# FALCON: Learning Force-Adaptive Humanoid Loco-Manipulation

**Yuanhang Zhang** YUANHANZ@ANDREW.CMU.EDU
**Yifu Yuan** YIFUY@ANDREW.CMU.EDU
**Prajwal Gurunath** PGURUNAT@ANDREW.CMU.EDU
**Ishita Gupta** ISHITAG@ANDREW.CMU.EDU
*Carnegie Mellon University*

**Shayegan Omidshafiei** SHAYEGAN@FIELDAI.COM
**Ali-akbar Agha-mohammadi** ALI@FIELDAI.COM
*Field AI*

**Marcell Vazquez-Chanlatte** MARCELL.VAZQUEZCHANLATTE@NISSAN-USA.COM
**Liam Pedersen** LIAM.PEDERSEN@NISSAN-USA.COM
*Nissan USA*

**Tairan He** TAIRANH@ANDREW.CMU.EDU
**Guanya Shi** GUANYAS@ANDREW.CMU.EDU
*Carnegie Mellon University*

**Editors:** G. Sukhatme, L. Lindemann, S. Tu, A. Wierman, N. Atanasov

## Abstract

Humanoid loco-manipulation holds transformative potential for daily service and industrial tasks, yet achieving precise, robust whole-body control with 3D end-effector force interaction remains a major challenge. Prior approaches are often limited to lightweight tasks or quadrupedal/wheeled platforms. To overcome these limitations, we propose FALCON, a dual-agent reinforcement-learning-based framework for robust force-adaptive humanoid loco-manipulation. FALCON decomposes whole-body control into two specialized agents: (1) a lower-body agent ensuring stable locomotion under external force disturbances, and (2) an upper-body agent precisely tracking end-effector positions with implicit adaptive force compensation. These two agents are jointly trained in simulation with a force curriculum that progressively escalates the magnitude of external force exerted on the end effector while respecting torque limits. Experiments demonstrate that, compared to the baselines, FALCON achieves 2× more accurate upper-body joint tracking, while maintaining robust locomotion under force disturbances and achieving faster training convergence. Moreover, FALCON enables policy training without embodiment-specific reward or curriculum tuning. Using the same training setup, we obtain policies that are deployed across multiple humanoids, enabling forceful loco-manipulation tasks such as transporting payloads (0-20N force), cart-pulling (0-100N), and door-opening (0-40N) in the real world. Videos: https://lecar-lab.github.io/falcon-humanoid/

**Keywords:** Humanoid Loco-Manipulation, Reinforcement Learning, Force Adaptivity

## 1. Introduction

Humanoid robots have demonstrated remarkable progress in locomotion and manipulation Gu et al. (2025); Darvish et al. (2023); Dynamics (2024); Unitree (2024); AI (2024); Robotics (2025). However, extending these capabilities to forceful loco-manipulation remains fundamentally challenging. Tasks such as door opening, highlighted in the 2015 DARPA Challenge Krotkov et al. (2017), require not only precise manipulation under dynamic, multi-directional forces but also maintaining

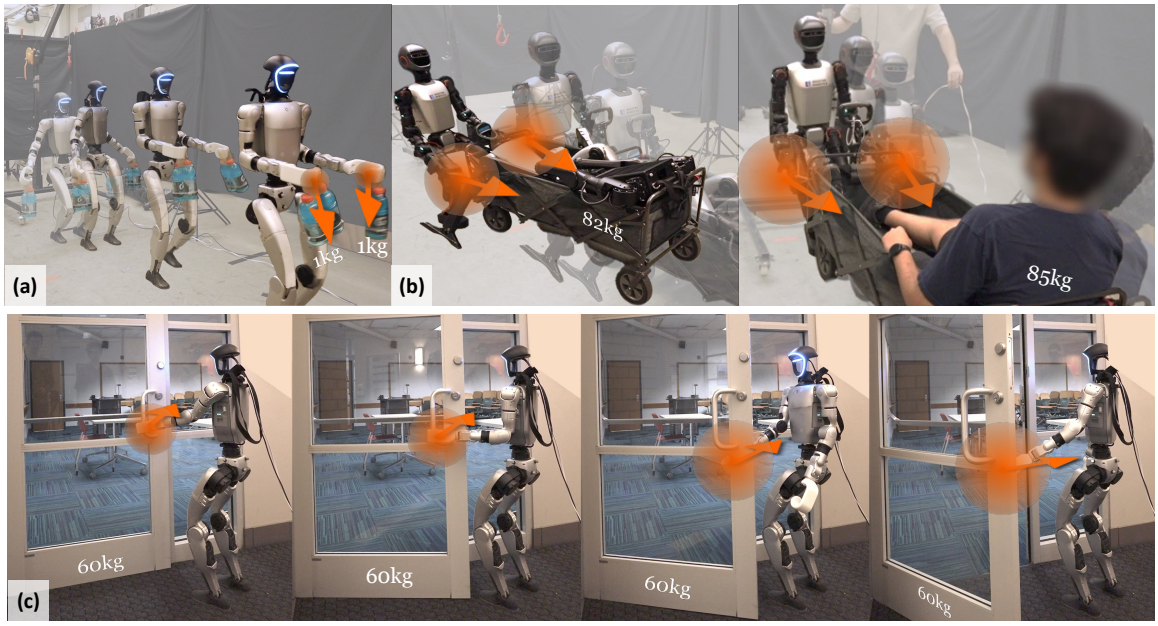



Figure 1: FALCON enables humanoids to perform versatile forceful loco-manipulation tasks: (a) **Transporting Payloads**: walk, squat, twist torso with payloads; (b) **Cart-Pulling** with significant longitudinal forces; (c) **Door-Opening** with multi-directional forces.

lower-body stability throughout the interaction. Meeting these demands calls for humanoid systems that can flexibly adapt to varying payloads and contact forces without compromising overall precision and robustness in loco-manipulation.

Reinforcement Learning (RL) has achieved impressive results for humanoid whole-body control Lu et al. (2024); Ben et al. (2025); Ji et al. (2024); Cheng et al. (2024); Sombolestan and Nguyen (2024); Fey et al. (2025); He et al. (2024a, 2025); Li et al. (2023); Murooka et al. (2021); Bouyarmane et al. (2018); Fu et al. (2024), yet existing RL approaches succeed mostly on lightweight tasks but do not consider significant interaction force during loco-manipulation tasks. Currently, there are two main paradigms: (1) *Lower-RL-Upper-IK*, which applies RL to lower-body locomotion while using kinematic solvers for upper-body control Lu et al. (2024); Ben et al. (2025), lacks whole-body dynamics modeling for forceful interaction and has limited whole-body coordination; (2) *Monolithic-Whole-body-RL*, which directly learns to control all degrees of freedom Ji et al. (2024); Cheng et al. (2024), suffers from inefficient exploration as a single policy must simultaneously learn weakly correlated locomotion and manipulation skills. Although some advances have been made in force adaptation for quadrupeds Sombolestan and Nguyen (2023, 2024); Fey et al. (2025); Cheng et al. (2025); Zhong et al. (2025), humanoids pose extra challenges like instability, higher complexity, and stricter torque limits, especially in certain joint configurations.

In this work, we aim to develop an RL framework that enables humanoid robots to perform a diverse set of force-adaptive loco-manipulation tasks. To this end, we introduce FALCON, a dual-agent RL architecture trained with a carefully designed 3D force curriculum respecting joint torque limits. Our key innovations include: (1) A *dual-agent learning decomposition* that separates lower-body and upper-body policy training with tailored rewards while sharing the same whole-body

proprioception and commands; (2) A *3D force curriculum* with joint torque feasibility that progressively scales applied 3D forces on both end-effectors (EEs) while enforcing joint torque constraints through inverse dynamics. FALCON enables efficient joint training of both stable locomotion and accurate EE tracking in forceful loco-manipulation tasks. We validate FALCON on Unitree G1 and Booster T1 humanoids, demonstrating its generalization across different platforms through: (1) Transporting Payloads, (2) Cart-Pulling, and (3) Door-Opening (Fig. 1), which require real-time adaptation to significant unknown 3D interaction force. In summary, our main contributions are:

- We introduce FALCON, a dual-agent reinforcement learning framework that enables humanoids to perform forceful loco-manipulation while adapting to substantial, unknown end-effector forces (0–100N, up to 30% of body weight). FALCON improves the upper-body joint tracking accuracy over prior methods by 100% while maintaining robust locomotion performance.
- To facilitate the efficient RL training, we design a 3D force curriculum with progressive force application while ensuring joint torque feasibility and maximizing its force adaptivity.
- We validate FALCON on two different humanoid platforms (Unitree G1, Booster T1), achieving strong cross-platform generalization with minimal tuning overhead.

