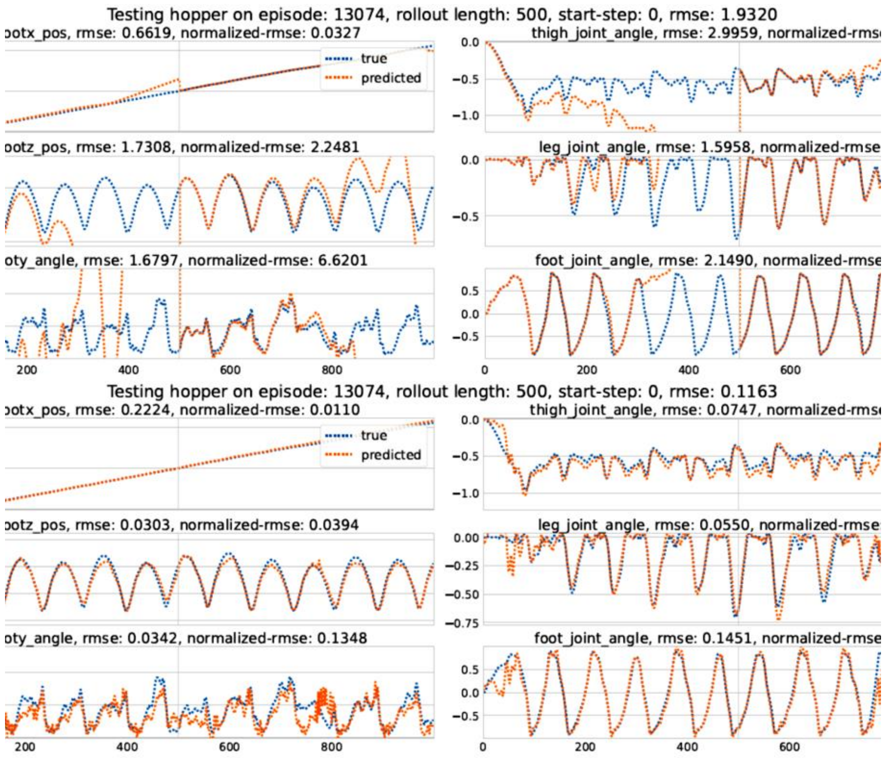


DRL Workshop NeurIPS 2022

Reinforcement learning in System identification

Presenter: Oscar Fernández Vicente

Jose Antonio Martin H., Oscar Fernández Vicente, Sergio Perez, Anas Belfadil, Cristina Ibanez-Llano, Freddy Perozo, Jose Javier Valle, Javier Arechalde

REPSOL TECHNOLOGY LAB

MOTIVATION

REINFORCEMENT LEARNING IN SYSTEM IDENTIFICATION

1. Learn forward models from data of Dynamical Systems trajectories

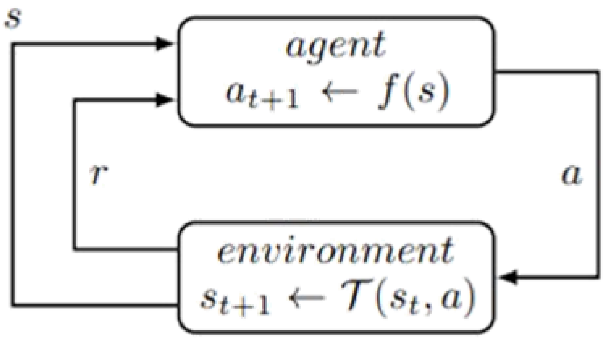
2. Propose the use of RL as a framework to learn forward models

