2026 IEEE/SICE International Symposium on System Integration (SII)
January 11-14, 2026. Cancún, México

Multi-Object Loco-Manipulation Using Body Holding Primitives for Humanoids

Zhongkai Gu, Wei Zhu* and Mitsuhiro Hayashibe*

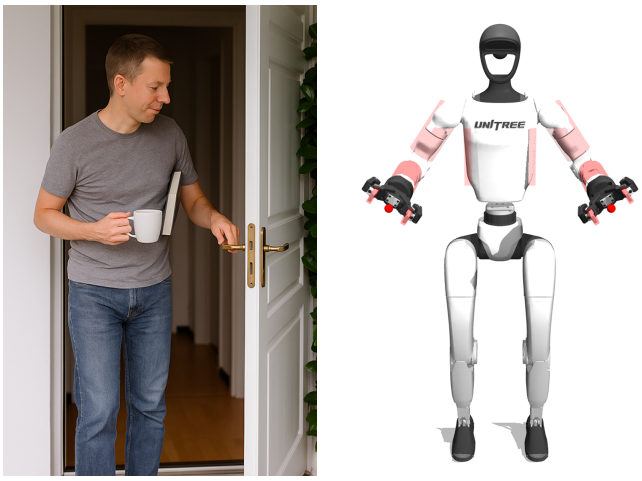
***Abstract*— Although considerable research in humanoid robotics focuses on developing general manipulation capabilities, whole-body manipulation—leveraging all body links for manipulation tasks—has received limited attention. To explore the potential of whole-body manipulation, we propose a framework enabling humanoid robots to carry multiple objects simultaneously using their entire body. We introduce two types of body holding primitives: under-arm holding and wrist-torso holding, significantly expanding humanoid robot manipulation capabilities. To achieve the task of multi-objects loco-manipulation, we leverage compliant control to effectively stabilizes object holding using body links. Furthermore, we integrate existing learning-based whole-body locomotion controllers with a dynamic feedforward PID controller to generate accurate upper-body movements for manipulation while maintaining stable locomotion.**

I. INTRODUCTION

Humans effortlessly exploit every available degree of freedom (DoF) to accomplish everyday tasks. For example, one might carry a cup of coffee in one hand, hold a book under the arm, and open a door with the other hand, as shown in Fig. 1-(a). This demonstrates the essence of *multi-object loco-manipulation*—the seamless integration of locomotion and manipulation through distributed contacts across the hands, arms, torso, and legs [1]. By leveraging all available body surfaces, humans are capable of holding more than two objects at once while navigating complex environments. This strategy improves task efficiency and avoids unnecessary intermediate hand-offs (e.g., putting objects down before manipulating the environment).

Humanoid robots, designed to mirror human morphology, appear ideally suited to replicate these human-level capabilities of multi-object loco-manipulation. However, achieving this level of integrated, contact-rich behavior is remarkably challenging and remains an open problem, despite significant, independent advances in both bipedal locomotion and arm manipulation.

Recent progress in deep reinforcement learning (RL) and imitation learning (IL) has produced controllers capable of stable and dynamic lower-body motion [2], [3]. A commonly adopted and effective strategy is to train a lower-body policy for balance maintenance and overlay arbitrary upper-body motions, enabling arm manipulation without compromising stability [3]–[5]. In parallel, manipulation policies have leveraged expert demonstrations [6], [7] or RL-only approaches [8] to achieve basic pick-and-carry behaviors. Whole body controller which combines IL and RL through Central Pattern Generators, are recently proposed [9].



(a) (b)

Fig. 1: Example of multi-object loco-manipulation and the simulation setup. (a) A person opening a door while holding a book and a cup. *(This image was generated using DALL·E, a generative AI model.)* (b) Simulated humanoid robot setup. Red blocks indicate touch sensor placement. A red dot on each gripper marks the inverse kinematics (IK) tracking point for grasping tasks.

A key limitation of these approaches is that they typically treat the robot in a decoupled manner—locomotion is restricted to the legs, while manipulation is confined to the end-effectors, with limited contribution from the rest of the body to force generation or contact modulation. This highlights three critical research gaps that prevent true multi-object loco-manipulation [10]:

Perception. Humans possess a continuous layer of compliant skin densely populated with tactile and thermal receptors. In contrast, most humanoids still rely on rigid surfaces equipped with only sparse six-axis force/torque sensors. The absence of large-area, conformal tactile skins prevents reliable estimation of distributed contact wrenches—information that is indispensable for safe, high-bandwidth whole-body interaction.

