges

Problem 1: Correspondence Problem

Even when the robot looks more like the human, its body does not have the same range and dynamics of motion.

Problem 1: Correspondence Problem

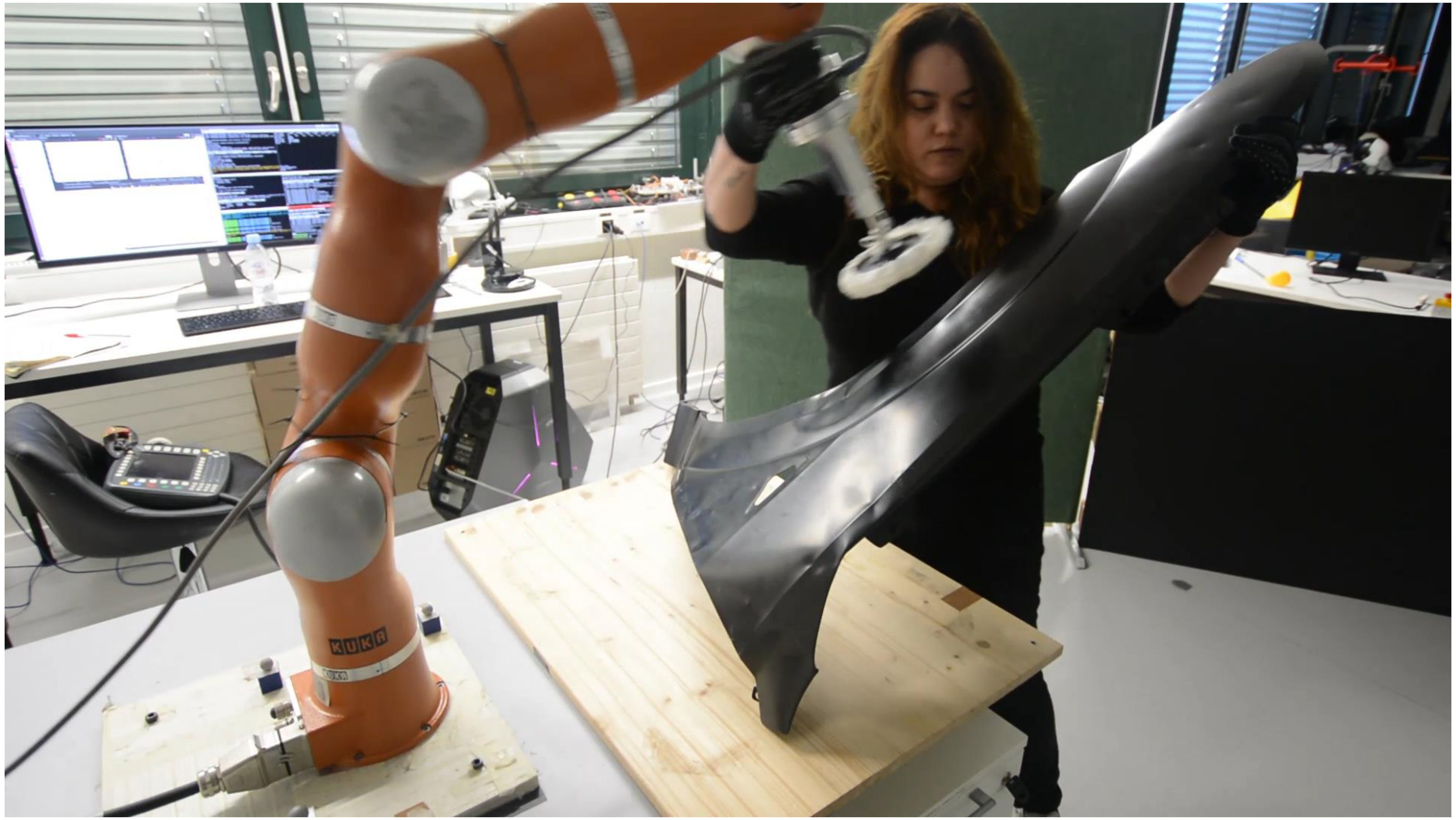


Robots do not perceive things like we do.

Sonars, infrared sensors, lasers are common on robots and easier to process than information from cameras.

Problem 1: Correspondence Problem

- Teachers need to train themselves before training the robots.



Problem 2: Learning is Data-Sensitive

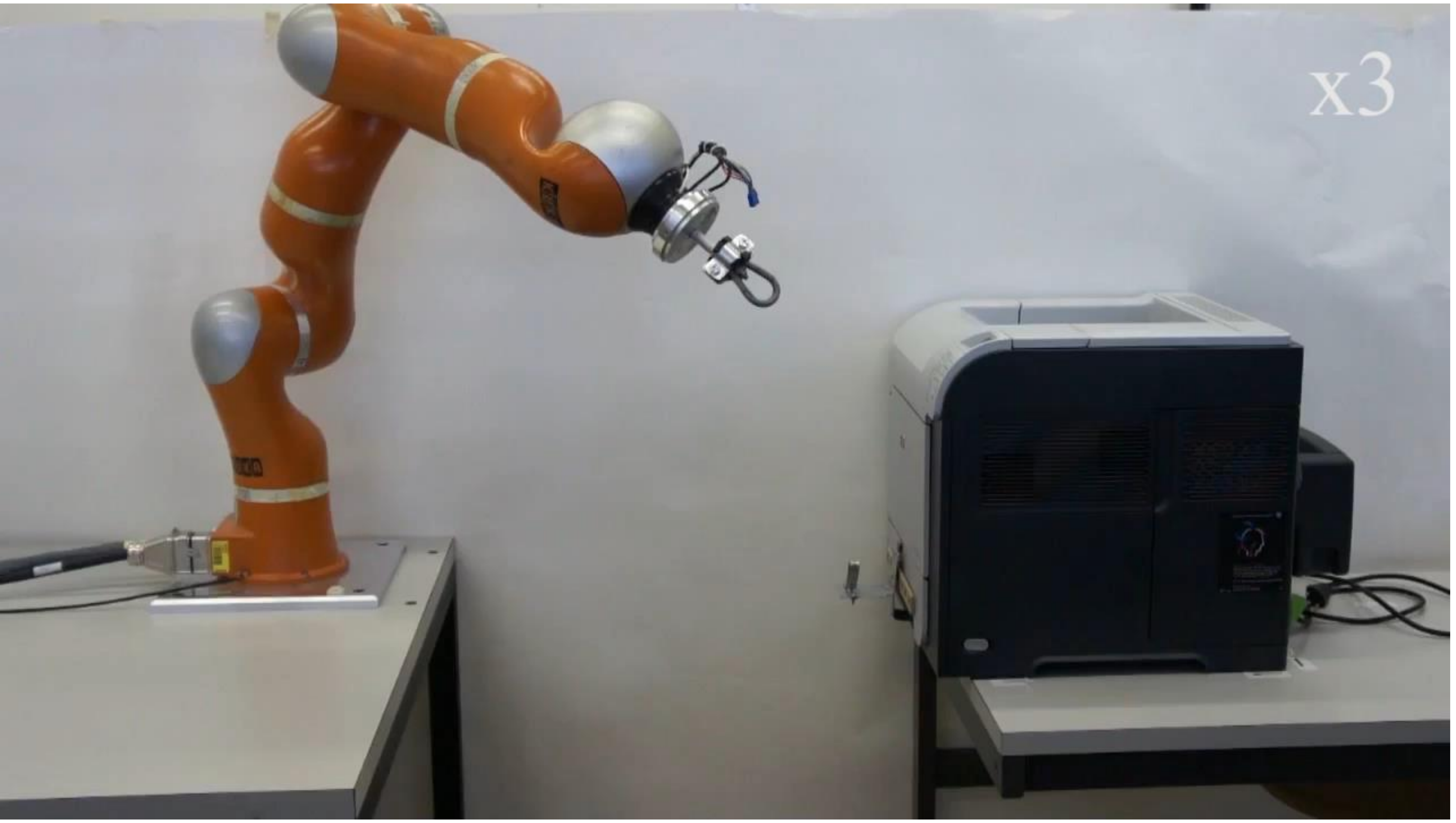
- Teaching a task will differ depending on the kinematics and dynamics of the robot; the same motion in Carthesian space will lead to different trajectories in joint spa

UR5: 6DOF

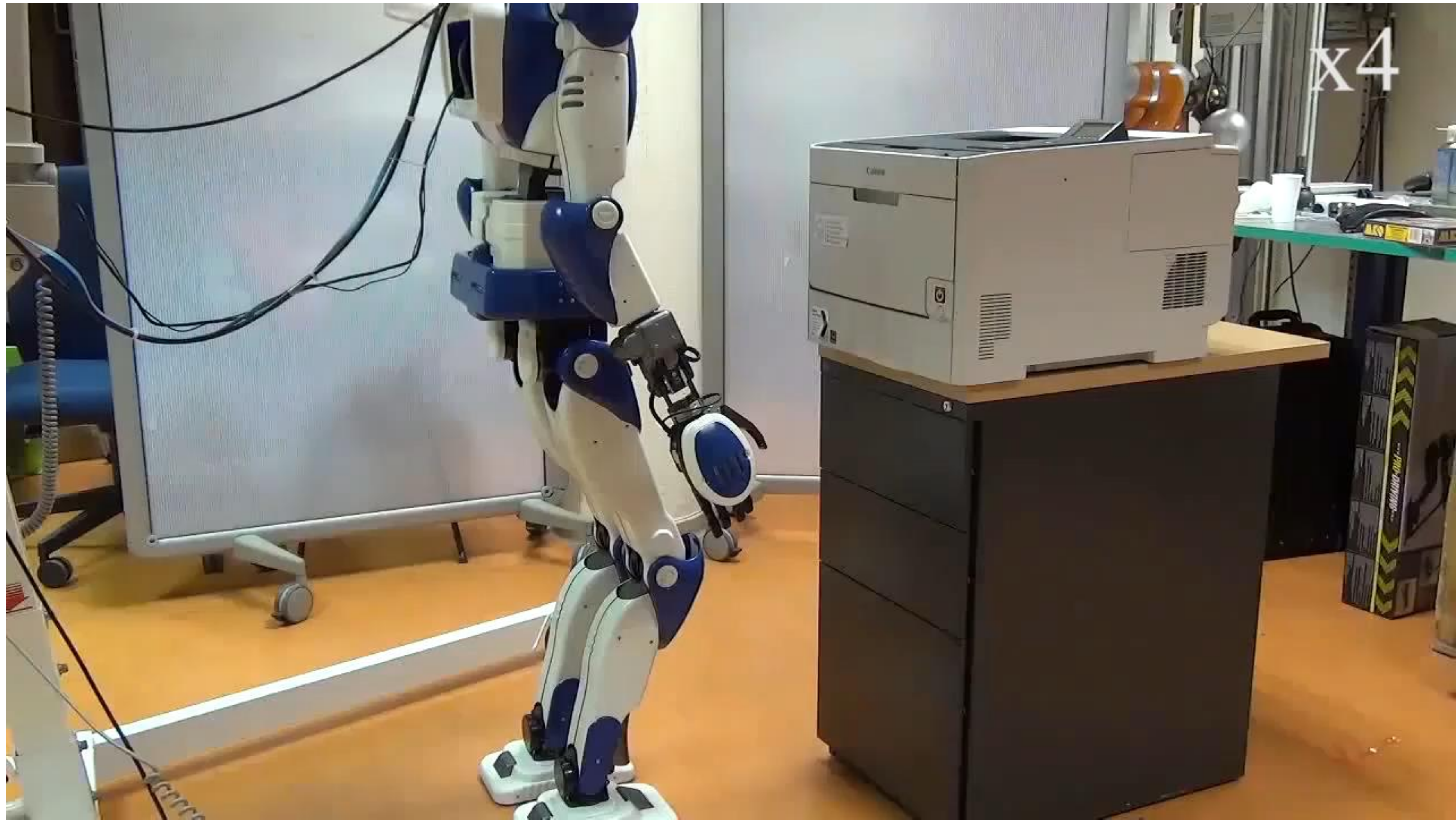
Franka Panda: 7DOF

Problem 2: Learning is Data-Sensitive

- Data is environment-dependent



Model Learned at EPFL



Model transferred at AIST/JRL

Problem 3: Variability in Task Definition

- Question: **What does it mean to perform a task?**
- Multiple ways to accomplish a task:
  - multiple motions

Problem 3: Variability in Task Definition

- Question: **What does it mean to perform a task?**

- Multiple ways to accomplish a task:
  - multiple motions
  - multiple tools

Problem 4: Generalizing Control Law – Beyond the Demonstrations

Infer that the task is composed of sequence of actions; each action is relative to the object the robot must manipulate; but a priori – look at several predefined frames of reference

