Knowledge Transfer Between Robots with Similar Dynamics for High-Accuracy Impromptu Trajectory Tracking

European Control Conference
June 26, 2019

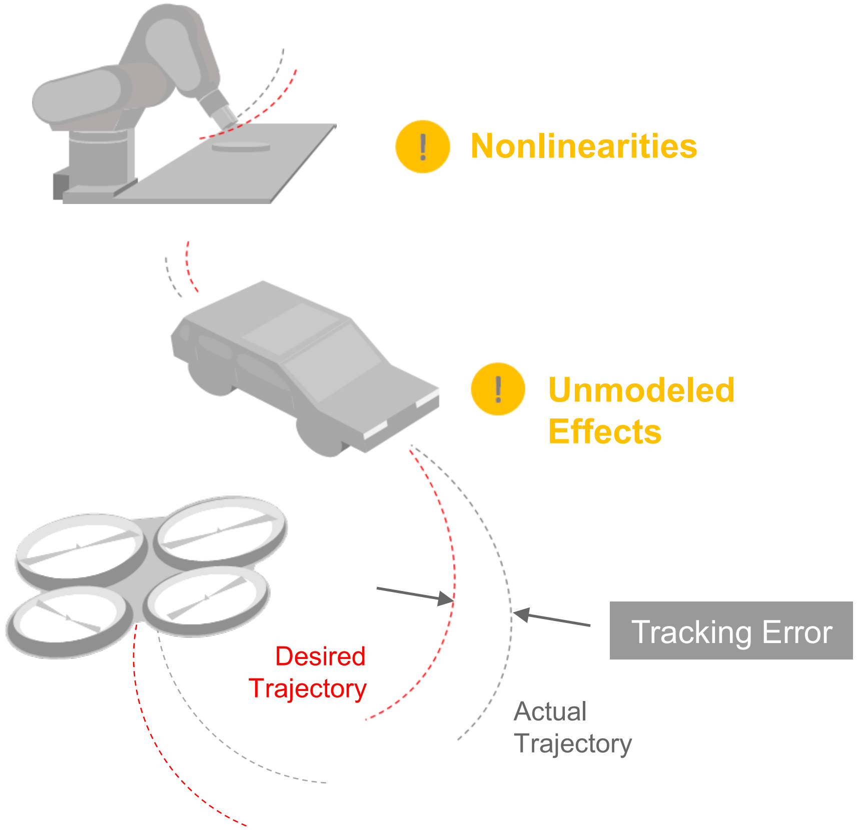
SiQi Zhou[1], Andriy Sarabakha[2], Erdal Kayacan[3], Mohamed K. Helwa[1], and Angela P. Schoellig[1]

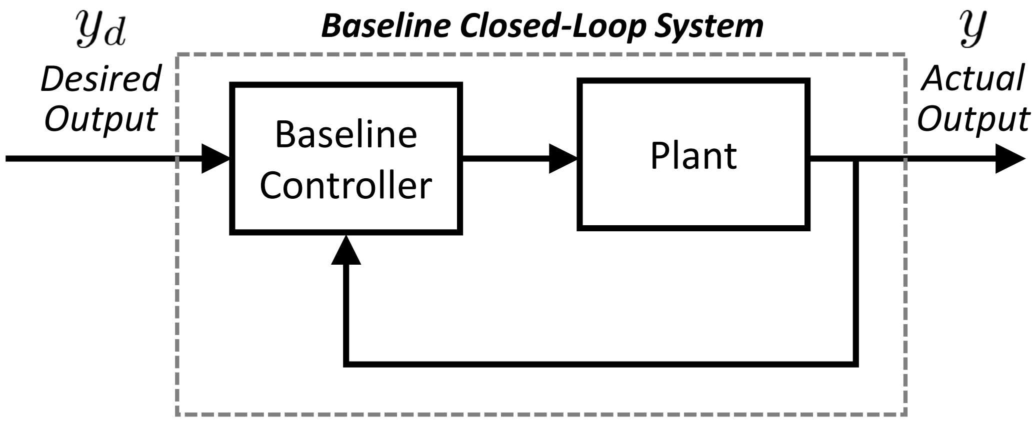
[1] *Dynamic Systems Lab, University of Toronto Institute for Aerospace Studies*
[2] *School of Mechanical and Aerospace Engineering, Nanyang Technological University*
[3] *Department of Engineering, Aarhus University*

Institute for Aerospace Studies UNIVERSITY OF TORONTO
VECTOR INSTITUTE | INSTITUT VECTEUR
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE
AARHUS UNIVERSITY
DYNAMIC SYSTEMS LAB

Introduction

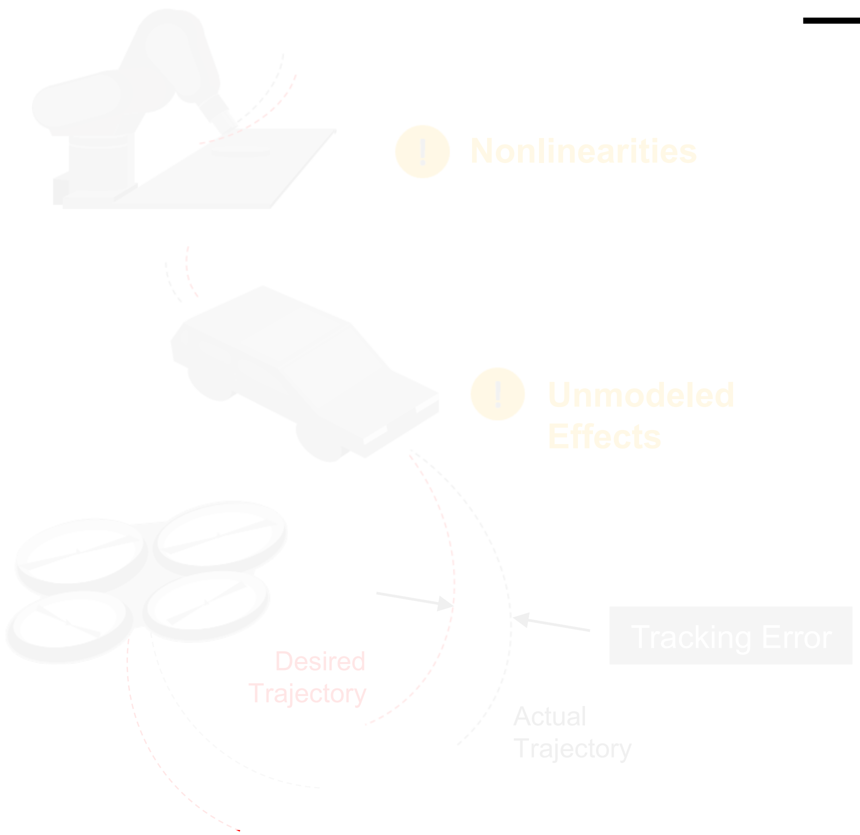
Designing control systems for high-accuracy tracking can be challenging