Control. Loco-manipulation requires controllers that can simultaneously coordinate motion and contact forces across numerous body links, yet current advanced locomotion policies are not integrated with the compliant control needed to

This work was supported by JSPS KAKENHI Grant JP24K00841.

Zhongkai Gu, Wei Zhu and Mitsuhiro Hayashibe are with Department of Robotics, Graduate School of Engineering, Tohoku University, Japan.
*Corresponding authors: zhu.wei.c1@tohoku.ac.jp, mitsuhiro.hayashibe.e6@tohoku.ac.jp

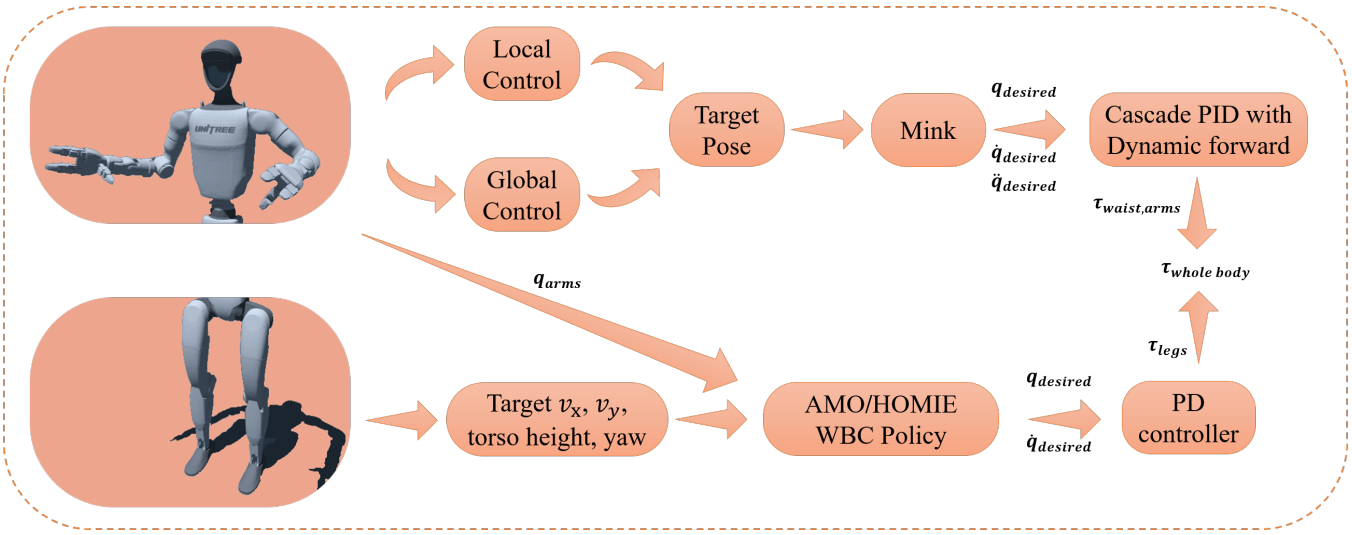


Fig. 2: Overview of the whole-body controller.

handle multiple objects simultaneously [3], [4].

Planning. Generating long-horizon, multi-contact sequences for real-world scenarios—e.g., carrying a box upstairs while opening a door—poses a significant planning challenge. Existing imitation-learning studies can replicate discrete tasks like box pickup [11] or under-arm grasping [12], but scaling to complex sequences is difficult, primarily due to the scarcity of high-quality demonstrations. To address this data scarcity, planner-guided data generation has emerged, where a model-based planner generates coarse, contact-feasible reference motions that learning algorithms subsequently refine [13], [14]. While promising, these frameworks build on global contact-planning [15] and diffusion-based extensions [14] that primarily focus on whole-body manipulation without incorporating locomotion. Similarly, synthesis frameworks such as HumanoidGen [16] can generate bimanual motions but do not support whole-body loco-manipulation. Although teleoperation methods [5], [7] can provide loco-manipulation data, they require extensive human effort and often lack the force feedback necessary for delicate whole-body tasks [17].