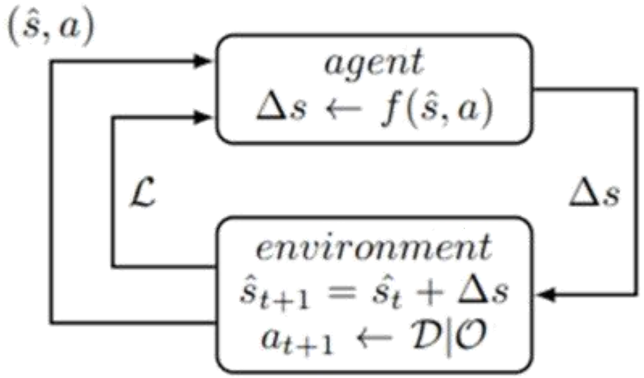
3. RL optimizes in the long run, so we aim to reduce compounding error

4. RL explores, so we expect this help to obtain robust models

5. RL provide us with a Q-function, which in this case learns the cumulative error the model will commit

INTRODUCTION
REINFORCEMENT LEARNING IN SYSTEM IDENTIFICATION

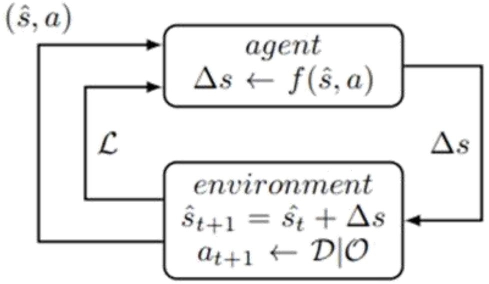




$$\hat{s}_{t+1} = s_t + f(s_t, a_t)$$

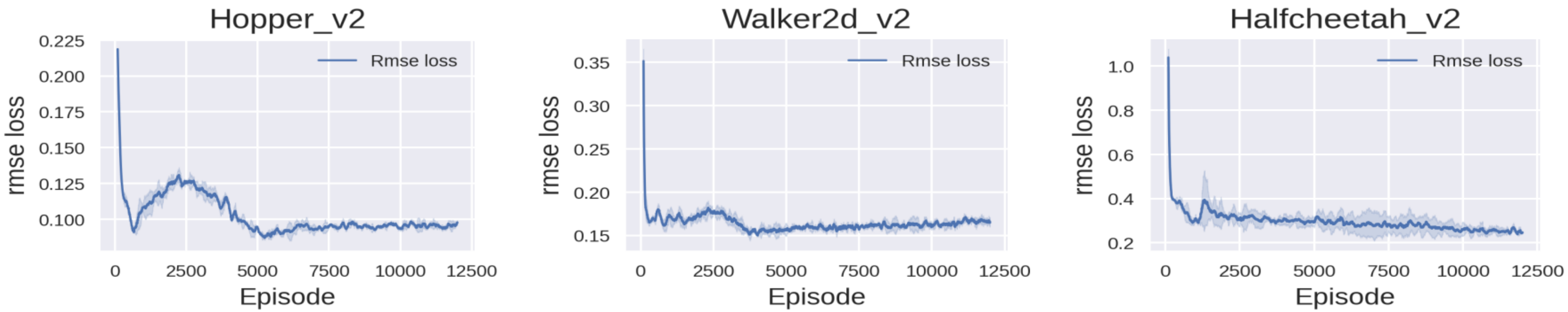
1. Transpose the datasets from {s,a,r,s’} to {(s,a),∆s,||s’-s||,(s’,a’)}
2. The reward is then the negative of the cumulative prediction error
3. Simulate episode from datasets to let the RL explore/exploit
4. Let RL optimize the trajectory error for the long run (Bellman)
5. Obtain a forward model: actor policy
6. Obtain an uncertainty model: Q-function