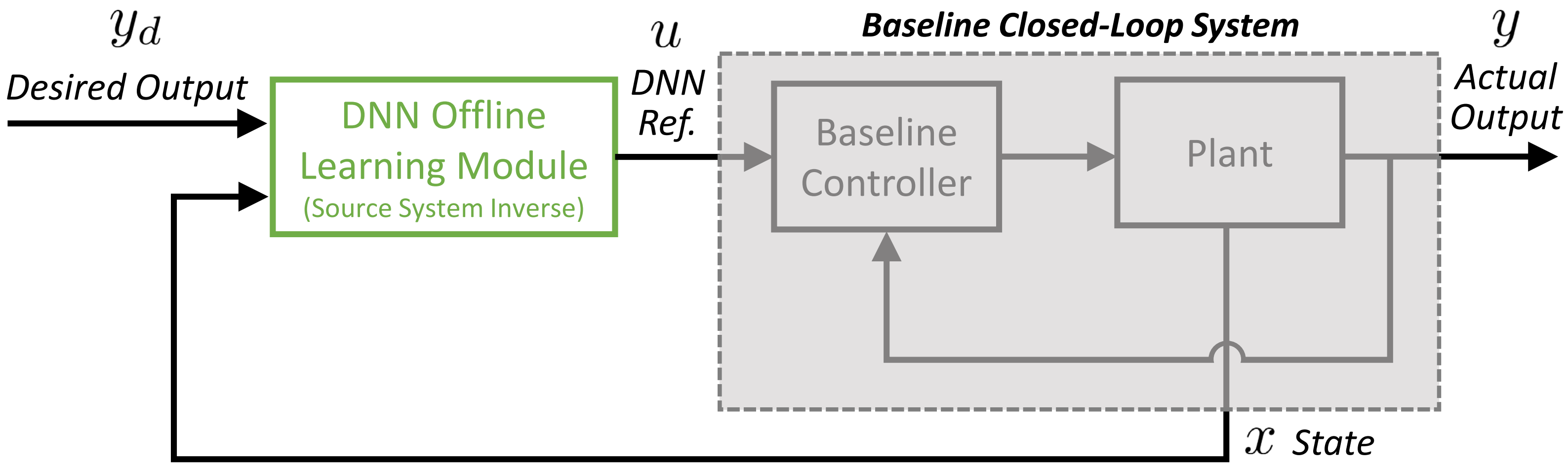




Introduction

Neural networks as add-on blocks to enhance ‘black-box’ systems





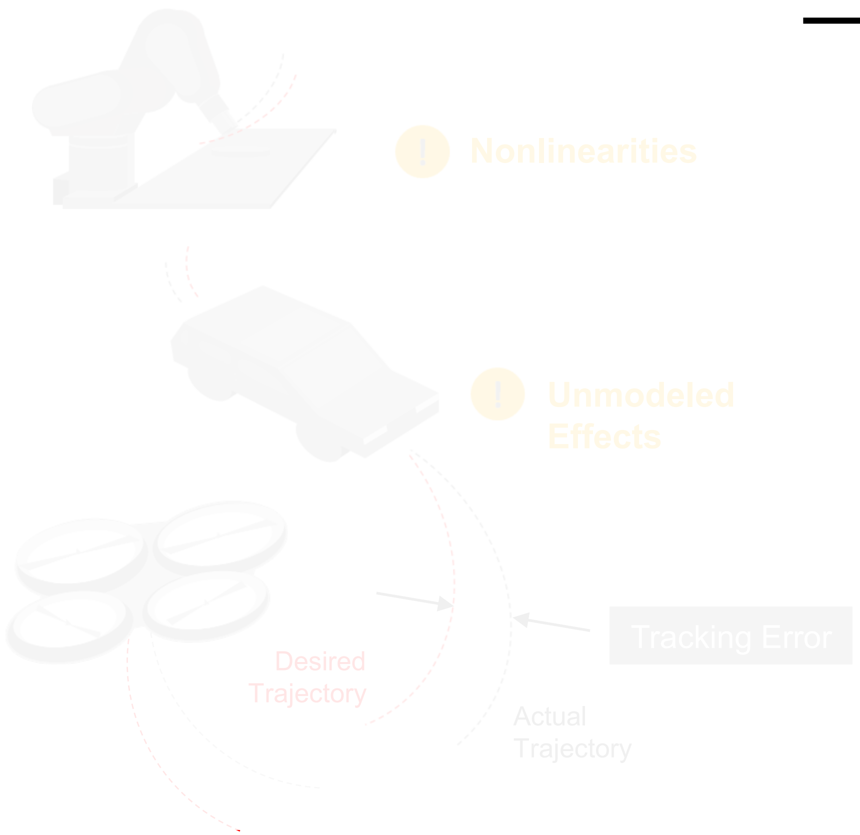
Note: If the video on previous slide has a problem, the full version of the video can be viewed here:

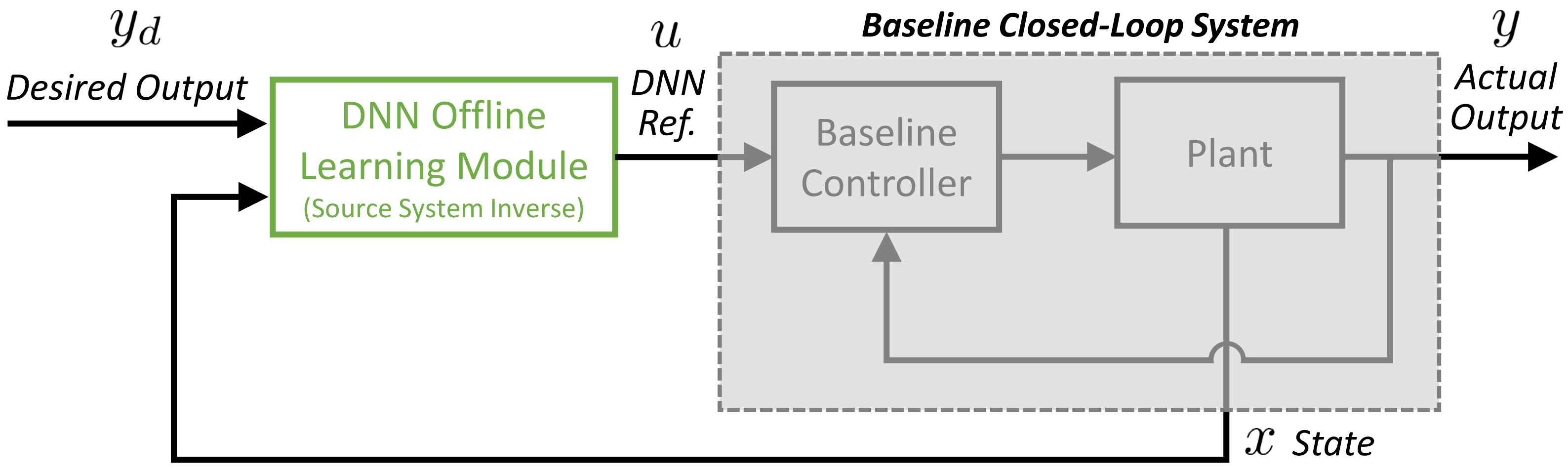
https://youtu.be/C_teLkJDq3Y

Introduction

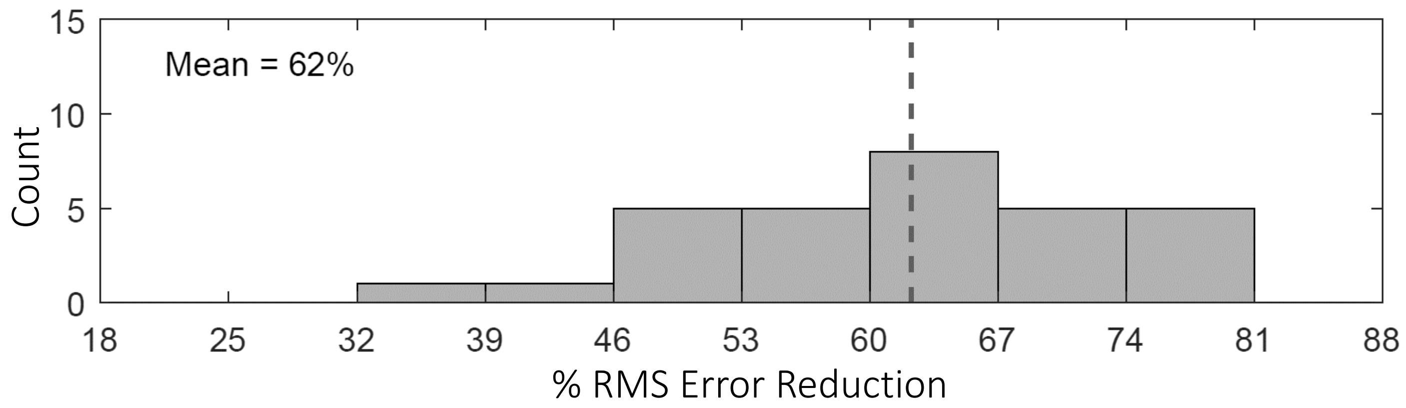
Neural networks as add-on blocks to enhance ‘black-box’ systems

DYNAMIC SYSTEMS LAB





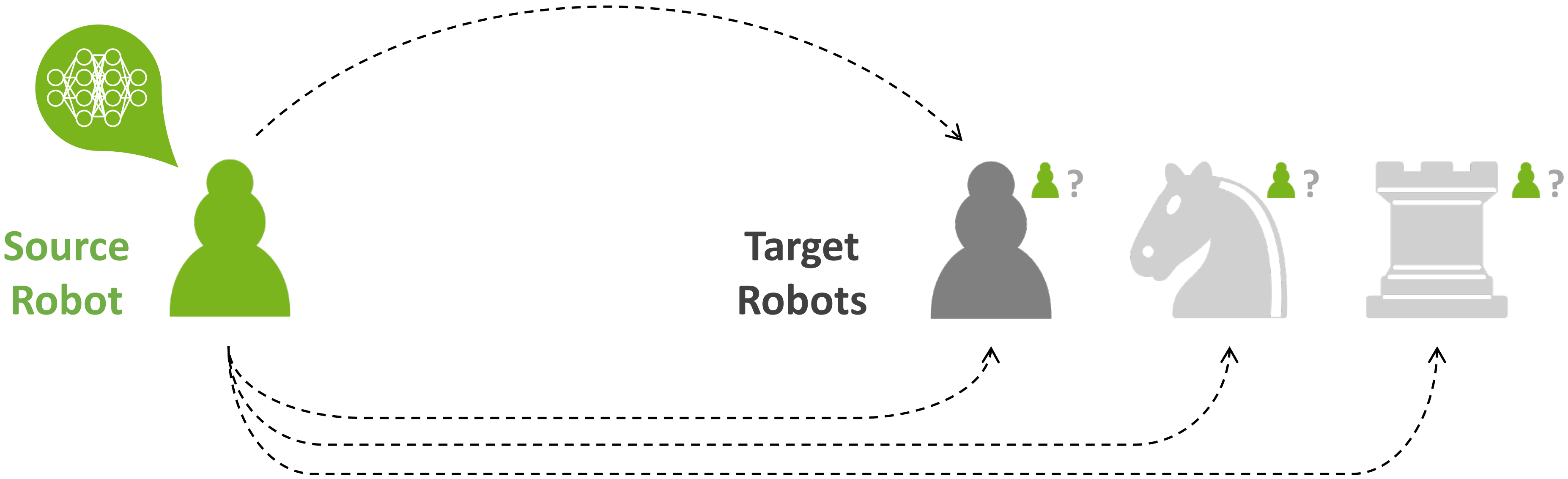
Average of 62% error reduction over 30 test trajectories



Research Question

What if we have a team of robots with different dynamics?

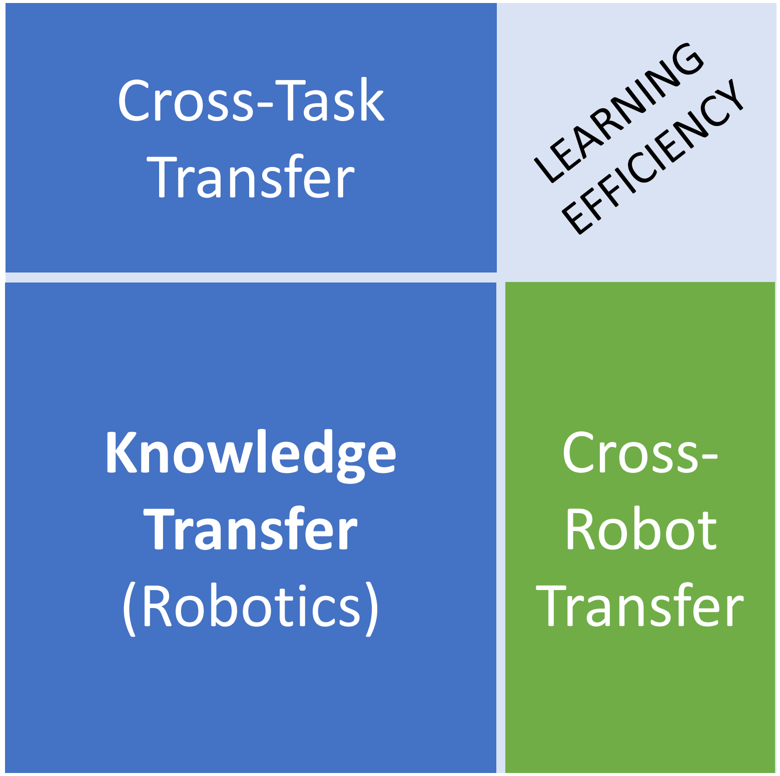
Research Question



Related Literature

Transfer experience to accelerate learning on new tasks or for new robots

Knowledge transfer: Leverage existing data or learned experience to accelerate or improve subsequent learning



Related Literature

Approaches for transferring data across robots

DYNAMIC SYSTEMS LAB

Knowledge transfer: Leverage existing data or learned experience to accelerate or improve subsequent learning

Cross-Task Transfer	LEARNING EFFICIENCY	
Knowledge Transfer (Robotics)	Cross-Robot Transfer	**Alignment-Based** Map from Source to Target (e.g., [Bócsi et al., 2013; Helwa & Schoellig, 2017])
		Invariant Feature Learning Exploiting Common Feature Space (e.g., [Gupta et al., 2017; Daftry et al., 2016])