In this paper, we address these limitations through a unified simulation framework for whole-body loco-manipulation that enables humanoid robots to carry multiple objects. To address the perception problem, we augment the humanoid robot with force sensors. For control, we introduce body holding primitive actions—such as under-arm holding and wrist-torso holding—that maximize the dexterity of the robot. We demonstrate how these primitives can be composed to accomplish complex multi-object loco-manipulation tasks. For planning, we define keypoints in each task, forming sequences of continuous and long-horizon whole-body motion planning.

The main contributions of our work are (1) a whole-body controller that enables humanoid robots to carry more than two objects simultaneously; (2) compliant control strategies for stable object holding during locomotion; (3) demonstrations of long-horizon multi-object loco-manipulation tasks.

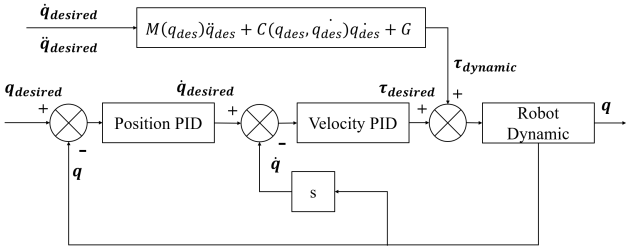


Fig. 3: Cascade position and velocity PID control with dynamic feedforward.

II. Method

In this section, we present our method for developing a comprehensive framework for multi-object loco-manipulation. We detail the simulation environment, whole-body control architecture, and the design of primitive actions.

A. Simulation Environment

The simulation environment is built using the MuJoCo physics engine [18], selected for its precision in modeling contact forces and frictional interactions—crucial for realistic manipulation tasks. We simulate the Unitree G1 humanoid robot (29 DoFs) equipped with Robotiq 2F-85 parallel grippers (Fig. 1-(b)), chosen due to their versatility in various grasping scenarios. We also provide support for other humanoid models and end-effectors, including the 23-DoF version of Unitree G1 and the Allegro hand.

To enable accurate whole-body manipulation, touch sensors are strategically placed on the lateral torso, forearms, and grippers to provide rich force feedback.

B. Whole-body Control Architecture

The whole-body controller is modularly divided into upper-body (17 DoFs, including waist and arms) and lower-body (12 DoFs, legs) modules. This modularity enhances

manipulation precision while maintaining stability during locomotion.

For the lower-body module, we incorporate two state-of-the-art humanoid locomotion controllers using RL and IL: AMO [7] and HOMIE [5]. AMO enables a broad range of torso orientations by jointly coordinating waist and leg joints, resulting in improved maneuverability. Prior to the open-sourcing of AMO's training pipeline, HOMIE served as a practical alternative due to its adaptability across different robot platforms. Both controllers receive current upper-body joint states as part of their observation space and output lower-body joint positions based on linear velocity commands, as shown in Fig. 2. Notably, AMO observes only the first four joints in each arm (shoulder yaw, roll, pitch, and elbow), whereas HOMIE uses all arm joint positions.

The upper-body module is governed by a cascaded position–velocity PID control loop with an dynamic feedforward term, as illustrated in Fig. 3. The target joint torque is calculated using the following equation:

$$\tau = \tau_{desired} + \tau_{dynamic}, \tag{1}$$

where τ_{desired} is the desired joint torque computed by the PID controller, and τ_{dynamic} is the dynamic feedforward term, defined as

$$\tau_{\text{dynamic}} = M(q_{\text{des}})\ddot{q}_{\text{des}} + C(q_{\text{des}}, \dot{q}_{\text{des}})\dot{q}_{\text{des}} + G,$$

where, $M(q)$ denotes the joint-space inertia matrix, $C(q,\dot{q})\dot{q}$ represents the Coriolis and centrifugal effects, and $G(q)$ is the gravity torque vector. The total torque τ is then sent to the motor actuators. This setup enables precise and quick joint trajectory tracking—essential for accurate manipulation.

