RANS: Highly-parallelised simulator for **R**einforcement Learning based **A**utonomous **N**avigating **S**pacecrafts

Matteo El Hariry, Antoine Richard, Miguel Olivares Mendez

Space Robotics Research Group (SpaceR)
SnT - University of Luxembourg

October 20, 2023

Overview

SpaceR
Space Robotics Research Group
uni.lu | SnT

Introduction

- Reinforcement Learning
- Robotic simulators
- Omniverse Isaac Gym

Methods

- Simulation engine
- Environments & tasks
- DRL agents

Preliminary results

- 3 DoF environment
- 6 DoF environment

Introduction

Introduction

Reinforcement Learning

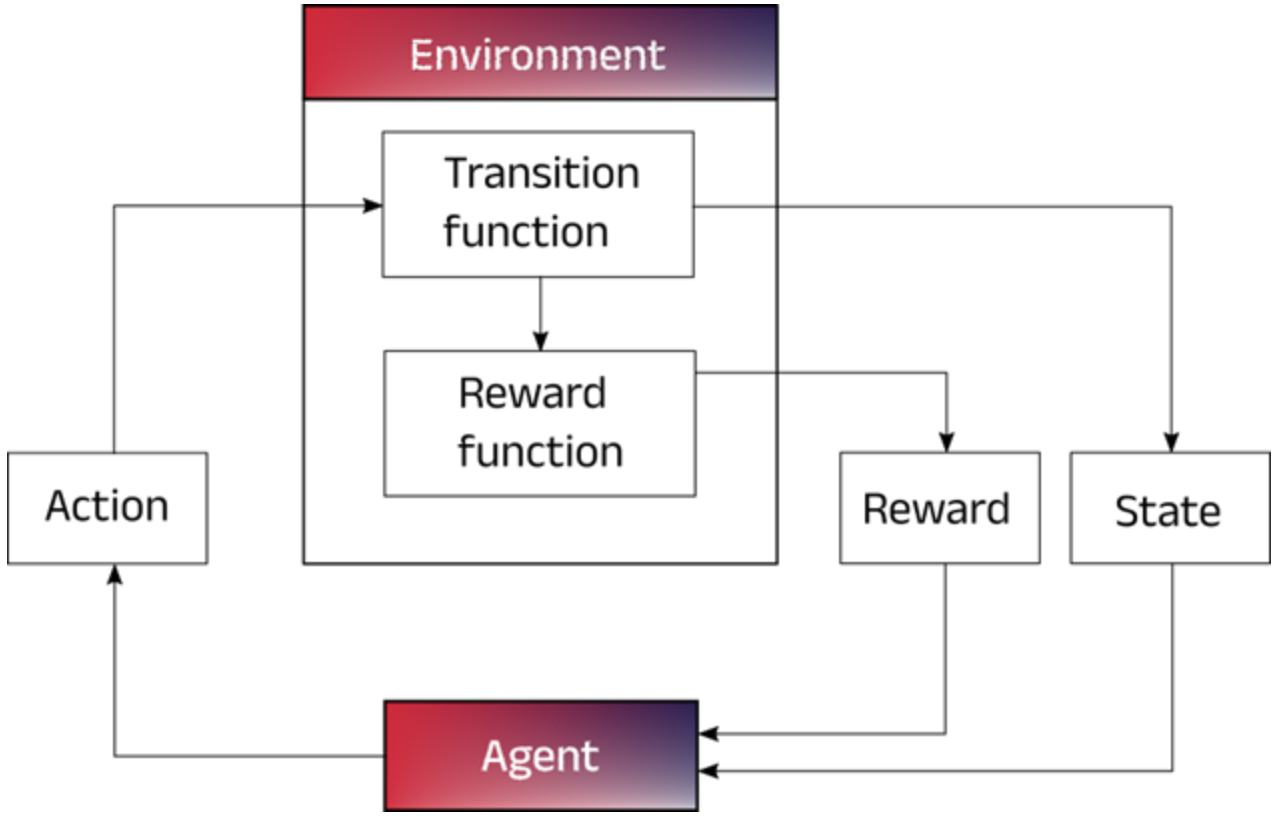
The science of learning decision making

- Environment
- Agent
- Policy π

Markov Decision Processes (MDPs).