Related Literature

Approaches for transferring data across robots

DYNAMIC SYSTEMS LAB

Knowledge transfer: Leverage existing data or learned experience to accelerate or improve subsequent learning

Cross-Task Transfer

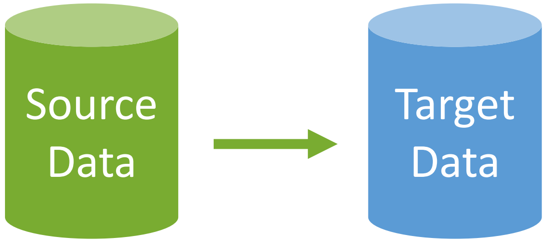
LEARNING EFFICIENCY

Knowledge Transfer (Robotics)

Cross-Robot Transfer

Alignment-Based
Map from Source to Target
(e.g., [Bócsi et al., 2013; Helwa & Schoellig, 2017])

Invariant Feature Learning
Exploiting Common Feature Space
(e.g., [Gupta et al., 2017; Daftry et al., 2016])



Related Literature

Approaches for transferring data across robots

Knowledge transfer: Leverage existing data or learned experience to accelerate or improve subsequent learning

Cross-Task Transfer

LEARNING EFFICIENCY

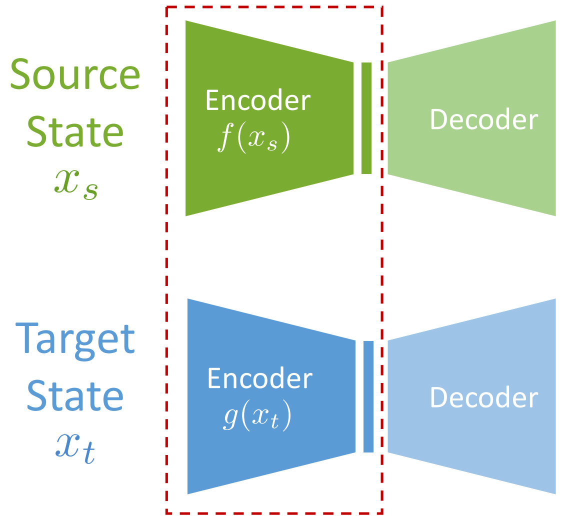
Knowledge Transfer (Robotics)

Cross-Robot Transfer

Alignment-Based
Map from Source to Target
(e.g., [Bócsi et al., 2013; Helwa & Schoellig, 2017])

Invariant Feature Learning
Exploiting Common Feature Space
(e.g., [Gupta et al., 2017; Daftry et al., 2016])

$$\min ||f(x_s) - g(x_t)||$$



[Gupta et al., 2017]

Related Literature

Maximizing learning efficiency on physical robots shares a broader interest

DYNAMIC SYSTEMS LAB

Knowledge transfer: Leverage existing data or learned experience to accelerate or improve subsequent learning

Cross-Task Transfer

LEARNING EFFICIENCY

Related Interests

- Sim-to-Real (e.g., [Marco et al., 2017])
- Meta-Learning (e.g., [Finn et al., 2017])
- Modularity (e.g., [Devin et al., 2017])
- …

Knowledge Transfer (Robotics)

Cross-Robot Transfer

Alignment-Based
Map from Source to Target
(e.g., [Bócsi et al., 2013; Helwa & Schoellig, 2017])

Invariant Feature Learning
Exploiting Common Feature Space
(e.g., [Gupta et al., 2017; Daftry et al., 2016])

Contributions

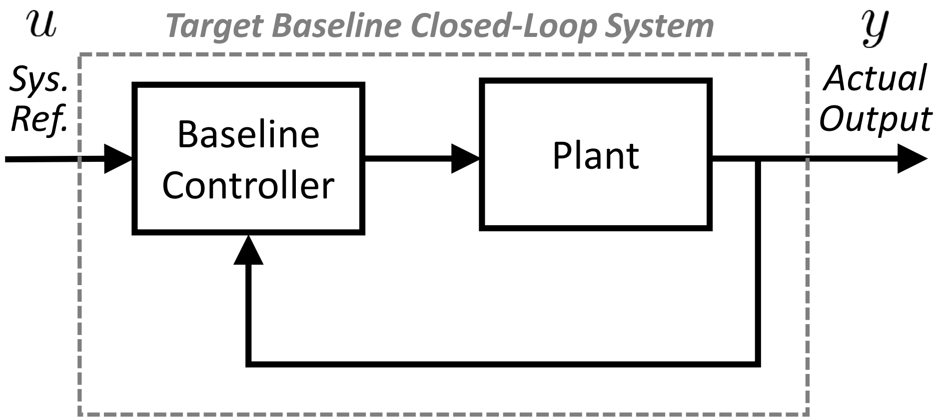
1. Impromptu knowledge transfer (i.e., without additional a-priori data collection on the robots)

2. Stability analysis of transfer-enhanced system and its connection to system similarity (linear case)

3. Verification of the knowledge transfer approach with quadrotors impromptu tracking experiments

Theoretical Results

Problem definition



Setup: Consider closed-loop source and target systems represented in the following form

$$x(k+1) = f(x(k)) + g(x(k))\,u(k)$$
$$y(k) = h(x(k))$$

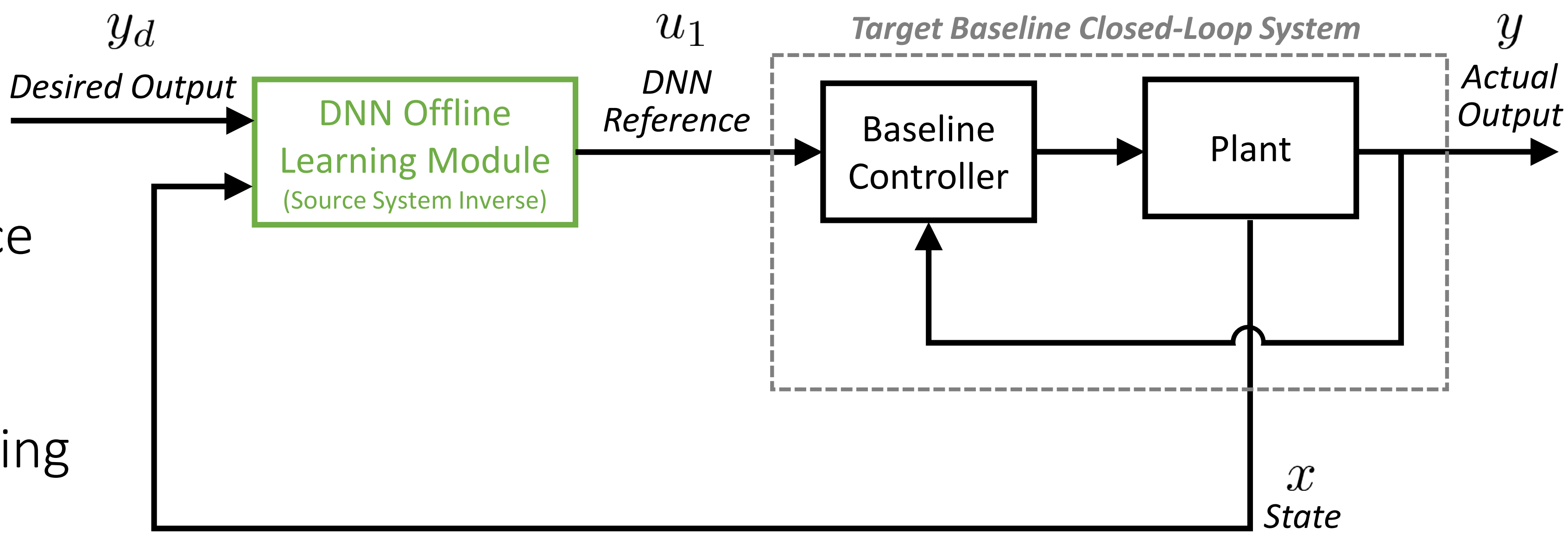
Assumption: The source and the target systems

a) are minimum phase
b) have well-defined and the same relative degree

Goal: To enhance the target baseline system with minimal amount of data (re)collection and training

Theoretical Results

Leveraging the DNN inverse module from the source system



How to leverage the source DNN model?

