NAZARBAYEV UNIVERSITY

SCHOOL OF ENGINEERING AND DIGITAL SCIENCES

PhD Thesis Defence

“Model Predictive Control and Imitation Learning Algorithms for Robot Motion Planning in Physical Human-Robot Interaction”

by

Aigerim Nurbayeva

August 7th, 2024 (Wednesday)

at 4.00 pm

via Microsoft Teams

NAZARBAYEV
UNIVERSITY

Model Predictive Control and Imitation Learning Algorithms for Robot Motion Planning in Physical Human-Robot Interaction

PhD candidate: Aigerim Nurbayeva

RCL LAB
NAZARBAYEV UNIVERSITY | Robot Control and Learning

NAZARBAYEV UNIVERSITY

RCL LAB
NAZARBAYEV UNIVERSITY | Robot Control and Learning

Supervisors

- Lead Supervisor

Matteo Rubagotti, Associate Professor,
Nazarbayev University, School of Engineering and Digital Sciences
Department of Robotics

- Co-Supervisor

Almas Shintemirov, Associate Professor,
Nazarbayev University, School of Engineering and Digital Sciences
Department of Robotics

- External Co-Supervisor

Gian Paolo Incremona, Associate Professor,
Politecnico di Milano, Milan, Italy
Department of Electronics, Information and Bioengineering
(Control Systems Group)

NAZARBAYEV UNIVERSITY

RCL LAB
NAZARBAYEV UNIVERSITY | Robot Control and Learning

Outline

- Introduction
- Imitation learning
- Nonlinear MPC with set terminal constraint
- Conclusions

Introduction

Introduction

Collaborative Robot (Cobot) - a robot intended for direct human robot interaction within a shared workspace

Figure 1. Robot workplace with collaborative (left figure) and traditional industrial robot (right figure) (Vysocky, A. L. E. S., & Novak, P. , 2016)

Speed and Separation Monitoring (SSM)

Figure 1. Speed and separation monitoring (Figure adapted from Villani et al., 2018)

The technical specification ISO/TS 15066:2016 defines principles and requirements for the design of Human Robot Collaboration (HRC) applications

In SSM, the robot maintains a certain maximum speed depending on its distance with the operator, to avoid any harmful physical contacts with the human

Model Predictive Control for Workspace Sharing

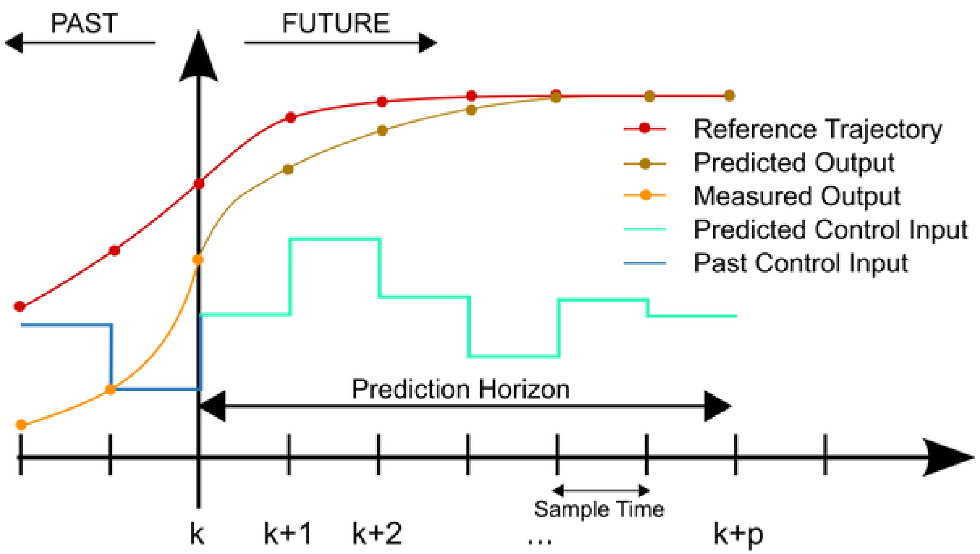
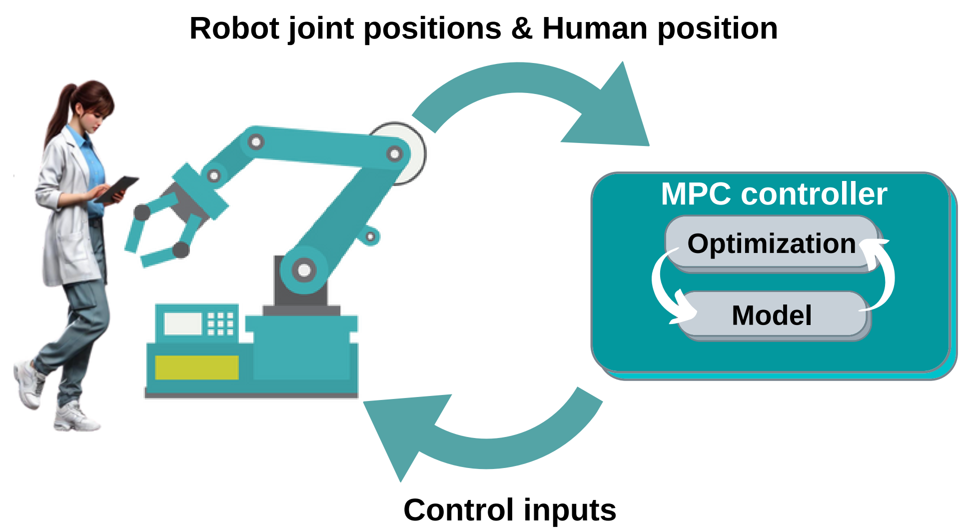


Figure 2. Model Predictive Control Scheme



Problem

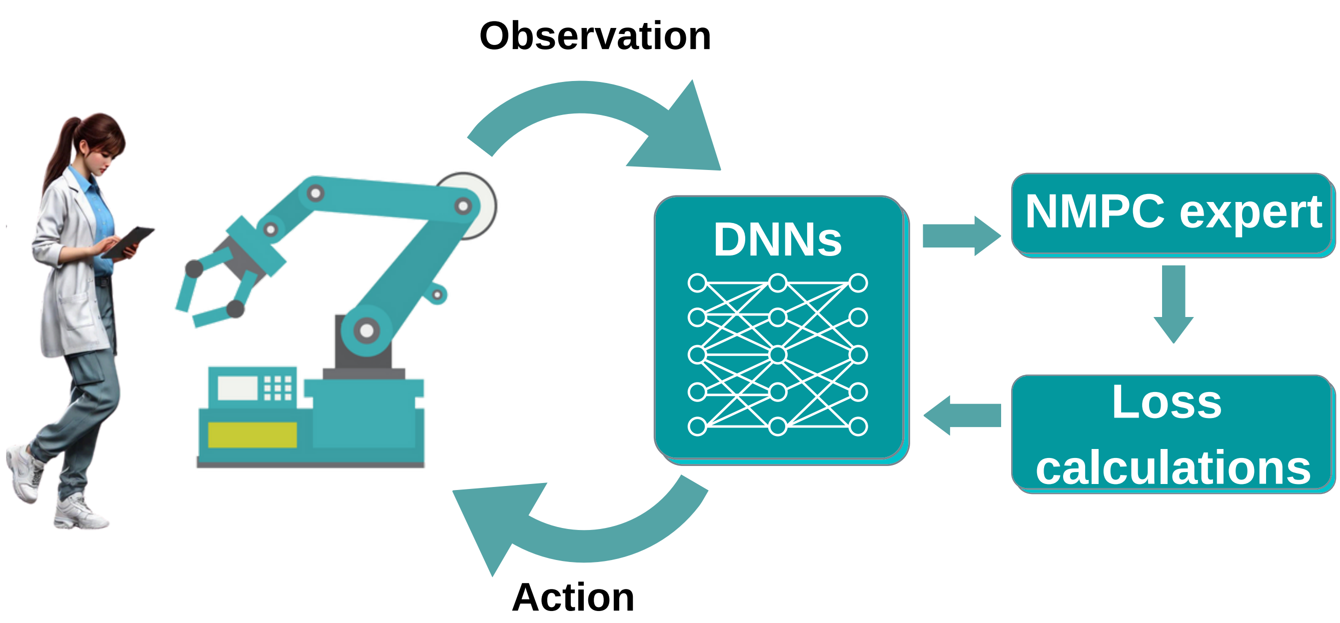
A robot has to complete point-to-point motions in the presence of a human, satisfying SSM constraints

Optimal Control Problem

- State space x
- Control inputs u
- Stage cost $\ell(x,u) = \|x - x_g\|_Q^2 + \|u\|_R^2 + \gamma\varphi^2(x, x_g)$
- Constraints
  limits on joint positions and speeds
  avoidance of fixed obstacles
  SSM constraints

Imitation learning

Robot manipulator learns how to act by observing the expert's demonstration, and skills can be generalized to other unseen scenarios