An MDP is a 5-tuple, $< S, A, T, R, \mu >$

- S is the set of all valid states,
- A is the set of all valid actions,
- T is the transition probability function $\rightarrow T(s'|s, a)$,
- R is the reward function $\rightarrow R(s, a, s')$,
- μ is the starting state distribution.



Trajectory (or episode) $\rightarrow \tau = (s, a, s', a', s'', \ldots)$

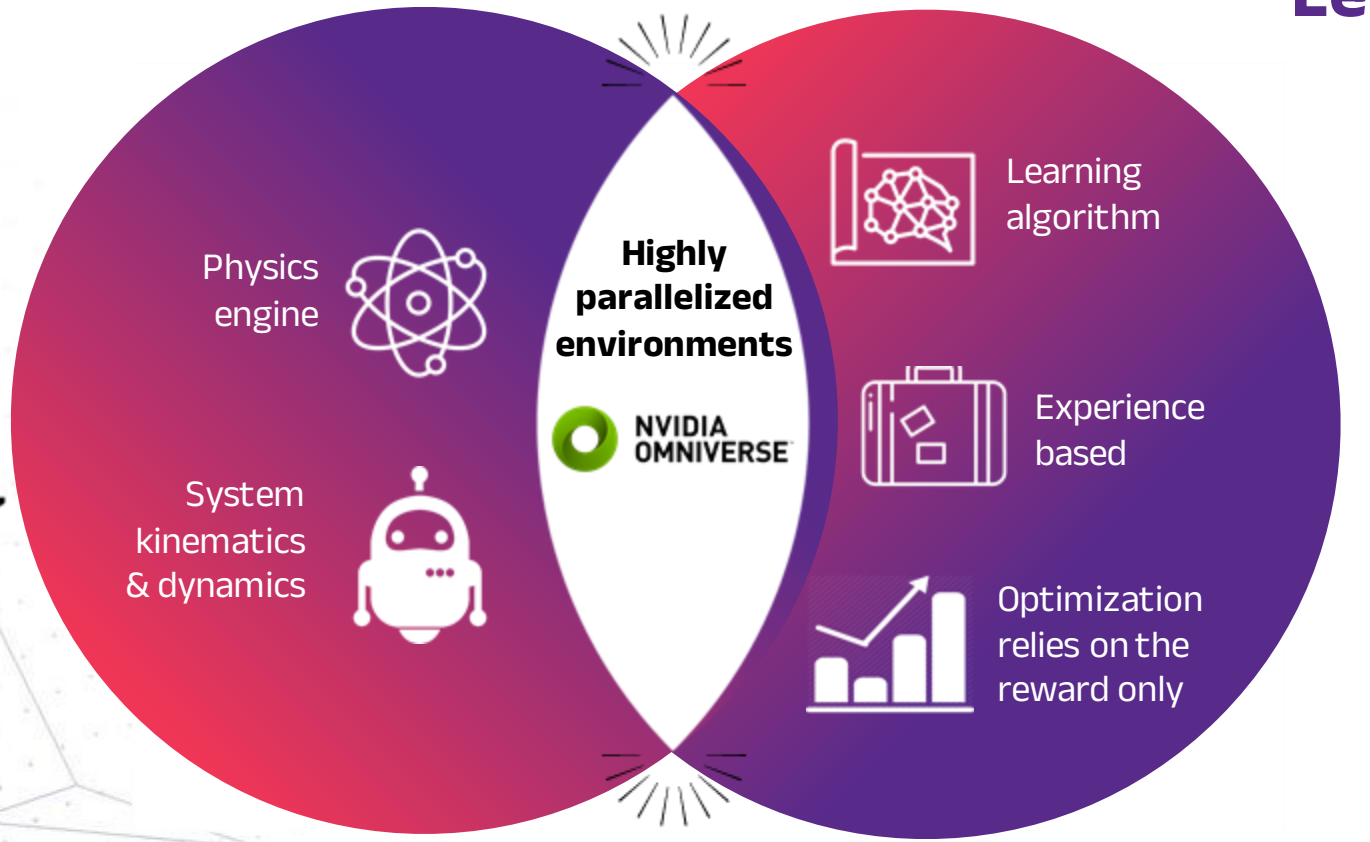
Goal $\rightarrow$ $\pi^* = \underset{\pi}{\text{argmax}}\ \mathbb{E}[R(\tau)]$