- Update source DNN
- Online correction learning

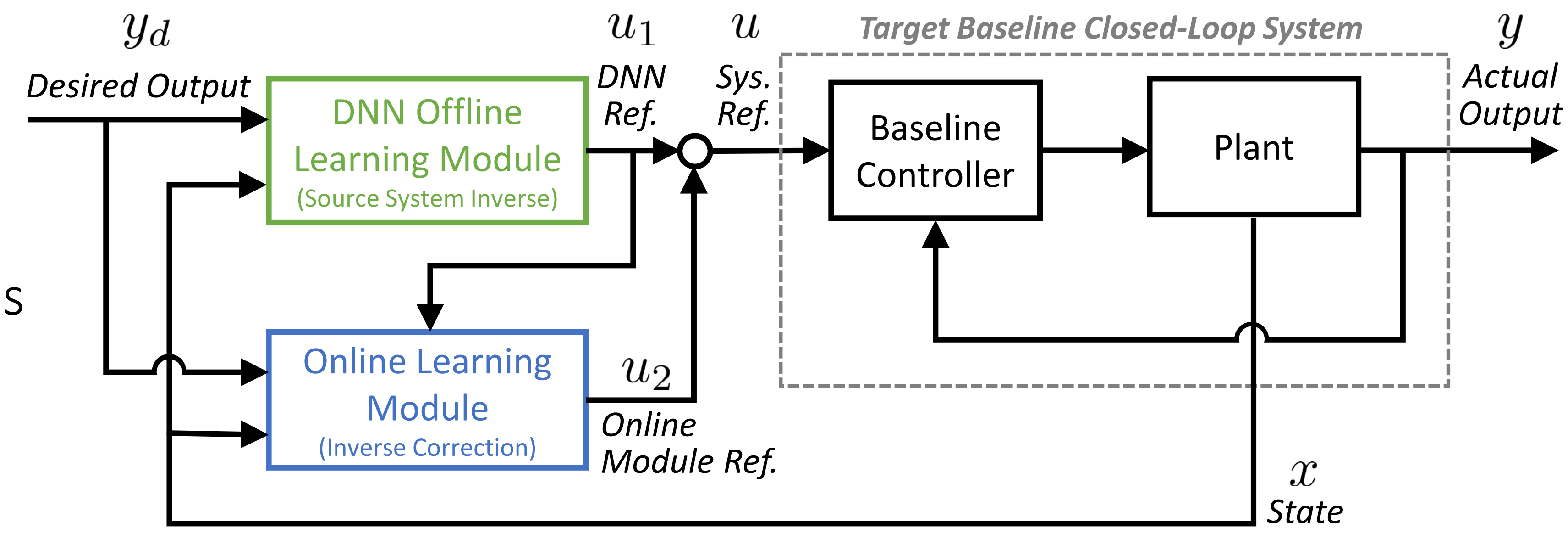
Offline Learning Module Approximates Inverse of the Source Robot System [CDC 17]

$$u_1(k) = \Big(\mathcal{G}_s\left(x(k)\right)\Big)^{-1}\left(y_d(k+r) - \mathcal{F}_s\left(x(k)\right)\right) \longrightarrow u_1(k) = F_{\text{nn}}(x(k),\ y_d(k+r))$$

approximated by a DNN
(when $\mathcal{F}_s$ and $\mathcal{G}_s$ are unknown)

Theoretical Results

Using online learning to adapt to the differences



Online Learning Module for Reference Adjustments

$$u_2(k) = \alpha \; e_p(k+r)$$

Adaptation Gain

Error Prediction

Ideal Expressions for Exact Tracking

$$\alpha^* = (\mathcal{G}_t(x(k)))^{-1}$$

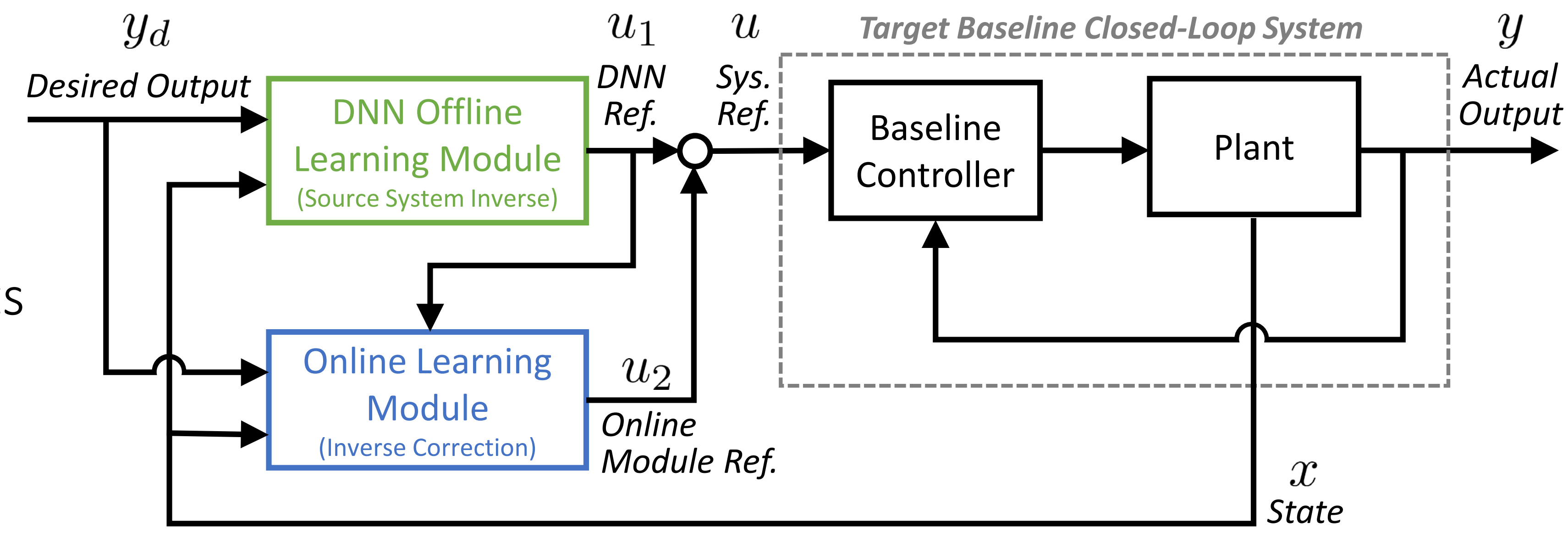
$$e_p^*(k+r) = y_d(k+r) - y_p(k+r)$$

$$y_p(k+r) = \mathcal{F}_t\,(x(k)) + \mathcal{G}_t\,(x(k))\,u_1(k)$$

Predicted output of target system when u_1 is sent to the system

Theoretical Results

Using online learning to adapt to the differences



Online Learning Module for Reference Adjustments

$$u_2(k) = \alpha \; e_p(k+r)$$

Adaptation Gain

Error Prediction

Ideal Expressions for Exact Tracking

$$\alpha^* = \left(\mathcal{G}_t(x(k))\right)^{-1}$$

$$e_p^*(k+r) = y_d(k+r) - y_p(k+r)$$

Online Learning of Error Predictor

Online Training Dataset (Based on Latest Observations)