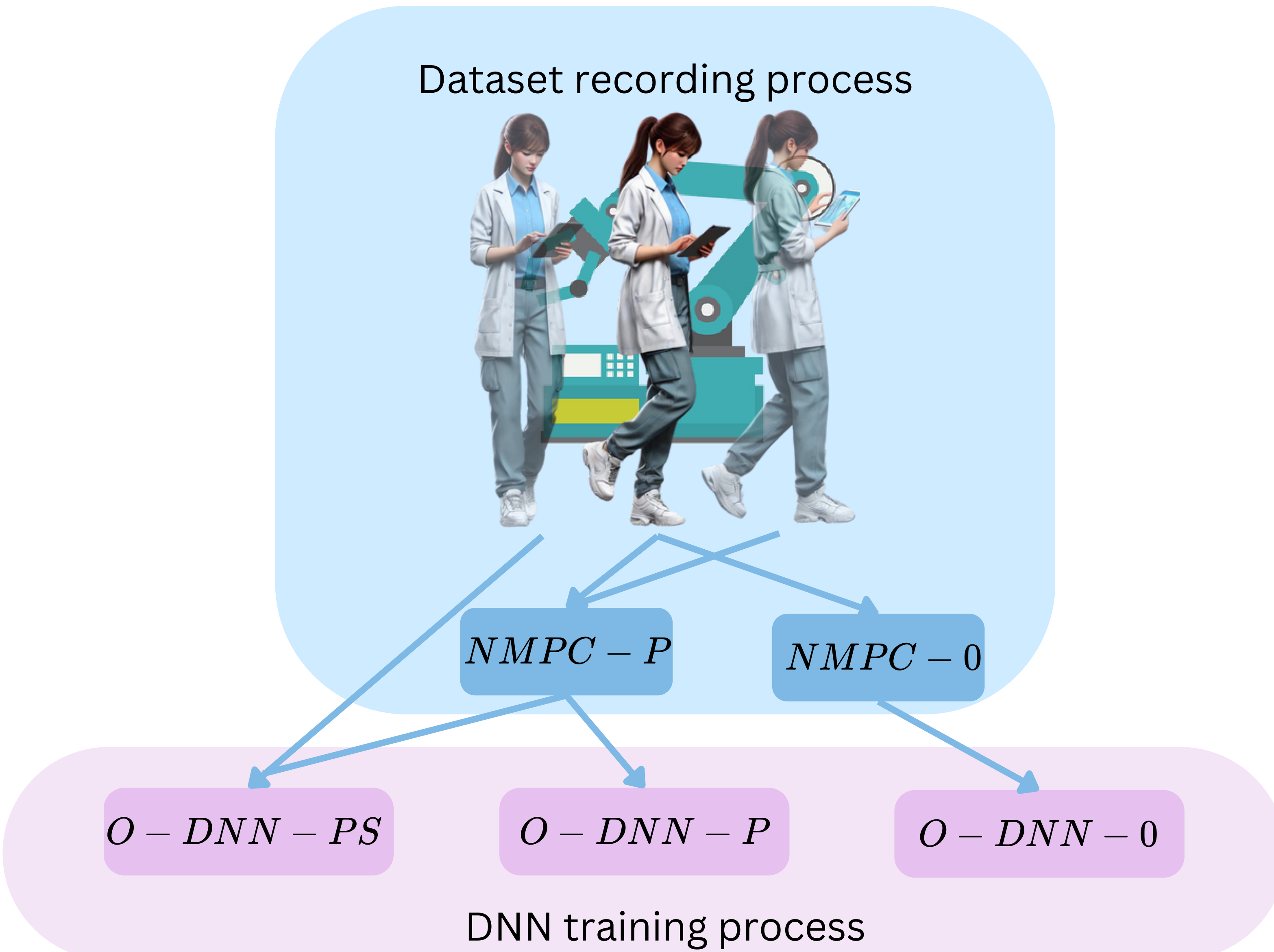
Research Questions

- Is it possible to improve the performance of SSM-based NMPC laws, at the same time guaranteeing human safety, by emulating their behavior through DNNs?

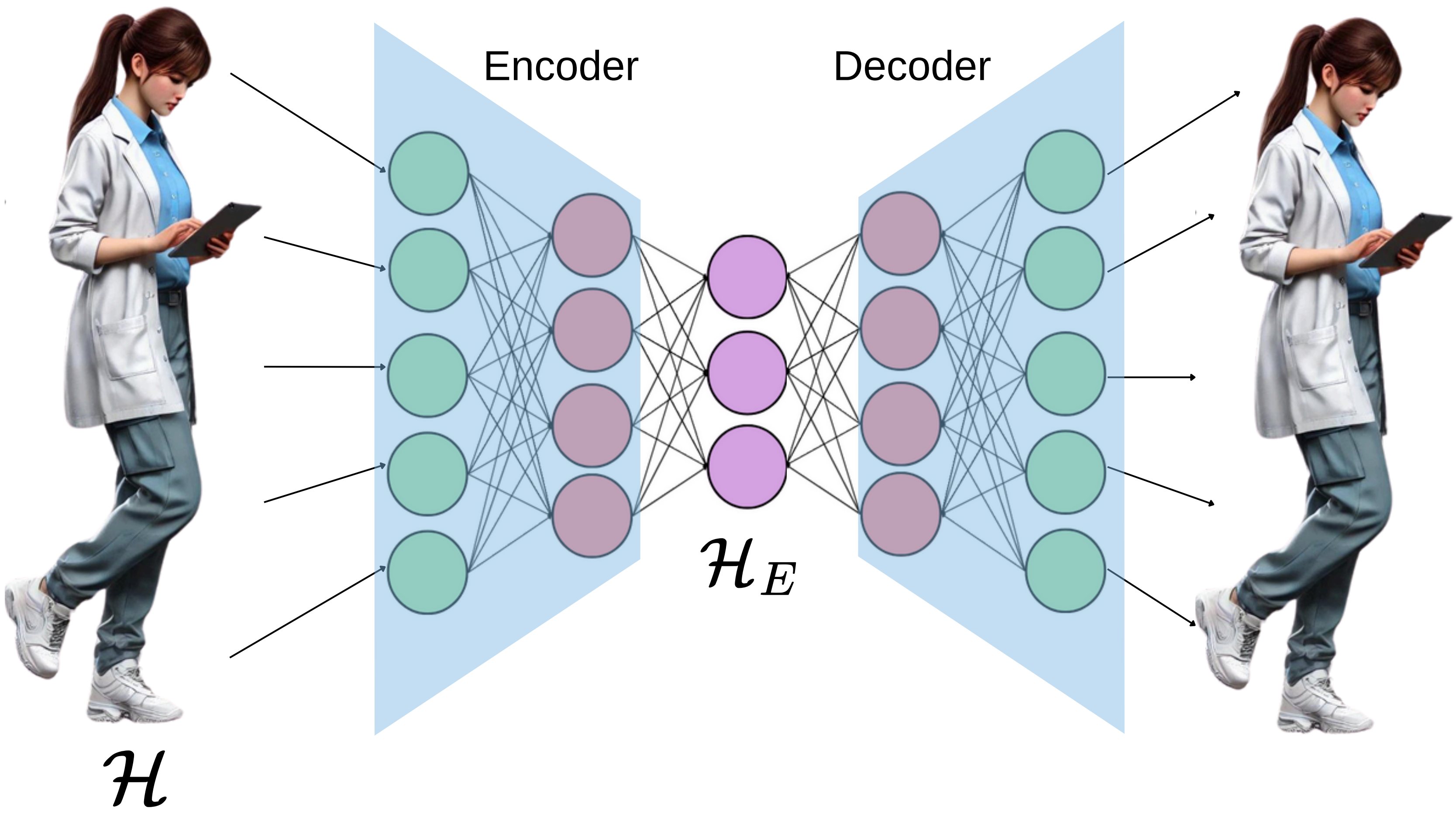
- Is it possible to define and design an SSM-based NMPC law with guaranteed stability properties and based on a set terminal constraint, which increases the domain of attraction compared to the point terminal constraint approach?