Upper-body inverse kinematics (IK) is solved using Mink [19], a MuJoCo-based differential IK toolkit. We extend the controller to support both local (robot-frame) and global (world-frame) tracking modes. This distinction is critical: whole-body manipulation typically relies on the relative configuration of body parts (best described in the robot frame), while object grasping demands accurate world-frame tracking. The upper-body loop executes at 1 kHz, enabling responsive hybrid position/force control. The overall control architecture is depicted in Fig. 2.

C. Primitive Actions

To modularize and simplify complex loco-manipulation tasks, we define four reusable primitive actions: gripper-grasp, gripper-place, under-arm hold, and wrist-torso hold (see Fig. 4). Each action encapsulates a fundamental manipulation behavior that can be composed to perform diverse tasks.

- **Gripper-Grasp:** Involves two keypoints: (1) move to a pre-grasp pose, and (2) close the gripper at the grasp pose. Object poses are assumed known. To avoid collisions between object and gripper before grasp aciton, an optional *pre-pre-grasp* waypoint may be inserted.
- **Gripper-Place:** Comprises three keypoints: (1) use the gripper-grasp action, (2) move to the placement position, and (3) gently release the object onto the table.

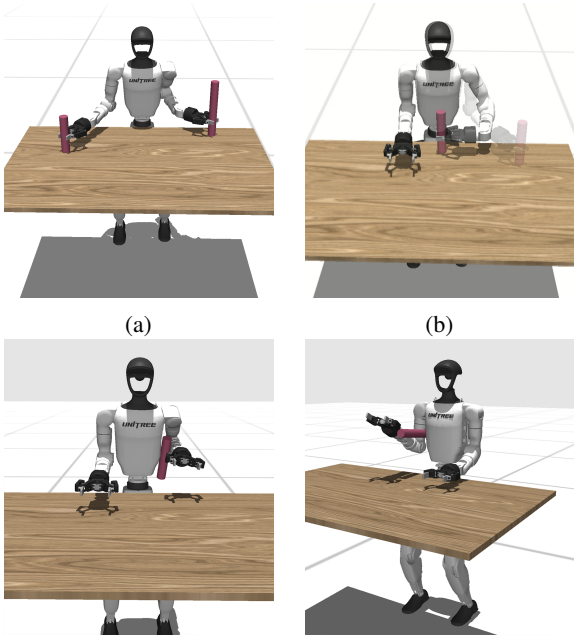


Fig. 4: Primitive Actions. (a) Gripper-grasp. (b) Gripper-place. (c) Wrist-torso hold. (d) Under-arm hold.

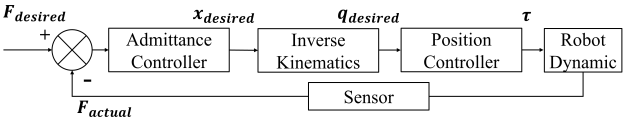


Fig. 5: Admittance control used in wrist-torso hold action.

- **Under-Arm Hold:** Example with right arm: (1) left gripper grasps the object, (2) left arm moves to a pre-insert pose, (3) right arm moves to a pre-hold pose, (4) left arm inserts the object, (5) right arm applies force control to hold the object against the torso, (6) left arm releases and retracts.
- **Wrist-Torso Hold:** Example with left arm: (1) use gripper-grasp and gripper-place to position the object near the torso, (2) incline the upper body to a pre-hold pose, and (3) press the wrist against the object using force control. Force regulation is implemented via admittance control, defined by:

$$M\ddot{x} + B\dot{x} + Kx = F_{\text{ext}}, \tag{2}$$

where M, B, and K represent the virtual mass, damping, and stiffness matrices, respectively. The wrist senses the external force F_{ext} via a touch sensor and adjusts its motion x to maintain a desired contact force. The complete control process is illustrated in Fig. 5.

III. Experiments

In this section, we first evaluate the performance of the proposed whole-body controller. We then analyze the efficiency of the primitive actions, followed by demonstrations of multi-object loco-manipulation tasks.