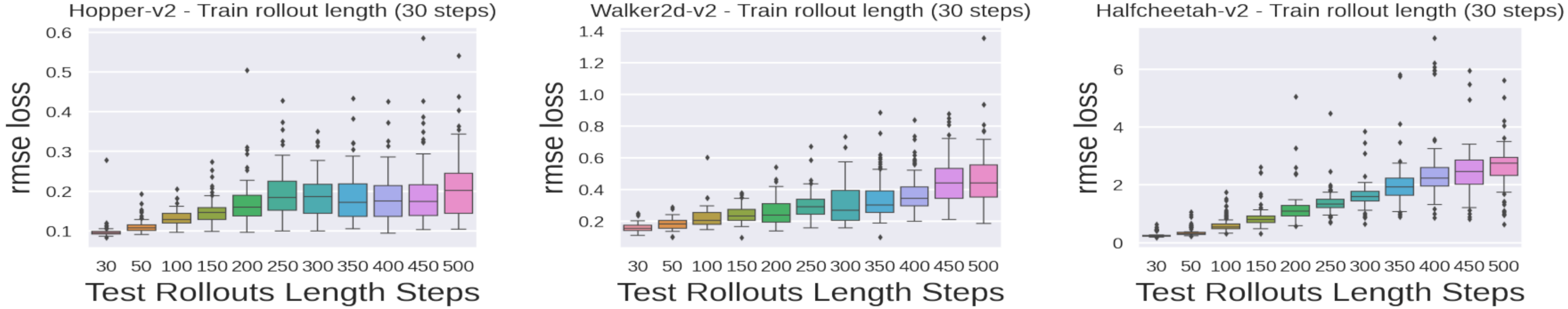
FORWARD MODELS WITH RL - MUJOCO
REINFORCEMENT LEARNING IN SYSTEM IDENTIFICATION



Key points:

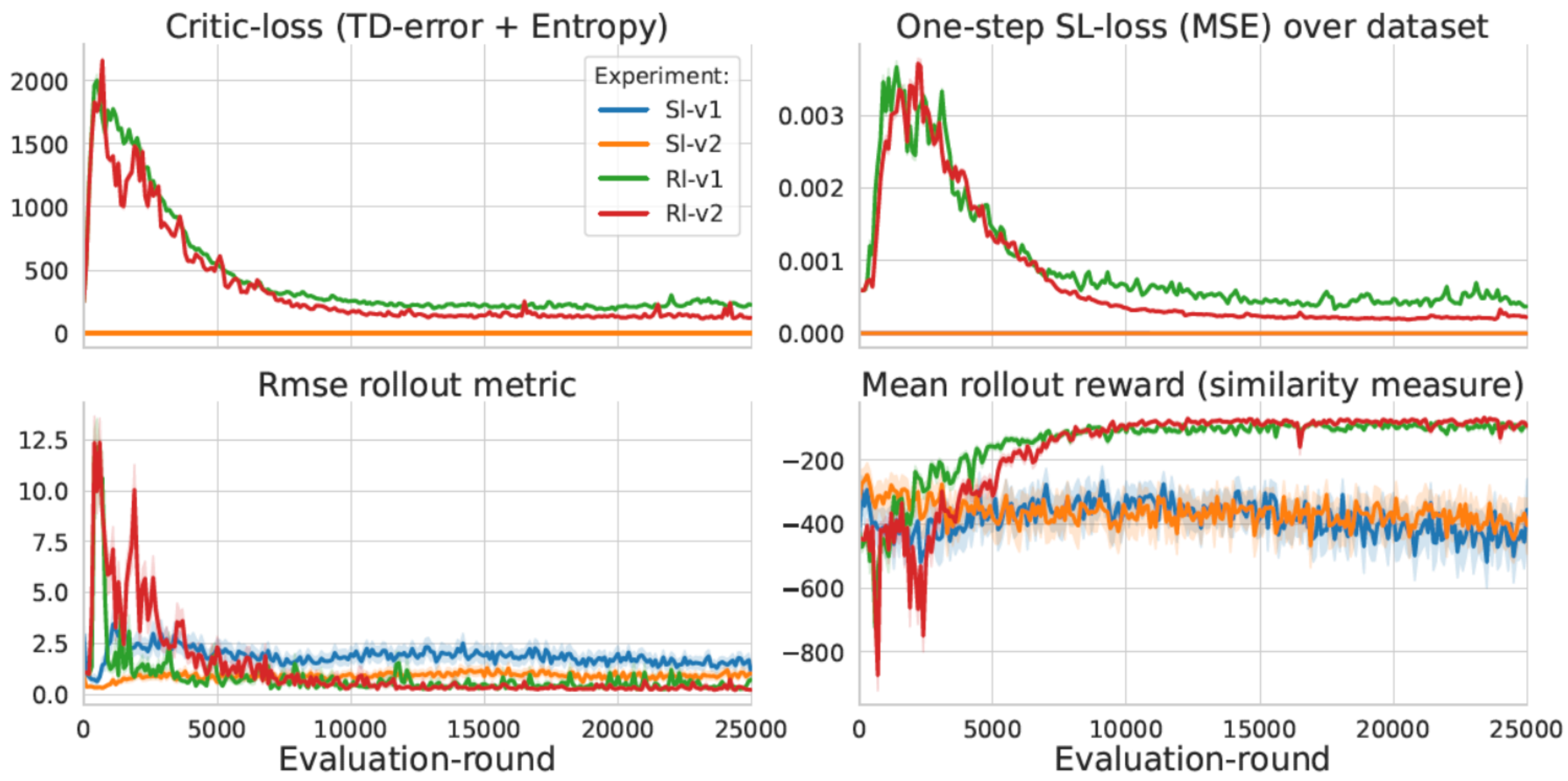
- SAC (Soft Actor-Critic) agent
- Train and test on 3 Mujoco's environments
- Test on different rollouts lengthts