Imitation learning

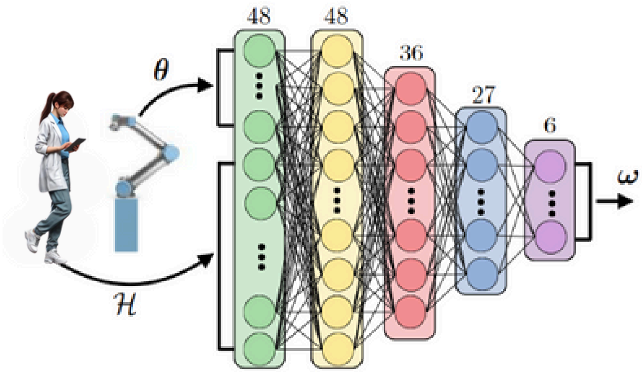
Imitation learning: an offline approach



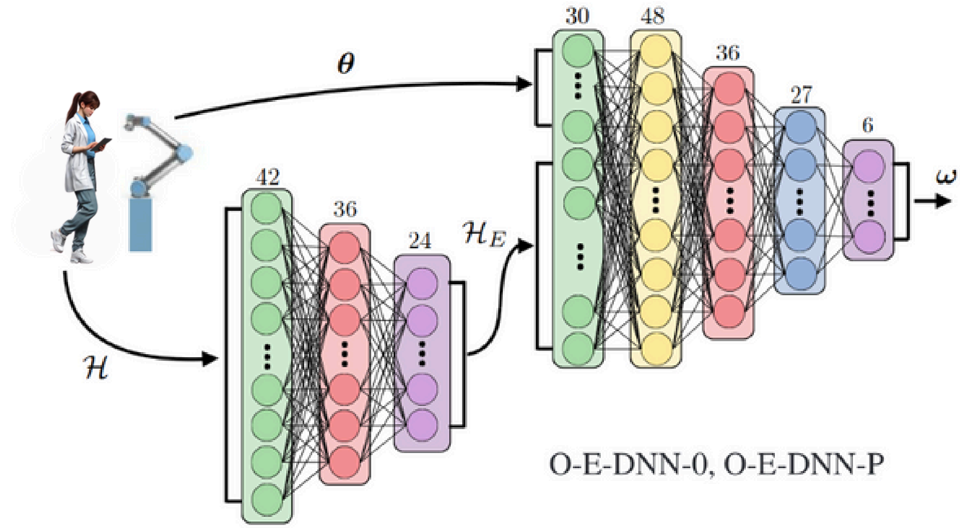
Autoencoders



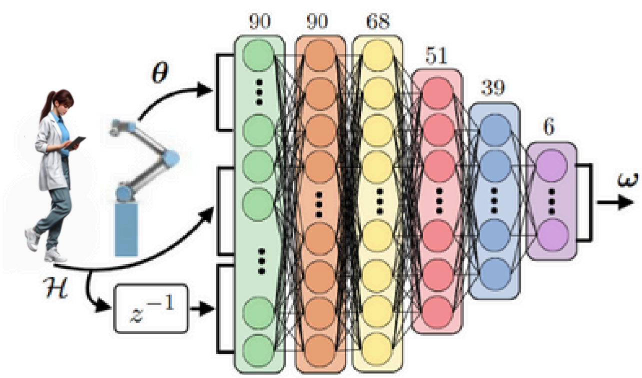
DNN structures



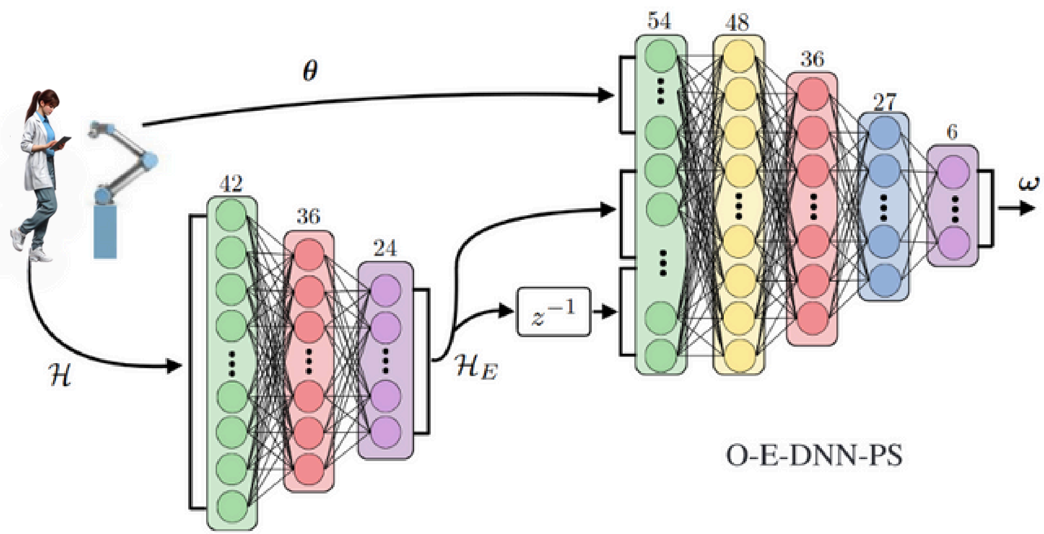
O-DNN-0, O-DNN-P



O-E-DNN-0, O-E-DNN-P



O-DNN-PS



O-E-DNN-PS

Case study

6 DOF UR5 robot

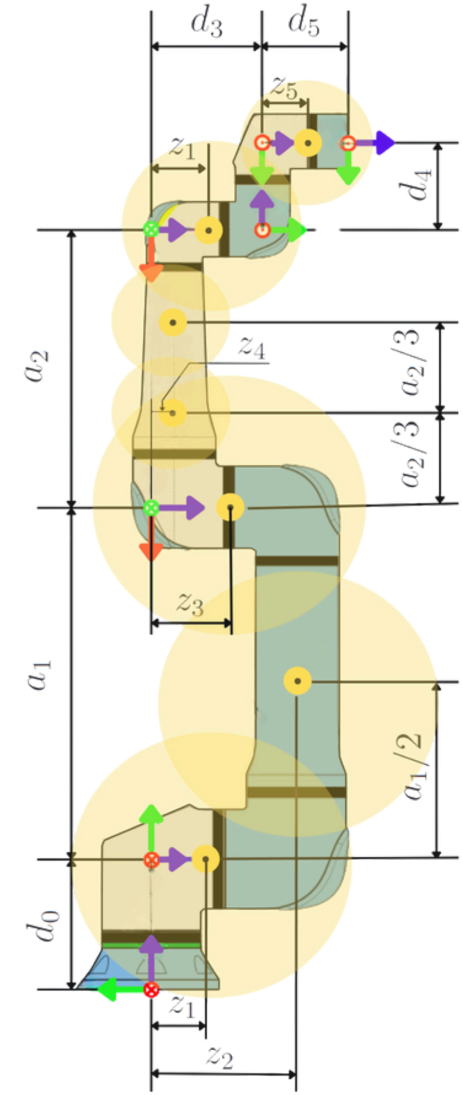
A robot has to complete point-to-point motions in the presence of a human, satisfying SSM constraints

$$x_A = [0.0 \quad -2.3 \quad -1.1 \quad -1.2 \quad -1.2 \quad 0.5]'$$

$$x_B = [3.0 \quad -1.6 \quad -1.7 \quad -1.7 \quad -1.7 \quad 1.0]'$$

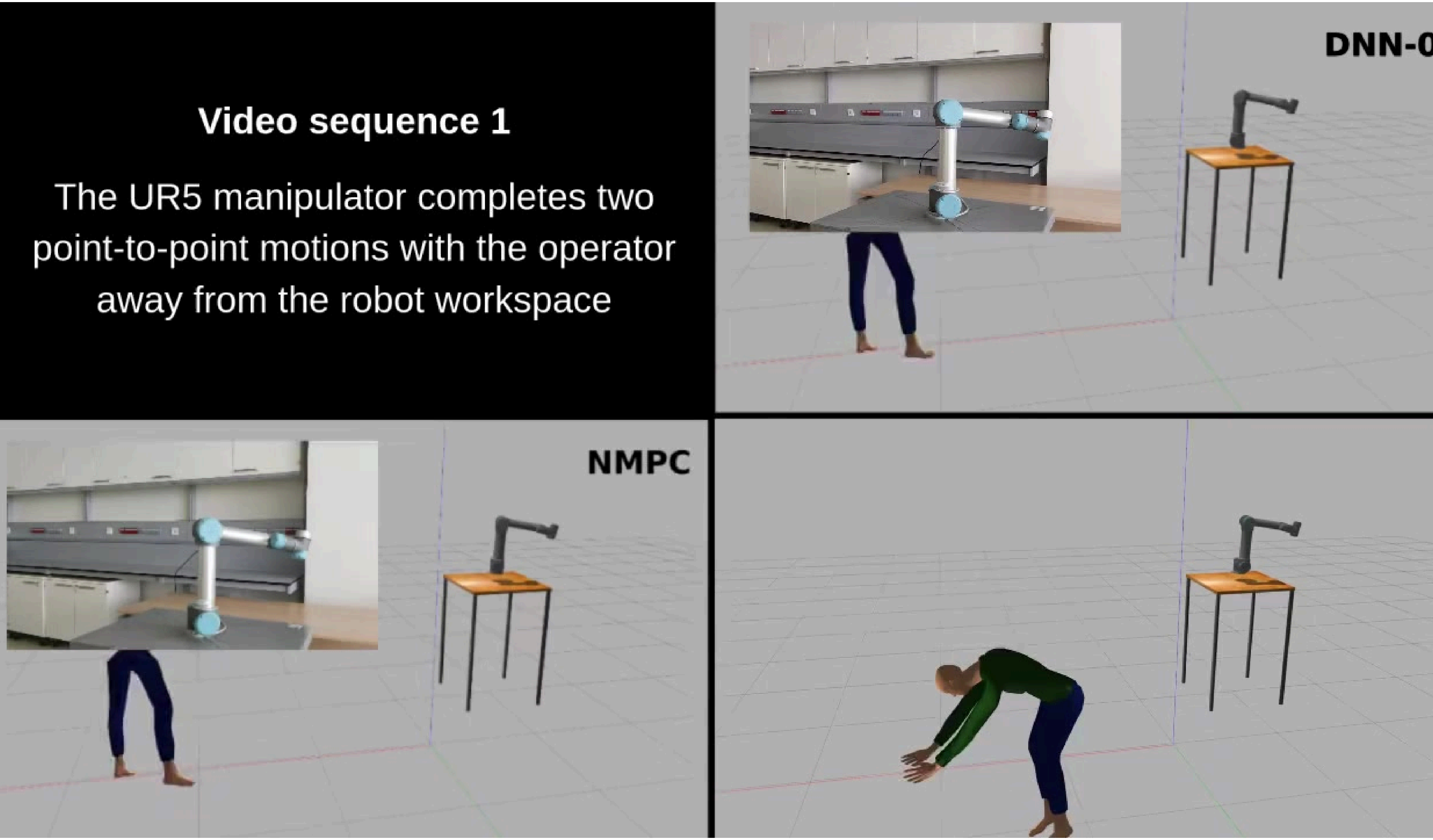
Robot - 7 test points

Human - 14 test points



Results

Controllers	Task completion time, s	Cost	Maximum solver time, ms	Average solver time, ms
NMPC (acado)	22.74	6588.8	**32.5**	**14.0**
NMPC (acados)	**17.58**	**5311.7**	162.2	39.9
O-DNN-O	15.71	5098.7	24.6	**1.7**
O-DNN-P	**15.6**	**5081.3**	**11.8**	1.8
O-DNN-PS	16.12	5348.4	22.9	1.9
O-E-DNN-O	15.74	**4871.8**	**26.0**	2.8
O-E-DNN-P	**15.56**	4993.8	26.5	**2.2**
O-E-DNN-PS	15.82	5057.3	39.6	3.2


Video sequence 1
The UR5 manipulator completes two point-to-point motions with the operator away from the robot workspace
DNN-0
NMPC