Introduction

SpaceR
Space Robotics Research Group
uni.lu | SnT

Robotics Simulator

Deep Reinforcement Learning



MuJoCo
Advanced physics simulation

GAZEBO

BULLET PHYSICS LIBRARY

Webots
robot simulation

Gym

Motivations

SpaceR
Space Robotics Research Group
uni.lu | SnT

uni.lu
UNIVERSITY OF LUXEMBOURG

Motivations

General advantages

- Flexible parametrization of environmental constraints.
- Easy to tune policies for control: Rewards/Penalties shaping.
- Domain randomization.

Advantages of our RL framework:

- Fast training:
  - ~10 minutes to train a model on a RTX 4090.
  - Approx 40.000 steps / seconds.
- Large scale testing:
  - Can evaluate over thousands of initial conditions in seconds.
- Comes pre-packaged with examples and a set of different tasks.
- Provides rich visualization:
  - Training: WandB with different metrics
  - Evaluation: Tables + plots.
- Comes with ROS bindings.

NVIDIA OMNIVERSE

Gym

Methods

SpaceR
Space Robotics Research Group
uni.lu | SnT

uni.lu
UNIVERSITY OF LUXEMBOURG

Methods

Simulation Engine

Isaac Sim: **PhysX engine**

Training fully on GPU

Substepping strategy

Control frequency < physics engine frequency (5-50, 10-100 Hz)

Free floating conditions

3 DoF (forces in the xy plane)

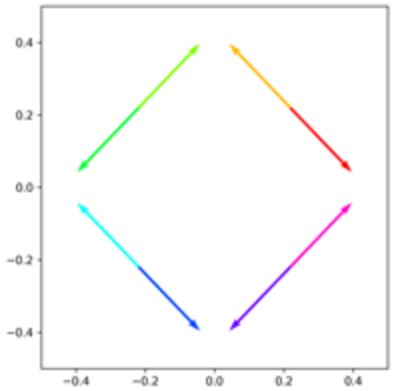
6 DoF (no constraint)

Methods

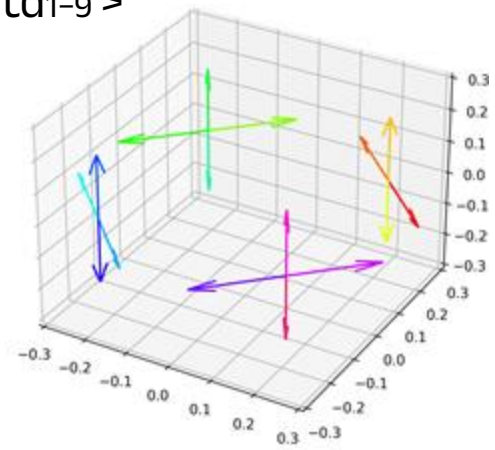
Environment and Tasks

Environments:

- 3 DoF
  - Observations:
    < cos(θ), sin(θ), vxy, ωz, tf, td_{1-4} >
  - Actions:
    $T_{1-8} \in [0,1]$



- 6 DoF
  - Observations:
    < 6D, v_{xyz}, ω_{xyz}, tf, td_{1-9} >
  - Actions:
    $T_{1-16} \in [0,1]$



Tasks:

- GoToXY
- GoToPose-2D
- TrackXYVelocity
- GoToXYZ
- GoToPose-3D
- TrackXYZVelocity

Task	tf	td_1	td_2	td_3	td_4	td_5	td_6	td_7	td_8	td_9
3DoF Go to position	0	Δx	Δy	-	-					
3DoF Go to pose	1	Δx	Δy	$\cos(\Delta\theta)$	$\sin(\Delta\theta)$					
3DoF Track velocity	2	Δv_x	Δv_y	-	-					
6DoF Go to position	0	Δx	Δy	Δz	-	-	-	-	-	-
6DoF Go to pose	1	Δx	Δy	Δz	$\Delta R[0,0]$	$\Delta R[0,1]$	$\Delta R[0,2]$	$\Delta R[1,0]$	$\Delta R[1,1]$	$\Delta R[1,2]$
6DoF Track velocity	2	Δv_x	Δv_y	Δv_z	-	-	-	-	-	-