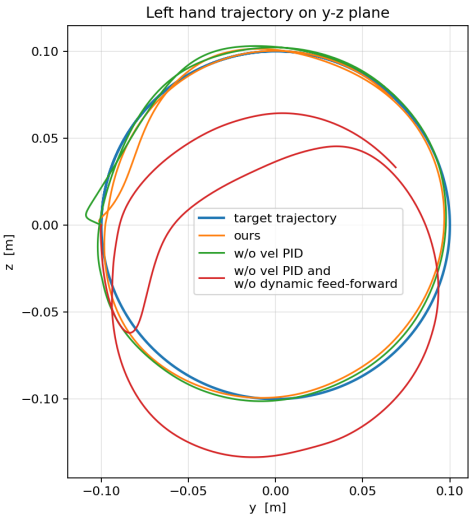


Fig. 6: Comparison of trajectory-tracking errors with baselines. Each trajectory was simulated for 7000 timesteps in MuJoCo.

TABLE I: Comparison of tracking errors and speed.

Metrics	E_x (mm)	$\sigma_x(mm)$	T_{min}	$\sigma_{T_{min}}$
Ours	**1.9**	**0.8**	**1864**	**828.3**
w/o vel PID	1.9	0.8	1968	864.2
w/o vel PID and w/o DFF	41.8	12.5	-	-

A. Performance of the Upper-Body PID Controller

We conducted two experiments to evaluate the dynamic trajectory-tracking performance and static target-reaching accuracy of the upper-body PID controller, operating under the AMO locomotion policy. Performance was compared against the following baseline controllers:

- **w/o vel PID:** A position PID controller with dynamic feedforward, as described in the previous section.
- **w/o vel PID and w/o dynamic feedforward (DFF):** A pure PD controller without feedforward compensation.

In the first experiment, the robot's left end-effector was commanded to follow a circular trajectory of radius 0.1 m in the y–z plane (Fig. 6). The mean tracking errors for our controller, the w/o vel PID baseline, and the pure PD controller were 3.18%, 3.20%, and 8.29%, respectively. These results confirm that the dynamic feedforward component is the primary contributor to improved trajectory tracking.

In the second experiment, the robot's left end-effector was instructed to reach 10 different static target poses. The following metrics were used to evaluate performance:

- **Tracking Accuracy (E_x):** Defined as the average Euclidean distance between the target and actual end-effector positions. Results demonstrate that adding dynamic feedforward significantly improves precision.
- **Accuracy Variability (σ_x):** Defined as the standard deviation of the tracking accuracy across all trials, indicating the consistency of the controller's performance.

TABLE II: Velocity tracking errors (Δv_x, Δv_y) for AMO and HOMIE under varying loads.

Controller	Load (kg)	Δv_x (m/s)	Δv_y (m/s)
AMO			
AMO	0	0.3788	0.0670
AMO	0.5	0.3489	0.0656
AMO	1.0	0.3227	0.0660
AMO	1.5	0.2997	0.0669
AMO	2.0	0.2808	0.0697
HOMIE			
HOMIE	0	**0.0243**	**0.0131**
HOMIE	0.5	**0.0108**	**0.0100**
HOMIE	1.0	**0.0112**	**0.0093**
HOMIE	1.5	**0.0497**	**0.0151**
HOMIE	2.0	**0.1131**	**0.0254**

- **Tracking Speed (T_{min}):** Defined as the timestep when the tracking error first falls below 0.01 m. Results show that incorporating a velocity PID block accelerates convergence to the target pose.
- **Speed Variability ($\sigma_{T_{min}}$):** Defined as the standard deviation of T_{min} across all trials, reflecting the consistency of the convergence speed.

TABLE I summarizes the reaching performance. The results highlight the benefit of using dynamic feedforward in improving both accuracy and responsiveness of the upper-body controller. In addition, the velocity PID block contributes to enhancing responsiveness, leading to faster convergence toward the target poses.

B. Performance Analysis of AMO and HOMIE Policies

We evaluated HOMIE and AMO controllers in MuJoCo. This allows us to examine their velocity-tracking performance under more realistic physical conditions. The Unitree G1 robot was initialized with all arm joints set to zero and leg joints configured as $[-0.1,\ 0.0,\ 0.0,\ 0.3,\ -0.2,\ 0.0]$, with a target base height of 0.75 m. The velocity command was set to $v_x = 1$ m/s and $v_y = 0$ m/s.