Results

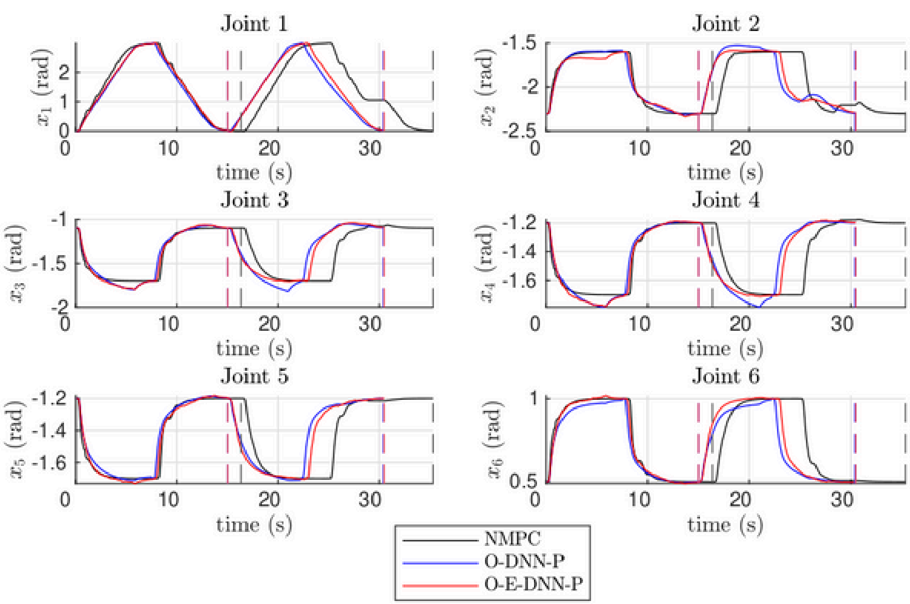


Figure 2. Time evolution of joint positions

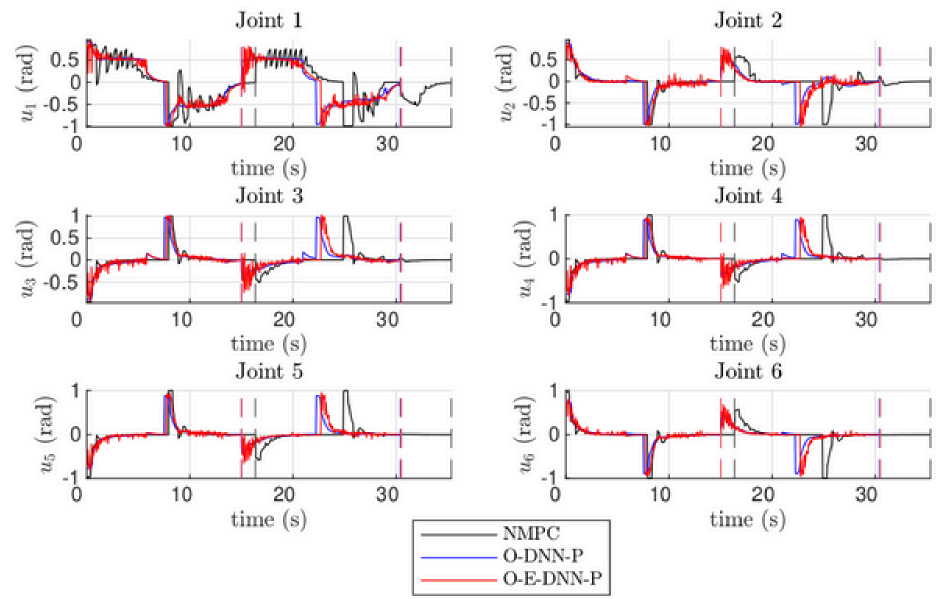


Figure 3. Time evolution of joint speeds

Results

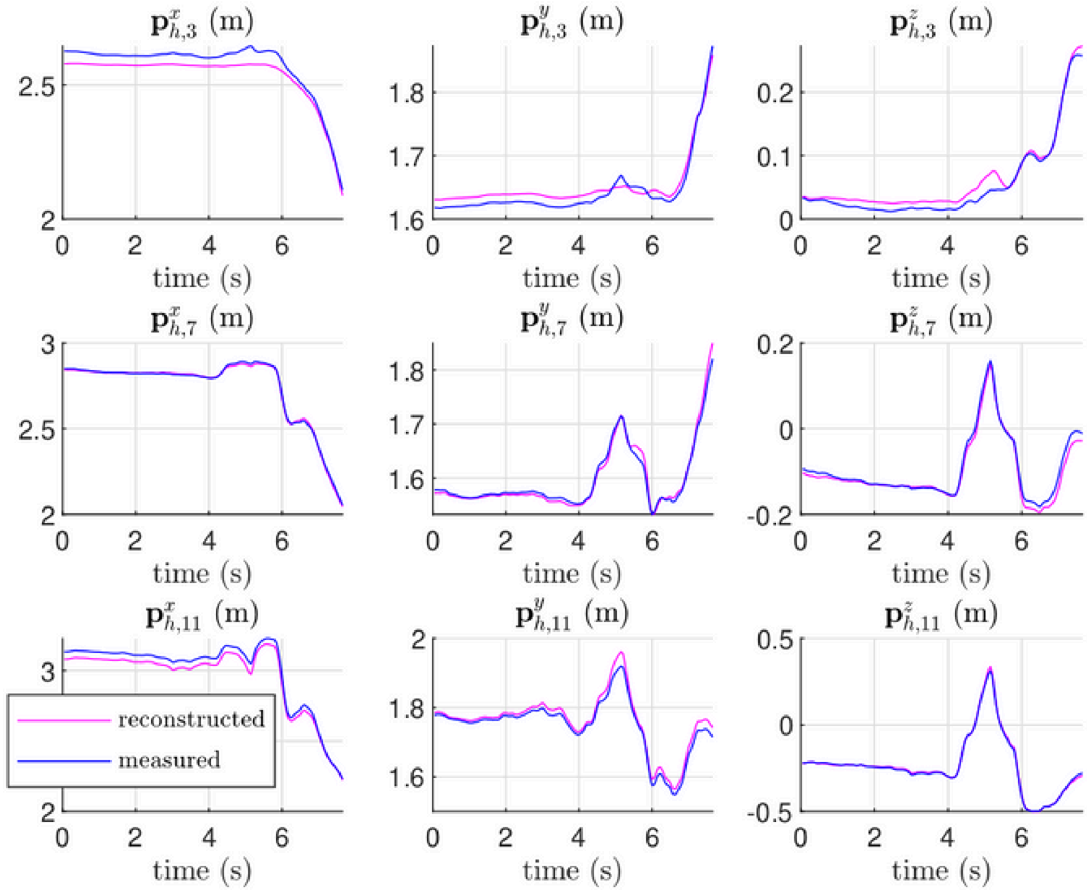


Figure 4. Time evolution of inputs and outputs of the autoencoders

Results

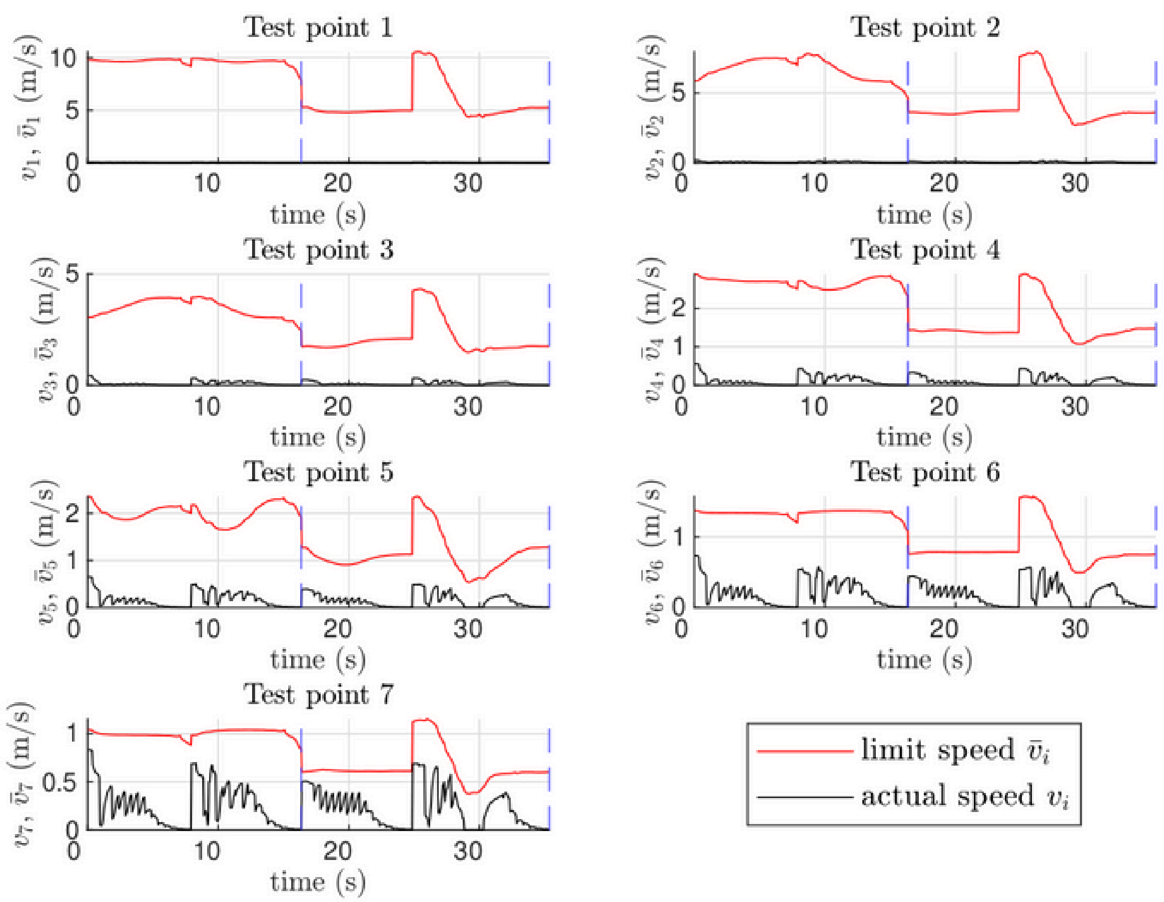


Figure 5. Time evolution of robot test points for the NMPC (acados)