Methods

DRL Agents

Based on the RL Games library

- **PPO** with multi-discrete action space
- Actor-critic architecture
- MLP network with:
  - 3DoF → 2 hidden layers of 128 neurons
  - 6DoF → 3 hidden layers of 256 neurons
- Training for 2000 epochs ~130M steps

Preliminary Results

SpaceR
Space Robotics Research Group
uni.lu | SnT

uni.lu
UNIVERSITY OF LUXEMBOURG

Experimental Setup & Results

3 DoF Evaluations

Evaluation for:

- **PPO** agent
- **GoToPose-2D** task
- Reward function used: exponential
- Training epochs: 2000
- No hyperparameter tuning
- Episodes number: 1024
- Initial spawning distance: [3, 4] m

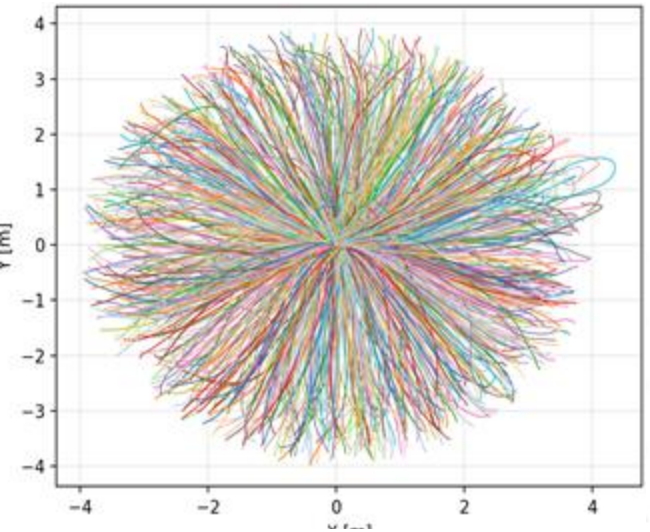
System properties:

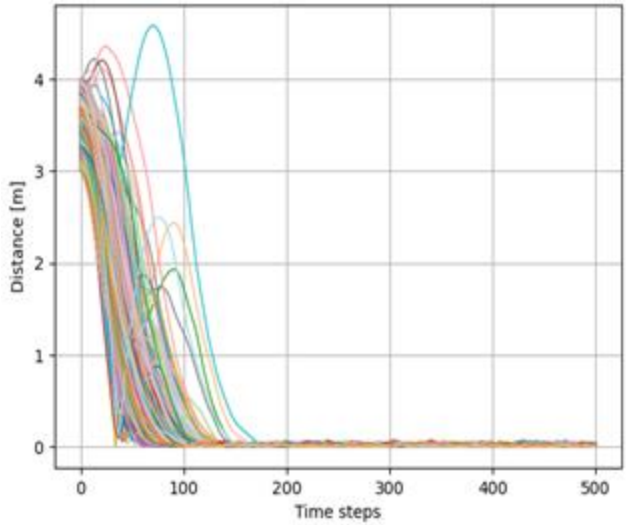
- Mass: ~5 kg
- CoM: [0, 0, 0]
- Thrust force: 1 N
- Radius: 31 cm
- Shapes: sphere or cylinder