Learning from Demonstration: Using dynamical systems

EPFL

Goal

- Learn Motion Representations from Task Demonstrations

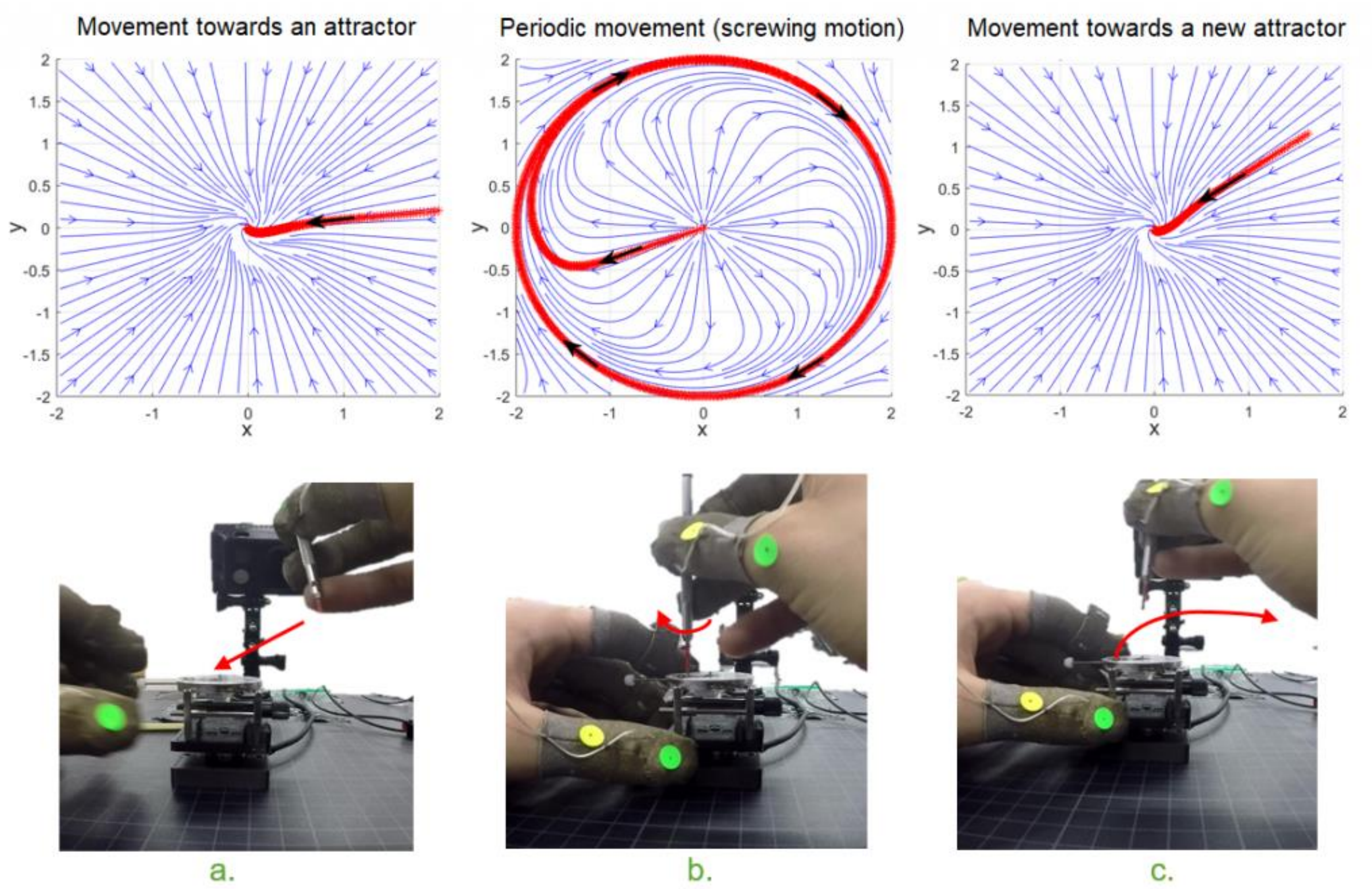


Image source: https://www.epfl.ch/labs/lasa/sahr/research/

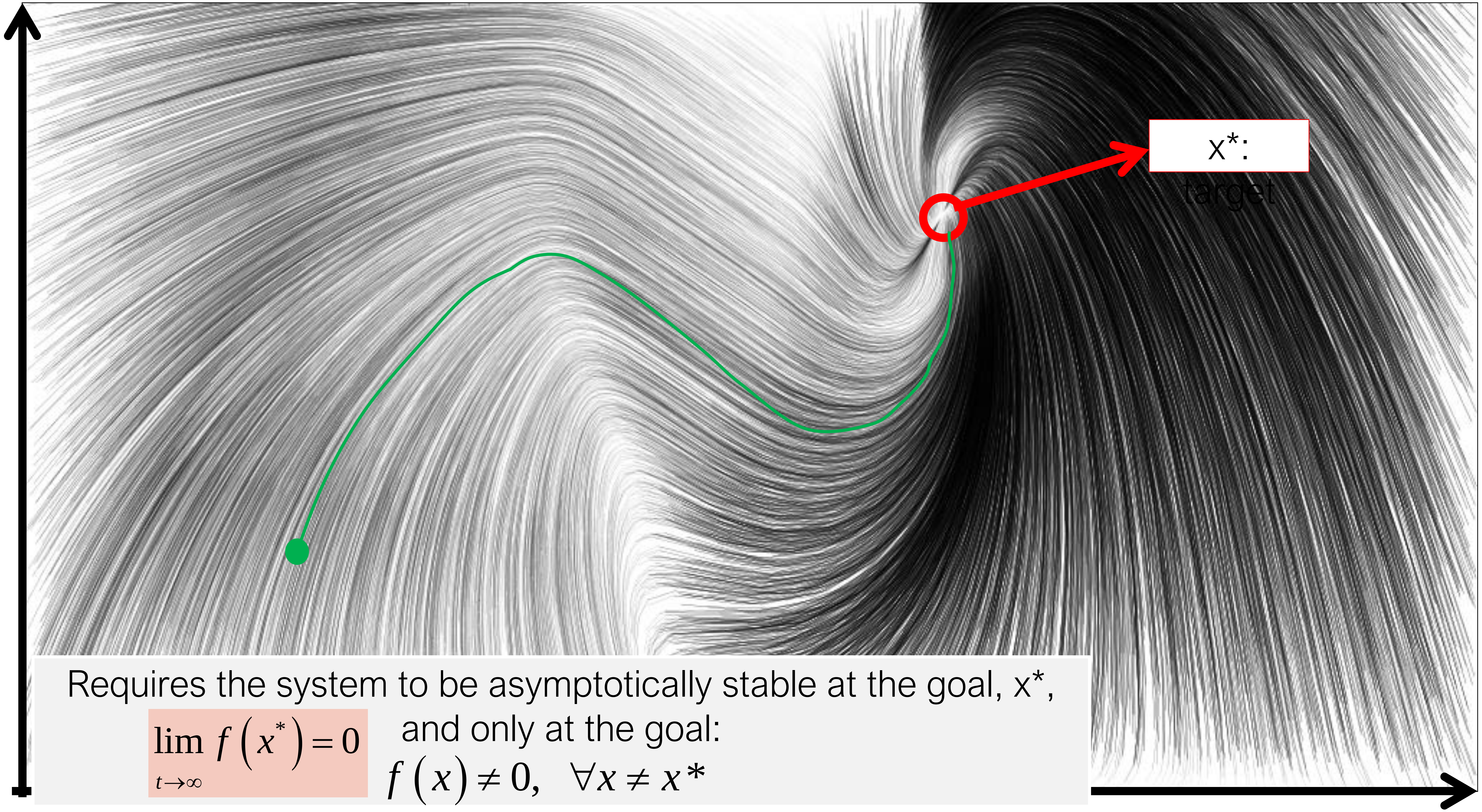
Learning from Demonstration: Examples

- How to use DS-based control to learn a hitting task
- How to extend with compliant control to improve task performance

Modelling Hitting Task using Dynamical Systems-Based Control

- Collect Demonstrations of hitting a golf ball using kinesthetic teaching
- Collect the recorded robot states and velocity at each time step
- We could generate a dynamical system representing this motion: $\dot{x} = f(x)$

Control using Dynamical Systems-Based Control - Recap



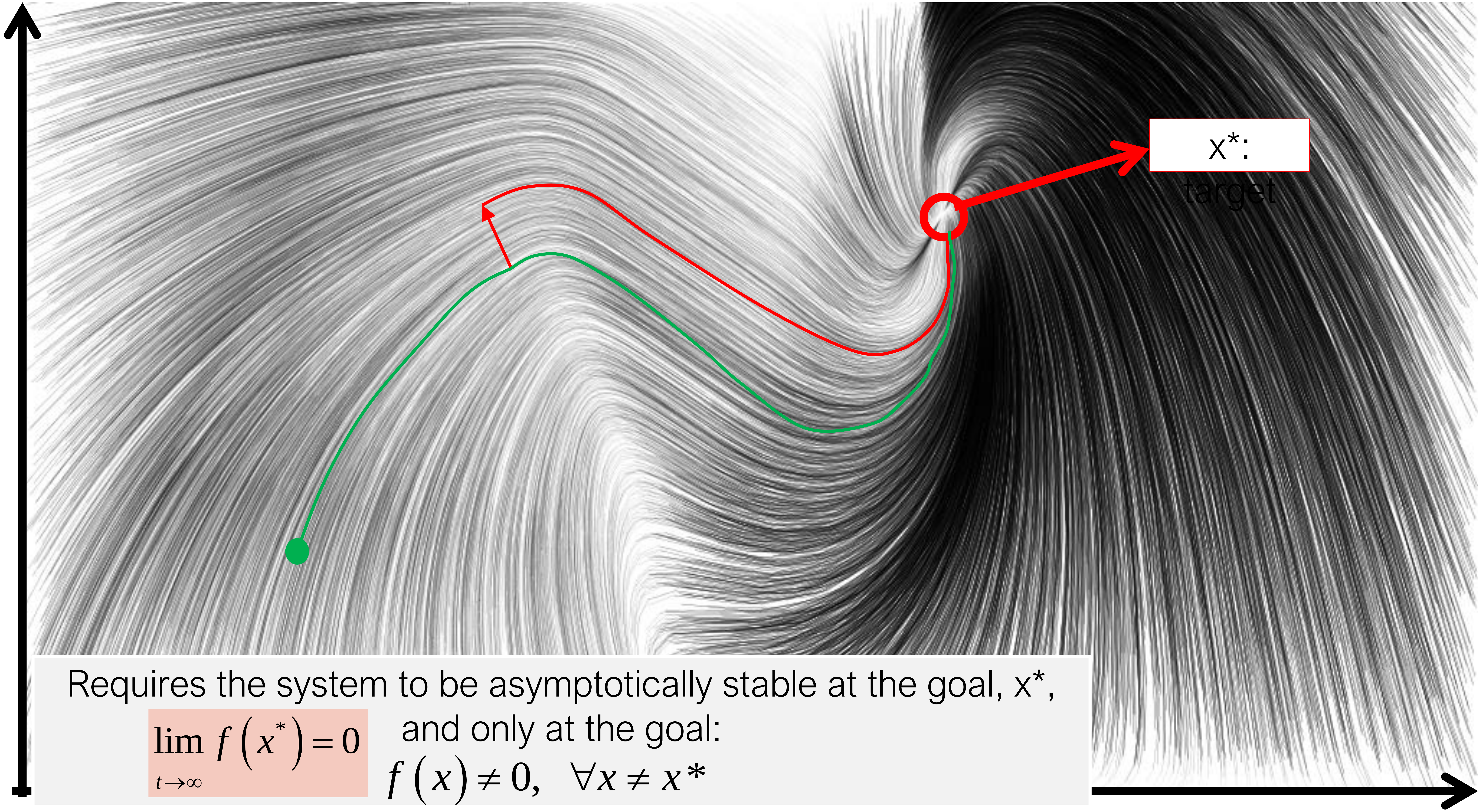
Requires the system to be asymptotically stable at the goal, x*, and only at the goal:

$$\lim_{t\to\infty} f\left(x^*\right)=0$$

$$f\left(x\right)\neq 0, \quad \forall x \neq x^*$$

EPFL

Control using Dynamical Systems-Based Control - Recap



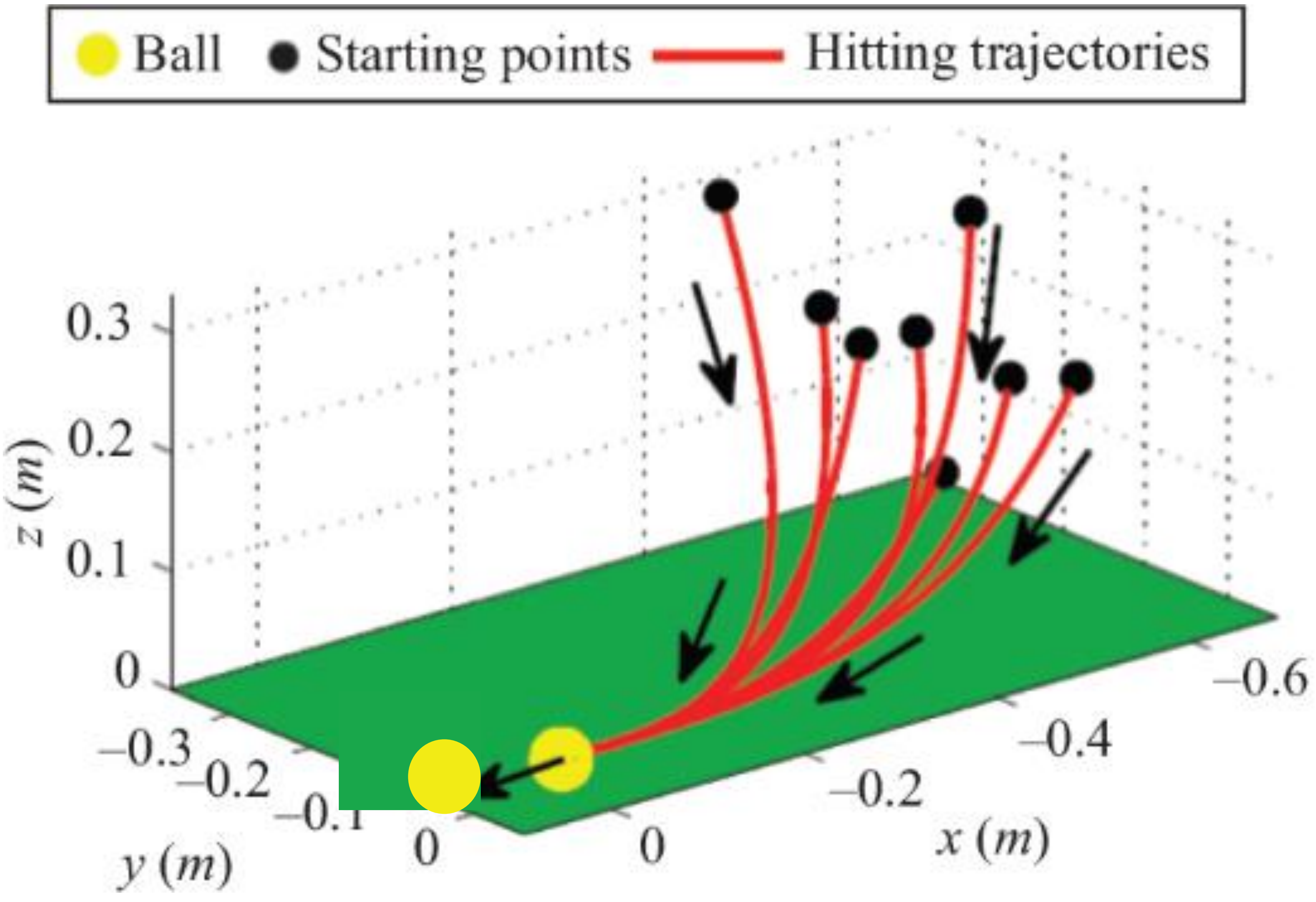
Requires the system to be asymptotically stable at the goal, x*, and only at the goal:

$$\lim_{t\to\infty} f\left(x^*\right)=0$$

$$f\left(x\right)\neq 0,\quad \forall x\neq x^*$$

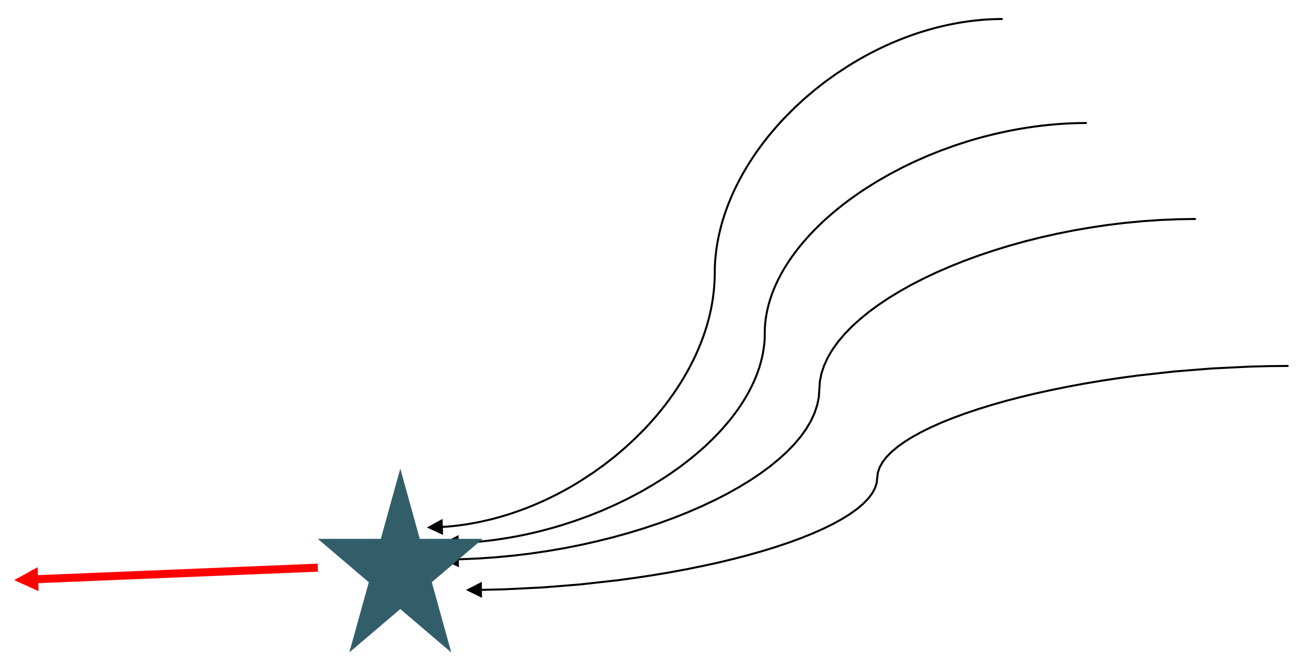
DS-based control of hitting tasks

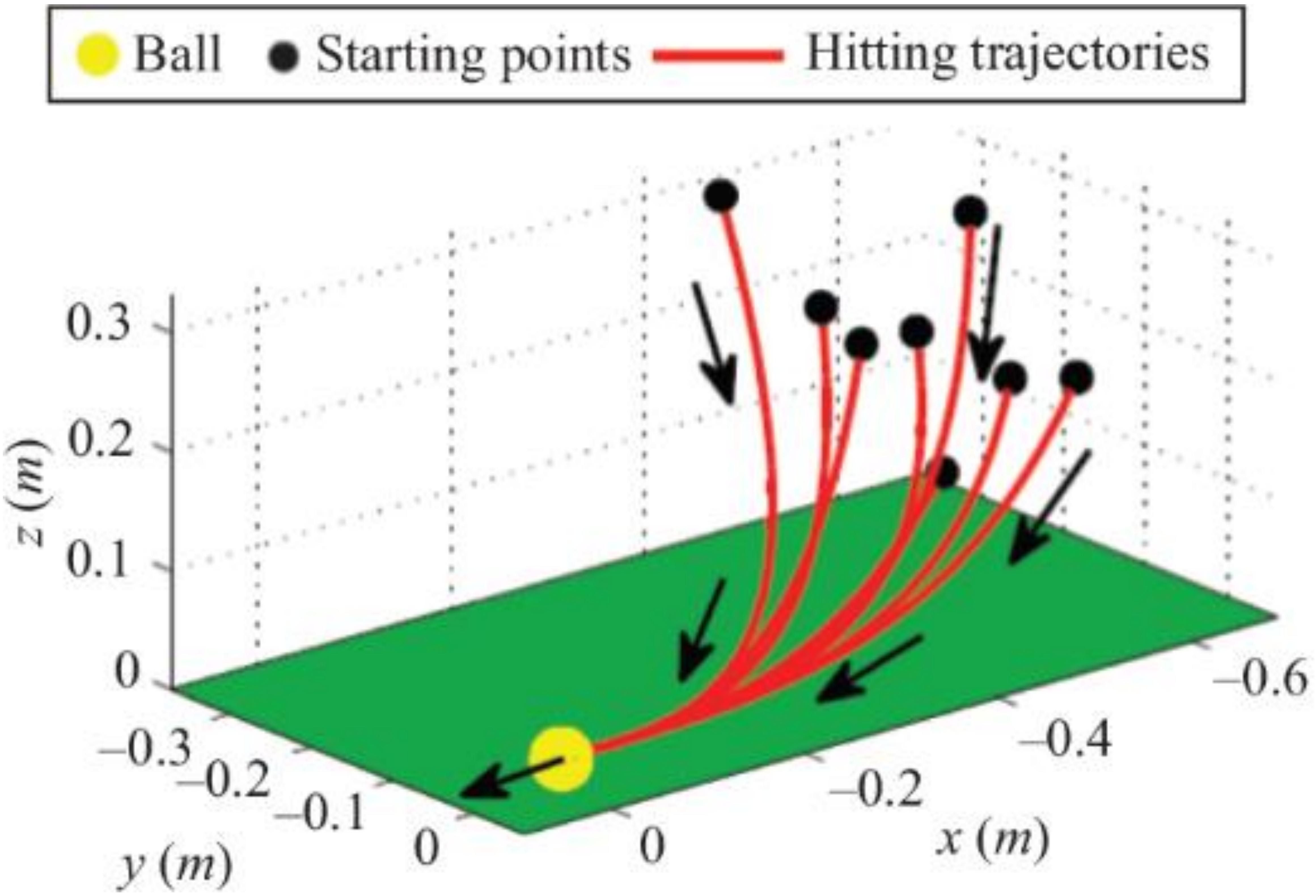
- Generate a dynamical system representing this motion: $\dot{x} = f(x)$
- Guarantee that the system asymptotically reaches and stabilizes at attractor: $\lim_{\{t\to\infty\}} x = x^*$, where x^*: Ball Location



DS-based control of hitting tasks

- $\dot{x} = f(x)$ with a fixed-point attractor $f(x^*) \neq 0$
- Desired velocity at attractor $\dot{x}^* \neq 0$