Results

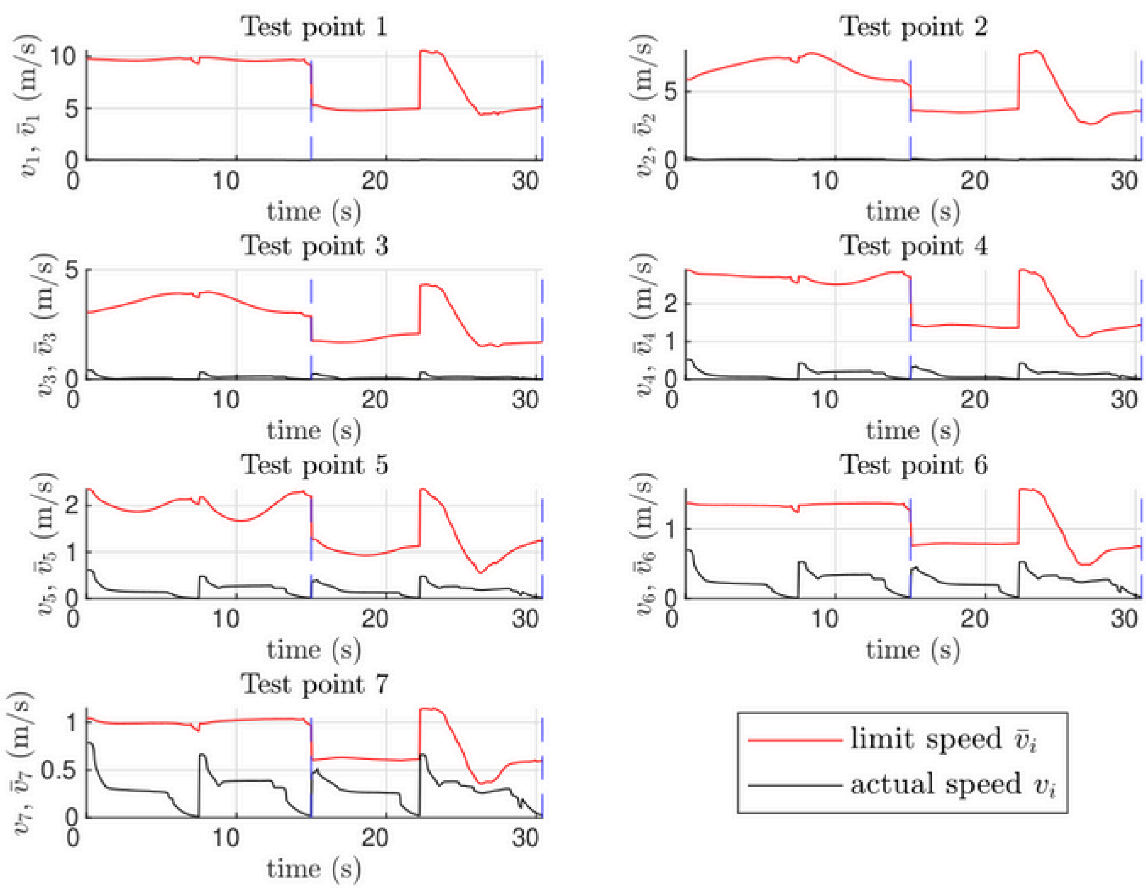


Figure 6. Time evolution of robot test points for the O-DNN-P

Results

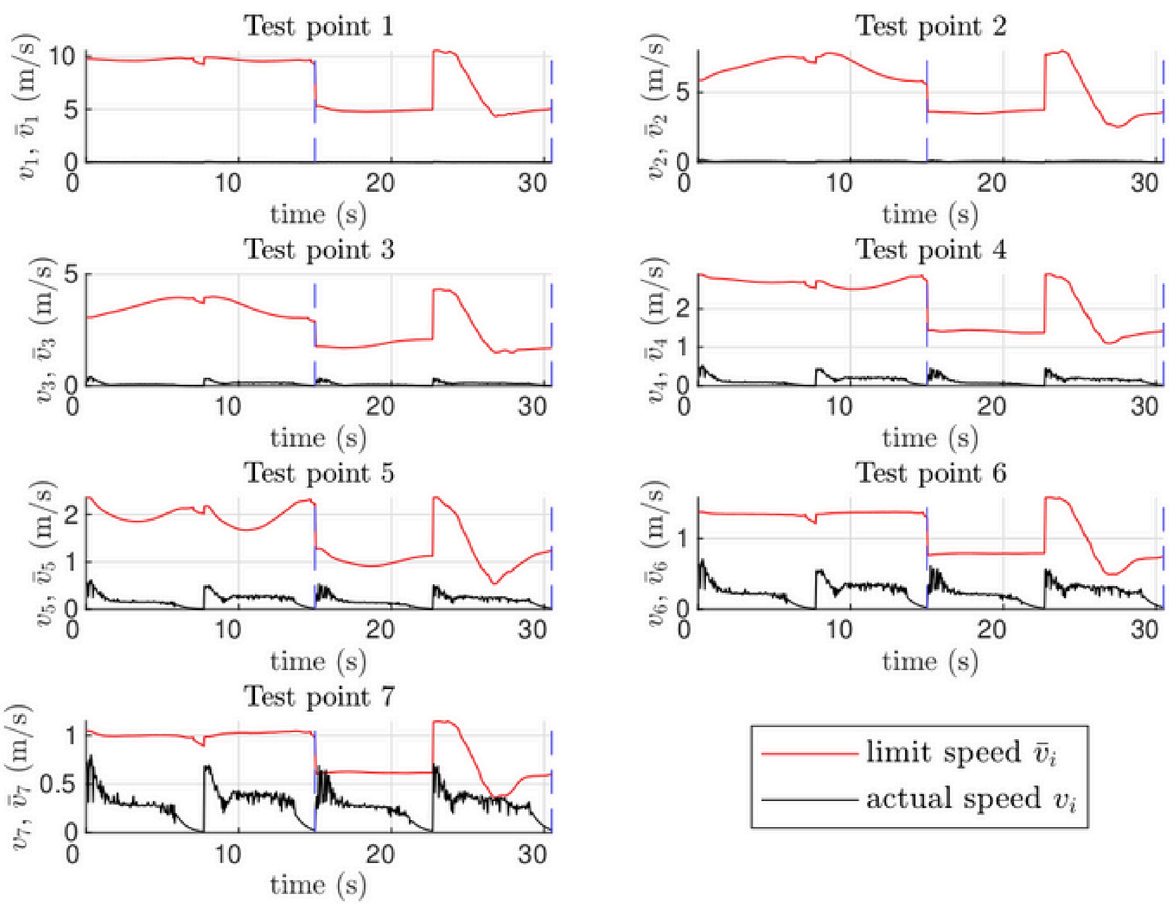
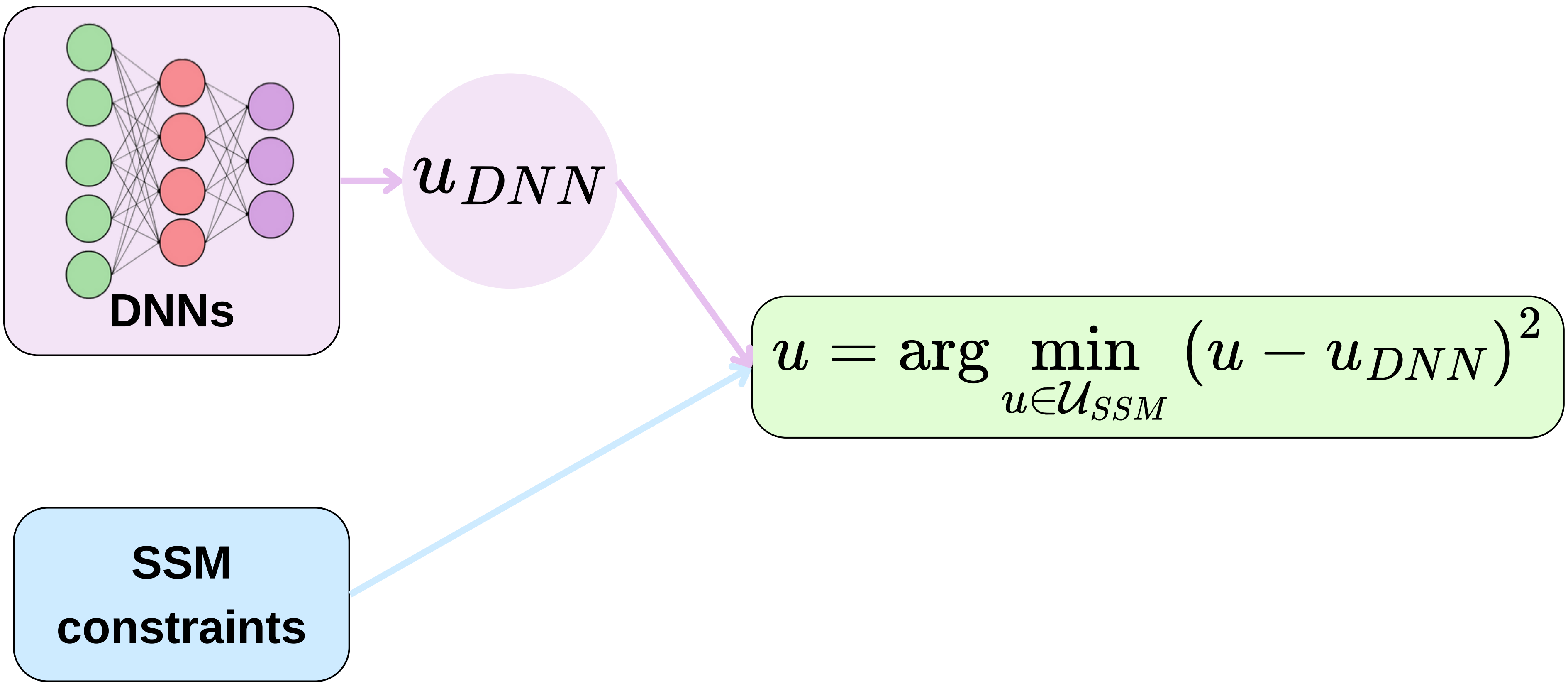
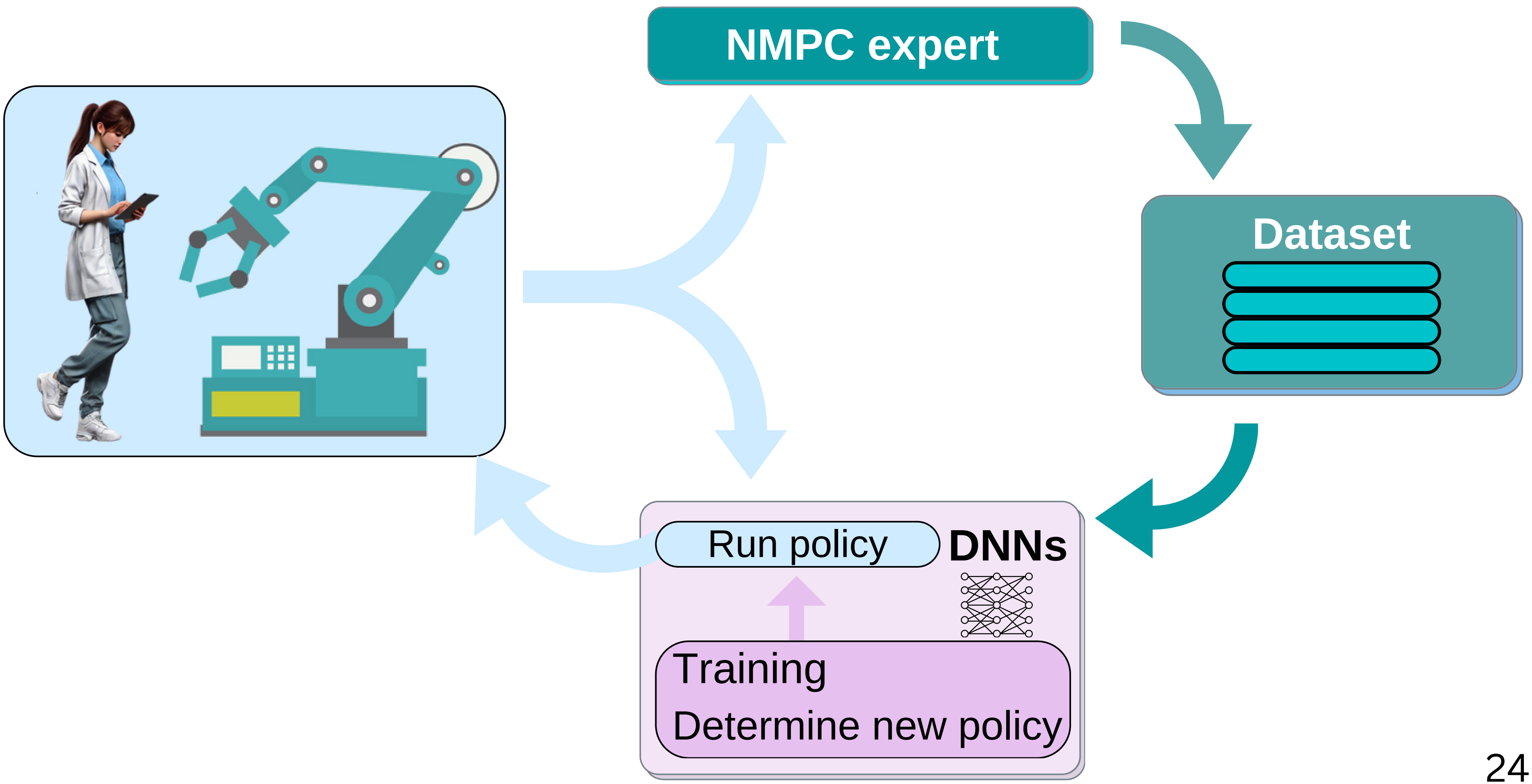