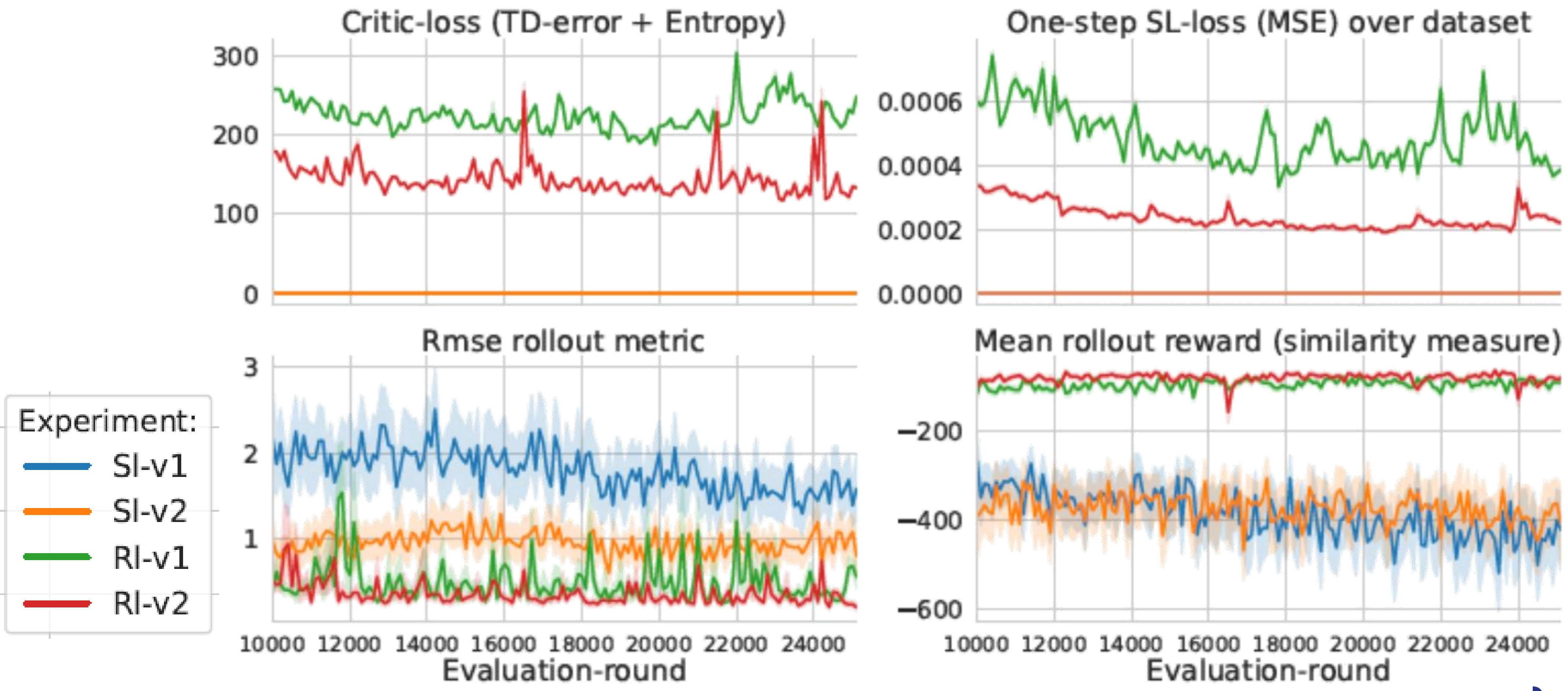
REINFORCEMENT LEARNING VS SUPERVISED LEANING