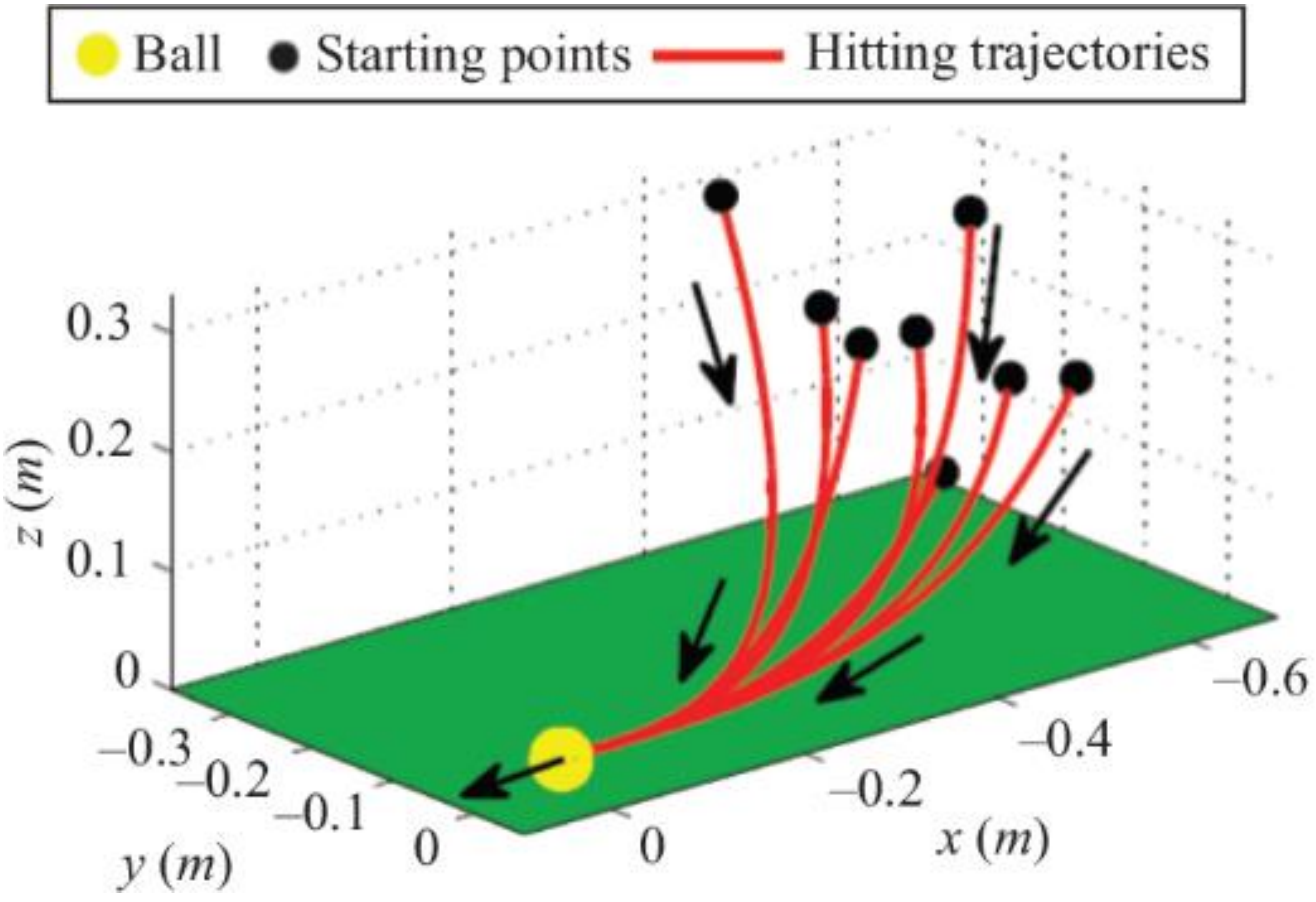




EPFL

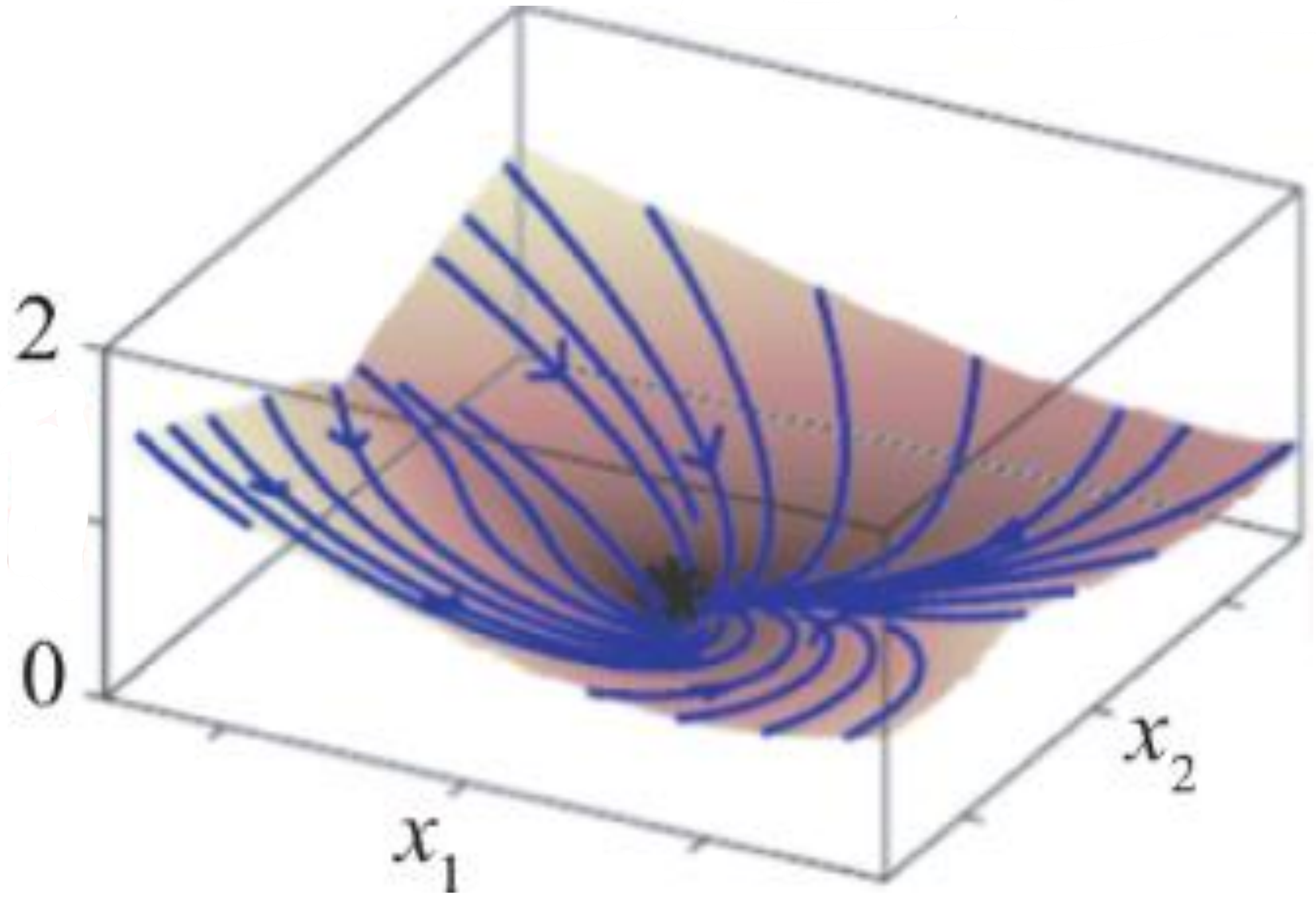
DS-based control of hitting tasks

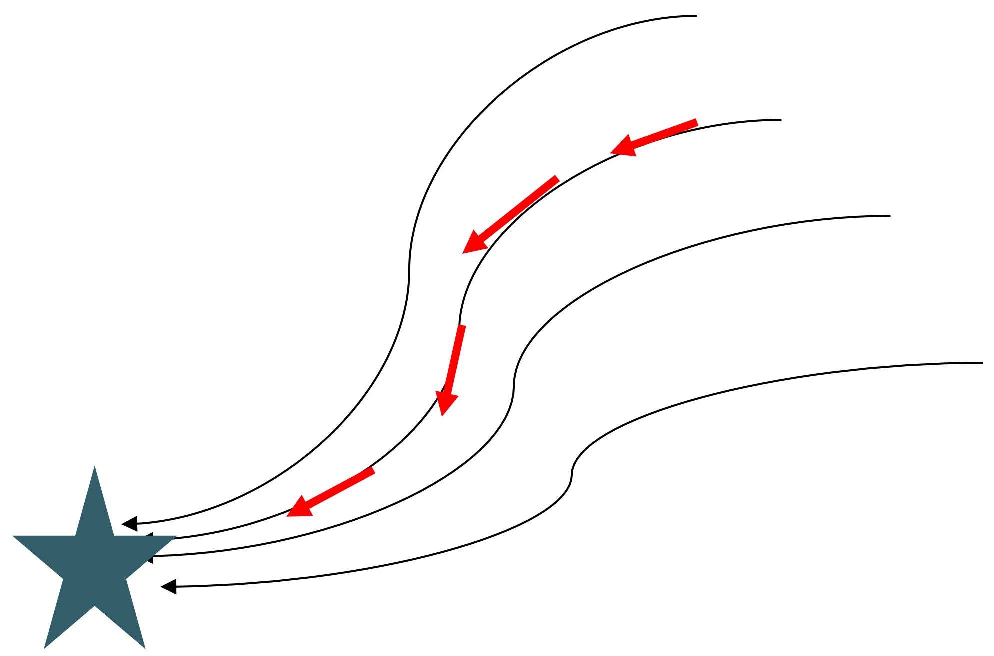
- Modulate the initial stable DS
- Dynamical system representing this motion: $\dot{x} = g(x), g(x^*) \neq 0$
- $f(x) = \dot{x} = M(x) * E(x)$
- Target field: $E(x)$
- Strength Factor: $M(x)$



DS-based control of hitting tasks

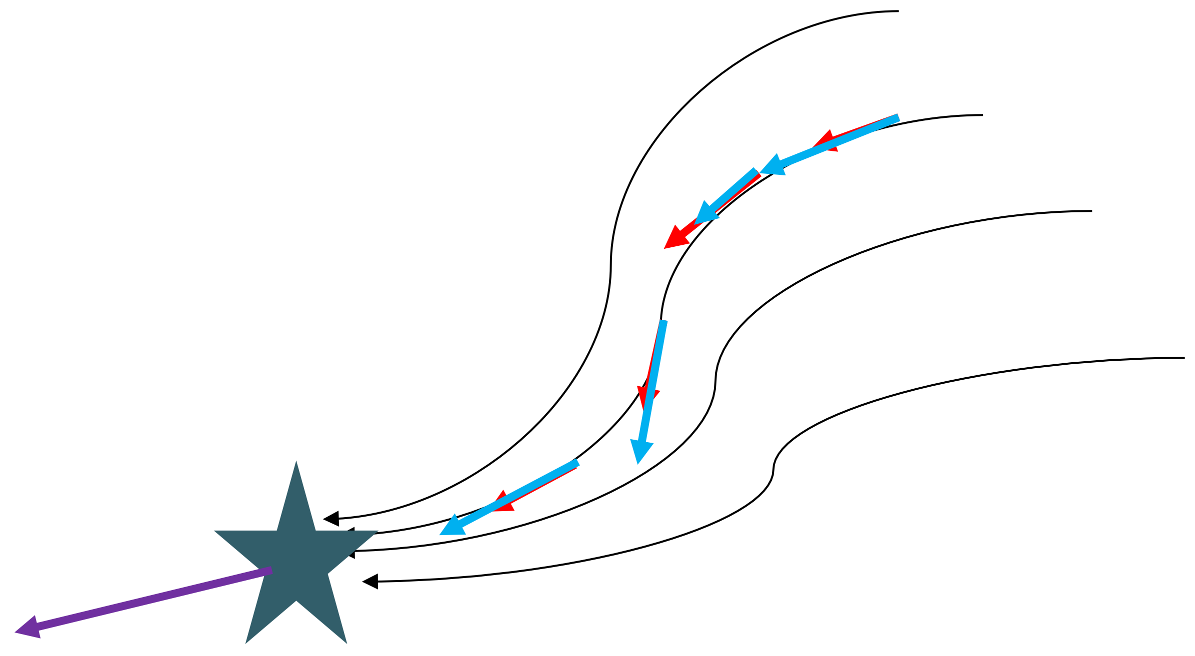
- Dynamical system representing this: $\dot{x} = g(x)$, $g(x^*) \neq 0$
- We only store the unitary vector field $\mathrm{E}(\mathrm{x}) := \|g(x)\| = 1\ \forall x$
- $E(x) = \frac{g(x)}{\|g(x)\|}$
- Embeds the orientation at target and asymptotic stability





DS-based control of hitting tasks

- We also learn a function $M(x)$ from the demonstration set
- Embeds the amplitude of the velocity when approaching the ball during demonstration: $M(x) = ||\dot{x}||$
- Can be learned using any ML regression technique