$$e_p(k+r) = F_{\text{gp}}(x(k),\, y_d(k+r),\, u_1(k))$$

$$\mathcal{D} = \left\{\left(x(p-r),\, y_d(p),\, u(p-r)\right);\; \left(\, y_d(p) - y(p)\right)\right\}_{p=k-N}^{p=k}$$

Theoretical Results

Characterizing similarity between the source and the target systems

DYNAMIC SYSTEMS LAB

Linear Case

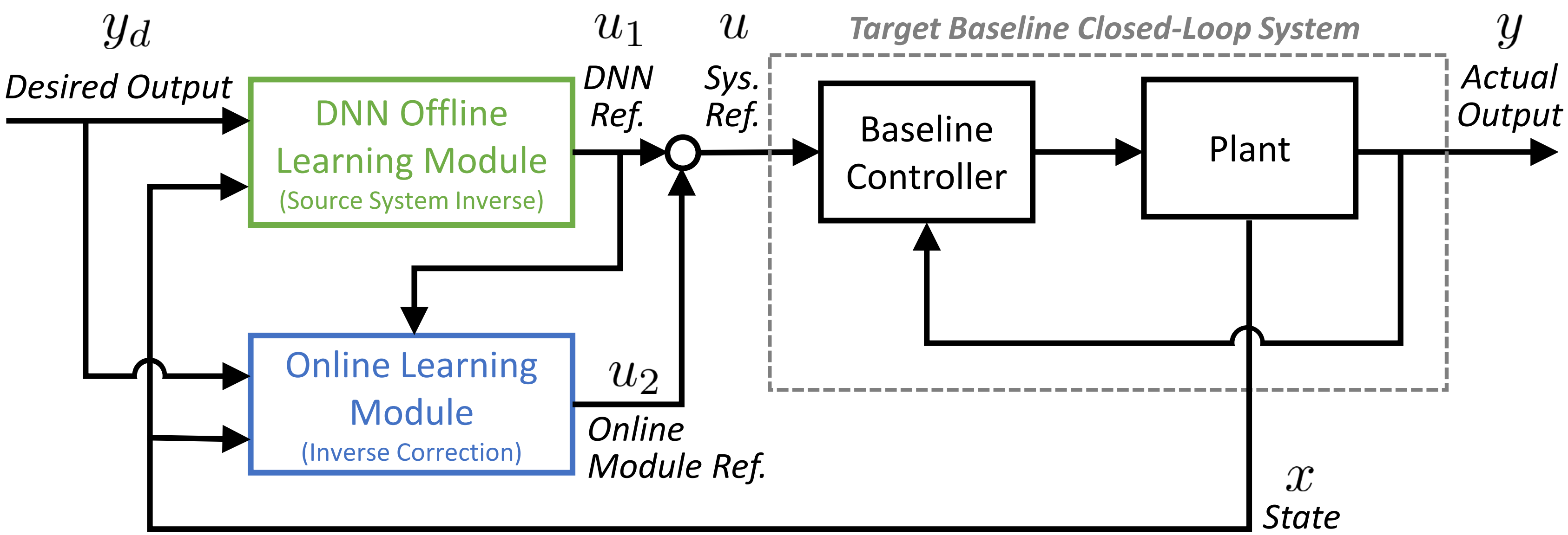
$$x(k+1) = Ax(k) + Bu(k)$$
$$y(k) = Cx(k)$$

Input-Output Equation

$$y(k+r) = \mathcal{A}x(k) + \mathcal{B}u(k),$$

where $\mathcal{A} = CA^r$ and

$\mathcal{B} = CA^{r-1}B$



Similarity Characterization

$$S = \left[1 - \frac{\mathcal{B}_t}{\mathcal{B}_s} \quad \underset{\text{Target}}{\mathcal{A}_t} - \frac{\mathcal{B}_t}{\mathcal{B}_s}\underset{\text{Source}}{\mathcal{A}_s}\right],$$