## 2. Related Works

### 2.1. Humanoid Loco-Manipulation

Humanoid loco-manipulation remains a challenging control problem in robotics. While traditional model-based methods (e.g., simplified dynamics models and MPC) Sombolestan and Nguyen (2024); Li et al. (2023); Murooka et al. (2021); Bouyarmane et al. (2018); Di Carlo et al. (2018); Kajita et al. (2003); Xue et al. (2024) offer real-time planning, their reliance on manual design limits flexibility and generalizability. In contrast, learning-based methods—particularly sim-to-real RL—have demonstrated promising results in versatile loco-manipulation tasks Zhang et al. (2024); He et al. (2024a); Cheng et al. (2024); Lu et al. (2024); Ben et al. (2025); Dao et al. (2024); Liu et al. (2024). For humanoids, two primary paradigms have emerged: *Lower-RL-Upper-IK* and *Monolithic-Whole-body-RL*. For *Lower-RL-Upper-IK*, Lu et al. (2024) introduce PMP, which uses inverse kinematics (IK) and PD control for upper body control while locomotion is trained and conditioned on a Conditional Variational Autoencoder (CVAE) representing upper-body motions. Then, Ben et al. (2025) propose HOMIE that follows the same decoupling framework but introduces an exoskeleton-based cockpit for more intuitive human teleoperation. For *Monolithic-Whole-body-RL*, Dao et al. (2024) adopted a unified RL approach for box pick-and-place tasks, training distinct skills (e.g., lifting, walking, stance) and orchestrating them via a finite state machine. He et al. (2024b,a) and Ji et al. (2024) employ a teacher-student training framework to mimic human motions for loco-manipulation tasks.

Despite these advances, few RL methods address significant unknown force disturbances on the EEs for humanoid loco-manipulation, and both paradigms exhibit critical shortcomings accordingly. *Lower-RL-Upper-IK* approaches suffer from delayed force compensation for upper-body control. *Monolithic-Whole-body-RL* methods face sample inefficiency from coarsely related task objectives between upper-body manipulation and lower-body locomotion, often leading to overfitting and the behavioral dominance of either upper or lower body. In this study, inspired by Zhang et al. (2021), we propose FALCON, a dual-agent RL framework employing task-specific reward formulations

for upper-lower body decomposition. Unlike separately trained architectures, the two agents in FALCON are jointly trained with shared proprioception and commands, allowing mutual awareness of each other's behaviors. This joint training prevents the agents from adapting in isolation and enables coordinated responses to external forces that affect the full-body dynamics.

### 2.2. Forceful Interaction in Legged Robots

Forceful interaction has been extensively studied for quadrupedal robots with mounted arms, through model-based approaches—particularly MPC combined with force planning and control for robust and adaptive locomotion and manipulation Sombolestan and Nguyen (2023, 2024); Rigo et al. (2024). Recent advances in RL have further enhanced adaptability, enabling quadrupeds to learn adaptive and agile force interactions including impedance control Portela et al. (2024) and aggressive force adaptation Fey et al. (2025). For humanoids, forceful interaction presents significantly greater challenges due to their more complex dynamics and stringent joint limits. Unlike quadrupeds with centralized mass distributions, humanoids exhibit coupled dynamics between their upper and lower bodies, making force adaptation particularly difficult. Recent model-based approaches have demonstrated force control for heavy-duty tasks Li and Nguyen (2023); Murooka et al. (2021), but these require prior knowledge of manipulated objects' mass, center of mass (CoM), or pre-defined force trajectories, limiting their applicability to unknown disturbances. While some works have attempted explicit force estimation for humanoids Mattioli and Vendittelli (2016), they are restricted to quasi-static scenarios and cannot handle force adaptation in dynamic loco-manipulation scenarios.

In this paper, FALCON learns to implicitly adapt to unknown external forces on the different EEs with a novel 3D EE force curriculum that considers humanoid joint torque limits. In this way, we can maximize the force adaptability of the learned loco-manipulation policy while ensuring the joint torque limits for robust and safe real-world deployment.

## 3. FALCON: Force-Adaptive Humanoid Loco-Manipulation

Humanoid loco-manipulation under external EE forces requires coordinated control of both the lower and upper body. We first formulate the problem as a unified *dual goal-conditioned* policy learning problem. Let the degrees of freedom (DoFs) of the humanoid be partitioned into lower-body joints and upper-body joints, with $n^l$ denoting the number of lower-body DoFs, $n^u$ the number of upper-body DoFs, and $n = n^l + n^u$ the total number of actuated joints.

The robot proprioception $\boldsymbol{s}_t^p \in \mathcal{S}_t$ is defined as $\boldsymbol{s}_t^p \triangleq [\boldsymbol{q}_{t-4:t}, \dot{\boldsymbol{q}}_{t-4:t}, \boldsymbol{\omega}_{t-4:t}^{\text{root}}, \boldsymbol{g}_{t-4:t}, \boldsymbol{a}_{t-5:t-1}]$, which contains five-step histories of joint positions $\boldsymbol{q}_t \in \mathbb{R}^n$, joint velocities $\dot{\boldsymbol{q}}_t \in \mathbb{R}^n$, root angular velocity $\boldsymbol{\omega}_t^{\text{root}} \in \mathbb{R}^3$, projected gravity $\boldsymbol{g}_t \in \mathbb{R}^3$, and previous actions $\boldsymbol{a}_{t-1} \in \mathbb{R}^n$. The goal space $\mathcal{G}_t$ consists of locomotion goals $\mathcal{G}_t^l \triangleq [\mathbf{v}_t^{\text{lin,ang}}, \phi_t^{\text{stance}}, h_t^{\text{root}}, w_t^{\text{yaw}}]$, specifying desired root linear and angular velocities, stance indicators, root heights, and waist yaw angles, and manipulation goals $\mathcal{G}_t^u \triangleq [\mathbf{q}_t^{\text{upper}*}]$, specifying target joint configurations for the upper body where $\mathbf{q}_t^{\text{upper}*} \in \mathbb{R}^{n^u}$. Under this unified formalism, conventional methods differ mainly in how they generate the action $\boldsymbol{a}_t \in \mathbb{R}^n$ that commands the robot joints:

- *Lower-RL-Upper-IK*: lower-body actions $\boldsymbol{a}_t^l \in \mathcal{A}_t^l \subset \mathbb{R}^{n^l}$ are generated by a policy $\pi^l : \boldsymbol{s}_t^p \times \langle \mathcal{G}_t^l, \mathcal{G}_t^u \rangle \mapsto \mathcal{A}_t^l$ conditioned on whole-body proprioception and goals, while upper-body actions $\boldsymbol{a}_t^u \in \mathcal{A}_t^u \subset \mathbb{R}^{n^u}$ are computed through inverse kinematics (IK) solvers based on $\mathcal{G}_t^u$.