As shown in TABLE II, HOMIE consistently outperforms AMO in velocity tracking across all load conditions. HOMIE maintains low tracking error (even slightly overcompensating at higher loads), while AMO consistently underperforms, especially along the x-axis. This performance discrepancy may be attributed to the fact that AMO was only trained with v_x commands up to 0.5 m/s, whereas HOMIE was trained on a larger velocity range.

Both controllers exhibit the trend that increased end-effector load leads to higher forward velocity. This is likely due to the shift in the robot's center of mass caused by the added mass, encouraging faster forward movement.

In summary, although AMO offers more flexibility in torso movement—making it well-suited for manipulation-intensive tasks—HOMIE achieves higher velocity tracking accuracy and is therefore preferable for tasks that demand precise locomotion. This suggests a task-aware controller selection

TABLE III: Success rates of gripper-grasp and placing actions for different control strategies across object weights.

Strategy	250 g	500 g	750 g	1000 g
Grasp				
Force Control	19/20	19/20	19/20	19/20
Position Control	19/20	19/20	19/20	19/20
Place				
Force Control	**19/19**	**18/19**	**16/19**	**15/19**
Position Control	15/19	14/19	13/19	13/19

TABLE IV: Maximum stable walking distance (in meters) for different control strategies across object weights.

Strategy	250 g	500 g	750 g	1000 g
Under-Arm Hold				
Force Control	**50+**	**50+**	**30.67**	**17.50**
Position Control	19.79	18.14	8.71	0
Wrist-Torso Hold				
Force Control	**50+**	**50+**	**50+**	**50+**
Position Control	0	0	0	0

strategy: use AMO when manipulation dexterity is critical, and HOMIE when velocity performance is the priority.

C. Performance Analysis of Primitive Actions

This subsection evaluates the stability and effectiveness of four primitive actions through two experiments.

The first experiment evaluated the success rate of the gripper-grasp and gripper-place action under varying object weights for two control strategies, as summarized in TABLE III. In each trial, a bottle is placed within the reachable workspace of the robot's left arm. The robot must grasp the bottle and place it securely in front of its torso. Both methods performed equally well in grasping, achieving a consistent success rate of 19/20. Grasping failures were primarily due to premature contact between the object and the gripper, which prevented proper alignment. For placing, the success rate declined as the object weight increased. The force control strategy consistently outperformed position control, particularly for heavier objects, demonstrating greater robustness and adaptability under dynamic conditions.

The second experiment investigates the maximum stable holding distance for the under-arm hold and wrist-torso hold actions with varying object weights. In each trial, the robot was required to hold an object and walk forward for as long as possible. We compared two control strategies: position control and our proposed force control method. As shown in TABLE IV, force control significantly outperformed position control across all conditions, especially in the wrist-torso hold action, where position control failed immediately. The result shows the potential of stable holding 2 objects together with force control. These results highlight the potential of force control to enable reliable multi-object holding during whole-body locomotion.

Fig. 7 illustrates wrist force measurements during the wrist-torso hold action. In the force control case, the average contact force (orange line) stabilized around the commanded 5 N shortly after contact at approximately 7 s, indicating a stable interaction. In contrast, the position control case produced a higher initial force (12 N) that rapidly decayed once walking began, suggesting unstable contact and object slippage.

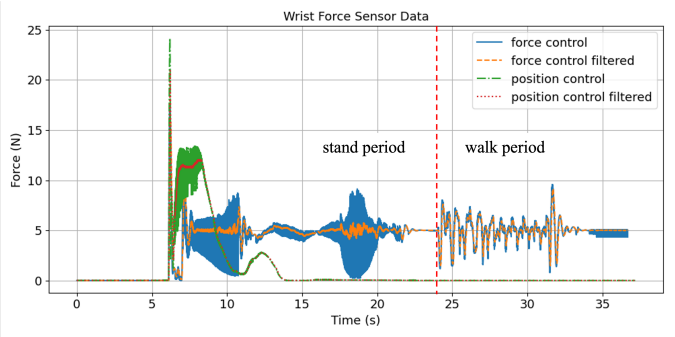
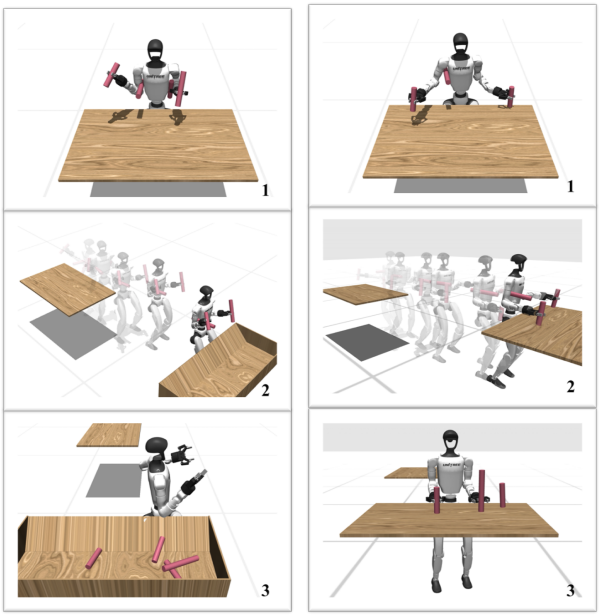


