Learning Whole-Body Loco-Manipulation Skills

Mayank Mittal
ETH Zürich and NVIDIA
Email: mittalma@ethz.ch

I. Motivation

Legged mobile manipulators are an intuitive robotic embodiment for operating in human environments. They allow using different parts of their body to assist in loco-manipulation and enhance their interaction capabilities. For instance, a robot can adjust its torso posture using its legs to help its arms apply more force to an object or use both arms and legs to open and traverse a spring-loaded pull door. Loco-manipulation, at its core, is a multi-contact planning and control problem, where the robot must carefully coordinate contacts with its surroundings to simultaneously move, maintain balance, and manipulate objects. However, prior works [5, 47, 19, 46] simplify this complex problem by separating locomotion from manipulation, limiting the latter to the robot's arm end-effector. Furthermore, existing approaches [30, 28, 49] focus on scenarios where switching between contact locations on the object is either non-existent or hand-crafted by a skilled engineer. Consequently, a more advanced framework for generating multi-modal behaviors is necessary to address a broader spectrum of real-world tasks.

Obtaining real-world robot data is costly, which limits the ability to generalize beyond a narrow set of tasks and operating conditions [37]. For legged systems, this challenge is even more pronounced due to their high Degrees-of-Freedom (DoFs), which complicates teleoperation, and their non-fixed base, which increases the risk of damaging hardware. As a result, physics-based simulation [27] offers a more scalable alternative to generate diverse training data, while also facilitating the testing of control policies for legged systems.

My vision is to achieve robust whole-body loco-manipulation skills for legged mobile manipulators primarily through simulation, and transferring them effectively to hardware.

To advance this vision, my research focuses on exploring: (a) how to scale synthetic data generation for robot learning and sim-to-real transfer of learned skills, (b) how to design a flexible and robust whole-body controller that integrates both mobility and manipulability, and (c) how to extend beyond single end-effectors to more complex, coordinated multi-limbed manipulation.

II. Research To Date And Related Work

A. Synthetic Data Generation for Robot Learning

A high simulation throughput often comes at the cost of simulation accuracy [13]. Popular simulators [44, 10] use CPU-based physics engines, which require massive clusters to meet the data demands of current learning algorithms. To overcome this limitation, we developed a **simulation framework for robot learning** [32] that leverages recent advances in GPU-accelerated physics simulation and hyper-realistic rendering. This framework allows users to train and deploy policies for complex tasks, such as rough-terrain locomotion, in just minutes. We integrated various robot learning workflows and environments, enabling systematic training, testing, and verification of robotic systems. This tool is now widely used in both industry [12, 36] and academia [25, 48] and serves as the foundation for my research on loco-manipulation skills.

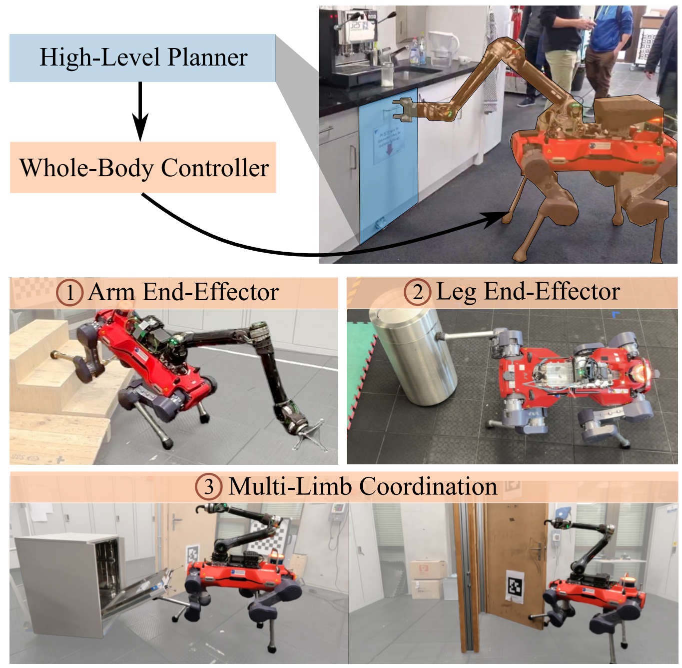

Fig. 1. Different paradigms for whole-body planning and control of legged mobile manipulators. The high-level planner may provide references to the robot in the task-space for the individual limbs (through teleoperation or as a learned policy) or in the joint-space (as offline motion plans).