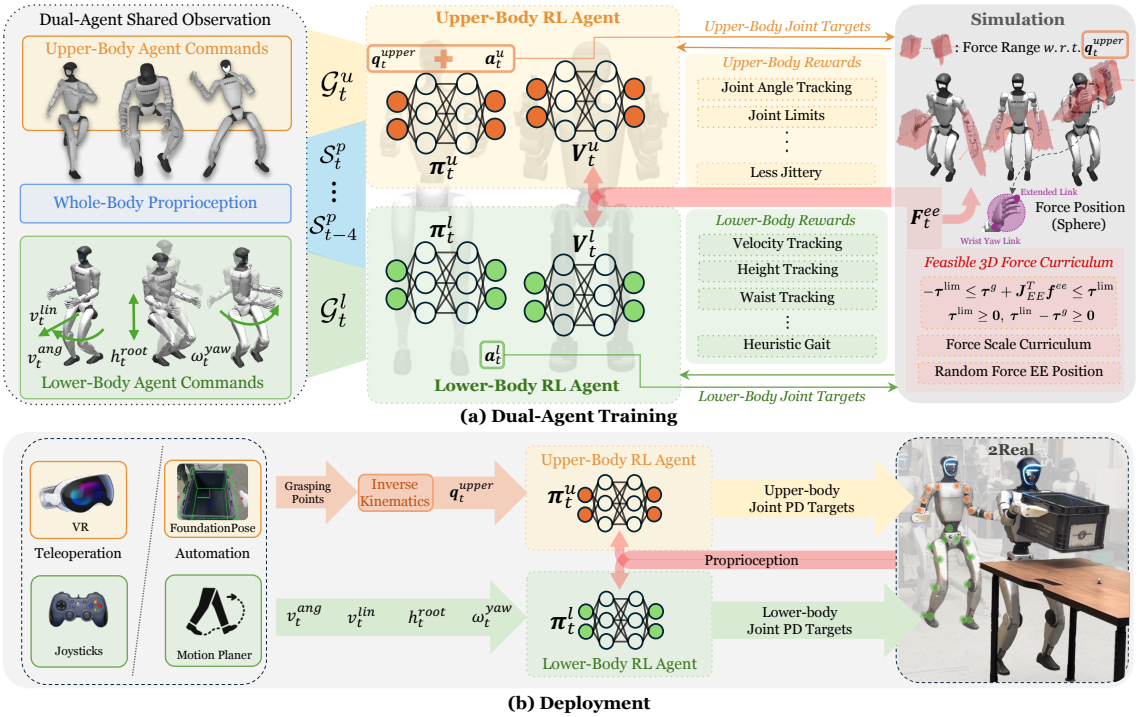



Figure 2: Overview of FALCON. (a) Two agents with different sub-tasks are jointly trained with shared whole-body proprioception. During training, we apply 3D external forces bounded by upper-body joint torque limits on the end-effectors; (b) FALCON is deployed with either teleoperation or an autonomy pipeline with FoundationPose Wen et al. (2024) for pose estimation and motion planning.

- *Monolithic-Whole-body-RL*: a single policy $\pi : \boldsymbol{s}_t^p \times \mathcal{G}_t \mapsto \boldsymbol{a}_t$ directly predicts the full-body action $\boldsymbol{a}_t \in \mathbb{R}^n$, attempting to satisfy both locomotion and manipulation objectives simultaneously.

While *Lower-RL-Upper-IK* methods are sample-efficient, they neglect upper-body force compensation and whole-body coupling under EE force disturbances. In contrast, *Monolithic-Whole-body-RL* methods improve expressiveness but suffer from exploration inefficiency due to the large action space spanning coarsely related locomotion and manipulation objectives. To overcome these challenges, we introduce FALCON, a dual-agent RL framework that achieves training efficiency and coordination through decomposition learning with shared whole-body observation.

### 3.1. Dual-Agent Learning Framework

As shown in Fig. 2, FALCON jointly trains two agents, each specialized for a different subtask. The lower-body locomotion agent learns a policy $\pi^l : \boldsymbol{s}_t^p \times \mathcal{G}_t^l \mapsto \mathcal{A}_t^l$ with value function $V^l(\cdot)$, while the upper-body manipulation agent learns a policy $\pi^u : \boldsymbol{s}_t^p \times \mathcal{G}_t^u \mapsto \mathcal{A}_t^u$ with value function $V^u(\cdot)$. Both agents observe the same proprioceptive input $\boldsymbol{s}_t^p$ but optimize independent goal-conditioned objectives:

$$r_t^l = \mathcal{R}^l(\boldsymbol{s}_t^p, \mathcal{G}_t^l) \text{ (locomotion)} \quad r_t^u = \mathcal{R}^u(\boldsymbol{s}_t^p, \mathcal{G}_t^u) \text{ (manipulation)} \tag{1}$$

These two policy parameters $\theta_l$ and $\theta_u$ are updated via proximal policy optimization (PPO Schulman et al. (2017)):

$$\max_{\theta_l} \mathbb{E}\left[\sum_{t=1}^{T} \gamma^{t-1} r_t^l\right] \text{ (Lower-body)} \quad \max_{\theta_u} \mathbb{E}\left[\sum_{t=1}^{T} \gamma^{t-1} r_t^u\right] \text{ (Upper-body)} \tag{2}$$

where $\gamma$ is the discount factor. During training, upper-body target joint angles $\mathbf{q}_t^{\text{upper}}$ (shoulders, elbows, wrists) are sampled from AMASS Mahmood et al. (2019), and computed via IK at deployment. The combined action $\boldsymbol{a}_t = [\boldsymbol{a}_t^l; \boldsymbol{a}_t^u]$ drives a joint-level PD controller. Given the partial observability of real-world control, we use asymmetric actor–critic training: critics access privileged data (root velocities, EE forces $\boldsymbol{F}_t^{ee}$) only during training. Reward terms and domain randomization details are given in A.1 and A.2.

### 3.2. Torque-Limit-Aware 3D Force Curriculum

For humanoid robots with limited joint torques—such as the wrist joints on the Unitree G1—it is essential to account for these constraints when large external forces act on the end-effectors (EEs). Ignoring them during training can cause unsafe behaviors from torque saturation or joint limit violations. Additionally, it's important to gradually increase the external force during training, allowing the policy to progressively learn effective force adaptation strategies. To achieve these, our force application framework follows through three principles:

**Torque-Aware Force Computation:** Before applying forces, we first need to estimate the maximum forces that we can exert on the left or right end-effector. Given the left or right end-effector Jacobian $\boldsymbol{J}_{EE} \in \mathbb{R}^{3\times\frac{n_u}{2}}$ at its Center of Mass (CoM), their joint torque limit $\boldsymbol{\tau}^{\lim} \in \mathbb{R}^{3\times\frac{n_u}{2}}$ (with $\boldsymbol{\tau}^{\lim} \geq \mathbf{0}$), and the gravity compensation torque $\boldsymbol{\tau}^g \in \mathbb{R}^{3\times\frac{n_u}{2}}$ which satisfies $-\boldsymbol{\tau}^{\lim} \leq \boldsymbol{\tau}^g \leq \boldsymbol{\tau}^{\lim}$ to ensure feasibility, we estimate the maximum and minimum admissible forces $\boldsymbol{f}^{\max}, \boldsymbol{f}^{\min}$ along each Cartesian axis $i \in x, y, z$ by analyzing the worst-case joint torque induced by a unit force applied in each direction. The element-wise force bound can be computed in parallel as:

$$-\boldsymbol{\tau}^{\lim} \leq \boldsymbol{\tau}^g + \boldsymbol{J}_{EE}^T \boldsymbol{f}^{ee} \leq \boldsymbol{\tau}^{\lim} \tag{3}$$

$$f_i^{\max} = \min_j \left(\frac{\tau_j^{\lim} - \tau_j^g}{|J_{EE}^{ji}| + \epsilon}\right), \quad f_i^{\min} = \max_j \left(\frac{-\tau_j^{\lim} - \tau_j^g}{|J_{EE}^{ji}| + \epsilon}\right), \tag{4}$$

where $J_{EE}^{ji}$ is the $(j,i)$-th element of the end-effector Jacobian, and $\epsilon$ is a small constant to avoid numerical singularities. Since the mapping $\boldsymbol{\tau} q = JEE^\top \mathbf{f}^{ee}$ between end-effector force $\mathbf{f}^{ee} = [\mathbf{f}, , \boldsymbol{\tau}]$ and joint torques is convex under typical torque limits, the feasible force space is also convex. Exploiting this, we adopt a probabilistic sampling scheme ensuring forces remain within admissible torque bounds: the directional ratios $\boldsymbol{\gamma} = [\gamma_x, , \gamma_y, , \gamma_z]$ are drawn from a Dirichlet distribution Ng et al. (2011) satisfying $\sum_i \gamma_i = 1$, thereby preserving convexity and physical feasibility. The feasible end-effector force is then uniformly sampled within the estimated magnitude range and expressed as:

$$\boldsymbol{f}_t^{ee} = \sum_{i\in\{x,y,z\}} F_i \cdot \boldsymbol{e}_i, \quad \text{where} \quad F_i \sim \mathcal{U}[\gamma_i \cdot f_i^{\min}, \gamma_i \cdot f_i^{\max}] \tag{5}$$

This approach maximizes force adaptivity while respecting torque limits, leading to more effective training than random sampling, as explained later in 3.2. Note that applied forces may differ between left and right EEs due to asymmetric upper-body configurations (2).

