ETH zürich

CRL Computational Robotics Lab

Sampling-based MPC and Diffusion-based Online Policy Distillation for Agile Control of Hybrid Mobile Robots

Motivation:

Hybrid mobile robots that combine wheels and legs offer powerful locomotion capabilities, merging the efficiency of wheeled travel with the versatility of legged movement. However, their control is particularly challenging due to the complex interaction between balancing and steering across different locomotion modes. Learning-based methods provide a way to train task-specific control policies, but they still suffer from the sim-to-real gap. Policies often degrade when deployed on hardware, and adjusting parameters typically involves a back-and-forth process of testing and retraining, which lacks real-time adaptability. Model-based methods, such as trajectory optimization and model predictive control (MPC), can produce high-quality control actions. Yet, they often accurate analytical dynamics models and state estimation, which are difficult to obtain in real-world, contact-rich environments.

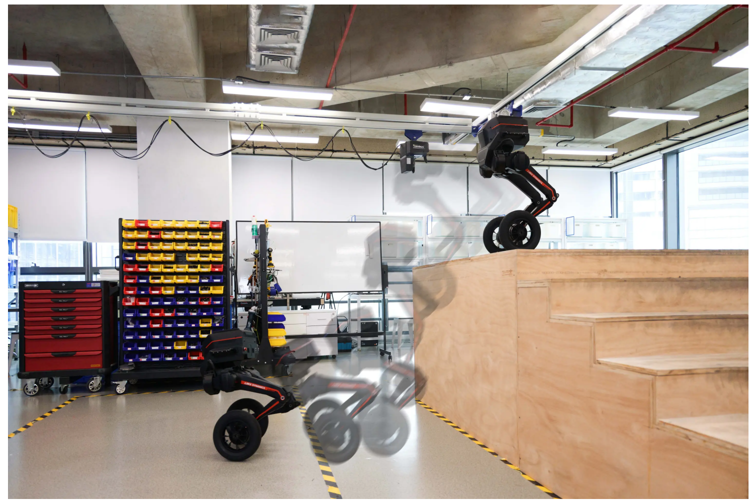

Figure 1: LIMX Dynamics TRON1 robot—a bipedal wheeled platform with 3 actuated joints and one wheel per leg (4 DoFs per leg total)

This project focuses on state-of-the-art sampling-based MPC and its deployment on hardware by exploiting fast, parallelized simulators to achieve real-time control action optimization. Furthermore, to reduce computational demands for future tasks, we aim to distill the generated state-action pairs into neural policies using diffusion models, preserving both the multi-modality and improve the efficiency of the sampling-based approach.

Approach:

The project will explore a control-learning pipeline that integrates sampling-based model predictive control with online policy distillation.

First, a sampling-based MPC controller, such as MPPI [1], CEM [2], or diffusion-based annealing [3, 4], will be implemented to generate high-quality control actions. These methods are well-suited for fast, parallelized simulators (e.g., MuJoCo Playground [5] or Brax [6]) and can handle the non-differentiable dynamics common in legged and wheeled locomotion. Task-specific cost functions will be designed to support stable locomotion, agile maneuvering, and robustness to disturbances.

While the MPC controller runs, its trajectories and actions will be logged in real time and used to train a neural policy through diffusion-based distillation [7]. Unlike conventional supervised distillation, diffusion models can better capture the multi-modality of the MPC-generated action distributions, leading to policies that are both efficient at inference and diverse in behavior [8, 9, 10]. The training process will ensure that the distilled policy generalizes across different tasks while remaining lightweight enough for real-time deployment on hardware.

General Details:

Applicants should have programming experience in both Python and C++, along with a basic understanding of robot simulation and trajectory optimization. Familiarity with reinforcement learning is a plus but not strictly required. We will be using LIMX Dynamics TRON1 robot as the main development platform Fig. 1.

Interested?

The project will be supervised by Jin Cheng and Prof. Dr. Stelian Coros. Please reach out to Jin (jin.cheng@inf.ethz.ch) with your CV and transcripts.

References

[1] Juan Alvarez-Padilla, John Z Zhang, Sofia Kwok, John M Dolan, and Zachary Manchester. Real-time whole-body control of legged robots with model-predictive path integral control. *arXiv preprint arXiv:2409.10469*, 2024.

[2] Taylor Howell, Nimrod Gileadi, Saran Tunyasuvunakool, Kevin Zakka, Tom Erez, and Yuval Tassa. Predictive sampling: Real-time behaviour synthesis with mujoco. *arXiv preprint arXiv:2212.00541*, 2022.

[3] 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.

[4] Chaoyi Pan, Zeji Yi, Guanya Shi, and Guannan Qu. Model-based diffusion for trajectory optimization. *Advances in Neural Information Processing Systems*, 37:57914–57943, 2024.

[5] Kevin Zakka, Baruch Tabanpour, Qiayuan Liao, Mustafa Haiderbhai, Samuel Holt, Jing Yuan Luo, Arthur Allshire, Erik Frey, Koushil Sreenath, Lueder A Kahrs, et al. Mujoco playground. *arXiv preprint arXiv:2502.08844*, 2025.

[6] C Daniel Freeman, Erik Frey, Anton Raichuk, Sertan Girgin, Igor Mordatch, and Olivier Bachem. Brax–a differentiable physics engine for large scale rigid body simulation. *arXiv preprint arXiv:2106.13281*, 2021.

[7] Miguel Angel Zamora Mora, Momchil Peychev, Sehoon Ha, Martin Vechev, and Stelian Coros. Pods: Policy optimization via differentiable simulation. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 7805–7817. PMLR, 18–24 Jul 2021. URL `https://proceedings.mlr.press/v139/mora21a.html`.

[8] Xiaoyu Huang, Yufeng Chi, Ruofeng Wang, Zhongyu Li, Xue Bin Peng, Sophia Shao, Borivoje Nikolic, and Koushil Sreenath. Diffuseloco: Real-time legged locomotion control with diffusion from offline datasets. *arXiv preprint arXiv:2404.19264*, 2024.

[9] Xiaoyu Huang, Takara Truong, Yunbo Zhang, Fangzhou Yu, Jean Pierre Sleiman, Jessica Hodgins, Koushil Sreenath, and Farbod Farshidian. Diffuse-cloc: Guided diffusion for physics-based character look-ahead control. *ACM Transactions on Graphics (TOG)*, 44(4):1–12, 2025.

[10] Qiayuan Liao, Takara E Truong, Xiaoyu Huang, Guy Tevet, Koushil Sreenath, and C Karen Liu. Beyondmimic: From motion tracking to versatile humanoid control via guided diffusion. *arXiv e-prints*, pages arXiv–2508, 2025.