REINFORCEMENT LEARNING IN SYSTEM IDENTIFICATION



REINFORCEMENT LEARNING VS SUPERVISED LEANING

REINFORCEMENT LEARNING IN SYSTEM IDENTIFICATION



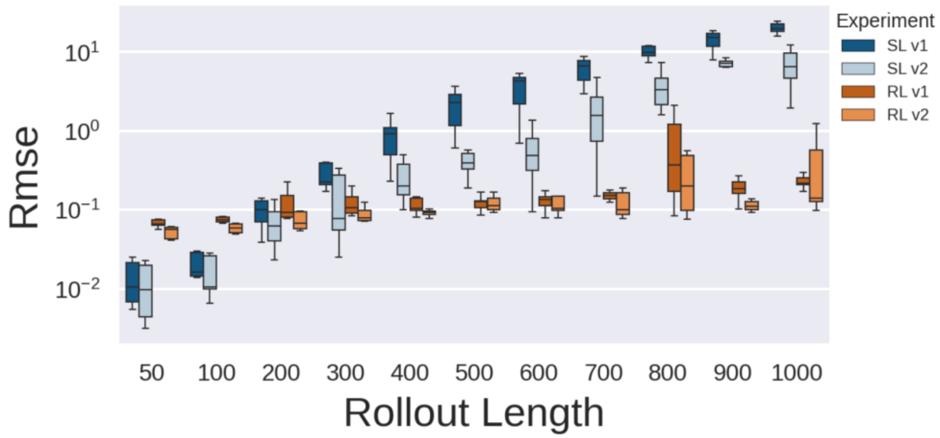
HR EXCELLENCE IN RESEARCH

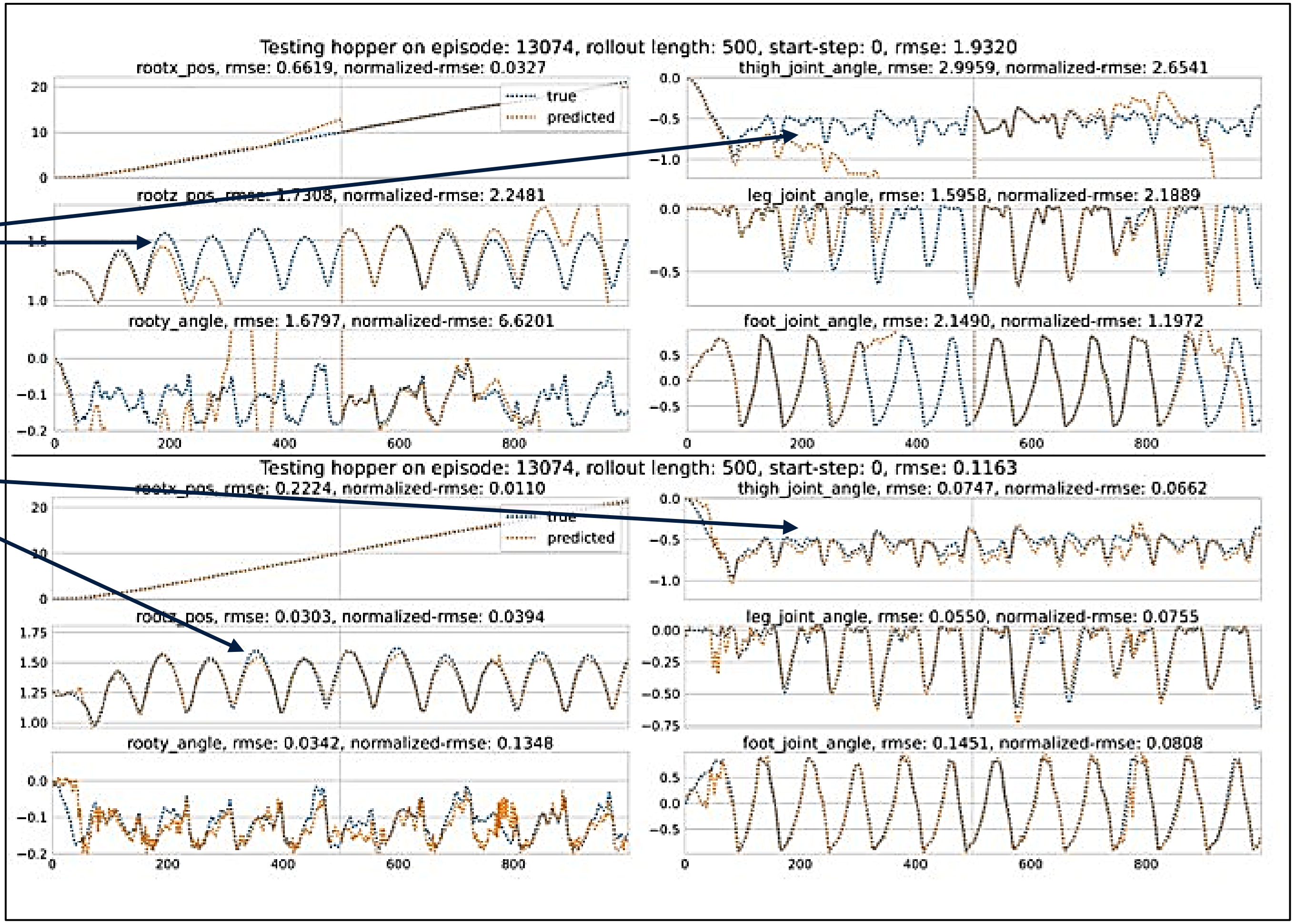
REINFORCEMENT LEARNING VS SUPERVISED LEANING

REINFORCEMENT LEARNING IN SYSTEM IDENTIFICATION

SL: compounding error

RL: no compounding error





REPSOL

HR EXCELLENCE IN RESEARCH

CONCLUSION & FUTURE WORK

REINFORCEMENT LEARNING IN SYSTEM IDENTIFICATION

Conclusions

RL has been proven as a valid Forward Model learning approach.

Strengths:
- Can generate forward-models from historical data of a system.
- Rollouts are more robust.
- Reduces the compounding error, especially in long trajectories.

Weaknesses:
- In shorts rollouts, SL seems performs better.
- In simulation stage, long trajectories can end in unexplored states, with undesired learned policies.
- Slower training than SL.

Future work

- Improve the RL forward model to become a robust training environment.
- Include constraints/controls on OOD state spaces.
- Develop similarity metrics with source systems.

REPSOL TECHNOLOGY LAB

THANK YOU!

Presenter: Oscar Fernández Vicente

Jose Antonio Martin H., Oscar Fernández Vicente, Sergio Perez, Anas Belfadil, Cristina Ibanez-Llano, Freddy Perozo, Jose Javier Valle, Javier Arechalde

HR EXCELLENCE IN RESEARCH