**Progressive Force Curriculum:** To facilitate progressive force adaptation, the estimated EE forces are scaled by a global factor $\alpha_g \in (0, 1)$, increasing over training, so the applied force becomes $\boldsymbol{F}_t^{ee} = \alpha_g \cdot \boldsymbol{f}_t^{ee}$. During walking, planar forces are projected opposite to the velocity. A low-pass filter is applied to reduce force jitter.

**Position Randomization of the Applied Force:** Learning-based force adaptation leverages proprioceptive history to implicitly compensate for external forces, removing the need for explicit force estimation Mattioli and Venditteli (2016) or sensing Guo et al. (2024). To improve robustness to variations in end-effector (EE) contact points—which alter the torque mapping via the EE Jacobian—we randomize force application along the EE link, from the wrist yaw to the distal segment, as illustrated in 2.

## 4. Simulation and Real-World Experiments

In this section, we present extensively quantitative comparison between FALCON and the baselines as well as qualitative results on real-world deployment. We choose Unitree Humanoid G1 and Booster T1 as our humanoid platforms. Specifically, we address the following key questions:

**Q1**: Can FALCON outperform other baselines in terms of both upper-body manipulation and lower-body locomotion performance?

**Q2**: Why does FALCON has better training-efficiency compared to Monolithic-Whole-body-RL (M-WB-RL) for force-adaptive loco-manipulation?

**Q3**: Does FALCON work for different humanoids to show cross-platform generalizability?

| | | $E^{\text{upper}}_{\text{tracking}} \downarrow$ | | | $E^{\text{root}}_{\text{tracking}} \downarrow$ | | |
|---|---|---|---|---|---|---|---|
| Methods | | N-Force | M-Force | L-Force | N-Force | M-Force | L-Force |
| Lower-RL Upper-IK | PD-w/o-Force-Curr. | 0.46 $\pm$ 0.04 | 0.94 $\pm$ 0.05 | 1.44 $\pm$ 0.06 | 0.38 $\pm$ 0.04 | 0.66 $\pm$ 0.05 | 1.14 $\pm$ 0.06 |
| | PD-Force-Curr. | 0.44 $\pm$ 0.03 | 0.93 $\pm$ 0.03 | 1.42 $\pm$ 0.04 | 0.33 $\pm$ 0.03 | 0.38 $\pm$ 0.03 | 0.46 $\pm$ 0.03 |
| | PID-Force-Curr. | 0.24 $\pm$ 0.03 | 0.31 $\pm$ 0.03 | 0.60 $\pm$ 0.04 | 0.32 $\pm$ 0.03 | 0.35 $\pm$ 0.03 | 0.46 $\pm$ 0.04 |
| | PD-ID-Force-Curr. | 0.29 $\pm$ 0.03 | 0.38 $\pm$ 0.03 | 0.53 $\pm$ 0.04 | 0.40 $\pm$ 0.03 | 0.42 $\pm$ 0.03 | 0.47 $\pm$ 0.04 |
| M-WB-RL | w/o-Force-Curr. | 0.46 $\pm$ 0.05 | 1.03 $\pm$ 0.07 | 1.65 $\pm$ 0.08 | 0.34 $\pm$ 0.04 | 0.67 $\pm$ 0.05 | 1.28 $\pm$ 0.07 |
| | with-Force-Curr. | 0.43 $\pm$ 0.04 | 0.50 $\pm$ 0.04 | 0.73 $\pm$ 0.05 | **0.28** $\pm$ 0.03 | **0.32** $\pm$ 0.03 | **0.44** $\pm$ 0.04 |
| **FALCON** | w/o-Force-Curr. | **0.14** $\pm$ 0.03 | 0.55 $\pm$ 0.04 | 1.06 $\pm$ 0.06 | 0.32 $\pm$ 0.03 | 0.66 $\pm$ 0.05 | 1.24 $\pm$ 0.06 |
| | with-Force-Curr. | 0.21 $\pm$ 0.03 | **0.24** $\pm$ 0.03 | **0.37** $\pm$ 0.04 | **0.27** $\pm$ 0.03 | **0.30** $\pm$ 0.03 | **0.45** $\pm$ 0.04 |

Table 1: Loco-Manipulation Evaluation of FALCON and Baselines in IsaacGym.

### 4.1. Evaluation Criterion

To evaluate the performance of the learned low-body locomotion and upper-body manipulation capabilities, we consider the following metrics under dynamically unknown and 3D EE forces $\mathbf{F}_t \in \mathbb{R}^3$, given a sequence of target upper-body joints $\boldsymbol{q}_t^{\text{upper}*}$, target root velocities $\boldsymbol{v}_t^{\text{lin,ang}*}$ and stance signal $\phi_t^{\text{stance}}$, where $t = 1, 2, ..., T$ and $T$ is the sequence length:

*(ii) Upper-Body Joints Tracking Error*: $E^{\text{upper}}_{\text{tracking}}(\boldsymbol{q}_t^{\text{upper}*}) = \frac{1}{T}\sum_{t=1}^{T}\left|\boldsymbol{q}_t^{\text{upper}} - \boldsymbol{q}_t^{\text{upper}*}\right|$

*(iii) Root Velocity Tracking Error:* $E^{\text{root}}_{\text{tracking}}(\boldsymbol{v}_t^{\text{lin,ang}*}) = \frac{1}{T}\sum_{t=1}^{T}\left|\boldsymbol{v}_t^{\text{lin,ang}} - \boldsymbol{v}_t^{\text{lin,ang}*}\right|$

### 4.2. Baselines

We consider two types of baseline methods for force adaptation, both trained under the same goal space (e.g., commands) in Sec. 3.1 and force curriculum described in Sec. 3.2, with each type further including relevant ablation variants.

**Decoupled Lower-body RL with Upper-body IK Controllers.** For all variants, RL is used for lower-body locomotion, and IK provides target upper-body joint angles from end-effector poses. The key differences lie in the use of force curriculum and the upper-body joint tracking strategy:

(a) *Upper-PD-w/o-Force-Curr.*: A baseline following Lu et al. (2024); Ben et al. (2025), using PD control for upper-body joint tracking without force randomization.

(b) *Upper-PD*: Extends (a) by incorporating force curriculum, enabling lower-body adaptation to external forces; upper-body remains PD-controlled.

(c) *Upper-PID*: Extends (b) by adding an integral term to the upper-body controller to reduce steady-state tracking error.

(d) *Upper-PD-ID*: Extends (a) with a learned force estimator Portela et al. (2024) and inverse dynamics-based torque compensation under quasi-static assumptions (details in Appendix A.3).

**Monolithic Whole-body RL**. RL is used for whole-body control.

(e) *Monolithic-WB-RL-w/o-Force-Curr.*: Built upon prior designs Cheng et al. (2024); He et al. (2024c), a single agent is trained with the same goal commands as FALCON, but without applying any force during training.

(f) *Monolithic-WB-RL-with-Force-Curr.*: Based on (e), we adopt force randomization into the training curriculum for force adaptation, while keeping the other training settings identical.

### 4.3. Simulation Results

To answer **Q1** (*Can FALCON outperform other baselines in terms of both upper-body manipulation and lower-body locomotion performance?*) and **Q2** (*Why does FALCON has better training-efficiency compared to Monolithic-Whole-body-RL (M-WB-RL) for force-adaptive loco-manipulation?*), we conduct quantitative comparisons of our method with other two baselines in IsaacGym on Unitree Humanoid G1.