B. Whole-Body Planning and Control 1 2

A *holistic* planning and control approach for mobile manipulation needs to account for the high DoFs and potentially conflicting concurrent tasks and constraints. Recent works [39, 24] design a real-time Model Predictive Controller (MPC) for legged mobile manipulators to track end-effector poses. We extended this **MPC formulation to incorporate constraints for avoiding unwanted self and environment collisions** and utilizing a real-time Signed Distance Field (SDF) representation of the environment [31, 9]. Combining this controller with a vision-based *object-centric* planner, we demonstrated that this approach enables interaction with unknown cabinets and drawers in static and dynamic kitchen scenes. It achieves higher performance compared to conventional methods that decompose the whole-body control or rely on inverse kinematics. These results underscore the importance of whole-body coordination for effective mobile manipulation.

While model-based optimal control techniques are often effective for manipulating simpler objects, such as cabinets and drawers, they are sensitive to modeling mismatches and need a pre-specified contact pattern. This reliance limits their robustness, particularly in handling large uncertainties in object dynamics and unmodeled terrains. To overcome these limitations, we investigated the application of Reinforcement Learning (RL) for robust tracking of pose commands, whether for a quadrupedal robot's foot [3, 42] or its arm end-effector [35]. The hardware results demonstrate that the **learned controllers can handle uneven and slippery terrains as well as external disturbances by actively adapting the robot's posture**. Compared to their model-based counterparts, these controllers also offer a larger reachable workspace and improved tracking. These advantages enable tasks previously considered unachievable– such as opening a door with a quadruped's foot or carrying an unknown heavy payload.

Concurrent research [16, 7, 26, 20], including ours [11], combine these RL-based tracking policies with learned high-level task-space planners for object manipulation. However, these methods still operate on a single end-effector for manipulation, whether on the leg or the arm.

C. *Multi-Limbed Object Manipulation* 3

A higher number of DoFs increases the exploration complexity for an RL agent, often resulting in sub-optimal behaviors [18]. For instance, when training a quadrupedal robot to push a large box to a target location, the agent may learn to use a single leg and rarely use other limbs, even when those are closer to the object. To address this asymmetry in the learned behavior, we investigated **symmetry-based augmentation** [1] for the on-policy algorithm, PPO [38]. Directly applying this augmentation causes numerical instabilities, as the augmented samples are off-policy. To resolve this issue, we proposed an alternative policy update rule to stabilize the learning process [33]. On quadrupedal box climbing and pushing tasks, we demonstrated that this approach accelerates convergence and leads to more optimal behaviors. It allows the robot to naturally use the closest leg for manipulation, **eliminating the need for hand-crafted rewards to enforce limb selection**.

An alternate direction to improve the quality of learned RL behaviors is **incorporating reference motion data**. This technique is widely adopted for animating physically simulated characters [4, 34, 43], and more recently humanoid robots [21, 17]. Although these works demonstrate multi-limb tracking, they follow the reference motions at a fixed rate and involve limited or no object interaction. These assumptions make these methods brittle, as they struggle to deal with slippages or failed grasps. We highlight this issue in our work on interacting with large articulated objects, such as spring-loaded doors and dishwashers, with a quadrupedal mobile manipulator [41]. To overcome this limitation, we proposed an adaptive scheme for updating the reference motions, synthesized offline using Trajectory Optimization (TO) [40, 6]. Compared to prior works [34, 14], our proposed state-dependent update scheme enables a higher success rate in simulation despite model mismatches and significant external disturbances. Furthermore, the learned policies transfer reliably on hardware, **enabling the robot to use its legs or arm for object interaction** as dictated by the reference motion, while also **exhibiting recovery behaviors**, such as handle regrasping, that were absent in the demonstrations.

III. Ongoing And Future Research

A. *Benchmark for Multi-Contact Loco-Manipulation Skills*

The TO methods for multi-contact reference synthesis [40, 6] offer a scalable approach to generating a large dataset for various loco-manipulation tasks entirely in simulation [32]. Unlike existing datasets [45, 29], which have limited or no object interaction, this new dataset will feature high multi-modality, both in sensory inputs (through simulated exteroceptive and proprioceptive sensors) and interaction strategies (by utilizing different limbs or contact locations on the object). This dataset, along with the corresponding simulated environments in [32], will open up several exciting research avenues for building and evaluating a generalizable loco-manipulation framework for unknown environments.