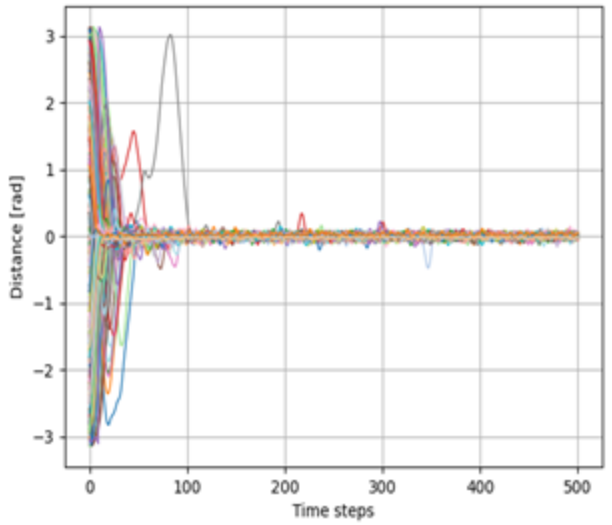
Experimental Setup & Results

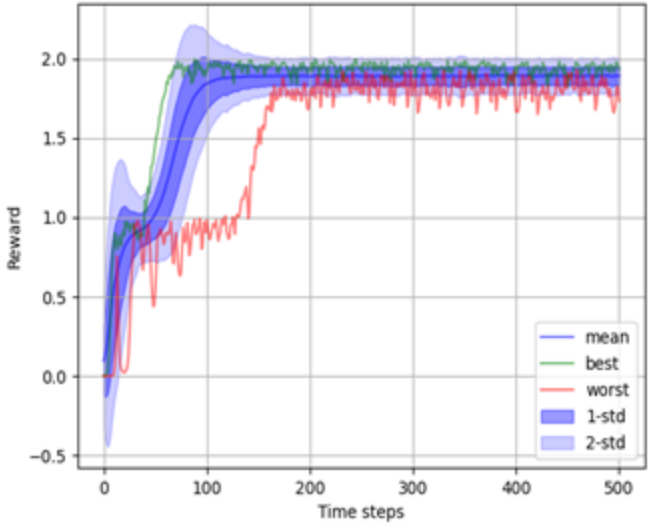
SpaceR
Space Robotics Research Group
uni.lu | SNT

