Benchmarking Model Predictive Control and Reinforcement Learning for Legged Robot Locomotion*

Shivayogi Akki[1], Zherong Pan[2] and Tan Chen[3]

I. Motivation and Problem Statement

Legged robots have gained significant attention in recent years because they can navigate over rough terrain, climb stairs, and move over obstacles that would be difficult or even impossible for wheeled robots [1]. Specifically, the quadrupedal robots can have a wide range of applications, including transportation tasks in industry, package delivery, search and rescue, etc. Currently model predictive control (MPC) and reinforcement learning (RL) are two major trends for controlling quadrupedal robots [2], [3]. However, selecting the most suitable controller for a specific application can be a daunting task for new researchers.

In this paper, we present a comparative study of MPC and RL controllers on the Unitree Go1 quadrupedal robot, evaluating their performance under perturbation and model uncertainty. Additionally, by assessing the controllers in different environments (flat/flat slippery/uneven terrain), we aim to provide valuable insights to aid researchers in making informed decisions when designing locomotion controllers.

II. Approach

A. Controllers

We utilize the predictive sampling algorithm for the MPC controller. It iteratively refines a nominal sequence of actions represented by spline parameters using random search. At each iteration, a set of N candidate splines is evaluated, including the nominal itself and $(N-1)$ noisy samples generated from a Gaussian distribution. Actions are clamped within control limits, and the best candidate is selected based on total return. For the RL control, we employ the Proximal Policy Optimization (PPO) algorithm. The reward system consists of a forward reward term, based on the change in the robot's x-coordinate, and a control cost term to discourage excessive control effort. The final reward is calculated by the summation of forward velocity reward, height reward, lateral position penalty and control penalty. In summary, a model-based controller relies on an explicit model of the system and environment, while a model-free RL controller learns from exploratory interaction.

B. Benchmark Experiments

In order to make the results comparable, we design the same task for both two controllers, for instance, commanding the robot to walk along one direction at a constant speed. The robot state can be represented by the joint and CoM states. After tuning the parameters for each controller, Experiment 1 will plot out the trajectories of robot states and compare their similarity for validation. Experiment 2 will compare the performance of the two controllers under perturbation. Experiment 3 will compare the performance of the two controllers under model uncertainty, *i.e.,* designing a controller on flat terrain and testing the performance of this controller on flat slippery terrain and uneven terrain.

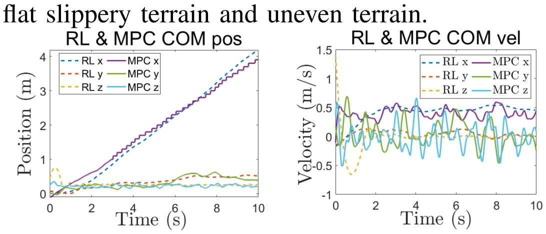

Fig. 1. Robot position and velocity trajectory plot of MPC and RL with target velocity of 0.5 m/s.

III. Main Results and Work Plan

Currently MPC and RL controllers can maintain approximate velocity of 0.5 m/s in x-axis (Figure 1). They have different gaits but similar result as per position and velocity of CoM. Figure 2 shows the CoM trajectories of the robot under major perturbation along x-axis. The RL controller has been implemented with an objective to maximize the speed. The flat/flat slippery/uneven terrains were built in MuJoCo. Here is the work plan for poster presentation:

- Apply perturbation to the robot with RL controller and compare the performance under perturbation
- Compare the performance under model uncertainty.

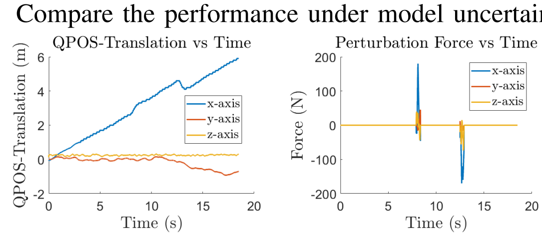

Fig. 2. Robot control with MPC under perturbation.

References

[1] Marc H Raibert. *Legged robots that balance*. MIT press, 1986.

[2] 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.

[3] Joonho Lee, Jemin Hwangbo, Lorenz Wellhausen, Vladlen Koltun, and Marco Hutter. Learning quadrupedal locomotion over challenging terrain. *Science robotics*, 5(47):eabc5986, 2020.

*This paper reports the work from a course project.

[1]Shivayogi Akki is with the Department of Mechatronics, Robotics and Automation Engineering, Michigan Technological University, Houghton, MI 49931 USA sakki@mtu.edu

[2]Zherong Pan is a Senior Researcher with Tencent America, Palo Alto, and received a Ph.D. degree from The University of North Carolina zherongpanusa@gmail.com

[3]Tan Chen is with the Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, MI 49931 USA tanchen@mtu.edu