B. *Generative Models for Online Whole-Body Planning*

As an initial step, I propose investigating generative models [8, 22] as a state-based high-level planner that provides whole-body references to the RL tracking policy from [41]. This approach offers the benefit of online kinematic planning via the high-level planner, combined with a robust tracking policy that can handle dynamic variations and external disturbances. The next step will involve utilizing image embeddings from on-board cameras as inputs to the model, instead of relying on the object's state. This will also require investigating different memory architectures to handle large occlusions that occur when the robot is near the object.

C. *Language-Conditioned Whole-Body Planning*

While my current research primarily focuses on quadrupedal mobile manipulators, I believe the insights gained from these works are also applicable to bipedal legged systems. Given their similar morphology to humans, additional data sources– such as internet-scale human videos– could also be leveraged for these platforms [15]. However, simply scaling datasets may not be sufficient to build a truly generalizable system [37]. Achieving this goal requires fundamental algorithmic improvements that enable these systems to reason about their embodiment in novel situations. Given the recent advancements in large language models for reasoning [2], one potential paradigm for generalization could be conditioning the high-level planner on language prompts. Through this approach, new situations could be addressed by providing a language description at inference time, which would influence the style and intent of the generated motion [23].

References

[1] Farzad Abdolhosseini, Hung Yu Ling, Zhaoming Xie, Xue Bin Peng, and Michiel van de Panne. On learning symmetric locomotion. In *Motion, Interaction and Games*, MIG '19, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450369947. doi: 10.1145/3359566.3360070. URL https://doi.org/10.1145/3359566.3360070.

[2] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

[3] Philip Arm, Mayank Mittal, Hendrik Kolvenbach, and Marco Hutter. Pedipulate: Enabling manipulation skills using a quadruped robot's leg. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 5717–5723. IEEE, 2024.

[4] Kevin Bergamin, Simon Clavet, Daniel Holden, and James Richard Forbes. Drecon: data-driven responsive control of physics-based characters. *ACM Transaction of Graphics (ToG)*, pages 206:1–206:11, 2019.

[5] Kenneth Blomqvist, Michel Breyer, Andrei Cramariuc, Julian Förster, Margarita Grinvald, Florian Tschopp, Jen Jen Chung, Lionel Ott, Juan Nieto, and Roland Siegwart. Go Fetch: Mobile Manipulation in Unstructured Environments. 2020.

[6] Jan Bruedigam, Ali Adeeb Abbas, Maks Sorokin, Kuan Fang, Brandon Hung, Maya Guru, Stefan Georg Sosnowski, Jiuguang Wang, Sandra Hirche, and Simon Le Cleac'h. Jacta: A Versatile Planner for Learning Dexterous and Whole-body Manipulation. In *Conference on Robot Learning (CoRL)*, 2024.

[7] Xuxin Cheng, Ashish Kumar, and Deepak Pathak. Legs as manipulator: Pushing quadrupedal agility beyond locomotion. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 5106–5112, 2023.

[8] Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. *The International Journal of Robotics Research (IJRR)*, 2023.

[9] Jia-Ruei Chiu, Jean-Pierre Sleiman, Mayank Mittal, Farbod Farshidian, and Marco Hutter. A collision-free mpc for whole-body dynamic locomotion and manipulation. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 4686–4693. IEEE, 2022.

[10] Erwin Coumans, , and Yunfei Bai. Pybullet, a python module for physics simulation for games, robotics and machine learning. https://pybullet.org/wordpress/, 2023. (accessed on May 2, 2023).

[11] Ioannis Dadiotis, Mayank Mittal, Nikos Tsagarakis, and Marco Hutter. Dynamic object goal pushing with mobile manipulators through model-free constrained reinforcement learning. *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2025.

[12] Boston Dynamics. Get Started with Reinforcement Learning for Spot, 2024.