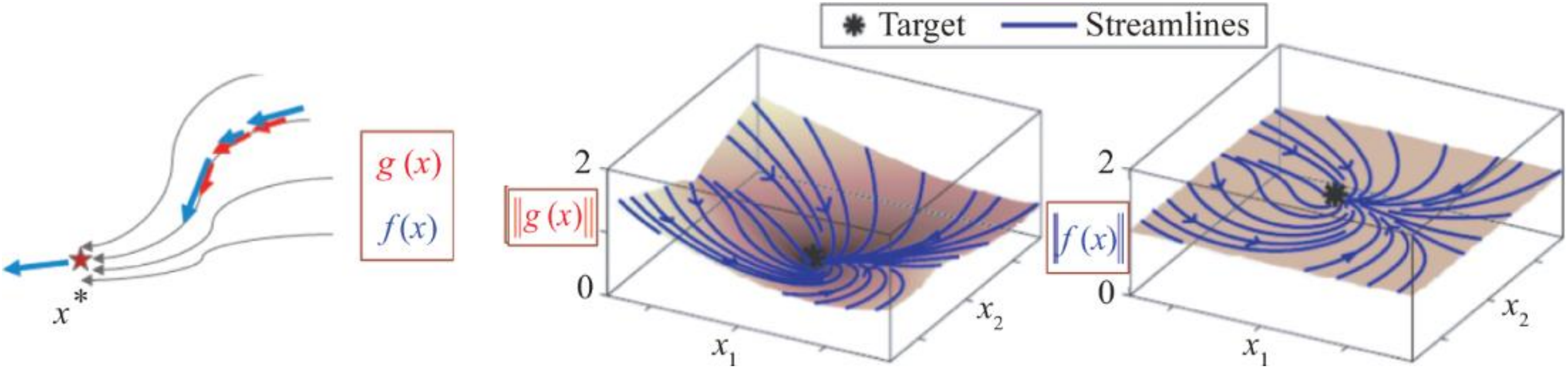


DS-based control of hitting tasks

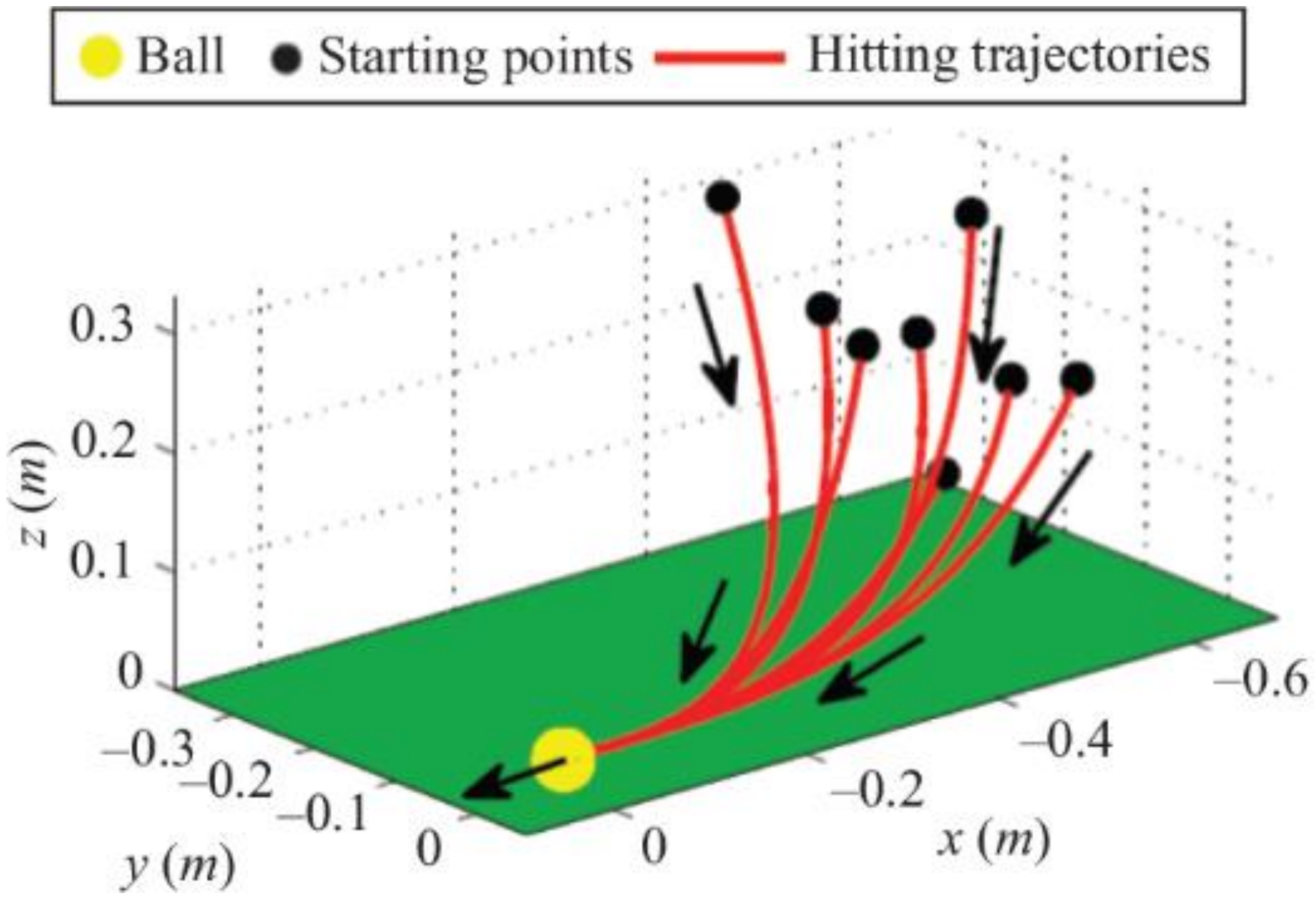
$$f(x) = M(x) * E(x)$$

Velocity at target

Orientation at target and asymptotic stability



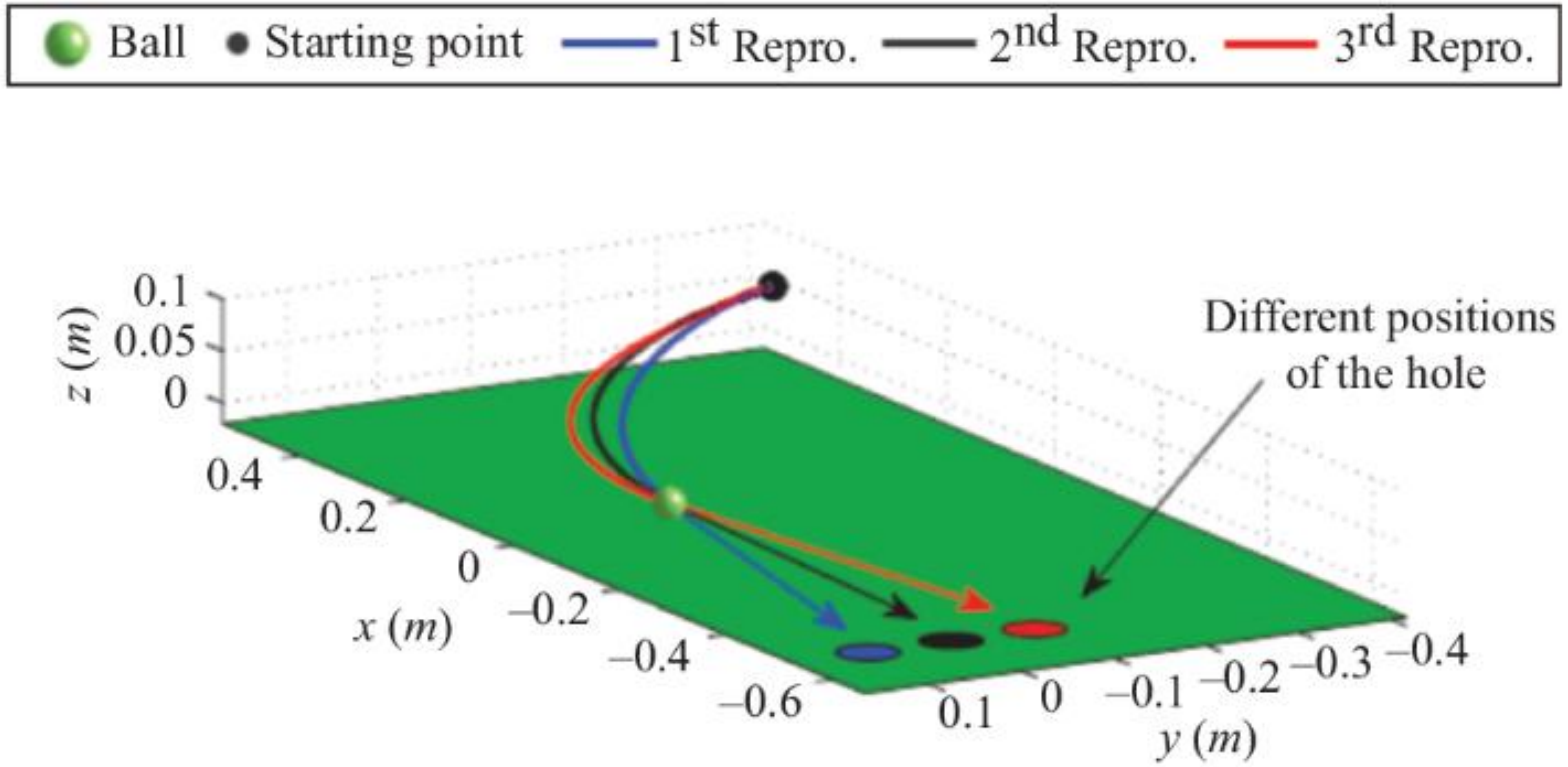
DS-based control of hitting tasks



Representation through DSs of trajectories to sink a golf ball.

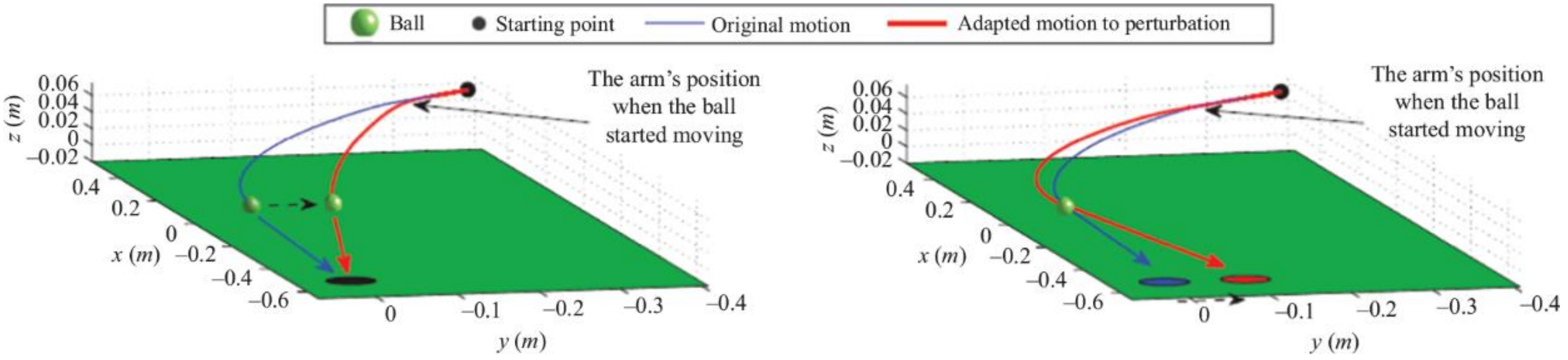
Generalization

- Expressing the dynamics from the relative position of the ball to the sink allows one to nicely generalize the orientation toward the ball without further demonstrations

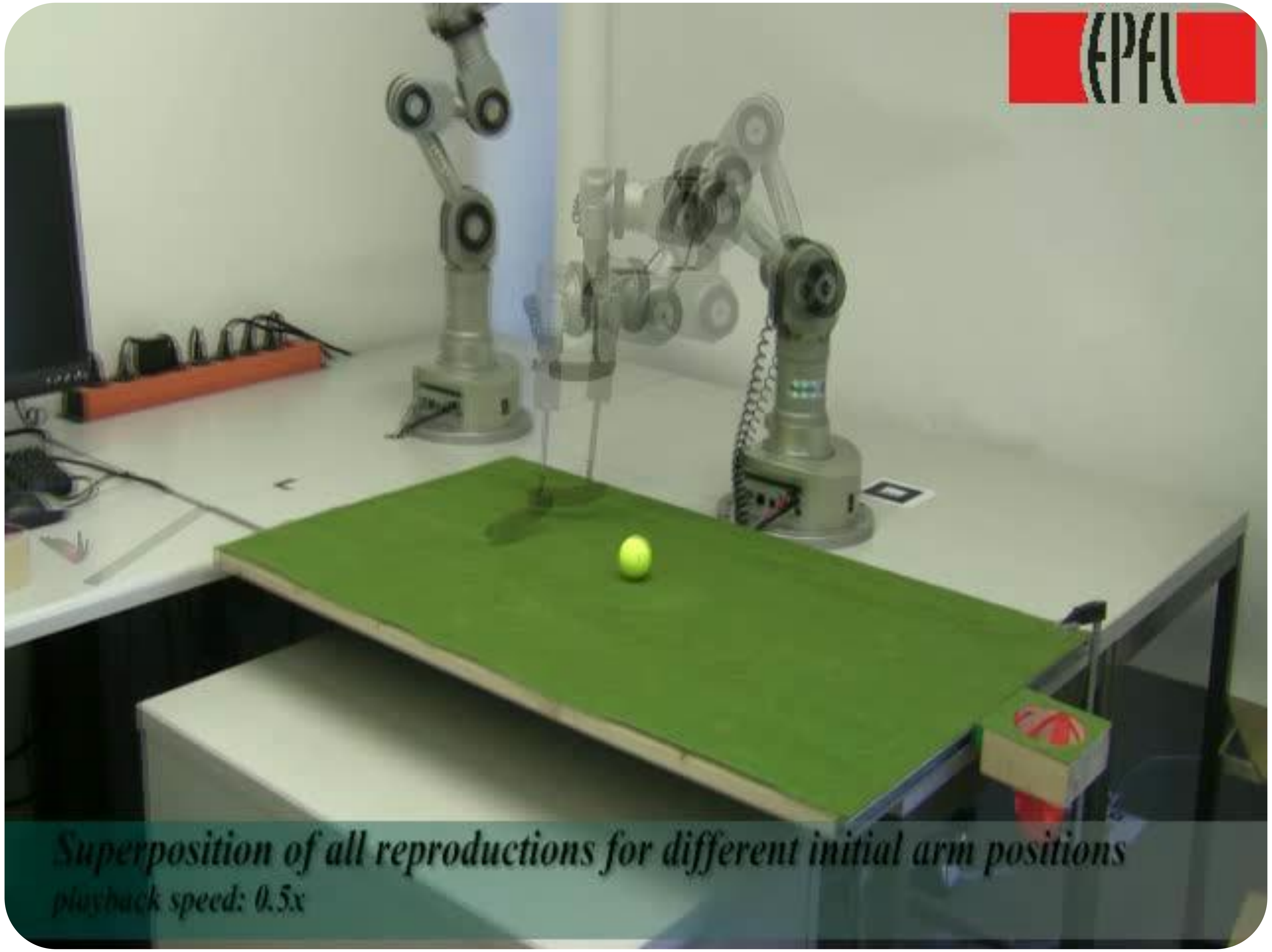


DS-based control of hitting tasks

- The learned DS can adapt at run time to perturbations, such as pushing the robot away from its trajectory (left) and moving the hole (right) by generating a new trajectory that reaches the target correctly

