A constrained RL control policy for contact-rich non-prehensile mobile manipulation

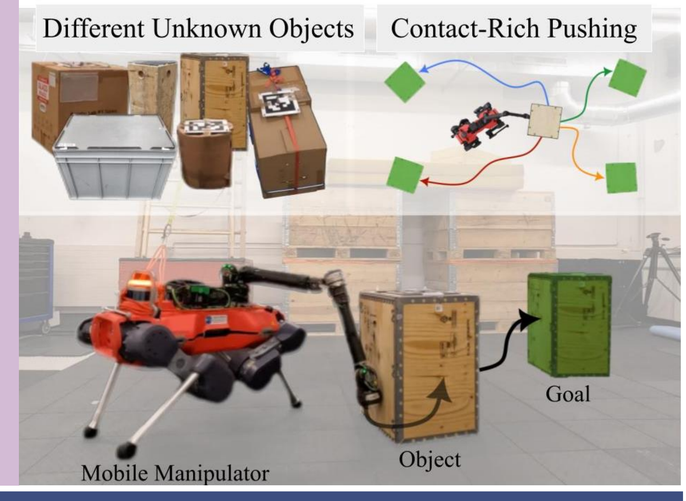
Ioannis Dadiotis,[1,2] Mayank Mittal,[3,4] Nikos Tsagarakis,[1] Marco Hutter[3,4]

[1] iit HHCM ISTITUTO ITALIANO DI TECNOLOGIA HUMANOIDS AND HUMAN CENTERED MECHATRONICS

[2] Università di Genova

[3] ETH zürich

[4] RSL Robotic Systems Lab

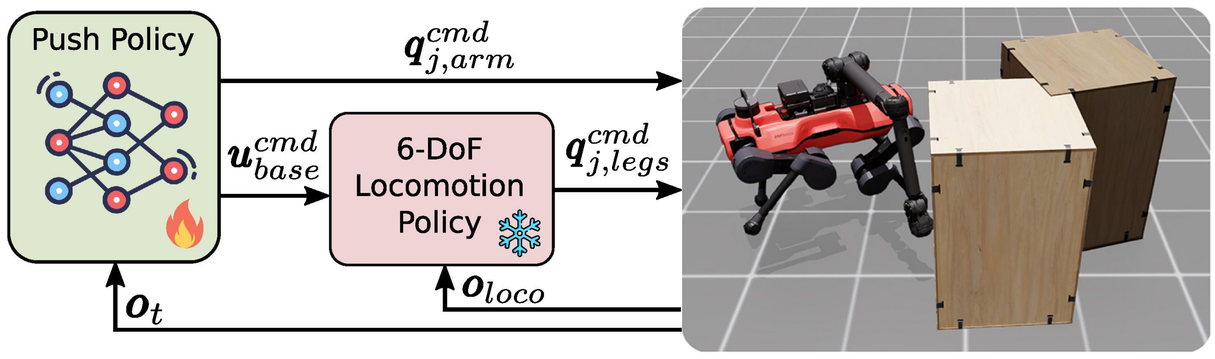


The challenge

- Push heavy/bulky objects to planar goal (x, y, yaw)
- Unknown object properties, only observe object 6D pose
- Complex, unknown frictional interactions
- Simultaneous locomotion and manipulation
- Need to dynamically make and break contact (contact-rich motion)
- Maintain object balance (avoid toppling)

Approach

- *Push policy* observations: proprioception, object 6D pose, α_{t-1}
- Actions: arm joint position targets $\boldsymbol{q}_{j,arm}^{cmd}$ + 6D base motion $\boldsymbol{u}_{base}^{cmd}$
- Pre-trained 6-DoF locomotion policy (frozen network) [Miki et al, 2024]
- Sim-to-real RL using Nvidia IsaacLab/IsaacSim

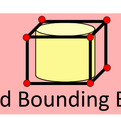


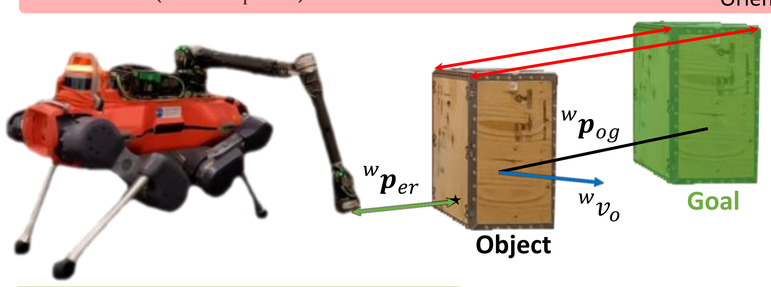
Rewards

Position and orientation matching through keypoints' distance

$$r_{1,t} = \exp\left(-\frac{\|{}^{w}\boldsymbol{K}_g - {}^{w}\boldsymbol{K}_o\|}{\sigma_1^2}\right)$$