**Loco-Manipulation Performance:** We evaluate FALCON and baselines on 252 ACCAD Advanced Computing Center for the Arts and Design motion targets under three force levels: (i) No-Force ($\alpha_g = 0$), (ii) Middle-Force ($\alpha_g = 0.5$), and (iii) Large-Force ($\alpha_g = 1.0$), applied to both end-effectors. As shown in Tab. 1, across all settings, FALCON with force curriculum achieves the lowest tracking errors in both upper-body motion ($E^{\text{upper}}_{\text{tracking}}$) and root velocity ($E^{\text{root}}_{\text{tracking}}$), demonstrating robust manipulation under disturbance. Under L-Force, it reduces upper-body error to 0.37, outperforming PID-Force-Curr. (0.60) and M-WB-RL (0.73). Root error remains low at 0.45, indicating stable locomotion. While force curriculum benefits all methods, FALCON gains most due to its decomposed learning structure.

**Torque-Limit-Aware Force Curriculum** To assess the effectiveness of the proposed torque-limit-aware force curriculum (Sec. 3), we compare it with a baseline that samples random forces from a wide clipping range ($X : [-100\text{N}, 100\text{N}]$, $Y : [-100\text{N}, 100\text{N}]$, $Z : [-100\text{N}, 5\text{N}]$) without enforcing torque feasibility. Training curves and quantitative results are shown in Fig. 3 and Tab. 2. During evaluation, applied forces remain bounded by the estimated admissible limits.

| | Methods | $E^{\text{upper}}_{\text{tracking}} \downarrow$ | | | $E^{\text{root}}_{\text{tracking}} \downarrow$ | | |
|---|---|---|---|---|---|---|---|
| | | N-Force | M-Force | L-Force | N-Force | M-Force | L-Force |
| FALCON | w/o-Torque-Limit-Aware. | 0.42 $\pm$ 0.07 | 0.45 $\pm$ 0.08 | 0.61 $\pm$ 0.10 | 0.45 $\pm$ 0.03 | 0.46 $\pm$ 0.05 | 0.54 $\pm$ 0.05 |
| | with-Torque-Limit-Aware. | **0.23** $\pm$ 0.03 | **0.26** $\pm$ 0.03 | **0.36** $\pm$ 0.04 | **0.30** $\pm$ 0.03 | **0.32** $\pm$ 0.03 | **0.39** $\pm$ 0.02 |

Table 2: Evaluation of FALCON with torque-limit-aware (Max-Force-Estimation) vs. *w/o* torque-limit-aware curriculum in IsaacGym. Our approach yields notably better tracking, especially for upper-body manipulation under large forces.

| | Methods | $E^{\text{upper}}_{\text{tracking}} \downarrow$ | | | $E^{\text{root}}_{\text{tracking}} \downarrow$ | | |
|---|---|---|---|---|---|---|---|
| | | N-Force | M-Force | L-Force | N-Force | M-Force | L-Force |
| FALCON | Smaller-Force-Clip. | **0.21** $\pm$ 0.03 | **0.24** $\pm$ 0.03 | **0.37** $\pm$ 0.04 | **0.27** $\pm$ 0.03 | **0.30** $\pm$ 0.03 | **0.45** $\pm$ 0.04 |
| | Larger-Force-Clip. | **0.23** $\pm$ 0.03 | **0.26** $\pm$ 0.03 | **0.36** $\pm$ 0.04 | **0.30** $\pm$ 0.03 | **0.32** $\pm$ 0.03 | **0.39** $\pm$ 0.02 |

Table 3: Evaluation of FALCON with a smaller clip range in 1 versus a larger clip range in 2.

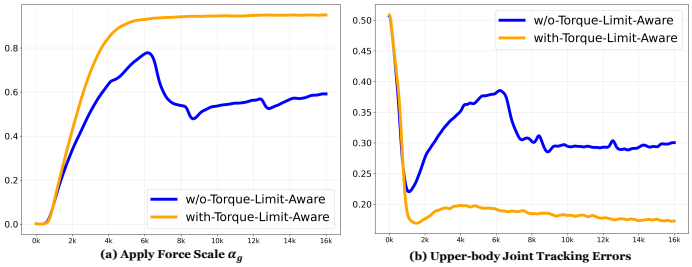



Figure 3: (a) Progression of the Apply Force Scale $\alpha_g$; (b) Upper-body Joint Tracking Errors During Training

Fig. 3 (a) shows that the force curriculum saturates at $\alpha_g = 0.6$ due to frequent torque-limit violations, which impede further progression. As shown in Fig. 3 (b), removing torque-limit awareness leads to larger upper-body tracking errors since excessive forces exceed feasible torque bounds, hindering effective compensation learning. In contrast, our torque-limit-aware curriculum enables stable learning of both joint and root velocity tracking under strong disturbances. Notably, Tab. 1 adopts a narrower force clipping range ($X, Y : [-50, 50]$N, $Z : [-60, 5]$N) than Tab. 2, yet Tab. 3 shows negligible performance loss, confirming the robustness of our curriculum.

### 4.4. Real-World Quantitative Tracking Results

We evaluate FALCON on Unitree G1 with each hand loaded with 1.2kg payload in a real-world task, which is walking at $(0.5, 0.0)$m/s with zero angular velocity, fixed height and waist, and keeping the upper body in its default position. We compare against two baselines: (i) *Upper-PD with Force Curriculum*, and (ii) *Monolithic-WB-RL with Force Curriculum*. As shown in Tab. 4, FALCON achieves the lowest tracking errors, and perform stable and natural motion in heavy-duty loco-manipulation.

### 4.5. Real-world Deployment with Teleoperation

To answer **Q3** (*Does FALCON work for different humanoids to show cross-platform generalizability?* ), we deploy simulation-trained policies on the Unitree G1 and Booster T1 humanoids ***without*** any reward or force curriculum modifications, enabled by FALCON 's dual-agent design and torque-limit-aware 3D force curriculum. As shown in Fig. 1, we evaluate three loco-manipulation tasks: (1)

| Method | $E^{\text{upper}}_{\text{tracking}}$ | $E^{\text{root}}_{\text{tracking}}$ |
|---|---|---|
| Upper-PD-Force-Curr. | 1.81 $\pm$ 0.13 | **0.40** $\pm$ 0.04 |
| M-WB-RL-Force-Curr. | 0.81 $\pm$ 0.11 | 0.58 $\pm$ 0.05 |
| **FALCON** | **0.39** $\pm$ 0.08 | **0.42** $\pm$ 0.03 |

Table 4: Real-world Tracking Errors.

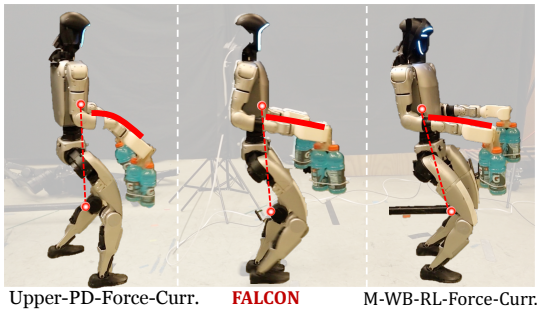



Figure 4: Real-World Payload Transportation.

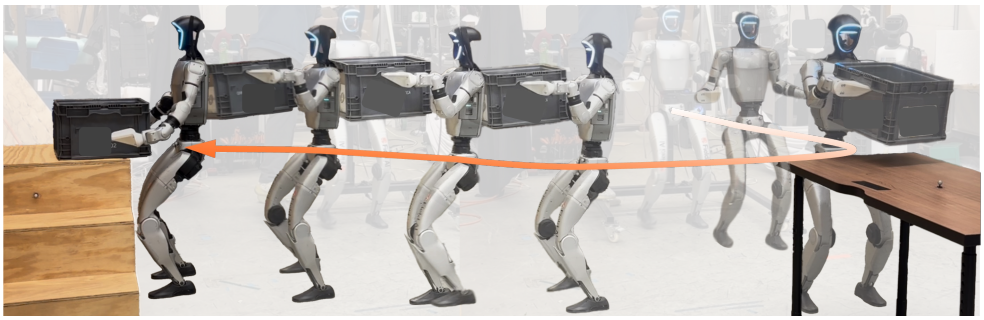


Figure 5: **Autonomous** Tote Logistics: a humanoid robot walks without a tote, picks up the tote, walks with the tote, and drops off the tote.