Figure 7. Time evolution of robot test points for the O-E-DNN-P

Safety filter



Imitation learning: DAgger approach



Imitation learning: DAgger approach

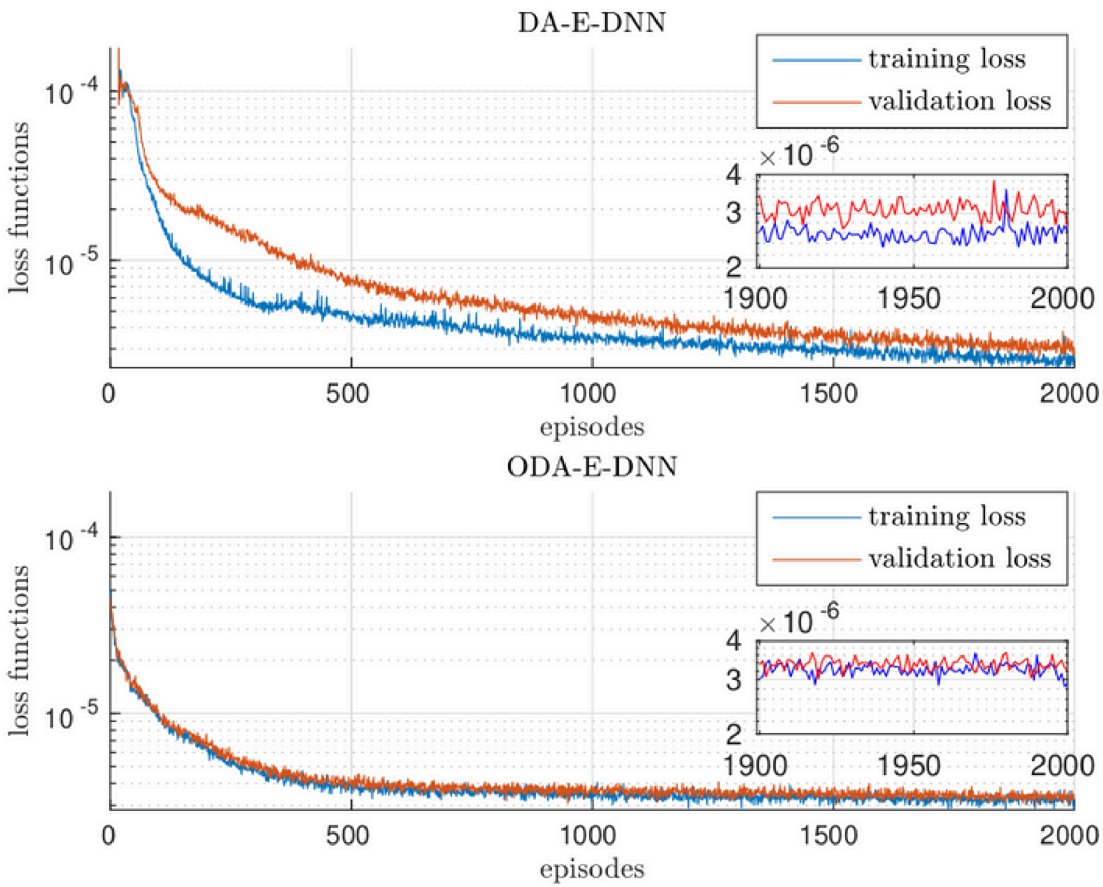


Figure 8. Learning curves for DA-E-DNN and ODA-E-DNN

Results

Controllers	Task completion time, s	Cost	Maximum solver time, ms	Average solver time, ms
NMPC	18.9	6257.2	119	43.8
O-DNN-P	17.9	6078.5	2.30	0.78
ODA-DNN	17.1	5680.8	**1.9**	**0.73**
DA-DNN	16.5	5850.5	3.0	0.81
O-E-DNN-P	18.2	5793.9	6.5	1.35
ODA-E-DNN	17.5	5429.7	6.2	1.3
DA-E-DNN	**15.3**	**5378.7**	29.4	1.4
O-S-DNN-P*	19.8	6913.6	2.3	0.7
ODA-S-DNN*	19.4	7254.0	3.6	0.67
DA-S-DNN	16.8	5875.2	**2.2**	**0.72**
O-SE-DNN-P*	20.4	6973.8	29.7	1.55
ODA-SE-DNN*	19.2	6793.6	31.1	1.6
DA-SE-DNN	**16.0**	**5722.8**	32.4	1.06

NMPC

- based on simplified system dynamics
- used the *acados* Toolkit to generate the optimization solver
- no need for training
- guarantees safety via SSM

DA-E-DNN

- uses a DNN to imitate NMPC
- trained on states generated first via the NMPC policy and then via DNN
- uses dataset aggregation
- does not guarantee safety

DA-SE-DNN

- uses a DNN to imitate NMPC
- trained on states generated first via the NMPC policy and then via DNN
- uses dataset aggregation
- safety filter imposes SSM