${}^{w}\boldsymbol{K}_i$: concatenated keypoints' position





Object velocity towards goal (only direction)

$$r_{3,t} = \exp\left(\frac{{}^{w}\hat{\boldsymbol{v}}_o \cdot {}^{w}\hat{\boldsymbol{p}}_{og}}{\sigma_3^2} - 1\right)$$

Interaction between the arm EE and object surface

$$r_{2,t} = \exp\left(-\frac{\|{}^{w}\boldsymbol{p}_{er}\|}{\sigma_2^2}\right)$$

Minimize difference between current and previous actions

$$r_{4,t} = \exp\left(-\frac{|\Delta\boldsymbol{u}_{base,t}^{cmd} - \Delta\boldsymbol{u}_{base,t-1}^{cmd}|}{\sigma_{4,b}^2}\right) + \exp\left(-\frac{|\Delta\boldsymbol{q}_{j,t}^{cmd} - \Delta\boldsymbol{q}_{j,t-1}^{cmd}|}{\sigma_{4,a}^2}\right)$$

Constrained RL formulation

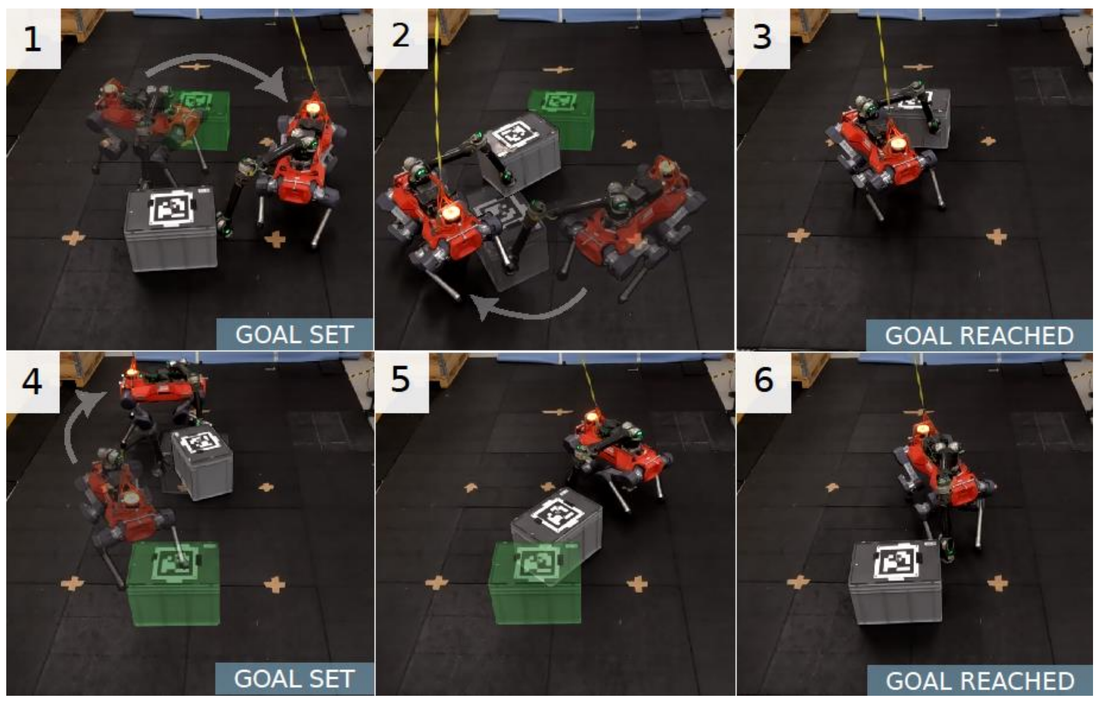
- Constrained PPO proposed in [Chane-Sane et al, 2024]
- Constraints: Action limits, actuation limits, undesired collisions, object balance
- Curriculum learning to progressively impose constraints