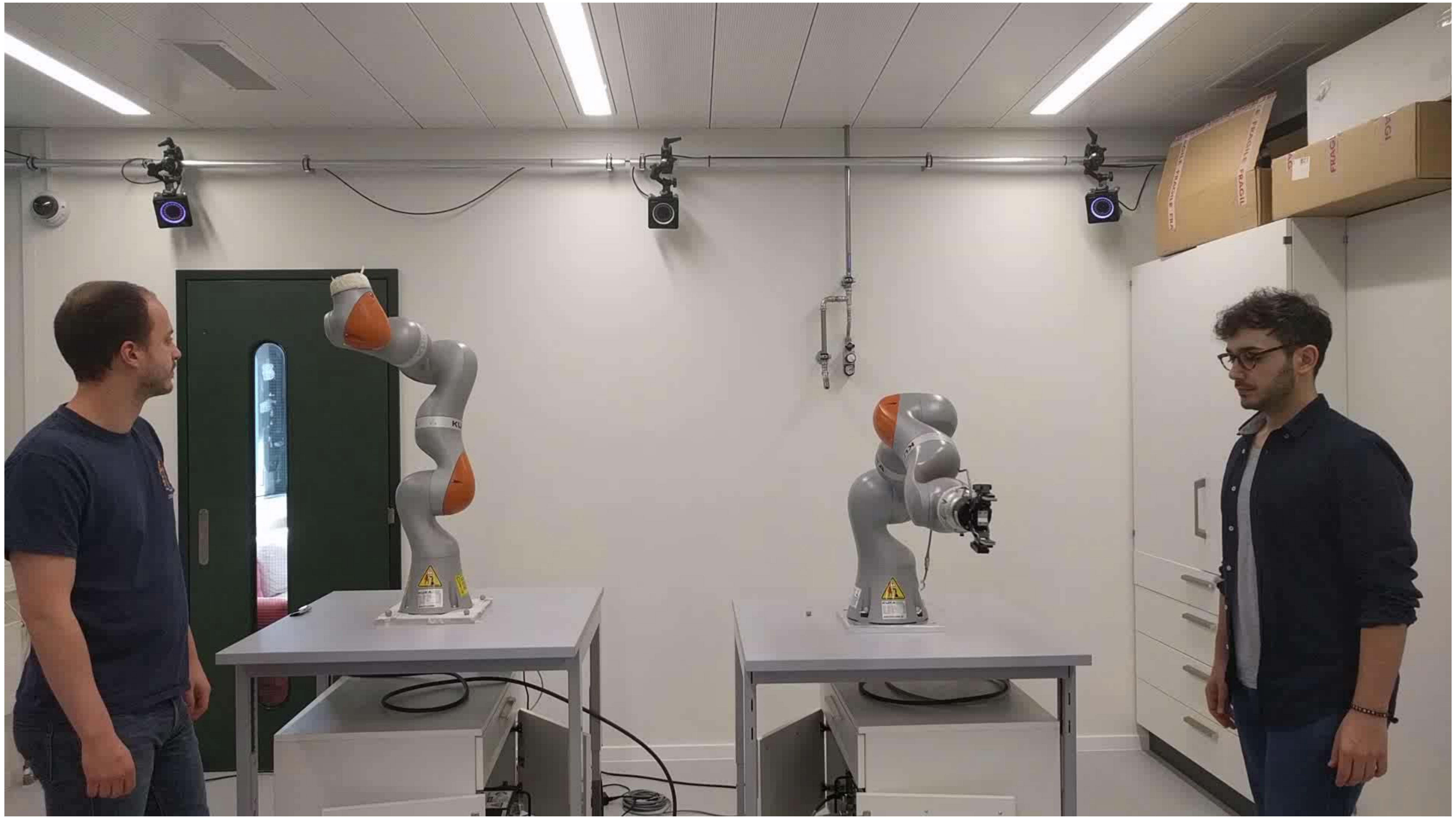


Robotic implementation

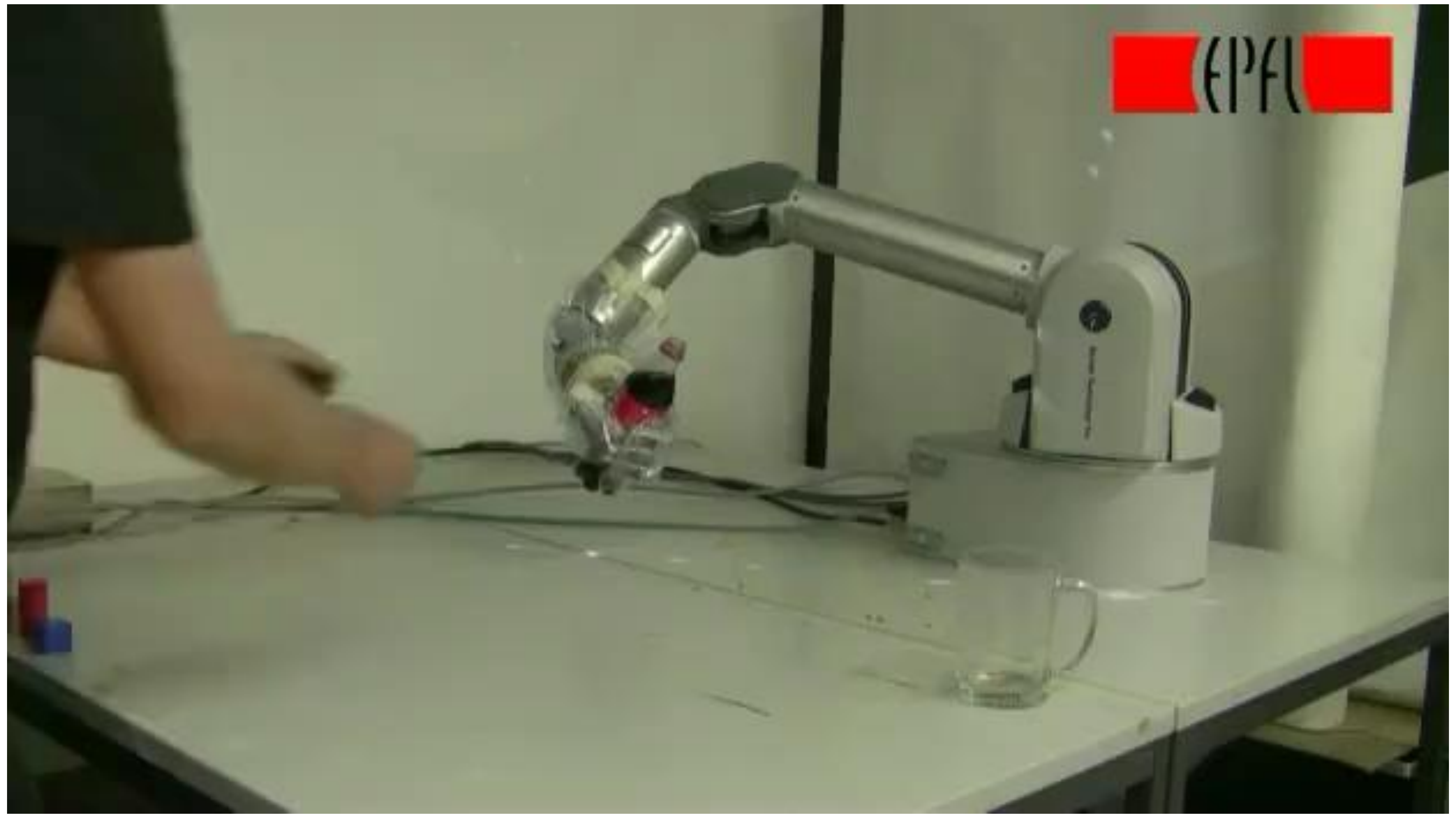


Teaching Compliant Control

Compliant Control



Teaching Compliant Control: What happens when stiffness not considered?



Too stiff: Liquid spills from jerking

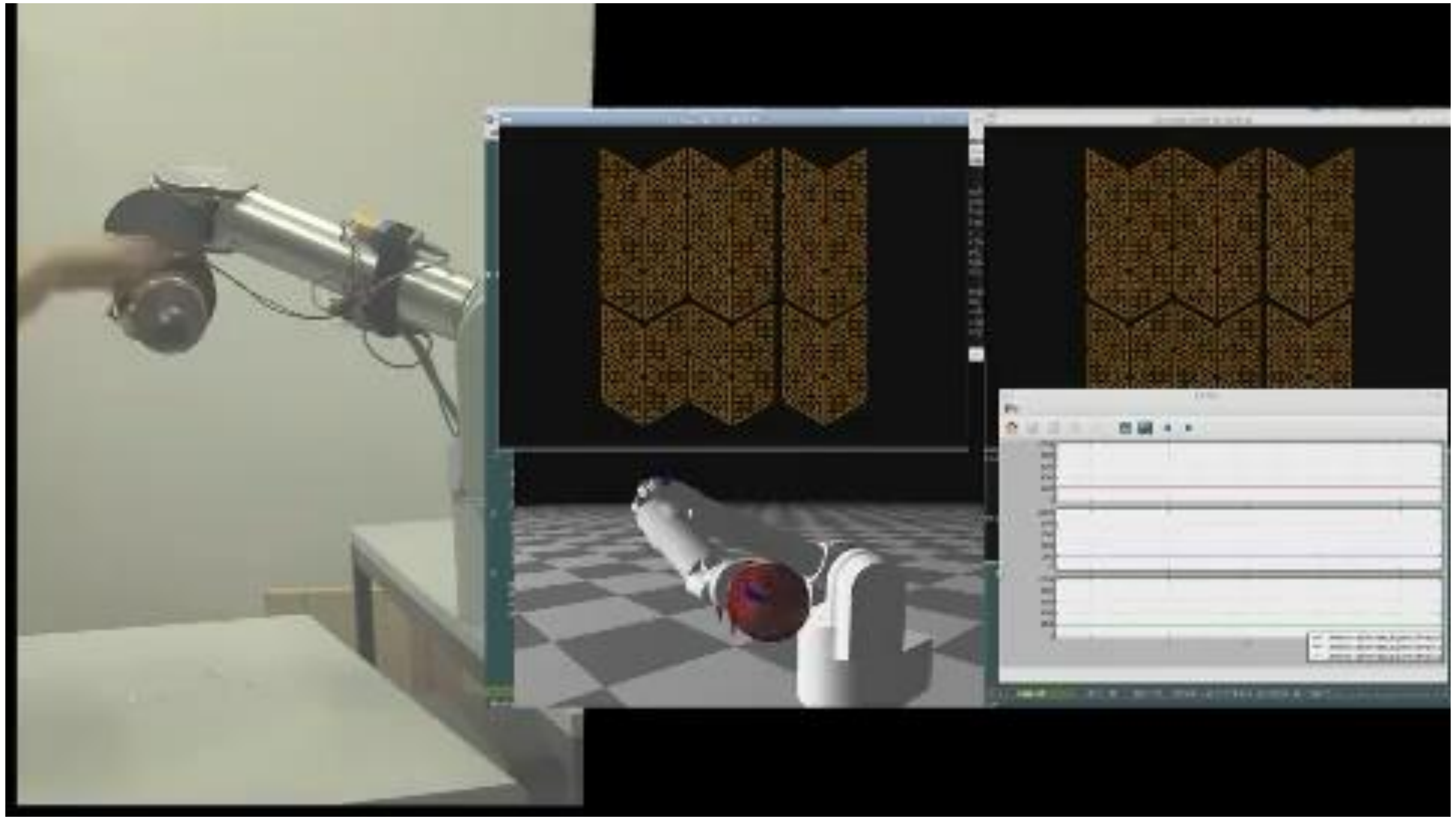
Too compliant: Liquid spills from glass

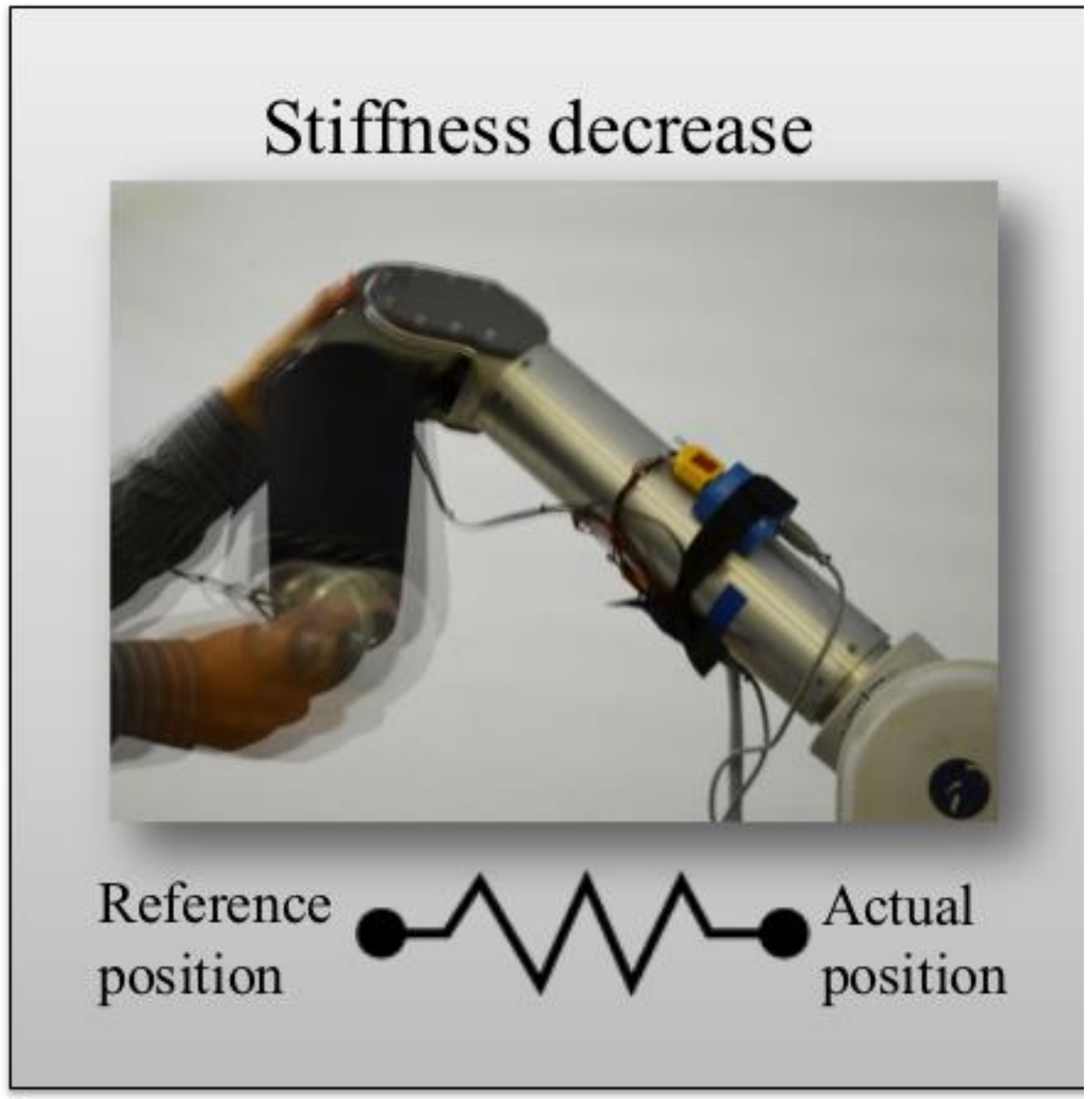
How can we teach robot when to increase and decrease compliance?