$\mathcal{A}$ = state-to-output gain

$\mathcal{B}$ = input-to-output gain

Theoretical Results

Higher similarity leads to higher tolerances for learning error

Assumptions

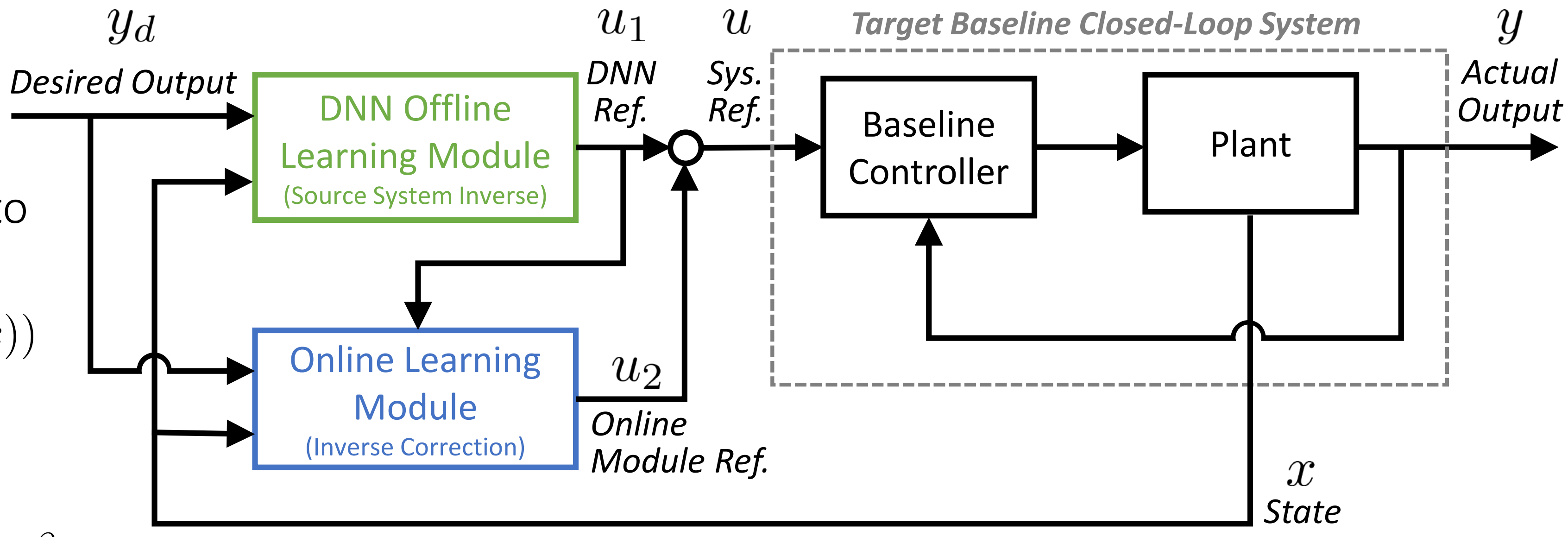
1. Input-to-state stable

2. Offline module corresponds to the source inverse

$u_1(k) = \mathcal{B}_s^{-1}\left(y_d(k+r) - \mathcal{A}_s x(k)\right)$

3. Error of the online learning module is bounded as

$\Lambda \leq \beta_1||y_d(k+r)|| + \beta_2||x(k)|| + \beta_3$



Stability of the Overall Learning-Enhanced Target System

$|\alpha|\left(||S_2|| + \beta_2\right) < \beta_4/L_1$

when $\alpha \neq 0$

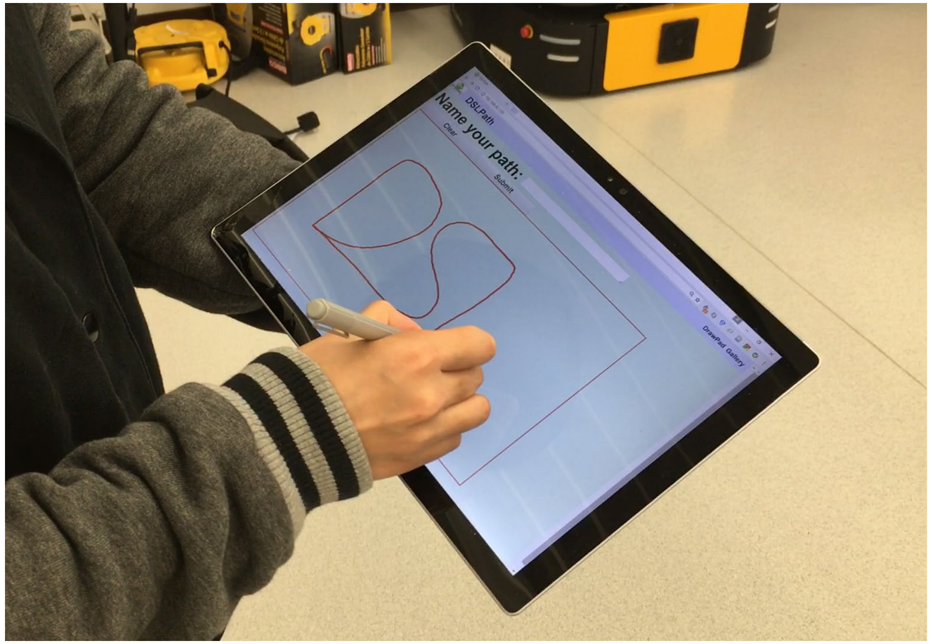
when $\alpha = 0$

(i.e., online learning module is not active)

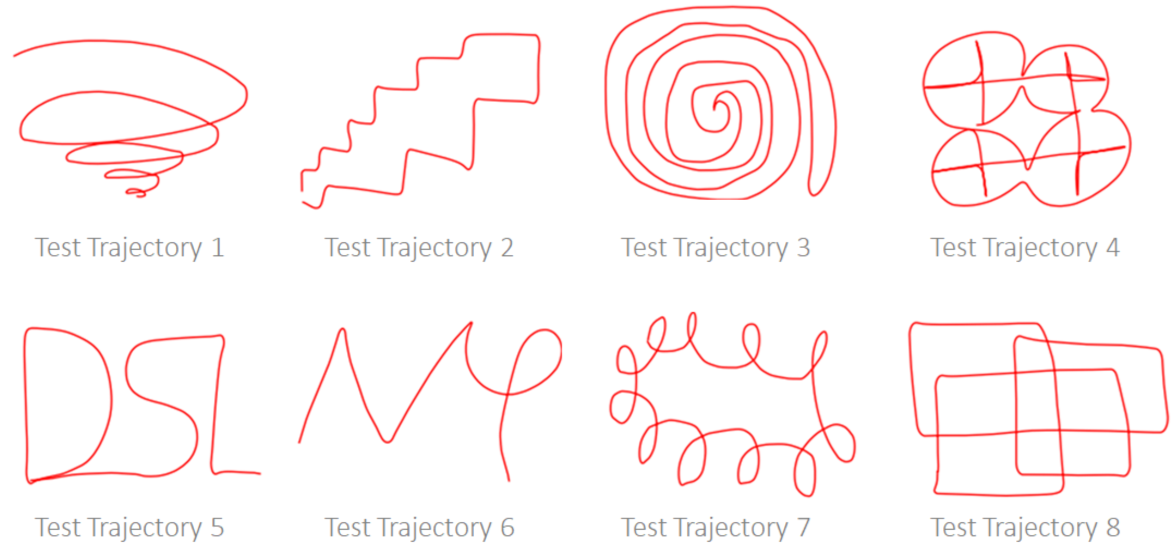
$L_1||\mathcal{A}_s/\mathcal{B}_s|| < 1$

Experiments

We test our online learning approach on arbitrary hand drawings



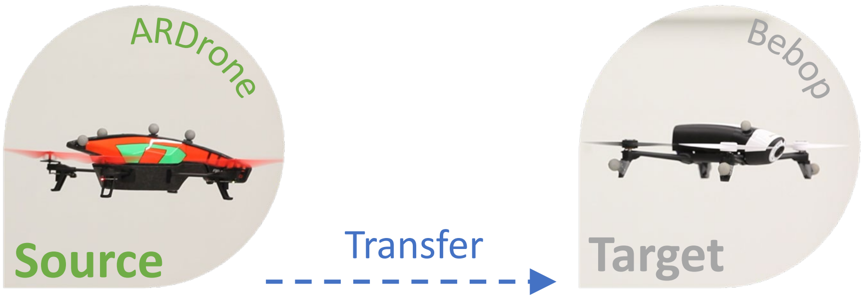
Samples of Arbitrary Hand-Drawn Test Trajectories