Domain randomization

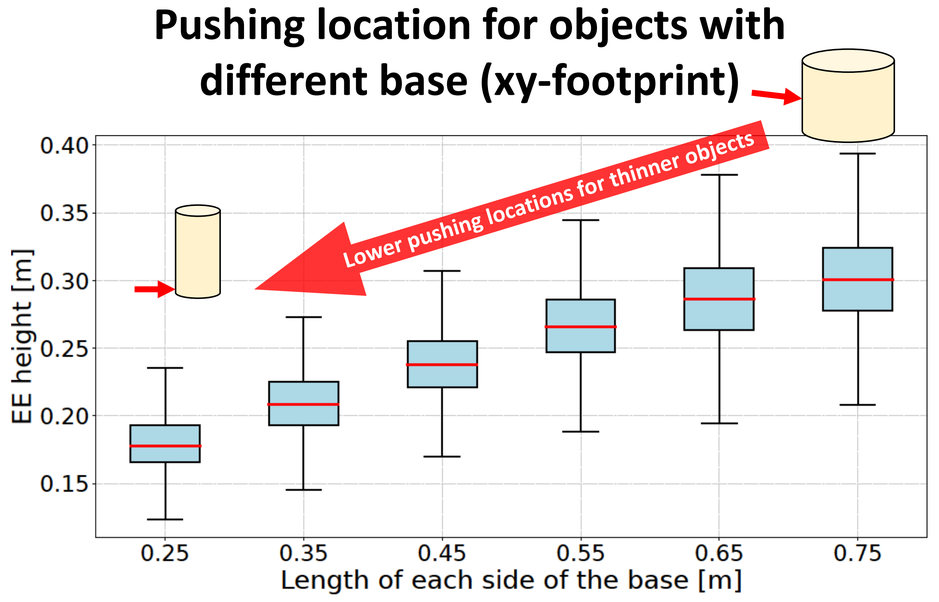
- Randomize object mass, shape, dimensions, friction, CoM position
- Episode start: randomly sample object pose, goal pose, robot pose, and arm joint configuration

Results

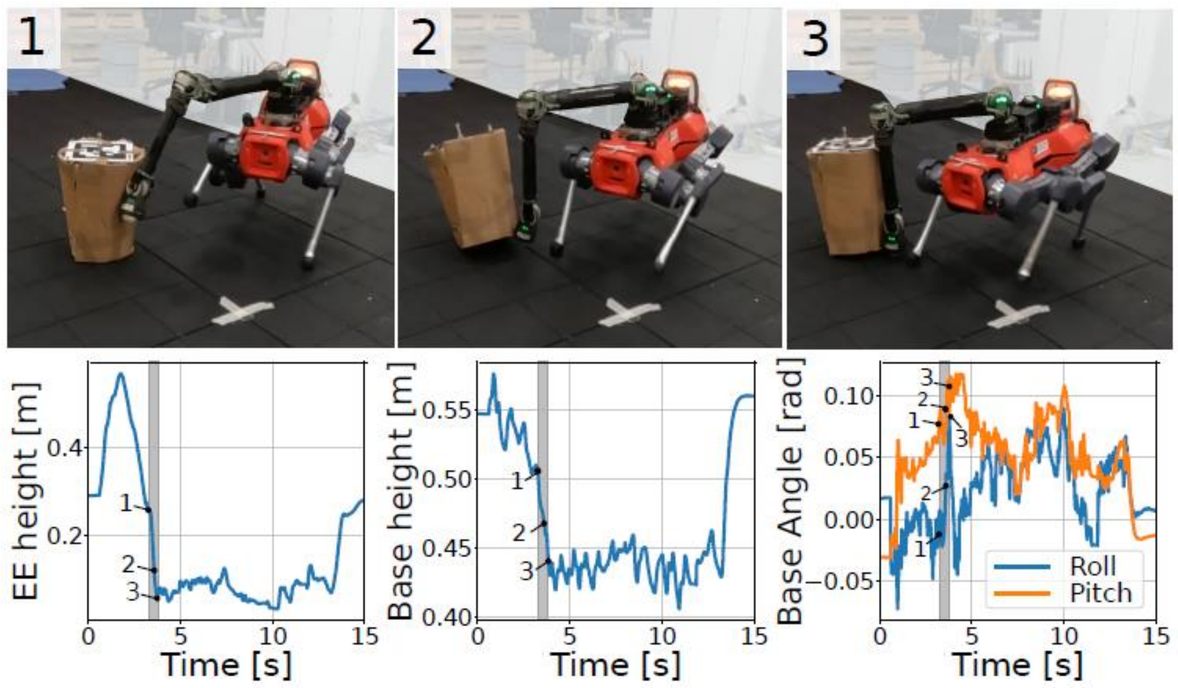
Contact-rich behavior



Pushing location for objects with different base (xy-footprint)



Reactively adapting the pushing location to keep the object balanced



Robustness with object of different material (Plastic, Cardboard, Wood), shape (CUboid, CYlinder) and size

Object	Mass [kg]	Size [cm^3]	$\Delta\theta_z$ [deg]	# of face switches / goal	Success rate [%]
P-CU	6.43	60x34x40	180	0.90	91.6
C-CU	5.30	50x50x53	0	0.23	92.9
C-CU	8.32	50x50x53	90	0.75	83.3
C-CU	4.5	100x50x53	0	0.14	80.0
W-CU	6.30	40x40x60	180	1.00	91.6
C-CU	13.30	50x50x60	0	4.80	83.3
C-CY	2.45	Φ30x40	0	-	83.3

ioannis.dadiotis@iit.it

Ioannis Dadiotis

Paper

Video

ICRA IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION ATLANTA 2025

euROBIN ETH RobotX