GLiDE: Generalizable Quadrupedal Locomotion in Diverse Environments with a Centroidal Model

Zhaoming Xie
University of British Columbia

Xingye Da
NVIDIA

Buck Babich
NVIDIA

Animesh Garg
University of Toronto

Michiel van de Panne
University of British Columbia

***Abstract*—Model-free reinforcement learning (RL) for legged locomotion commonly relies on a physics simulator that can accurately predict the behaviors of every degree of freedom of the robot. In contrast, approximate reduced-order models are often sufficient for many model-based control strategies. In this work we explore how RL can be effectively used with a centroidal model to generate robust control policies for quadrupedal locomotion. Advantages over RL with a full-order model include a simple reward structure, reduced computational costs, and robust sim-to-real transfer. We further show the potential of the method by demonstrating stepping-stone locomotion, two-legged in-place balance, balance beam locomotion, and sim-to-real transfer without further adaptations. Additional Results: https://sites.google.com/view/centroidal-rl.**

I. Introduction

Tremendous progress has been made recently in the field of legged locomotion, achieved using both model predictive control (MPC) and reinforcement learning (RL) methods. MPC methods leverage modern optimization techniques and known models of the physics, possibly simplified, to synthesize responsive control at run time. However, they can be prone to local minima, can require substantial manual tuning, and are difficult to generalize to complex terrains and rich perceptual streams. Moreover, the real-time MPC usually uses a linear model to reduce the computation time, which is hard to represent the legged system's nonlinear and hybrid nature. Alternatively, model-free methods such as reinforcement learning (RL) utilize Monte-Carlo sampling strategies and can learn control policies for general tasks. This comes at the expense of extensive offline physics simulations required during training, careful system modeling, and detailed reward design to produce results that are feasible for physical robots.

In this paper, we seek to realize some of the key benefits of both approaches. Instead of relying on an accurate simulation of the robot model, we use a strongly-abstracted centroidal model. We model the robot as a single rigid body with massless legs that is controlled via the ground reaction forces (GRF) applied at the legs that are in contact with the environment, as shown in Fig. 1. We further assume a specified gait pattern and a foot-placement function. Taken together, this allows for a simple task reward specification, in contrast to the more complex reward structures commonly required for full-model RL. Additional constraints such as no-slip constraints and leg lengths are enforced via quadratic programming (QP) and foot placement strategies. With the above in place, we then synthesize control policies for this simplified model via reinforcement learning. The resulting policy actions are realized on the full robot model by converting the GRFs to joint torques using the Jacobian transpose for the stance legs and a simple trajectory-tracking approach for the swing legs. The resulting control policies are validated on simulations of the Laikago and A1[1] quadruped robots as well as on a physical A1 robot.

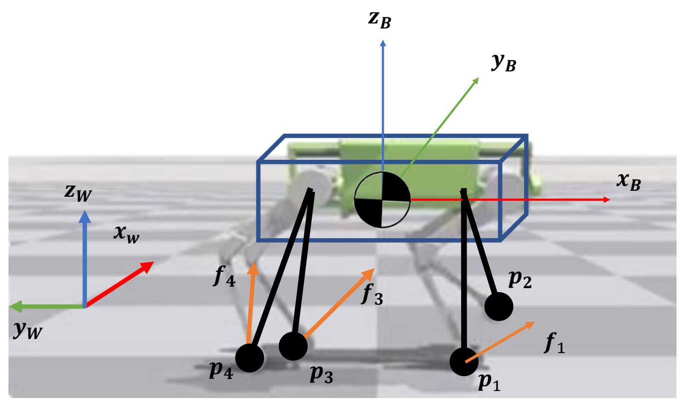

Fig. 1: Learning to generate a whole-body locomotion controller with reinforcement learning is both inefficient and brittle. A key insight of this paper is to combine the strengths of reduced-order modeling with learning through RL in Centroidal Model Space. The centroidal model consists of a single rigid body with virtual legs attached as illustrated above. Control is realized by generating ground reaction forces at foot locations in contact with the ground.

Our core contributions are as follows:

- We introduce a framework for learning control policies suited for centroidal dynamics models, enabling the anticipatory behavior required for quadrupedal locomotion and balance tasks, without the complexities of working with the full model. This allows for simple reward design, enables efficient simulation during training, and yields flexible and robust motion control.
- We demonstrate the effectiveness of this framework for multiple gaits, stepping stone scenarios, balance beam locomotion, and two-legged in-place balancing. We further show successful transfer to a physical robot.

References

[1] Laikago and A1 are quadrupedal robots made by Unitree Robotics.