Results

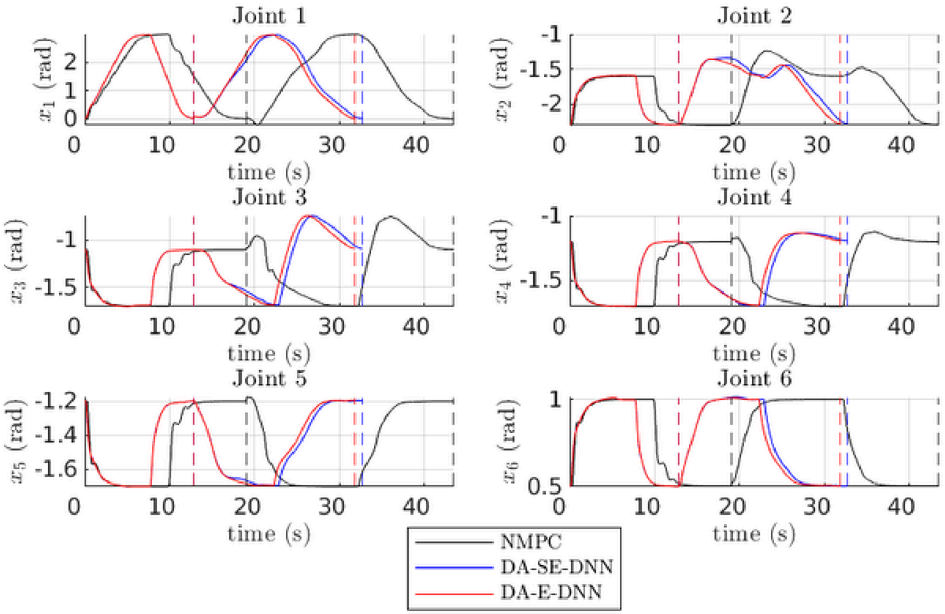


Figure 9. Time evolution of joint positions

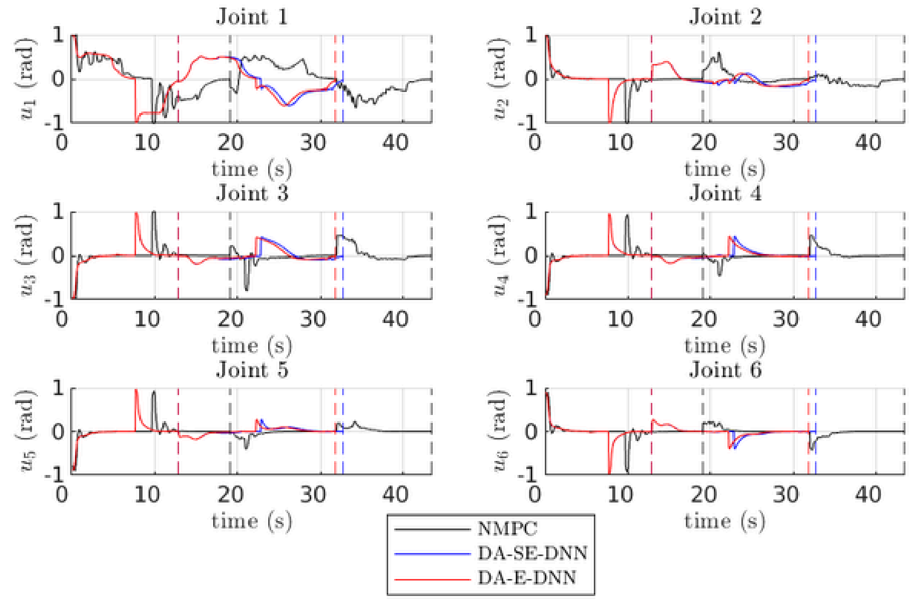


Figure 10. Time evolution of joint speeds

Results

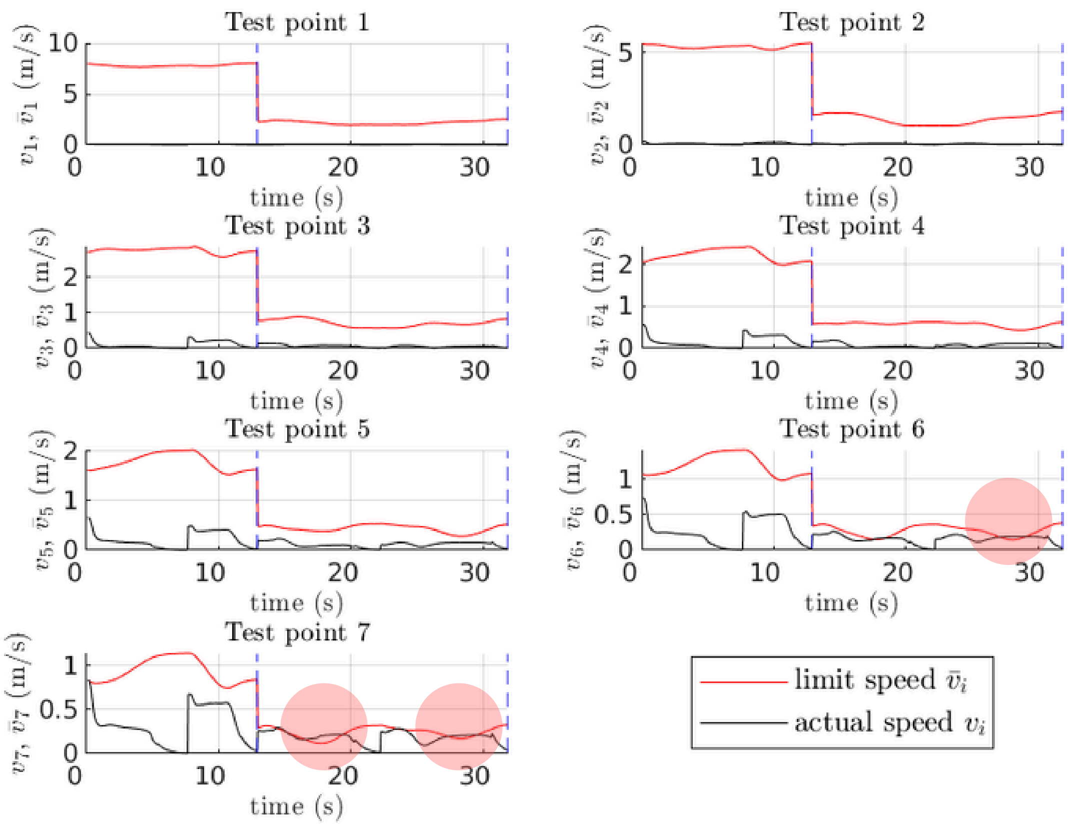


Figure 11. Time evolution of robot test points for the DA-E-DNN

Results

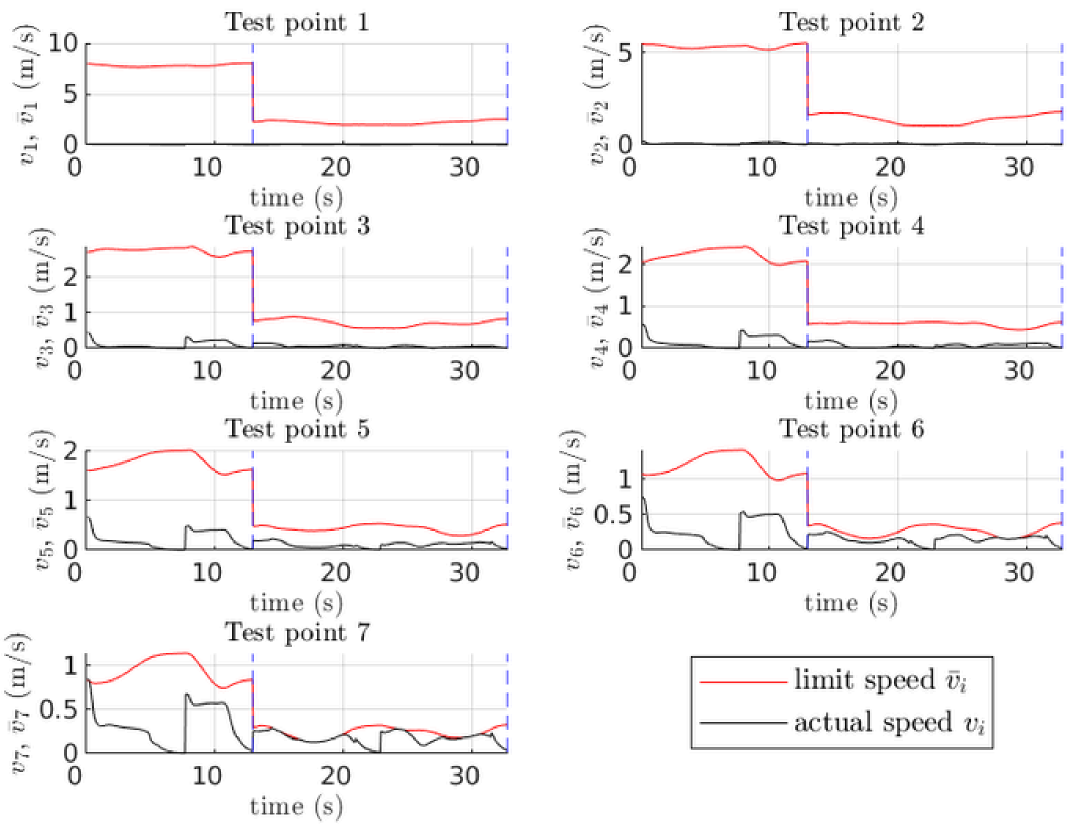
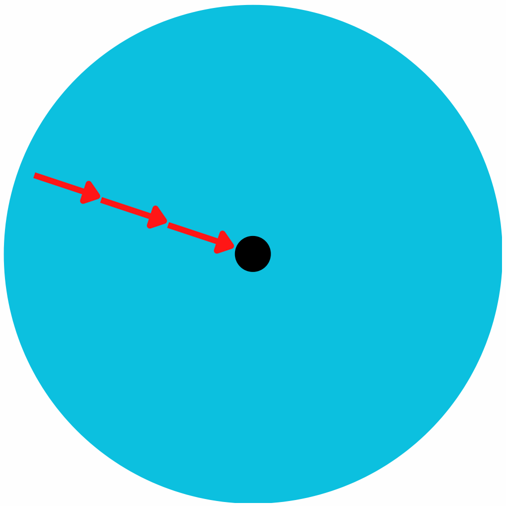


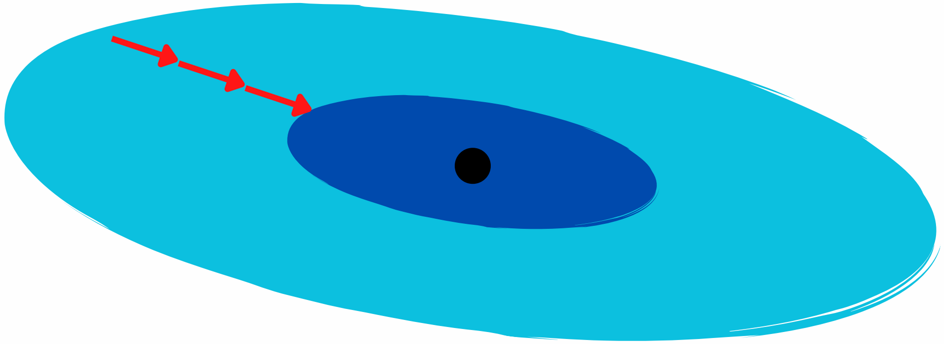
Figure 12. Time evolution of robot test points for the DA-SE-DNN

NMPC with set terminal constraint

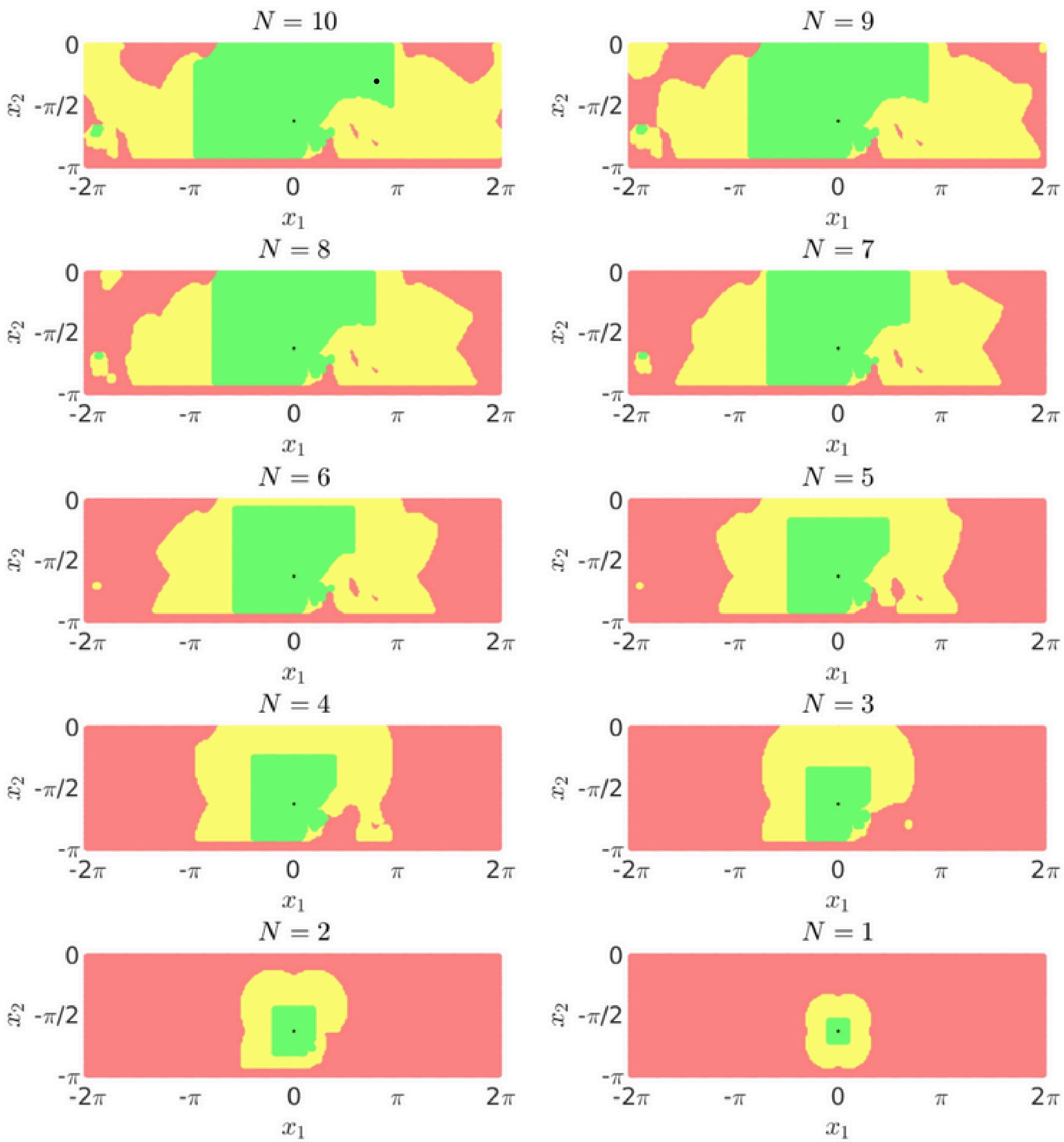
Point Terminal Constraint
(PTC)



Set Terminal Constraint
(STC)



Domain of attraction



$$x_g = \begin{bmatrix} 0.0 \\ -2.0 \\ -1.22 \\ -1.52 \\ -1.59 \\ 0.0 \end{bmatrix} \qquad x_i = \begin{bmatrix} 2.62 \\ -0.96 \\ -1.22 \\ -1.52 \\ -1.59 \\ 0.0 \end{bmatrix}$$

- feasible for STC and PTC
- feasible for STC
- no feasible solution

Domain of attraction

.