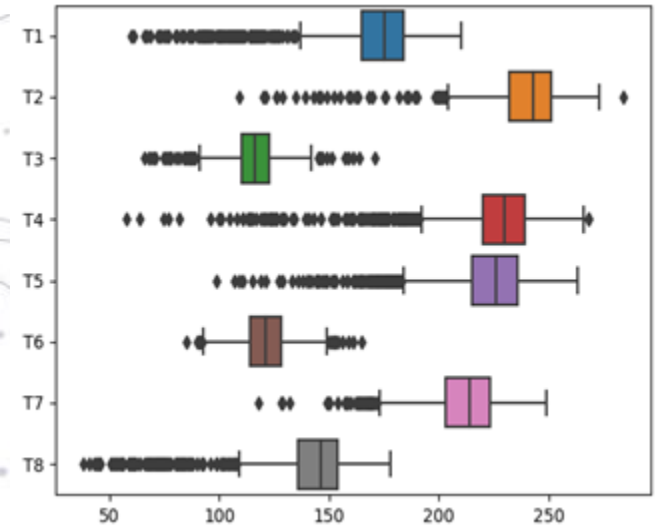
3 DoF Evaluations

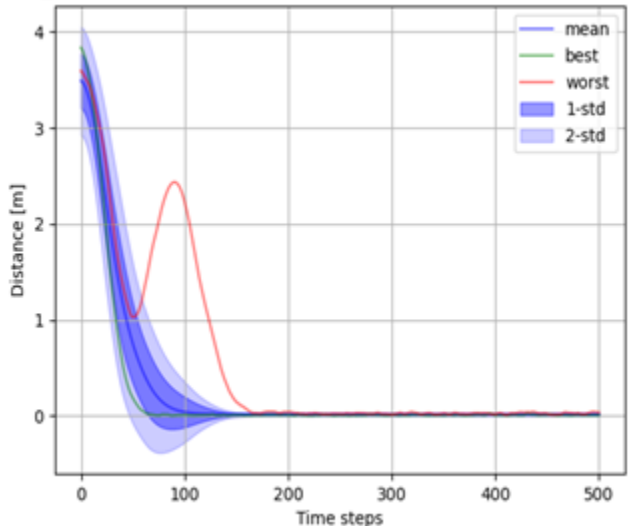


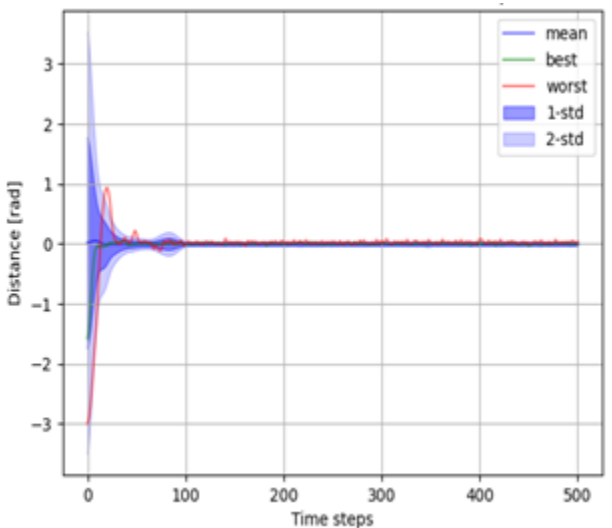


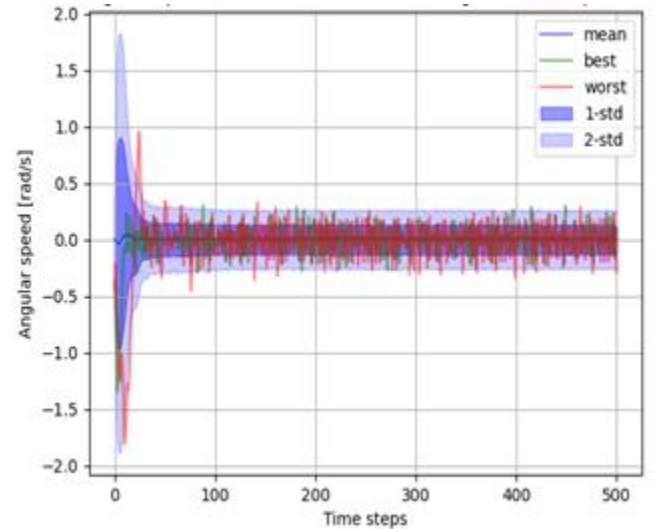








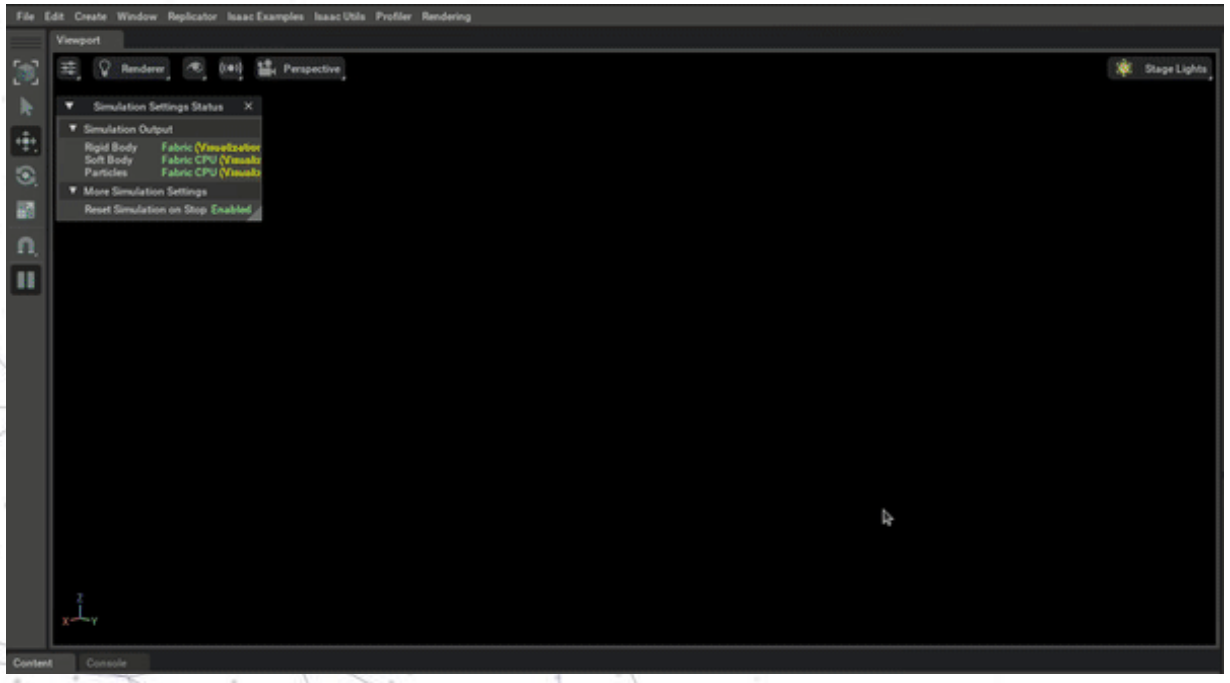


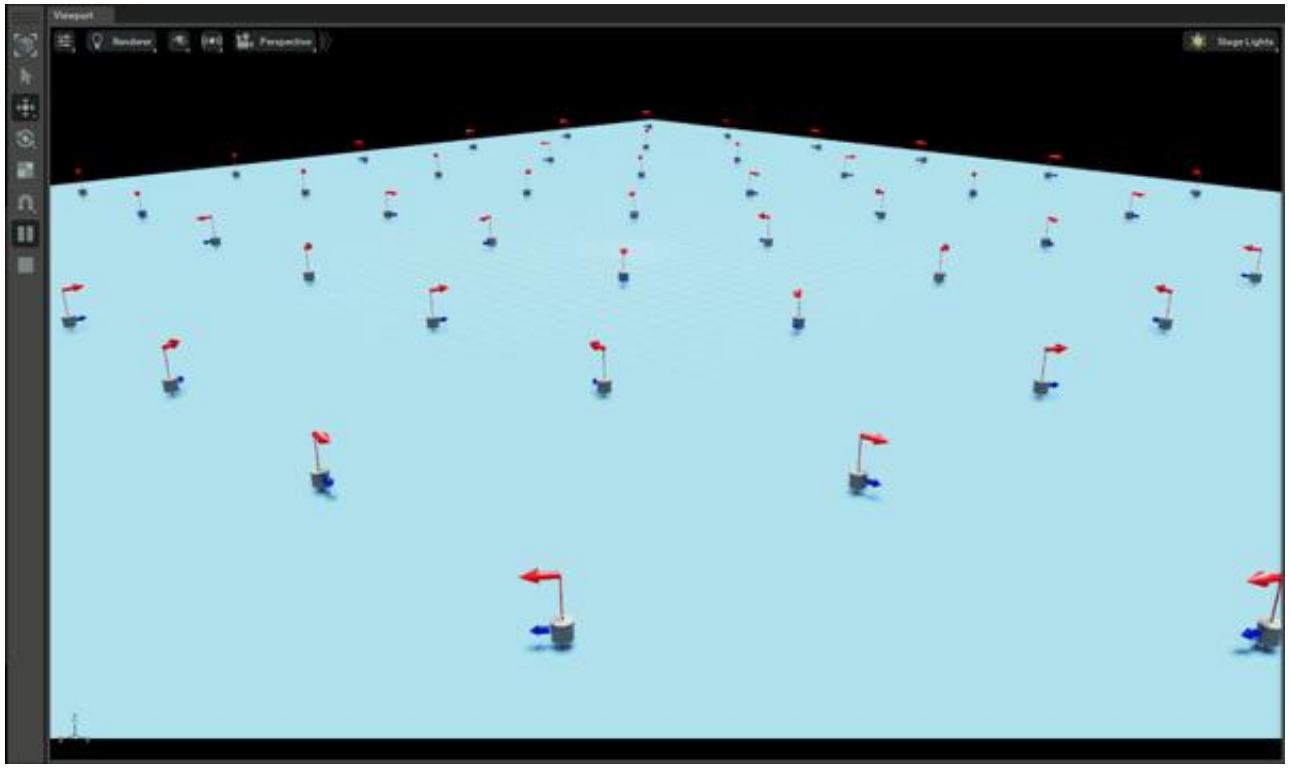