**Payload Transport** (0–20 N vertical forces), (2) **Cart-Pulling** (up to 100 N longitudinal forces), and (3) **Door-Opening** (up to 40 N 3D forces), with force ranges verified using a gauge, as shown in Appendix A.4.

These results highlight FALCON 's strong cross-platform robustness: policies adaptively compensate 3D forces through coordinated upper–lower body control, leaning into longitudinal forces and maintaining stable base height under heavy loads.

### 4.6. Real-World Deployment with Autonomy

We also deploy FALCON on the Unitree G1 for autonomous tote logistics, a representative warehouse task. As illustrated in Fig. 5, the robot is required to walk from an initial location to a pickup station, lift a tote of unknown weight, and transport it to a designated area for precise placement. The detailed implementation of the autonomous pipeline can be found in Appendix A.5.

## 5. Conclusion

We present FALCON, a dual-agent reinforcement learning framework for force-adaptive humanoid loco-manipulation. By decoupling upper and lower body learning with torque-limit aware force curriculum, FALCON effectively adapts to 3D end-effector forces in complex tasks. Extensive experiments show that FALCON surpasses both Lower-RL-Upper-IK and Monolithic-Whole-body-RL baselines with faster convergence, lower tracking errors, and greater stability across force regimes. Furthermore, FALCON transfers robustly from simulation to real humanoids, performing tasks such as payload transport, cart-pulling, and door-opening—demonstrating strong real-world potential for forceful interactions.

## Acknowledgments

We would like to thank our other CMU MRSD Capstone teammate Shivang Vijay, for his valuable contributions to the autonomy part in this project. We are also grateful to our Capstone advisors, Prof. John Dolan and Prof. Dimitrios Apostolopoulos, for their continuous guidance and support. We acknowledge the hardware support from Unitree Robotics and Booster Robotics. Finally, we thank Haoyang Weng, Wenli Xiao and Yitang Li for their insightful discussions that helped shape the direction of this work.

## References

Advanced Computing Center for the Arts and Design. ACCAD MoCap Dataset. URL https://accad.osu.edu/research/motion-lab/mocap-system-and-data.

Figure AI. Figure 02 humanoid robot, 2024. URL https://www.figure.ai/.

Qingwei Ben, Feiyu Jia, Jia Zeng, Junting Dong, Dahua Lin, and Jiangmiao Pang. HOMIE: humanoid loco-manipulation with isomorphic exoskeleton cockpit. CoRR, abs/2502.13013, 2025. doi: 10.48550/ARXIV.2502.13013. URL https://doi.org/10.48550/arXiv.2502.13013.

Karim Bouyarmane, Kevin Chappellet, Joris Vaillant, and Abderrahmane Kheddar. Quadratic programming for multirobot and task-space force control. IEEE Transactions on Robotics, 35(1): 64–77, 2018.

Jin Cheng, Dongho Kang, Gabriele Fadini, Guanya Shi, and Stelian Coros. Rambo: Rl-augmented model-based optimal control for whole-body loco-manipulation. arXiv preprint arXiv:2504.06662, 2025.

Xuxin Cheng, Yandong Ji, Junming Chen, Ruihan Yang, Ge Yang, and Xiaolong Wang. Expressive whole-body control for humanoid robots. arXiv preprint arXiv:2402.16796, 2024.

Jeremy Dao, Helei Duan, and Alan Fern. Sim-to-real learning for humanoid box loco-manipulation. In 2024 IEEE International Conference on Robotics and Automation (ICRA), pages 16930–16936. IEEE, 2024.

Kourosh Darvish, Luigi Penco, Joao Ramos, Rafael Cisneros, Jerry Pratt, Eiichi Yoshida, Serena Ivaldi, and Daniele Pucci. Teleoperation of humanoid robots: A survey. IEEE Transactions on Robotics, 39(3):1706–1727, 2023.

Jared Di Carlo, Patrick M Wensing, Benjamin Katz, Gerardo Bledt, and Sangbae Kim. Dynamic locomotion in the mit cheetah 3 through convex model-predictive control. In 2018 IEEE/RSJ international conference on intelligent robots and systems (IROS), pages 1–9. IEEE, 2018.

Boston Dynamics. Atlas® and beyond: the world's most dynamic robots, 2024. URL https://bostondynamics.com/atlas/.

Nolan Fey, Gabriel B Margolis, Martin Peticco, and Pulkit Agrawal. Bridging the sim-to-real gap for athletic loco-manipulation. arXiv preprint arXiv:2502.10894, 2025.

Zipeng Fu, Qingqing Zhao, Qi Wu, Gordon Wetzstein, and Chelsea Finn. Humanplus: Humanoid shadowing and imitation from humans. In Conference on Robot Learning (CoRL), 2024.

Zhaoyuan Gu, Junheng Li, Wenlan Shen, Wenhao Yu, Zhaoming Xie, Stephen McCrory, Xianyi Cheng, Abdulaziz Shamsah, Robert Griffin, C Karen Liu, et al. Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning. arXiv preprint arXiv:2501.02116, 2025.

Xiaofeng Guo, Guanqi He, Jiahe Xu, Mohammadreza Mousaei, Junyi Geng, Sebastian Scherer, and Guanya Shi. Flying calligrapher: Contact-aware motion and force planning and control for aerial manipulation. IEEE Robotics and Automation Letters, 2024.

Tairan He, Zhengyi Luo, Xialin He, Wenli Xiao, Chong Zhang, Weinan Zhang, Kris Kitani, Changliu Liu, and Guanya Shi. Omnih2o: Universal and dexterous human-to-humanoid whole-body teleoperation and learning. arXiv preprint arXiv:2406.08858, 2024a.

Tairan He, Zhengyi Luo, Wenli Xiao, Chong Zhang, Kris Kitani, Changliu Liu, and Guanya Shi. Learning human-to-humanoid real-time whole-body teleoperation. arXiv preprint arXiv:2403.04436, 2024b.

Tairan He, Wenli Xiao, Toru Lin, Zhengyi Luo, Zhenjia Xu, Zhenyu Jiang, Changliu Liu, Guanya Shi, Xiaolong Wang, Linxi Fan, and Yuke Zhu. Hover: Versatile neural whole-body controller for humanoid robots. arXiv preprint arXiv:2410.21229, 2024c.

Tairan He, Jiawei Gao, Wenli Xiao, Yuanhang Zhang, Zi Wang, Jiashun Wang, Zhengyi Luo, Guanqi He, Nikhil Sobanbab, Chaoyi Pan, et al. Asap: Aligning simulation and real-world physics for learning agile humanoid whole-body skills. arXiv preprint arXiv:2502.01143, 2025.

Mazeyu Ji, Xuanbin Peng, Fangchen Liu, Jialong Li, Ge Yang, Xuxin Cheng, and Xiaolong Wang. Exbody2: Advanced expressive humanoid whole-body control. arXiv preprint arXiv:2412.13196, 2024.

Shuuji Kajita, Fumio Kanehiro, Kenji Kaneko, Kiyoshi Fujiwara, Kensuke Harada, Kazuhito Yokoi, and Hirohisa Hirukawa. Biped walking pattern generation by using preview control of zero-moment point. In 2003 IEEE international conference on robotics and automation (Cat. No. 03CH37422), volume 2, pages 1620–1626. IEEE, 2003.

Eric Krotkov, Douglas Hackett, Lawrence D. Jackel, Michael Perschbacher, James Pippine, Jesse Strauss, Gill A. Pratt, and Christopher Orlowski. The DARPA robotics challenge finals: Results and perspectives. J. Field Robotics, 34(2):229–240, 2017. doi: 10.1002/ROB.21683. URL https://doi.org/10.1002/rob.21683.

Junheng Li and Quan Nguyen. Kinodynamics-based pose optimization for humanoid loco-manipulation. arXiv preprint arXiv:2303.04985, 2023.