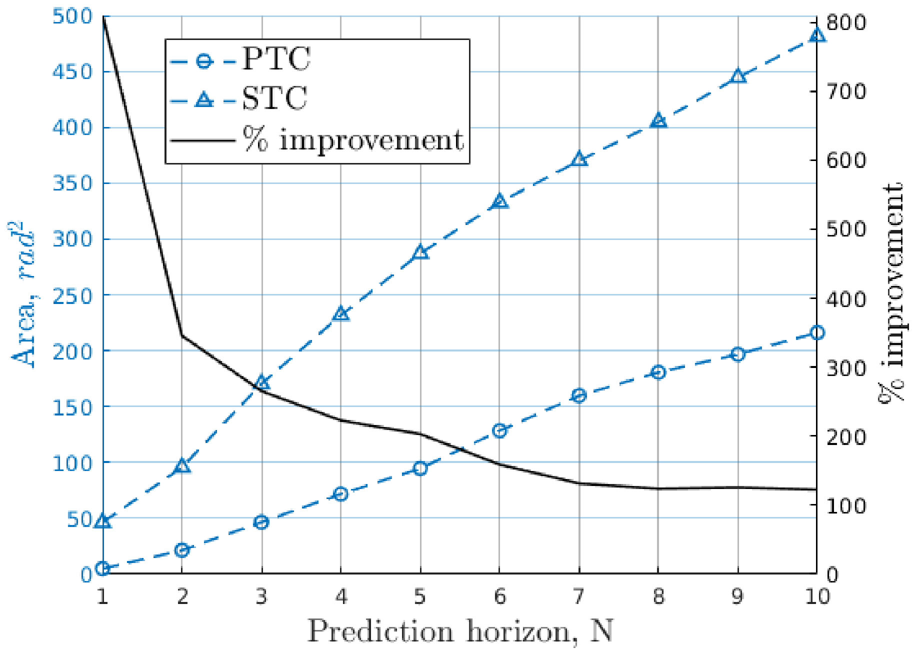


Figure 13. Evolution of the feasibility area

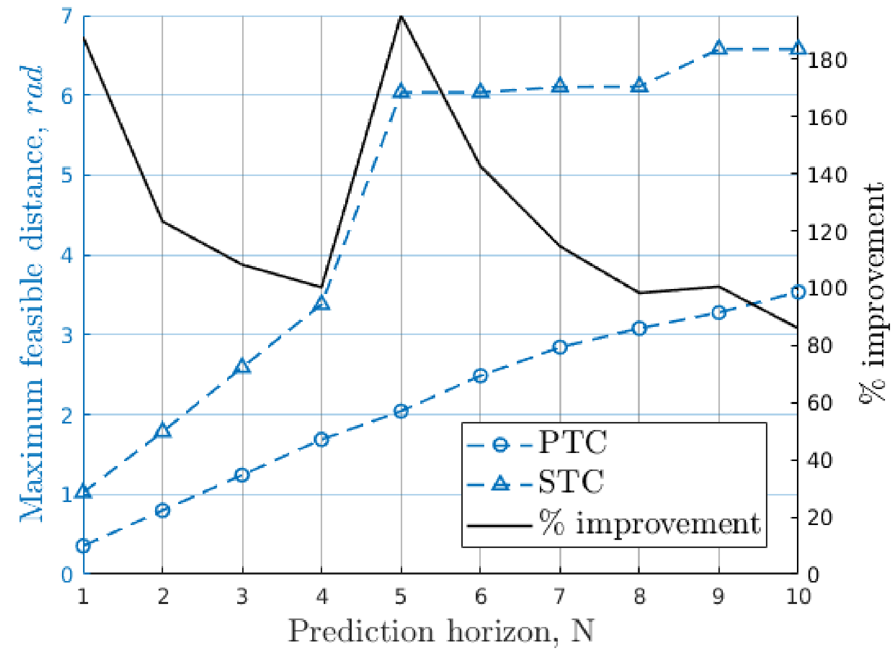


Figure 14. Evolution of the maximum feasible distance

Experiment with a UR5 robot-manipulator operating in a shared workspace

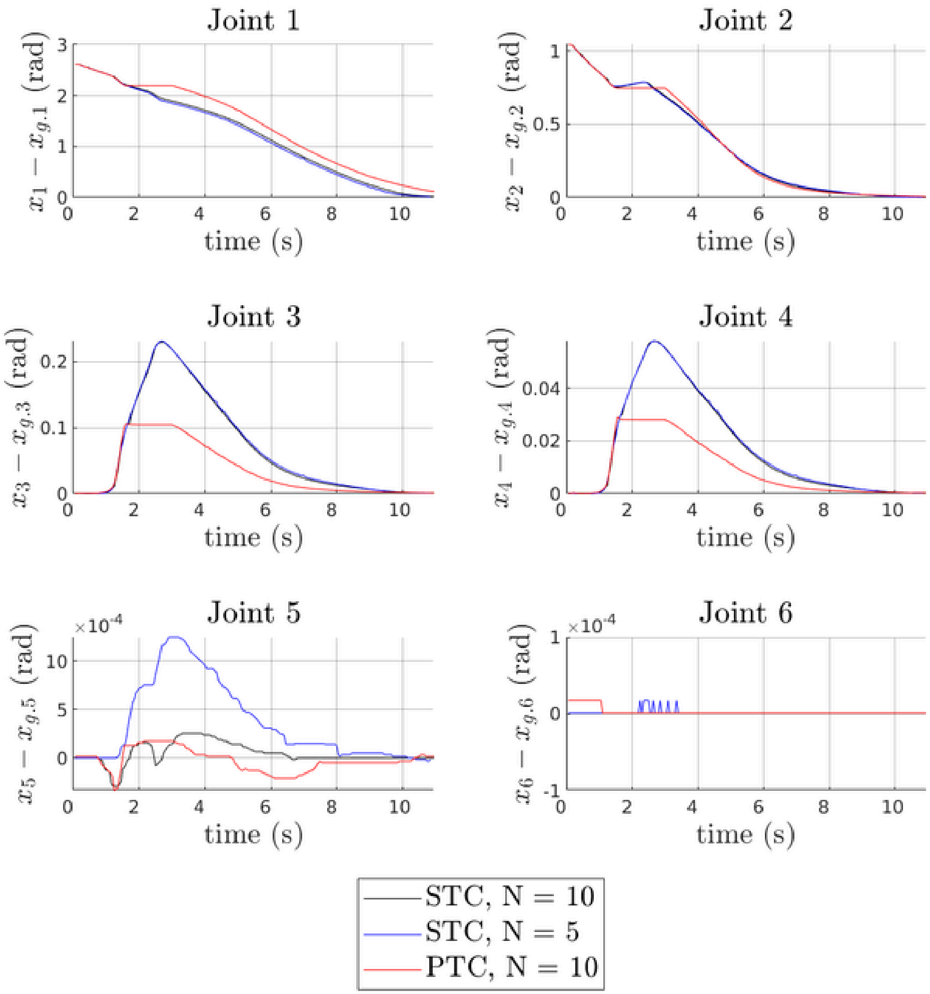


Figure 15. Time evolution of joint positions

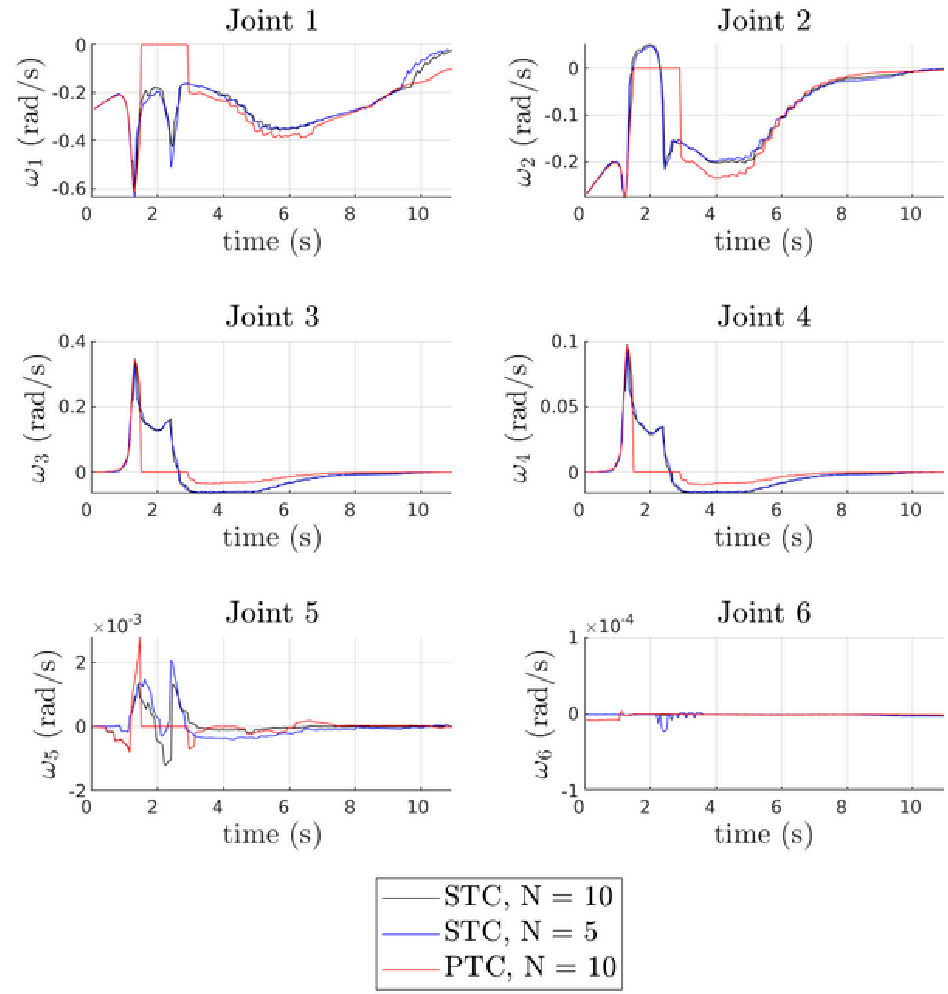


Figure 16. Time evolution of joint speeds

Results

Controllers	Task completion time, s	Cost	Maximum solver time, ms	Average solver time, ms
STC - 5	**10.85**	**5690.3**	**32.324**	**16.5839**
STC - 10	10.95	5861.3	69.341	48.3845
PTC - 10	12.45	6802.2	95.376	46.7844

Conclusions

- *Is it possible to improve the performance of SSM-based NMPC laws, at the same time guaranteeing human safety, by emulating their behavior through DNNs?*

Yes, 2 Imitation learning approaches were designed, and tested on a real robot. The DA-SE-DNN method provided the best trade-off between safety and performance. Presence of the safety filters ensured the satisfaction of the SSM constraints.

- *Is it possible to define and design an SSM-based NMPC law with guaranteed stability properties and based on a set terminal constraint, which increases the domain of attraction compared to the point terminal constraint approach?*

Yes, The NMPC with STC guarantees closed-loop stability, and the domain of attraction can be considerably extended, based on the considered case study.

Future work

- DNN-based approaches that provide strict guarantees of constraint satisfaction and stability

- To guarantee stability in a NMPC control strategy based on the assumption that a human will move

Thank you for attention!

References

- Nurbayeva, A., Shintemirov, A., & Rubagotti, M. (2022). Deep Imitation Learning of Nonlinear Model Predictive Control Laws for a Safe Physical Human–Robot Interaction. IEEE Transactions on Industrial Informatics, 19(7), 8384-8395.

- ISO/TS 15066: 2016: Robots and Robotic Devices–Collaborative Robots. International Organization for Standardization: Geneva, Switzerland.

- Oleinikov, A., Kusdavletov, S., Shintemirov, A., & Rubagotti, M. (2021). Safety-Aware Nonlinear Model Predictive Control for Physical Human-Robot Interaction. IEEE Robotics and Automation Letters.

- Ross, S., Gordon, G., & Bagnell, D. (2011, June). A reduction of imitation learning and structured prediction to no-regret online learning. In Proceedings of the fourteenth international conference on artificial intelligence and statistics (pp. 627-635). JMLR Workshop and Conference Proceedings.

SSM formulation

$$\|v_i\|^2 \leq \alpha^2 \left(d_{ij}^2 - \rho_{ij}^2\right),\; i,j \in \mathbb{N}_r \times \mathbb{N}_h$$

$$\|\mathscr{K}_{d,i}(x,u)\|^2 \leq \alpha^2 \left(\|\mathscr{K}_{f,i}(x) - p_{h,j}\|^2 - \rho_{ij}^2\right)$$

Stage cost function

$$\ell(x,u) \;= \|x - x_g\|_Q^2 + \|u\|_R^2 + \gamma\varphi^2(x,x_g)$$

$$\varphi(x,x_g) \triangleq \exp\left\{-\beta \frac{\left\|\mathscr{K}_{f,n_r}(x) - p_{h,j_\gamma}\right\|^2}{\left\|\mathscr{K}_{f,n_r}(x) - p_{r,n_r}^*\right\|^2}\right\}$$

NMPC with PTC

$$J_N\left(x_0, u(\cdot)\right) \triangleq \sum_{k=0}^{N-1} \ell\left(x_u(k), u(k)\right)$$

with respect to $u(\cdot) \in \mathcal{U}^N_{x_g}(x_0)$

subj. to $x_u(0) = x_0,\ x_u(k+1) = f(x_u(k), u(k))$

NMPC with STC

$$J_{NF}\left(x_0, u(\cdot)\right) \triangleq \sum_{k=0}^{N-1} \ell\left(x_u(k), u(k)\right) + F(x_u(N))$$

with respect to $u(\cdot) \in \mathcal{U}^N_{X_T}(x_0)$

subj. to $x_u(0) = x_0,\ x_u(k+1) = f(x_u(k), u(k))$

$$X_T \triangleq \{x \in \mathbb{R}^n \ :\ (x - x_g)^T P(x - x_g) \leq \eta\}$$

$$F(x) \triangleq \sigma(x - x_g)^T P(x - x_g).$$

minimize $(x - x_g)^T P(x - x_g)$

with respect to $x \in \mathbb{R}^n$

subj. to $g_i(x) \leq 0$.