Experimental Setup & Results

SpaceR
Space Robotics Research Group
uni.lu | SNT

3 DoF Evaluations

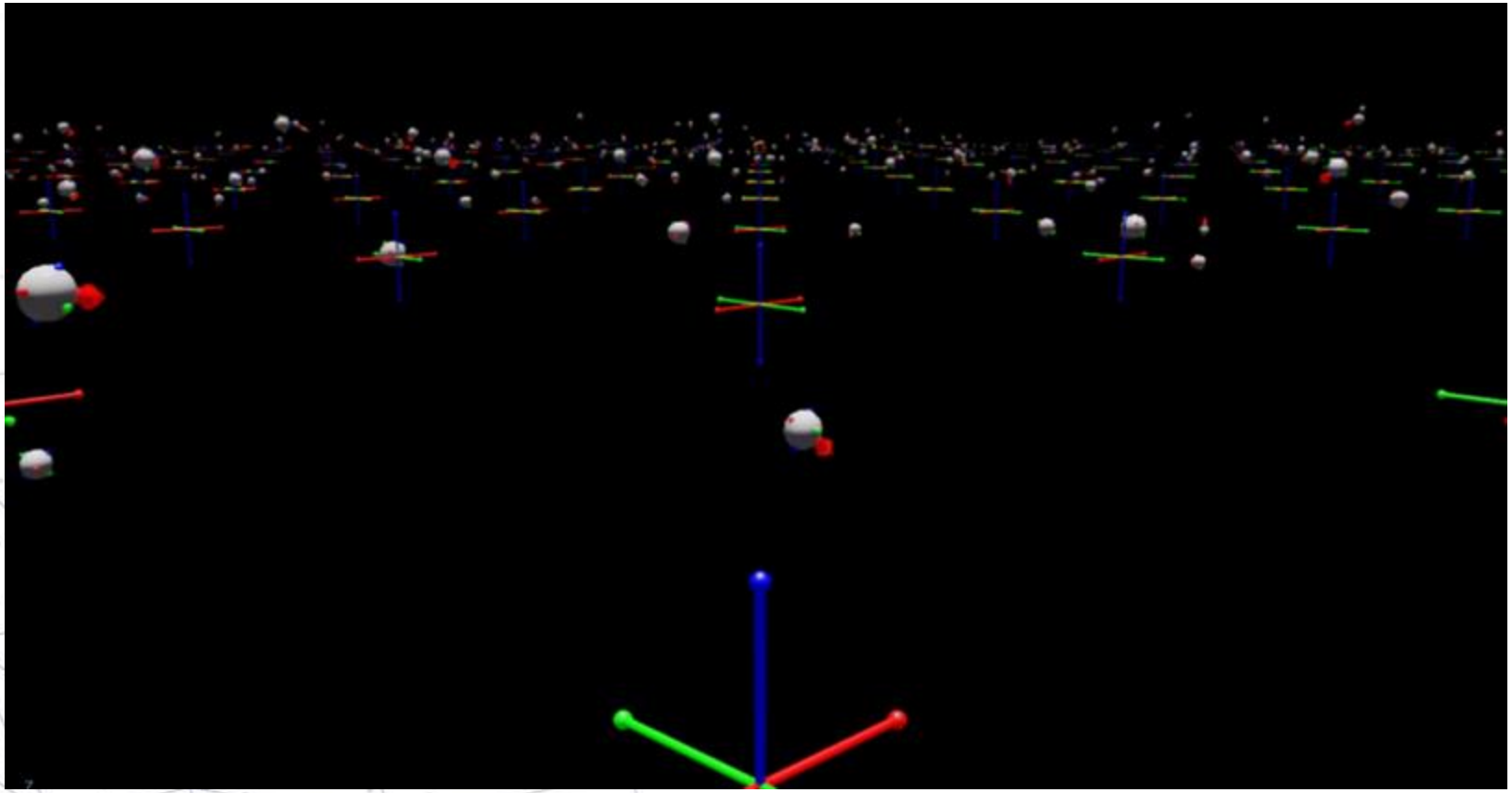




Experimental Setup & Results

SpaceR
Space Robotics Research Group
uni.lu | SNT

6 DoF Evaluations



Thank you

Questions ?

More information and videos are available https://www.spacer.lu/

SpaceR
Space Robotics Research Group
uni.lu | SNT

iSpaRo 2024

June 24-27, Luxembourg

SAVE the DATE

https://github.com/elharirymatteo/RANS

isparo.space

uni.lu
UNIVERSITY OF LUXEMBOURG