[13] Tom Erez, Yuval Tassa, and Emanuel Todorov. Simulation tools for model-based robotics: Comparison of bullet, havok, mujoco, ode and physx. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 4397–4404, 2015. doi: 10.1109/ICRA.2015.7139807.

[14] Yuni Fuchioka, Zhaoming Xie, and Michiel van de Panne. Opt-mimic: Imitation of optimized trajectories for dynamic quadruped behaviors. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 5092–5098, 2023.

[15] 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.

[16] Huy Ha, Yihuai Gao, Zipeng Fu, Jie Tan, and Shuran Song. Umi on legs: Making manipulation policies mobile with manipulation-centric whole-body controllers. In *Conference on Robot Learning (CoRL)*, 2024.

[17] Tairan He, Wenli Xiao, Toru Lin, Zhengyi Luo, Zhenjia Xu, Zhenyu Jiang, Jan Kautz, Changliu Liu, Guanya Shi, Xiaolong Wang, et al. Hover: Versatile neural whole-body controller for humanoid robots. *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2025.

[18] Nicolas Heess, Dhruva TB, Srinivasan Sriram, Jay Lemmon, Josh Merel, Greg Wayne, Yuval Tassa, Tom Erez, Ziyu Wang, S. M. Ali Eslami, Martin A. Riedmiller, and David Silver. Emergence of Locomotion Behaviours in Rich Environments. *CoRR*, abs/1707.02286, 2017.

[19] Daniel Honerkamp, Tim Welschehold, and Abhinav Valada. Learning kinematic feasibility for mobile manipulation through deep reinforcement learning. *IEEE Robotics and Automation Letters (RA-L)*, pages 6289–6296, 2021.

[20] Seunghun Jeon, Moonkyu Jung, Suyoung Choi, Beomjoon Kim, and Jemin Hwangbo. Learning whole-body manipulation for quadrupedal robot. *IEEE Robotics and Automation Letters (RA-L)*, pages 699–706, 2023.

[21] 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.

[22] Seungjae Lee, Yibin Wang, Haritheja Etukuru, H. Jin Kim, Nur Muhammad Mahi Shafiullah, and Lerrel Pinto. Behavior generation with latent actions. *arXiv preprint arXiv:2403.03181*, 2024.

[23] Jiaman Li, Alexander Clegg, Roozbeh Mottaghi, Jiajun Wu, Xavier Puig, and C Karen Liu. Controllable human-object interaction synthesis. In *Proc. of the EuropeanConf. on Computer Vision (ECCV)*, pages 54–72, 2024.

[24] Junheng Li and Quan Nguyen. Multi-contact mpc for dynamic loco-manipulation on humanoid robots. In *IEEE*