Kronander and Billard, Robot Learning from Demonstration for Compliant Manipulation, ICRA 2012

Teaching Compliant Control: Adding Compliance

Teaching ***decrease*** in stiffness by wiggling the robot

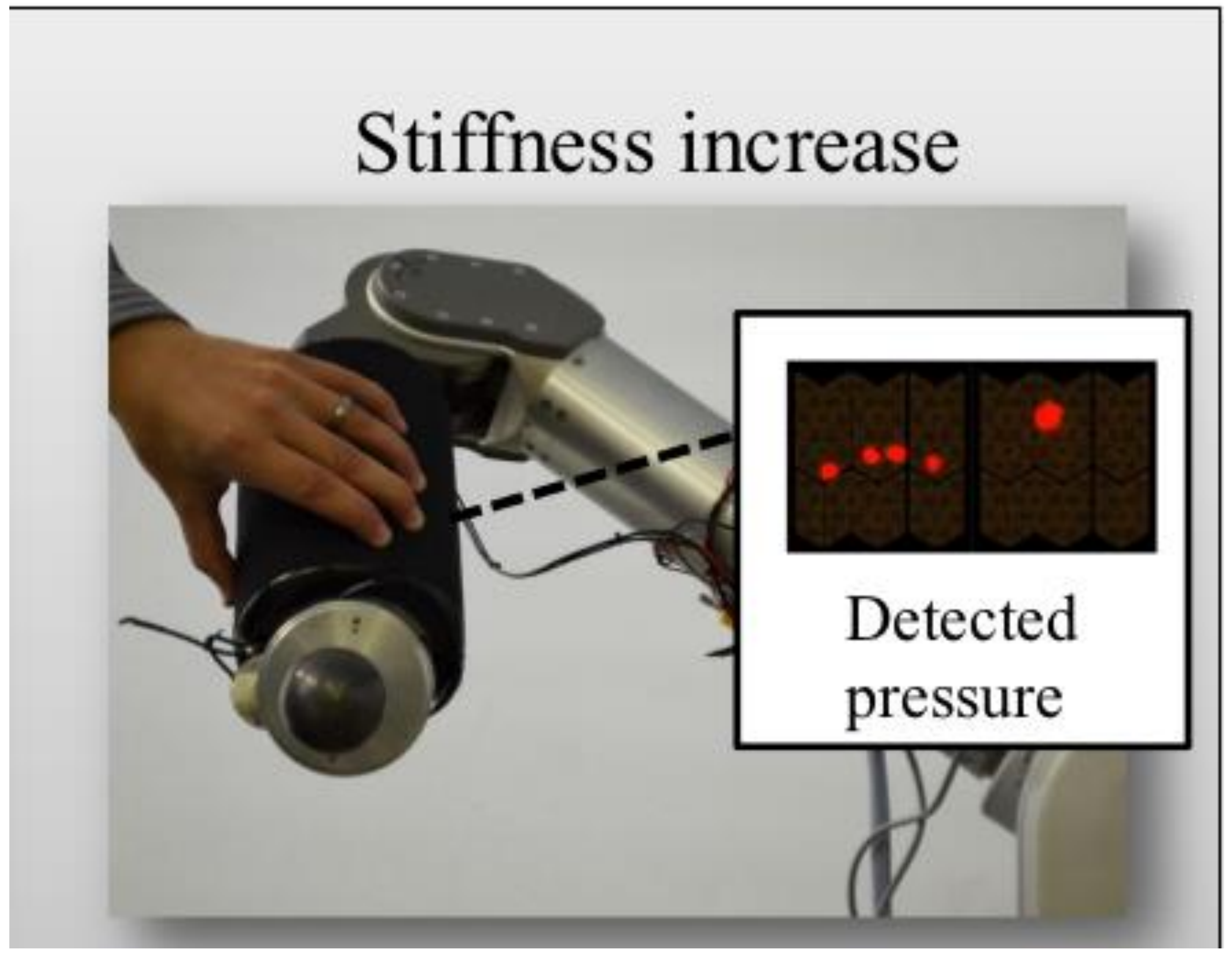




Kronander and Billard, Robot Learning from Demonstration for Compliant Manipulation, ICRA 2012

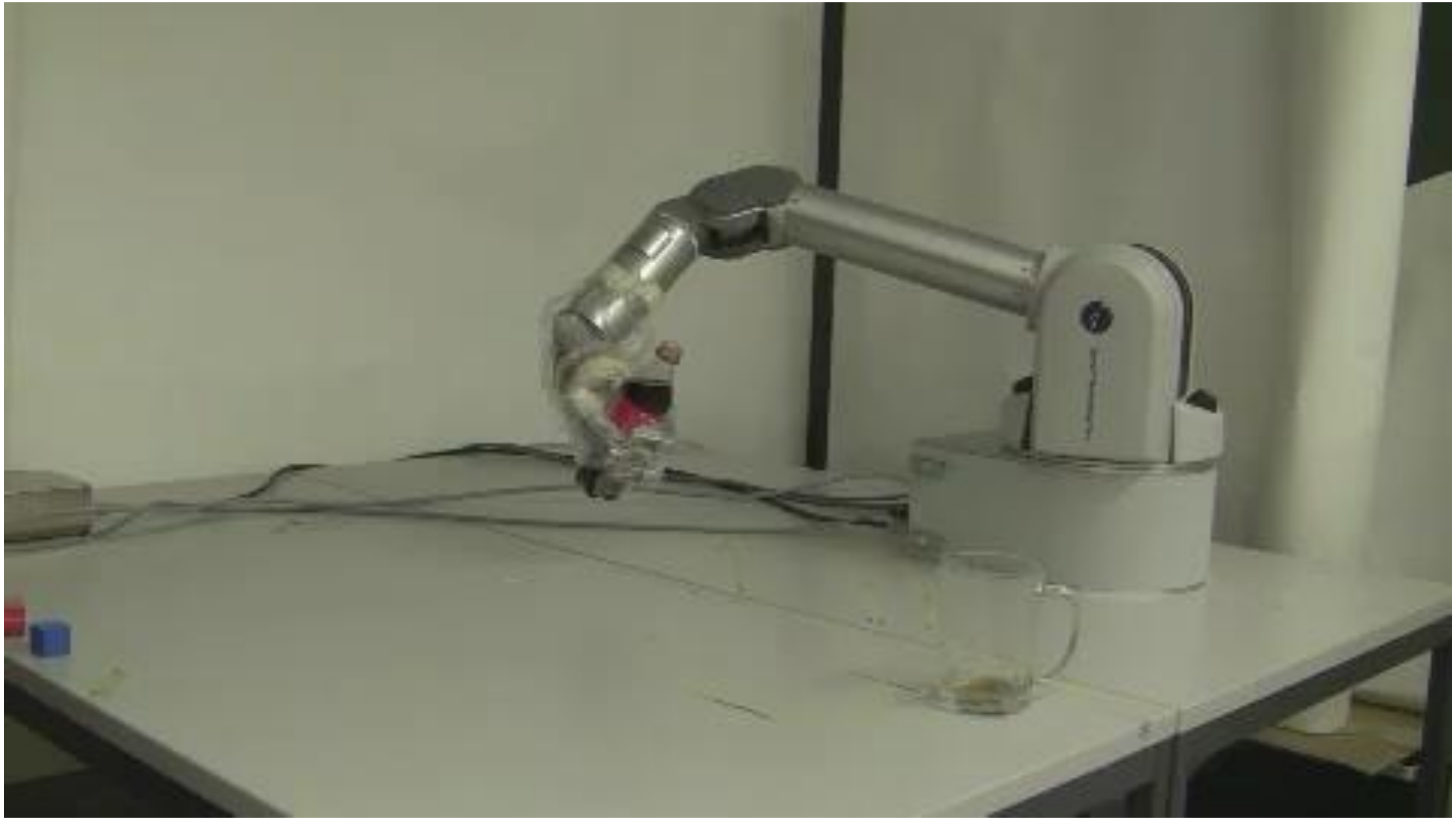
Teaching Compliant Control: Adding Stiffness

Teaching *increase* in stiffness
by exploiting tactile sensing on robot arm



Kronander and Billard, Robot Learning from Demonstration for Compliant Manipulation, ICRA 2012

Teaching Compliant Control: Final Result



Kronander and Billard, Robot Learning from Demonstration for Compliant Manipulation, ICRA 2012

Remaining Challenges and General Considerations

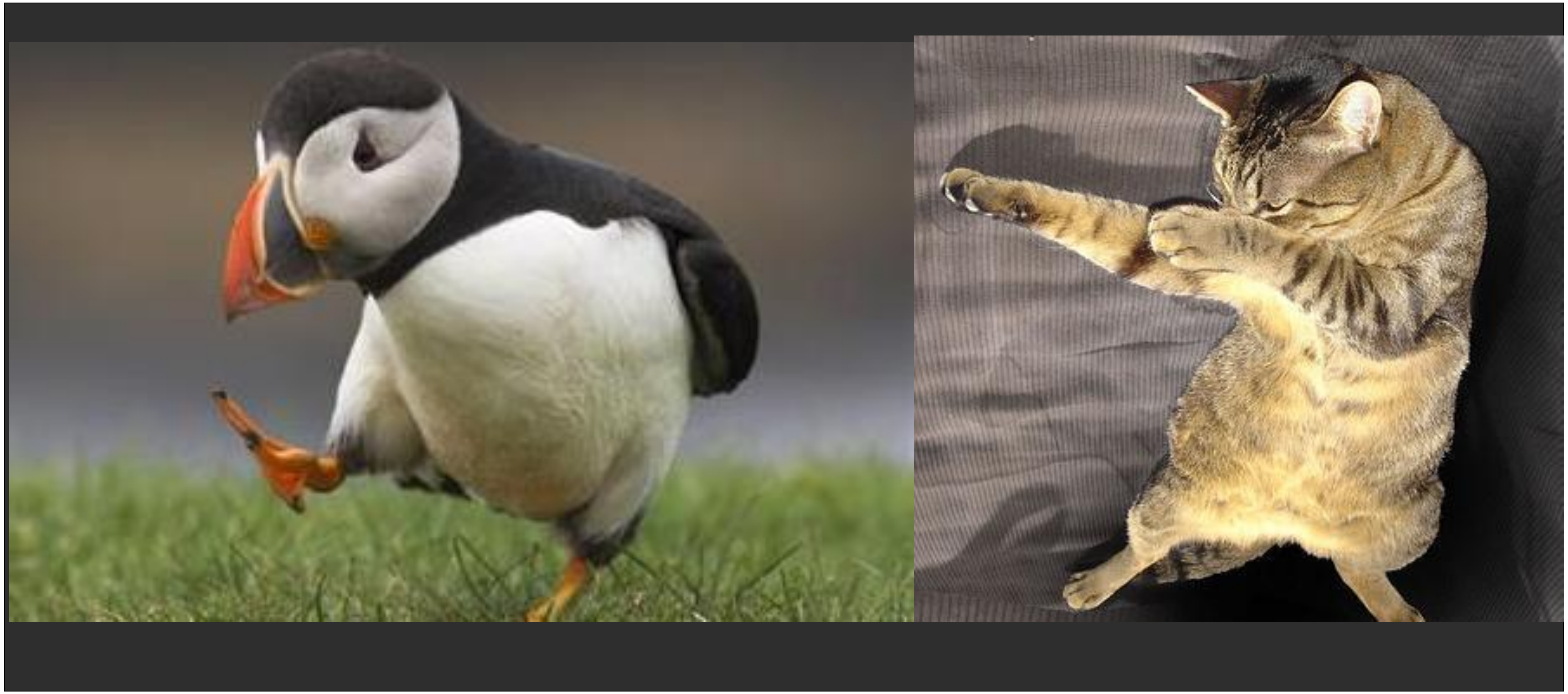
Open challenges

Learning human skills through LFD requires the following questions:

- What/Who to imitate?
- How to imitate?
- When to imitate?