Junheng Li, Junchao Ma, Omar Kolt, Manas Shah, and Quan Nguyen. Dynamic loco-manipulation on hector: Humanoid for enhanced control and open-source research. arXiv preprint arXiv:2312.11868, 2023.

Fukang Liu, Zhaoyuan Gu, Yilin Cai, Ziyi Zhou, Shijie Zhao, Hyunyoung Jung, Sehoon Ha, Yue Chen, Danfei Xu, and Ye Zhao. Opt2skill: Imitating dynamically-feasible whole-body trajectories for versatile humanoid loco-manipulation. arXiv preprint arXiv:2409.20514, 2024.

Chenhao Lu, Xuxin Cheng, Jialong Li, Shiqi Yang, Mazeyu Ji, Chengjing Yuan, Ge Yang, Sha Yi, and Xiaolong Wang. Mobile-television: Predictive motion priors for humanoid whole-body control. arXiv preprint arXiv:2412.07773, 2024.

Naureen Mahmood, Nima Ghorbani, Nikolaus F. Troje, Gerard Pons-Moll, and Michael J. Black. AMASS: Archive of motion capture as surface shapes. In International Conference on Computer Vision, pages 5442–5451, October 2019.

Tommaso Mattioli and Marilena Vendittelli. Interaction force reconstruction for humanoid robots. IEEE Robotics and Automation Letters, 2(1):282–289, 2016.

Masaki Murooka, Kevin Chappellet, Arnaud Tanguy, Mehdi Benallegue, Iori Kumagai, Mitsuharu Morisawa, Fumio Kanehiro, and Abderrahmane Kheddar. Humanoid loco-manipulations pattern generation and stabilization control. IEEE Robotics and Automation Letters, 6(3):5597–5604, 2021.

Kai Wang Ng, Guo-Liang Tian, and Man-Lai Tang. Dirichlet and related distributions: Theory, methods and applications. 2011.

Tifanny Portela, Gabriel B Margolis, Yandong Ji, and Pulkit Agrawal. Learning force control for legged manipulation. In 2024 IEEE International Conference on Robotics and Automation (ICRA), pages 15366–15372. IEEE, 2024.

Alberto Rigo, Muqun Hu, Satyandra K Gupta, and Quan Nguyen. Hierarchical optimization-based control for whole-body loco-manipulation of heavy objects. In 2024 IEEE International Conference on Robotics and Automation (ICRA), pages 15322–15328. IEEE, 2024.

Booster Robotics. Booster t1 humanoid robot, 2025. URL https://www.boosterobotics.com/.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. CoRR, abs/1707.06347, 2017. URL http://arxiv.org/abs/1707.06347.

Mohsen Sombolestan and Quan Nguyen. Hierarchical adaptive loco-manipulation control for quadruped robots. In 2023 IEEE International Conference on Robotics and Automation (ICRA), pages 12156–12162. IEEE, 2023.

Mohsen Sombolestan and Quan Nguyen. Adaptive force-based control of dynamic legged locomotion over uneven terrain. IEEE Transactions on Robotics, 2024.

Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2012, Vilamoura, Algarve, Portugal, October 7-12, 2012, pages 5026–5033. IEEE, 2012. doi: 10.1109/IROS.2012.6386109. URL https://doi.org/10.1109/IROS.2012.6386109.

Unitree. Humanoid agent ai avatar, 2024. URL https://www.unitree.com/g1.

Bowen Wen, Wei Yang, Jan Kautz, and Stan Birchfield. Foundationpose: Unified 6d pose estimation and tracking of novel objects. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 17868–17879, 2024.

Haoru Xue, Chaoyi Pan, Zeji Yi, Guannan Qu, and Guanya Shi. Full-order sampling-based mpc for torque-level locomotion control via diffusion-style annealing. arXiv preprint arXiv:2409.15610, 2024.

Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. Multi-agent reinforcement learning: A selective overview of theories and algorithms. Handbook of reinforcement learning and control, pages 321–384, 2021.

Yuanhang Zhang, Tianhai Liang, Zhenyang Chen, Yanjie Ze, and Huazhe Xu. Catch it! learning to catch in flight with mobile dexterous hands. arXiv preprint arXiv:2409.10319, 2024.

Yichao Zhong, Chong Zhang, Tairan He, and Guanya Shi. Bridging adaptivity and safety: Learning agile collision-free locomotion across varied physics. arXiv preprint arXiv:2501.04276, 2025.

# Appendix A.

## A.1. Reward Terms

We adopt the similar reward terms from He et al. (2024a); Cheng et al. (2024), but introduce some important penalties to ensure the locomotion stability under significant external forces, and other tracking rewards for squat and waist twist. The additioanl reward terms are summerized in 5:

Table 5: Additional Reward components and weights: penalty rewards for preventing undesired behaviors for sim-to-real transfer, and task rewards to achieve desired loco-manipulation capability.

| Term | Expression | Weight |
|---|---|---|
| | Penalty | |
| Hip pos | $\|\boldsymbol{q}^{hip}_{roll,pitch}\|$ | -2.5 |
| Negative knee joint | $\sum_j \mathbb{1}[q_j < q_j^{\min}]$ | -1.0 |
| Stance tap feet | $\|(\boldsymbol{p}_{left\_foot} - \boldsymbol{p}_{right\_foot})_x\|$ | -5.0 |
| Stance root | $\|(\boldsymbol{p}_{\text{root}} - \text{mid}(\boldsymbol{p}_{\text{feet}}))^y\|$ | -5.0 |
| Stand still | $\mathbb{1}[\text{no contact}]$ | -0.15 |
| Ankle roll | $\sum_j \|q_j^{\text{roll}}\|$ | -2.0 |
| | Task Reward | |
| Root linear velocity x | $\exp(-4.0\|\boldsymbol{v}_t^x - \boldsymbol{v}_t^{x*}\|_2)$ | 2 |
| Root linear velocity y | $\exp(-4.0\|\boldsymbol{v}_t^x - \boldsymbol{v}_t^{x*}\|_2)$ | 1.5 |
| Root angular velocity | $\exp(-4.0\|\boldsymbol{v}_t^{ang} - \boldsymbol{v}_t^{ang*}\|_2)$ | 4 |
| Root walk height | $\exp\left(-\frac{\|\text{command}_z - p_z^{\text{root}}\|}{0.05}\right)$ | 2 |
| Waist dofs | $\exp\left(-\frac{\sum_{\theta\in\{\text{yaw, roll, pitch}\}}(\theta^{\text{sim}} - \theta^{\text{cmd}})^2}{0.05}\right)$ | 2 |
| Upper body dofs | $\exp\left(-\frac{\|\boldsymbol{q}_{\text{upper}} - \boldsymbol{q}_{\text{ref}}\|_2^2}{0.01}\right)$ | 4 |

## A.2. Domain Randomization

We apply the following domain randomization terms during training, which are important for successful sim-to-real transfer.

Table 6: Domain randomization terms including dynamics randomization and external perturbation.

| Term | Value |
|---|---|
| **Dynamics Randomization** | |
| Friction | $\mathcal{U}(0.5, 1.25)$ |
| Link mass | $\mathcal{U}(0.9, 1.2) \times$ default kg |
| Base mass | $\mathcal{U}(-1.0, 3.0)$ kg |
| P Gain | $\mathcal{U}(0.9, 1.1) \times$ default |
| D Gain | $\mathcal{U}(0.9, 1.1) \times$ default |
| Control delay | $\mathcal{U}(0, 20)$ms |
| **External Perturbation** | |
| Push robot | interval $= 5s$, $v_{xy} = 1$m/s |

### A.3. Lower-RL-Upper-IK with Force Estimator

We jointly train a 3D force estimator, following a similar approach to Portela et al. (2024), using the robot's proprioception as input $s_t^{\mathrm{p}} \triangleq [q_{t-4:t}, \dot{q}_{t-4:t}, \omega^{\mathrm{root}}_{t-4:t}, g_{t-4:t}, a^{l}_{t-5:t-1}]$. As illustrated in 6 (a), the estimator predicts the end-effector forces $\tilde{F}_t^{ee}$, which are then concatenated with full-body proprioception and fed into the lower-body RL policy. Meanwhile, the upper-body joint torques with force compensation are computed as $\tau = K_p(q_t^{\mathrm{upper}} - q_t^{\mathrm{upper*}}) + K_d \dot{q}_t^{\mathrm{upper}} + J_{EE}^T \tilde{F}_t^{ee}$.