American Control Conference (ACC), pages 1215–1220, 2023.
[25] Qiayuan Liao, Bike Zhang, Xuanyu Huang, Xiaoyu Huang, Zhongyu Li, and Koushil Sreenath. Berkeley Humanoid: A Research Platform for Learning-based Control, 2024.
[26] Changyi Lin, Xingyu Liu, Yuxiang Yang, Yaru Niu, Wenhao Yu, Tingnan Zhang, Jie Tan, Byron Boots, and Ding Zhao. LocoMan: Advancing Versatile Quadrupedal Dexterity with Lightweight Loco-Manipulators. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2024.
[27] C. Karen Liu and Dan Negrut. The Role of Physics-Based Simulators in Robotics. *Annual Review of Control, Robotics, and Autonomous Systems*, 4(1):35–58, 2021.
[28] Minghuan Liu, Zixuan Chen, Xuxin Cheng, Yandong Ji, Ruihan Yang, and Xiaolong Wang. Visual whole-body control for legged loco-manipulation. *Conference on Robot Learning (CoRL)*, 2024.
[29] Naureen Mahmood, Nima Ghorbani, Nikolaus F. Troje, Gerard Pons-Moll, and Michael J. Black. AMASS: Archive of Motion Capture as Surface Shapes. In *Proc. of the Int. Conf. on Computer Vision (ICCV)*, Oct 2019.
[30] W. Meeussen, M. Wise, S. Glaser, S. Chitta, C. McGann, P. Mihelich, E. Marder-Eppstein, M. Muja, V. Eruhimov, T. Foote, J. Hsu, R. B. Rusu, B. Marthi, G. Bradski, K. Konolige, B. Gerkey, and E. Berger. Autonomous door opening and plugging in with a personal robot. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2010.
[31] Mayank Mittal, David Hoeller, Farbod Farshidian, Marco Hutter, and Animesh Garg. Articulated object interaction in unknown scenes with whole-body mobile manipulation. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 1647–1654. IEEE, 2022.
[32] Mayank Mittal, Calvin Yu, Qinxi Yu, Jingzhou Liu, Nikita Rudin, David Hoeller, Jia Lin Yuan, Ritvik Singh, Yunrong Guo, Hammad Mazhar, et al. Orbit: A unified simulation framework for interactive robot learning environments. *IEEE Robotics and Automation Letters (RA-L)*, 8(6):3740–3747, 2023.
[33] Mayank Mittal, Nikita Rudin, Victor Klemm, Arthur Allshire, and Marco Hutter. Symmetry considerations for learning task symmetric robot policies. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*. IEEE, 2024.
[34] Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel van de Panne. Deepmimic: example-guided deep reinforcement learning of physics-based character skills. *ACM Transaction of Graphics (ToG)*, 37(4):143, 2018.
[35] Tifanny Portela, Andrei Cramariuc, Mayank Mittal, and Marco Hutter. Whole-body end-effector pose tracking. *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2025.
[36] Agility Robotics. Crossing the sim2real gap with NVIDIA Isaac Lab, 2024.
[37] Nicholas Roy, Ingmar Posner, Tim Barfoot, Philippe Beaudoin, Yoshua Bengio, Jeannette Bohg, Oliver Brock, Isabelle Depatie, Dieter Fox, Dan Koditschek, et al. From machine learning to robotics: Challenges and opportunities for embodied intelligence. *arXiv preprint arXiv:2110.15245*, 2021.
[38] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.
[39] Jean-Pierre Sleiman, Farbod Farshidian, Maria Vittoria Minniti, and Marco Hutter. A unified mpc framework for whole-body dynamic locomotion and manipulation. *IEEE Robotics and Automation Letters (RA-L)*, pages 4688–4695, 2021.
[40] Jean-Pierre Sleiman, Farbod Farshidian, and Marco Hutter. Versatile multicontact planning and control for legged loco-manipulation. *Science Robotics*, 2023.
[41] Jean-Pierre Sleiman, Mayank Mittal, and Marco Hutter. Guided reinforcement learning for robust multi-contact loco-manipulation. In *Conference on Robot Learning (CoRL)*, 2024.
[42] Jonas Stolle, Philip Arm, Mayank Mittal, and Marco Hutter. Perceptive pedipulation with local obstacle avoidance. In *Proc. of the IEEE-RAS Int. Conf. on Humanoid Robots (ICHR)*, pages 157–164. IEEE, 2024.
[43] Chen Tessler, Yunrong Guo, Ofir Nabati, Gal Chechik, and Xue Bin Peng. Maskedmimic: Unified physics-based character control through masked motion inpainting. *ACM Transaction of Graphics (ToG)*, 2024.
[44] Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 5026–5033, 2012. doi: 10.1109/IROS.2012.6386109.
[45] Carnegie Mellon University. Cmu graphics lab motion capture library, 2002. URL http://mocap.cs.cmu.edu/.
[46] Nikolaus Vahrenkamp and Tamim Asfour. Representing the Robot's Workspace through Constrained Manipulability Analysis. *Autononous Robots*, page 17–30, jan 2015.
[47] Fei Xia, Chengshu Li, Roberto Martín-Martín, Or Litany, Alexander Toshev, and Silvio Savarese. Relmogen: Leveraging motion generation in reinforcement learning for mobile manipulation. *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2021.
[48] Qinxi Yu, Masoud Moghani, Karthik Dharmarajan, Vincent Schorp, William Chung-Ho Panitch, Jingzhou Liu, Kush Hari, Huang Huang, Mayank Mittal, Ken Goldberg, et al. Orbit-surgical: An open-simulation framework for learning surgical augmented dexterity. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 15509–15516. IEEE, 2024.
[49] Mike Zhang, Yuntao Ma, Takahiro Miki, and Marco Hutter. Learning to open and traverse doors with a legged manipulator. *Conference on Robot Learning (CoRL)*, 2024.