Experiments

We test our online learning approach on arbitrary hand drawings

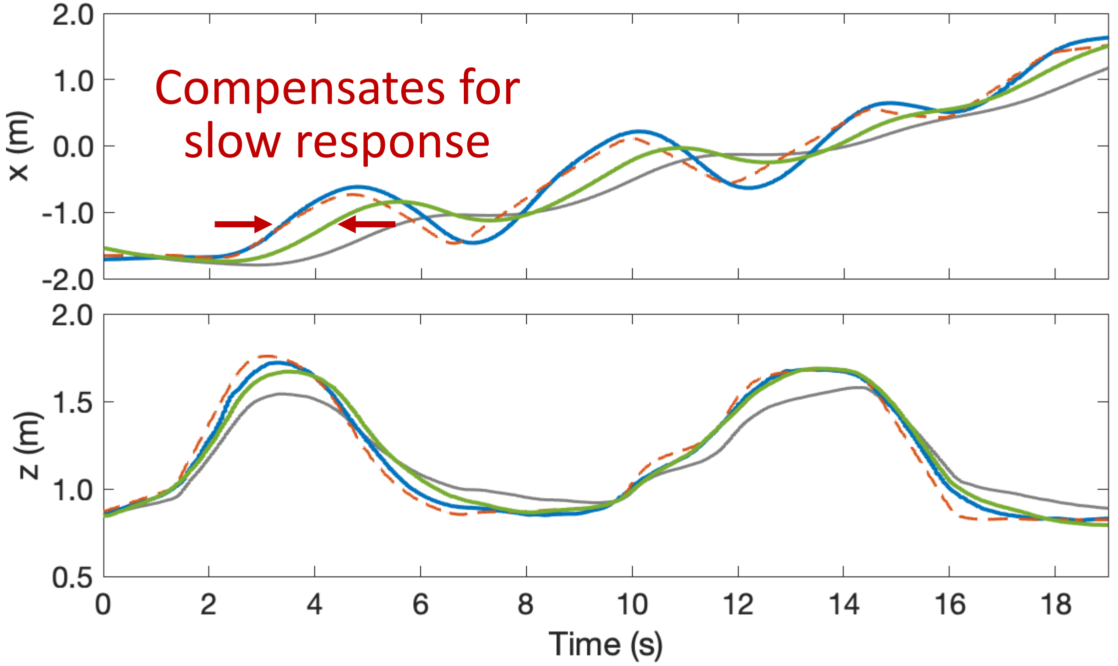
DYNAMIC
SYSTEMS LAB



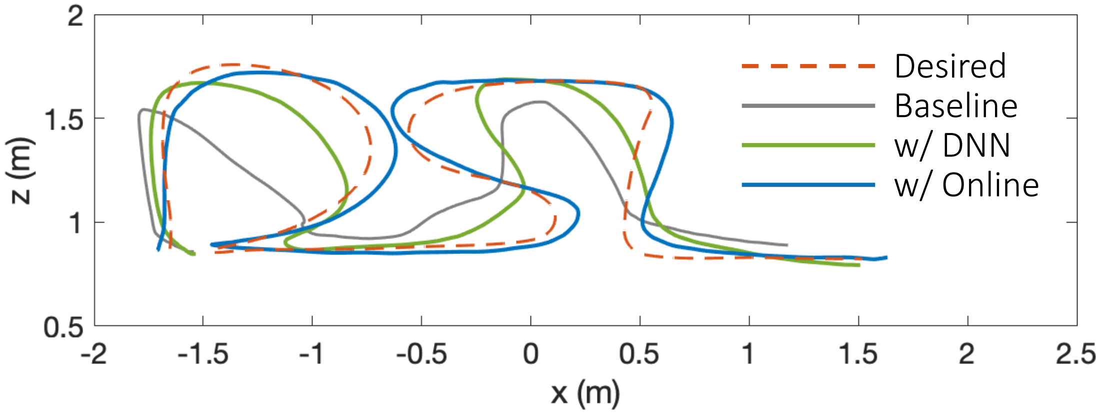
Trajectories in the x and z Directions

With offline transfer alone: 38% error reduction

With online transfer: 67% error reduction

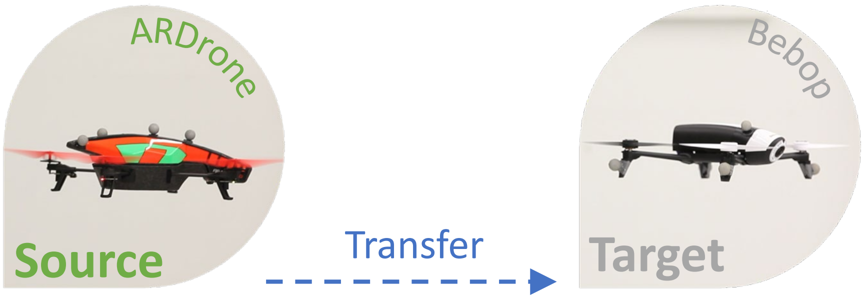


Path in the x-z Plane



Experiments

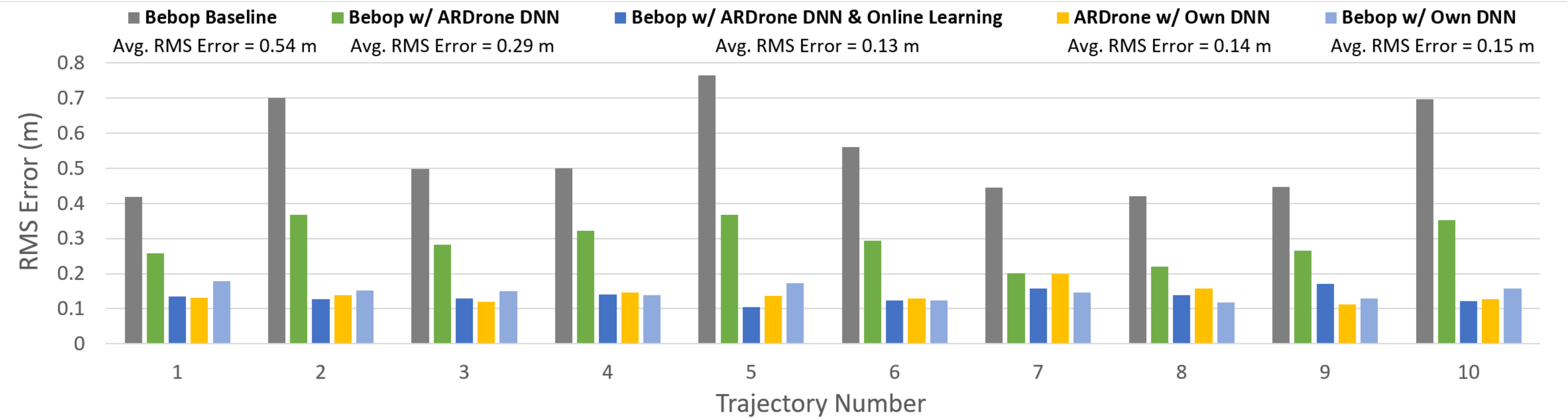
We can effectively reduce the amount of data required for training robots



With offline transfer alone: 46% error reduction

With online transfer: 74% error reduction

(Comparable to fully-trained DNNs)



Summary

DNN inverse for tracking performance enhancement of single robots [ICRA17, CDC17]

Online learning approach for impromptu cross-robot transfer of previously trained DNNs

Connection between system similarity and stability of target system enhanced with online learning

Performance improvement of 74% with online learning in quadrotor impromptu tracking tasks

Summary

DYNAMIC SYSTEMS LAB

DNN inverse for tracking performance enhancement of single robots [ICRA17, CDC17]

Online learning approach for impromptu cross-robot transfer of previously trained DNNs

Connection between system similarity and stability of target system enhanced with online learning

Performance improvement of 74% with online learning in quadrotor impromptu tracking tasks

Thanks

a light-writing by

Institute for Aerospace Studies
UNIVERSITY OF TORONTO