Fig. 7: Force sensor recordings during the wrist-torso hold action using force and position control.



(a) (b)

Fig. 8: Multi-object loco-manipulation tasks using G1. (a) G1 holds four objects, walks to a container, and drops all of them. (b) G1 holds three objects, walks to a table, and places each object sequentially.

These results highlight the importance of compliant control strategies in dynamic whole-body manipulation. While position control is prone to loss of contact and object slippage during motion, force control offers the adaptability and stability required for reliable object retention under varying loads and movements.

D. Multi-Object Loco-Manipulation Tasks

To validate the effectiveness of our system and its capacity for complex whole-body coordination, we conduct experiments on two long-horizon multi-object loco-manipulation tasks that demand hyper-dexterous control.

The first task, **carrying and dropping four objects**, is shown in Fig. 8-(a). It consists of three stages: (1) holding four objects, (2) walking to the drop point, and (3) dropping the objects into a container. The robot uses a combination of body-holding and gripper-grasping primitives—holding two objects on either side of the torso and grasping two additional objects with the grippers. After stabilizing all objects, the robot walks to the designated location and initiates the drop sequence. First, the grippers release their objects. Then, the torso rotates in yaw and roll to drop the two side-held objects. We use the AMO controller for the lower body in this task, as it supports torso roll manipulation. Without torso roll capability, aligning the side-held objects with the container becomes infeasible.

The second task, **carrying and placing three objects**, is illustrated in Fig. 8-(b). This task also comprises three stages: (1) holding three objects, (2) walking to the target table, and (3) sequentially placing the objects. The robot holds one object using the under-arm hold and two objects with the grippers. After reaching the table, the robot places the left gripper object first. Next, it retrieves and places the under-arm hold object. Finally, the right gripper places the third object. HOMIE is used as the lower-body controller in this task, as torso roll or pitch motion is not required.

These experiments demonstrate the ability of our system to coordinate whole-body manipulation effectively across diverse and dynamic multi-object scenarios. Moreover, the proposed body-holding primitives integrate robustly with both AMO and HOMIE lower-body controllers, highlighting their generality and adaptability.

IV. CONCLUSION

In this paper, we presented a whole-body control framework that enables a humanoid robot to carry up to four objects while walking, demonstrating the system's ability to execute complex loco-manipulation tasks using body holding primitives. Our framework composes a compact library of reusable primitive actions—gripper grasp/place, under-arm hold, and wrist-torso hold—to perform long-horizon, multi-object behaviors.

Precise upper-body motion is achieved using a dynamic feedforward PID controller, while compliant force control ensures stability at each body-object contact interface. Together, these components enable robust and adaptable whole-body manipulation in dynamic, contact-rich scenarios.

However, the current framework relies on hand-tuned keypoints to execute complex tasks, which makes generalization challenging. In the future, IL could be used to learn skills from demonstrations generated by our framework, and RL could then be used to generalize these skills. To reduce hardware dependencies, force control methods that do not require contact sensors, such as [20], could be explored for achieving stable holding. Furthermore, interactive mesh retargeting methods presented in [21] could be implemented to transfer the current whole-body holding skills to diverse robot morphologies.