Answering these questions requires us to better address *the correspondence problem*.

The body shapes our movements

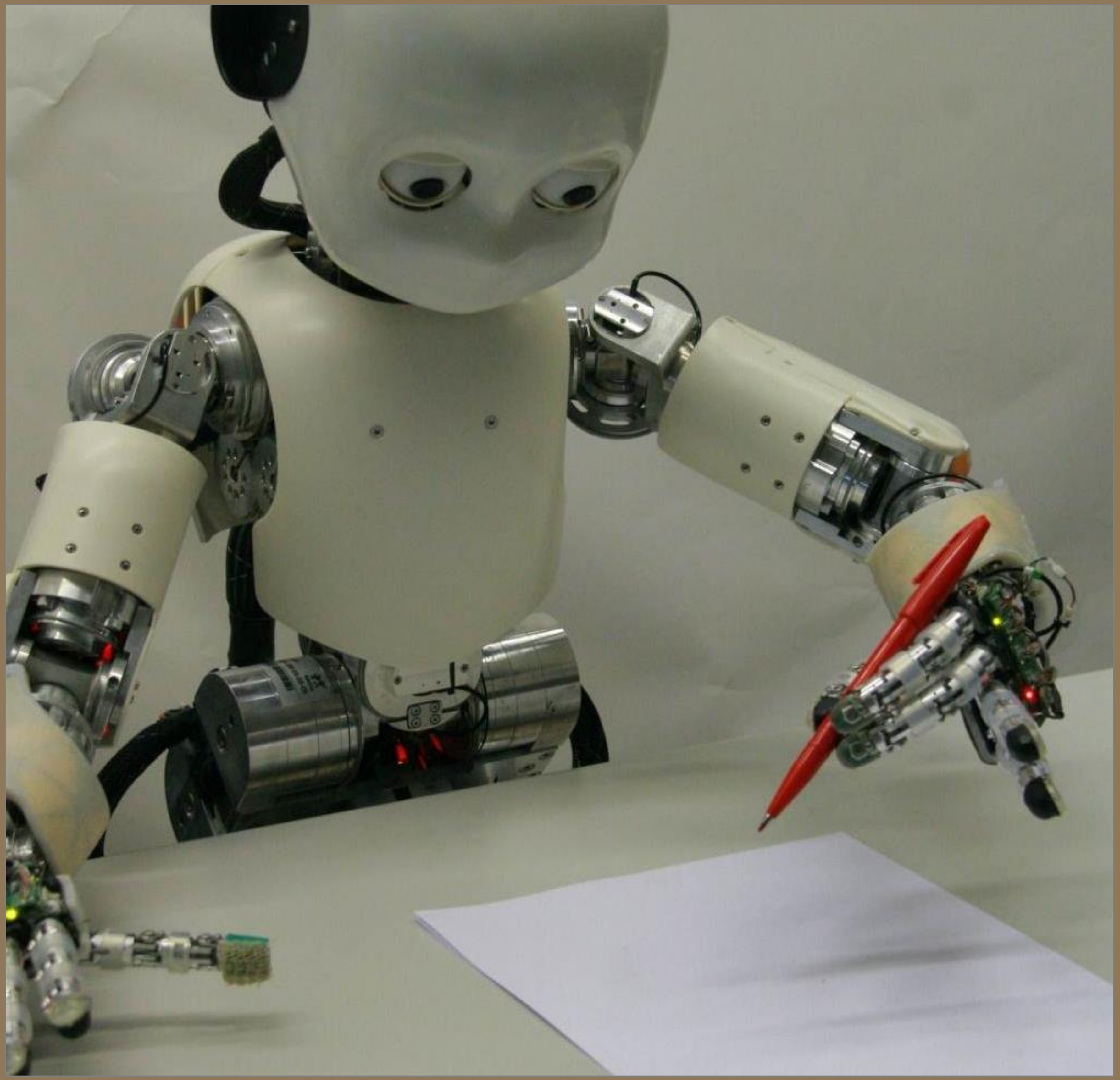


Evolution has shaped the body and the control system simultaneously so as to optimize the animal's overall motor control system

With enough training….

…we can get bodies to do things for which they were not designed for in the first place, but this requires tedious and long training periods.

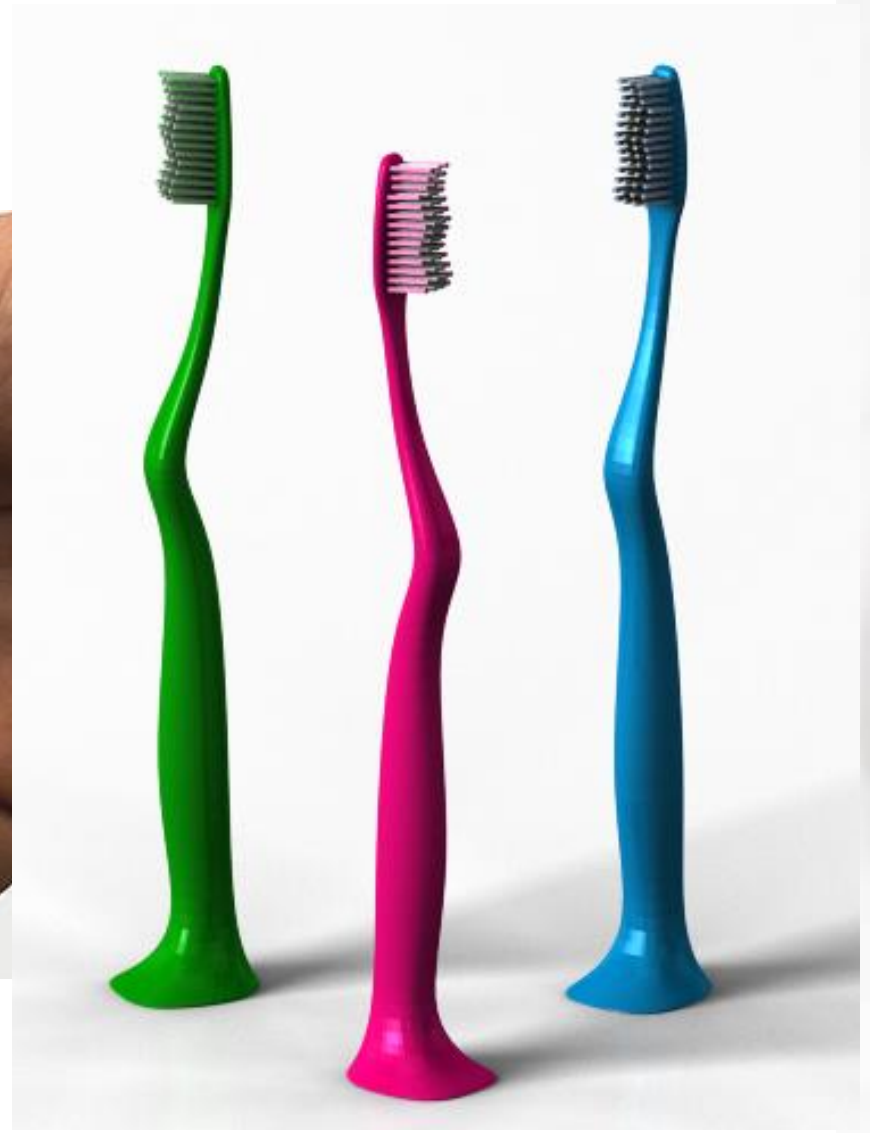
Which body for our robots?



Should robots have arms and hands that are similar to human hands and arms?

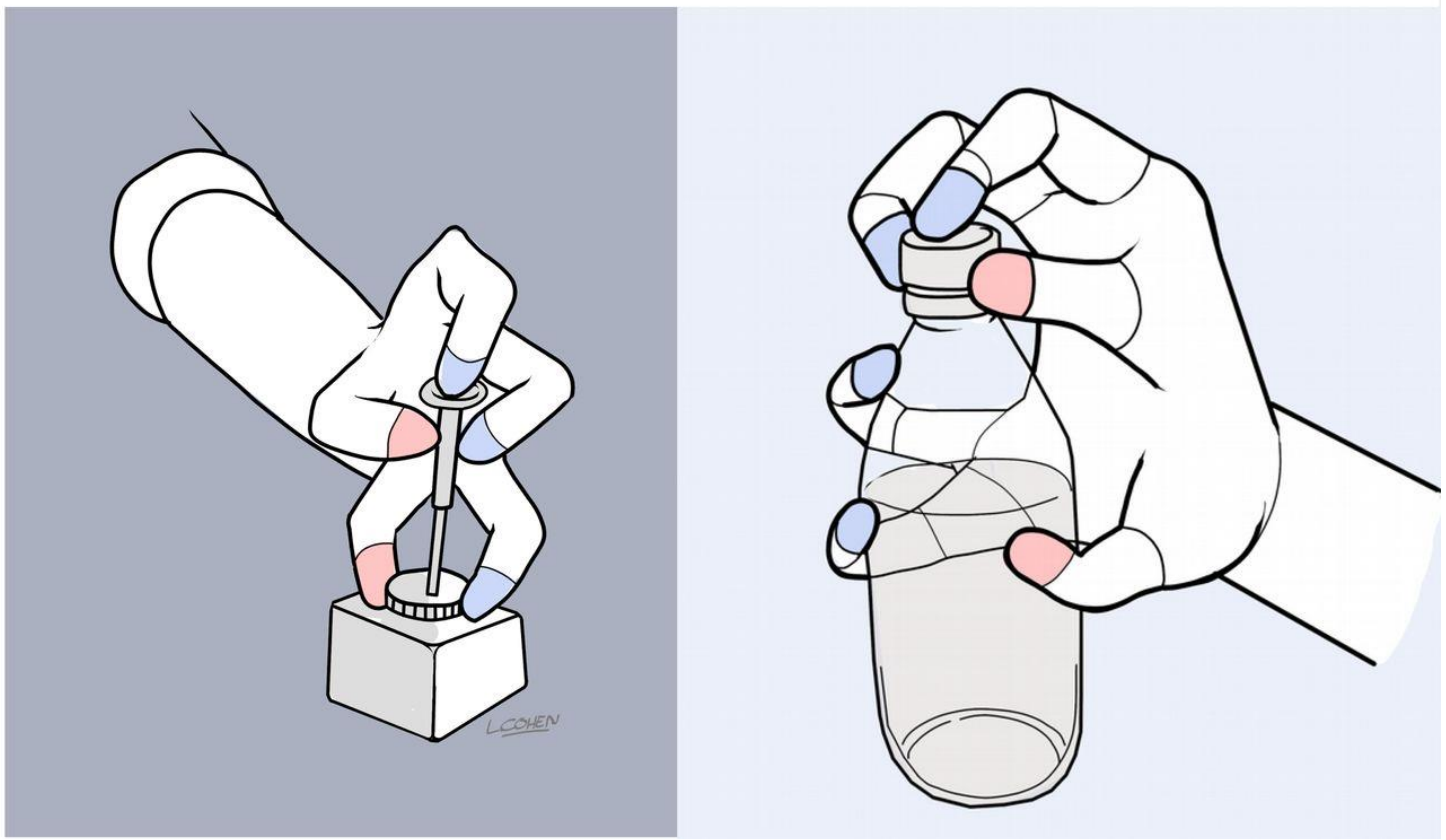
How the body shapes the design of tools

Robots should help us in our daily tasks
- manipulate the same objects
- objects are designed for the human hand


Bent to our needs!

Different bodies open the door for more



Billard, Aude, and Danica Kragic. "Trends and challenges in robot manipulation." Science 364.6446 (2019): eaat8414.

Current/Future Research Directions: Learn from Human Demonstrations

- How to ease training of robots from humans?
  - Learn from small datasets: Reduce the number of demonstrations needed
  - Combine heterogeneous data types: Use multiple interfaces at once
  - Improve teaching interactions – easier, cheaper and more user-friendly interfaces

- How to teach robots to do complex tasks without showing all the details?
  - Use Large Language Models to ease transfer of knowledge
  - Develop tools for automatically convert demonstrations that the robots can understand
  - Generate in simulation more examples, sufficient number of examples to generalize
  - Have robots query users for more information where needed (active learning)

EPFL

Summary

How can we learn controllers from data and how do we get the data?

- How can we can use demonstrations to learn task controllers?
- Different approaches to Learning from Demonstration
- How interface design affects data gathering and demonstration quality
- Examples of LfD
- General Considerations & Future Directions