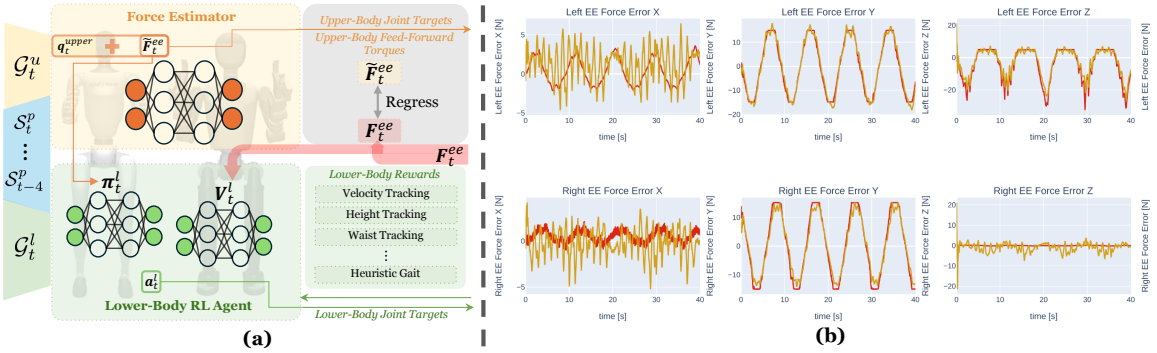
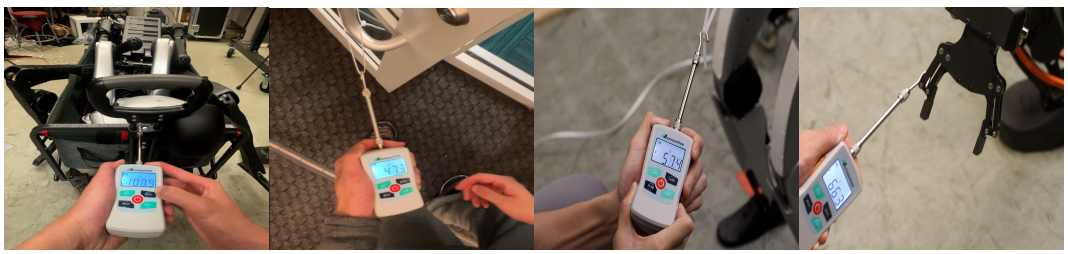



Figure 6: (a) Lower-RL-Upper-IK with Force Estimator; (b) Force Estimator Results: yellow lines are the estimated forces while the red lines are the actual forces.

We compare the estimated and applied forces in Fig. 6 (b), showing close alignment. However, variations in the contact point change the effective Jacobian $J'^{\top}_{EE}$, making the compensation term $J_{EE}^{\top} \tilde{F}_t^{ee}$ unreliable. Hence, a force sensor remains essential to localize contact and compute the correct $J'^{\top}_{EE}$. The quasi-static assumption further introduces errors during upper-body motion.

### A.4. Force Measurement

We use an Mxmoonfree Digital 500 N Force Gauge to measure the peak forces in three force-adaptive tasks: (1) **Cart-Pulling** (Booster T1 with Unitree G1 and H1 in the cart), (2) **Door-Opening**, and (3) **Stance-Pulling**. In **Stance-Pulling**, the robot applies longitudinal forces in the X–Y plane while maintaining a static stance, and we record the maximum resistive force before losing balance. The Booster T1 achieves a higher peak resistive force than the Unitree G1, mainly due to its lower center of mass, which enhances longitudinal stability.



(a) Cart-Pulling (peak: 107.9 N) (b) Door-Opening (peak: 47.3 N) (c) Stance-Pulling: Unitree G1 (peak: 57.4 N) (d) Stance-Pulling: Booster T1 (peak: 66.3 N)

Figure 7: Maximum force readings captured during different force-adaptive tasks using a handheld force gauge. Subfigures (a)–(d) show peak force values during individual tasks.

## A.5. Autonomy Pipeline

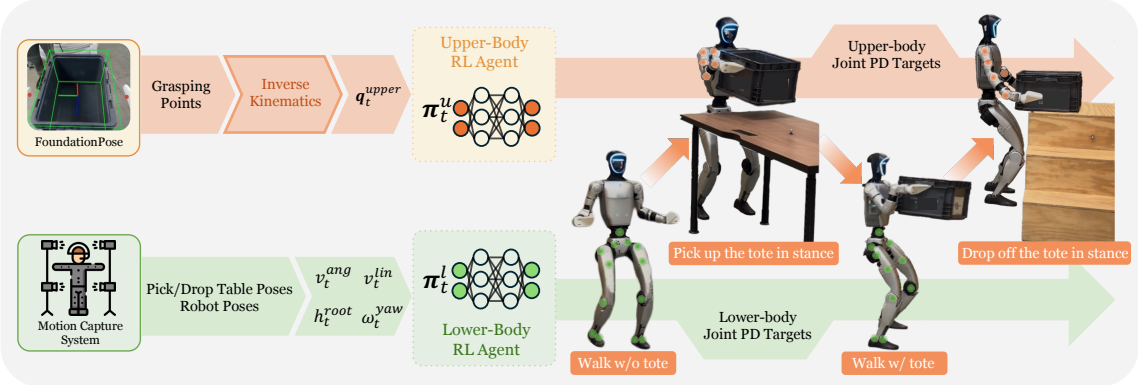



Figure 8: Overview of the autonomy pipeline integrating FALCON with 6-DoF object pose estimation (FoundationPose), MoCap-based localization, and inverse-kinematics grasp planning, enabling a humanoid to perform tote logistics: walking, picking up, carrying, and dropping off a tote.

We develop a hierarchically autonomous pipeline for tote logistics, leveraging a Motion Capture (MoCap) system to localize positions of the robot and desks. The robot is controlled by a state-machine framework with four states: (1) walking without the tote, (2) picking up the tote in stance, (3) walking with the tote, and (4) dropping off the tote in stance, as illustrated in Fig. 8. To estimate the tote's pose relative to the camera, we use FoundationPose Wen et al. (2024), a state-of-the-art method for accurate and reliable 6-DoF pose estimation.

## A.6. Hardware Limits in Real-World Deployment

During sim-to-real deployment, the FALCON-trained humanoid experiences rapid wrist motor overheating under sustained high torques, limiting payloads to $< 2$ kg per arm (Fig. 9a). In contrast, in MuJoCo Todorov et al. (2012), the same policy—without thermal modeling—transports $> 3$ kg per end effector while tracking $-1$ m/s velocity (Fig. 9 (b)–(c)), revealing an actuator endurance gap. For short, high-torque tasks like cart-pulling, however, overheating is negligible, enabling successful real-world execution.

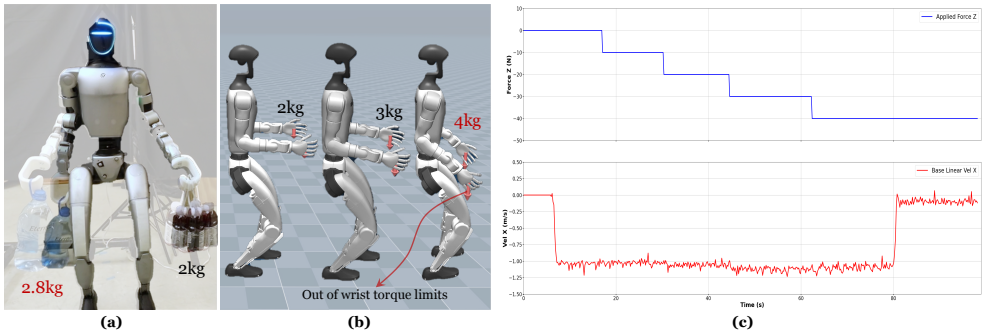



Figure 9: Transporting 0-4kg Payloads in Mujoco