REFERENCES

[1] V. Levé, J. Moura, S. Fujita *et al.*, "Scaling whole-body multi-contact manipulation with contact optimization," in *2025 IEEE-RAS 24th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2025, pp. 920–927.

[2] X. He, R. Dong, Z. Chen *et al.*, "Learning getting-up policies for real-world humanoid robots," *arXiv preprint arXiv:2502.12152*, 2025.

[3] T. He, J. Gao, W. Xiao *et al.*, "Asap: Aligning simulation and real-world physics for learning agile humanoid whole-body skills," *arXiv preprint arXiv:2502.01143*, 2025.

[4] M. Ji, X. Peng, F. Liu *et al.*, "Exbody2: Advanced expressive humanoid whole-body control," *arXiv preprint arXiv:2412.13196*, 2024.

[5] Q. Ben, F. Jia, J. Zeng *et al.*, "Homie: Humanoid loco-manipulation with isomorphic exoskeleton cockpit," *arXiv preprint arXiv:2502.13013*, 2025.

[6] M. Seo, S. Han, K. Sim *et al.*, "Deep imitation learning for humanoid loco-manipulation through human teleoperation," in *2023 IEEE-RAS 22nd International Conference on Humanoid Robots (Humanoids)*. IEEE, 2023, pp. 1–8.

[7] J. Li, X. Cheng, T. Huang *et al.*, "Amo: Adaptive motion optimization for hyper-dexterous humanoid whole-body control," *arXiv preprint arXiv:2505.03738*, 2025.

[8] J. Dao, H. Duan, and A. Fern, "Sim-to-real learning for humanoid box loco-manipulation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 16 930–16 936.

[9] G. Li, A. Ijspeert, and M. Hayashibe, "Ai-cpg: Adaptive imitated central pattern generators for bipedal locomotion learned through reinforced reflex neural networks," *IEEE Robotics and Automation Letters*, vol. 9, no. 6, pp. 5190–5197, 2024.

[10] Z. Gu, J. Li, W. Shen *et al.*, "Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning," *arXiv preprint arXiv:2501.02116*, 2025.

[11] F. Liu, Z. Gu, Y. Cai *et al.*, "Opt2skill: Imitating dynamically-feasible whole-body trajectories for versatile humanoid loco-manipulation," *arXiv preprint arXiv:2409.20514*, 2024.

[12] K. Hsiao and T. Lozano-Perez, "Imitation learning of whole-body grasps," in *2006 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2006, pp. 5657–5662.

[13] M. Zhang, J. Barreiros, and A. O. Onol, "Plan-guided reinforcement learning for whole-body manipulation," *arXiv preprint arXiv:2310.12263*, 2023.

[14] X. Li, T. Zhao, X. Zhu *et al.*, "Planning-guided diffusion policy learning for generalizable contact-rich bimanual manipulation," *arXiv preprint arXiv:2412.02676*, 2024.

[15] T. Pang, H. T. Suh, L. Yang *et al.*, "Global planning for contact-rich manipulation via local smoothing of quasi-dynamic contact models," *IEEE Transactions on robotics*, vol. 39, no. 6, pp. 4691–4711, 2023.

[16] Z. Jing, S. Yang, J. Ao *et al.*, "Humanoidgen: Data generation for bimanual dexterous manipulation via llm reasoning," *arXiv preprint arXiv:2507.00833*, 2025.

[17] J. J. Liu, Y. Li, K. Shaw *et al.*, "Factr: Force-attending curriculum training for contact-rich policy learning," *arXiv preprint arXiv:2502.17432*, 2025.

[18] E. Todorov, T. Erez, and Y. Tassa, "Mujoco: A physics engine for model-based control," in *2012 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2012, pp. 5026–5033.

[19] K. Zakka, "Mink: Python inverse kinematics based on MuJoCo," May 2025. [Online]. Available: https://github.com/kevinzakka/mink

[20] P. Zhi, P. Li, J. Yin *et al.*, "Learning a unified policy for position and force control in legged loco-manipulation," in *Conference on Robot Learning*. PMLR, 2025, pp. 652–669.

[21] L. Yang, X. Huang, Z. Wu *et al.*, "Omniretarget: Interaction-preserving data generation for humanoid whole-body loco-manipulation and scene interaction," *arXiv preprint arXiv:2509.26633*, 2025.