Leveraging Spring Mass Locomotion to Guide Learned Walking Controllers

Kevin Green, Yesh Godse, Jeremy Dao, Ross L. Hatton, Alan Fern, and Jonathan Hurst

{greenkev, godsey, daoje, ross.hatton, alan.fern, jhurst}@oregonstate.edu

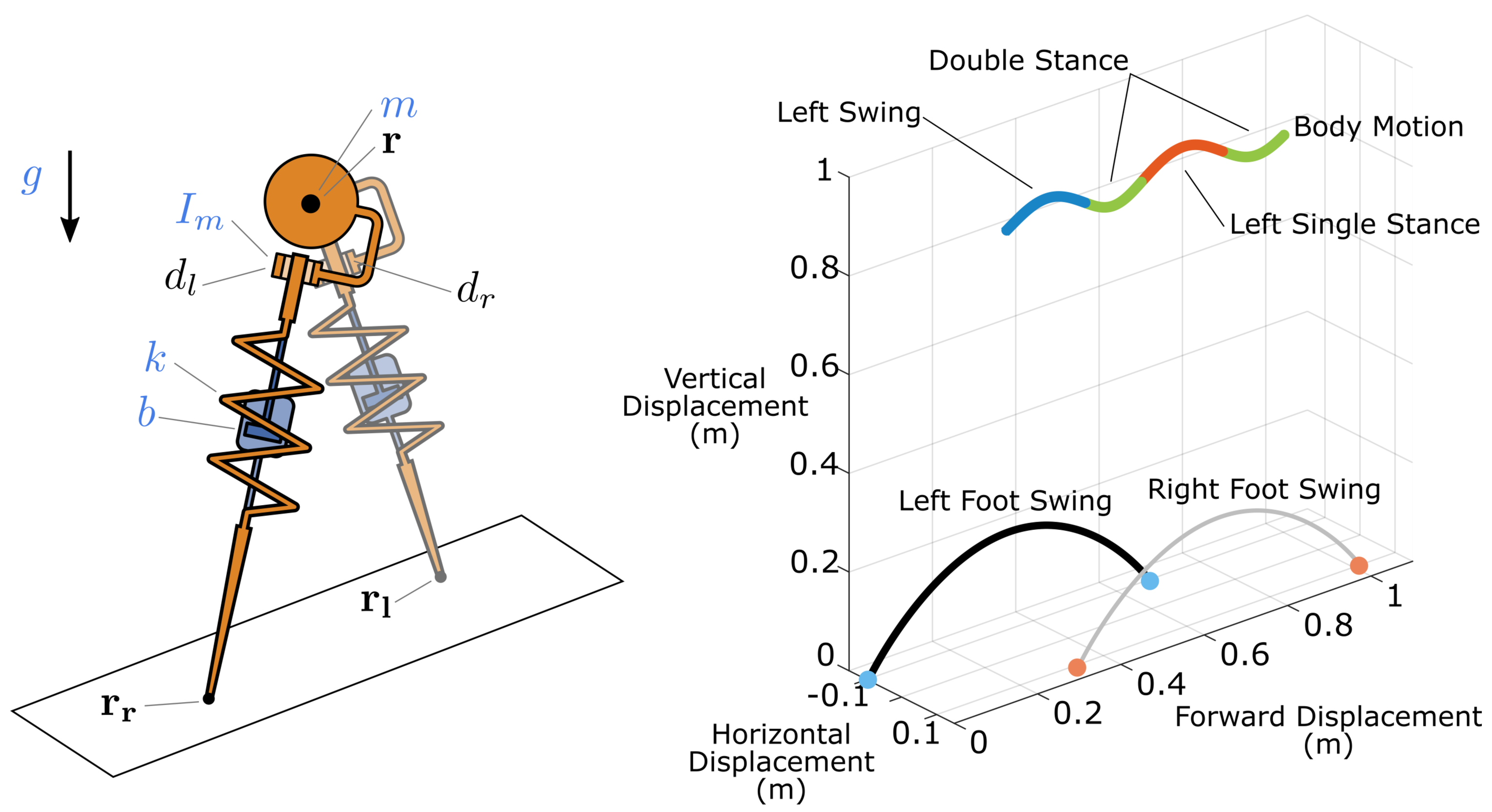
Oregon State University

Abstract

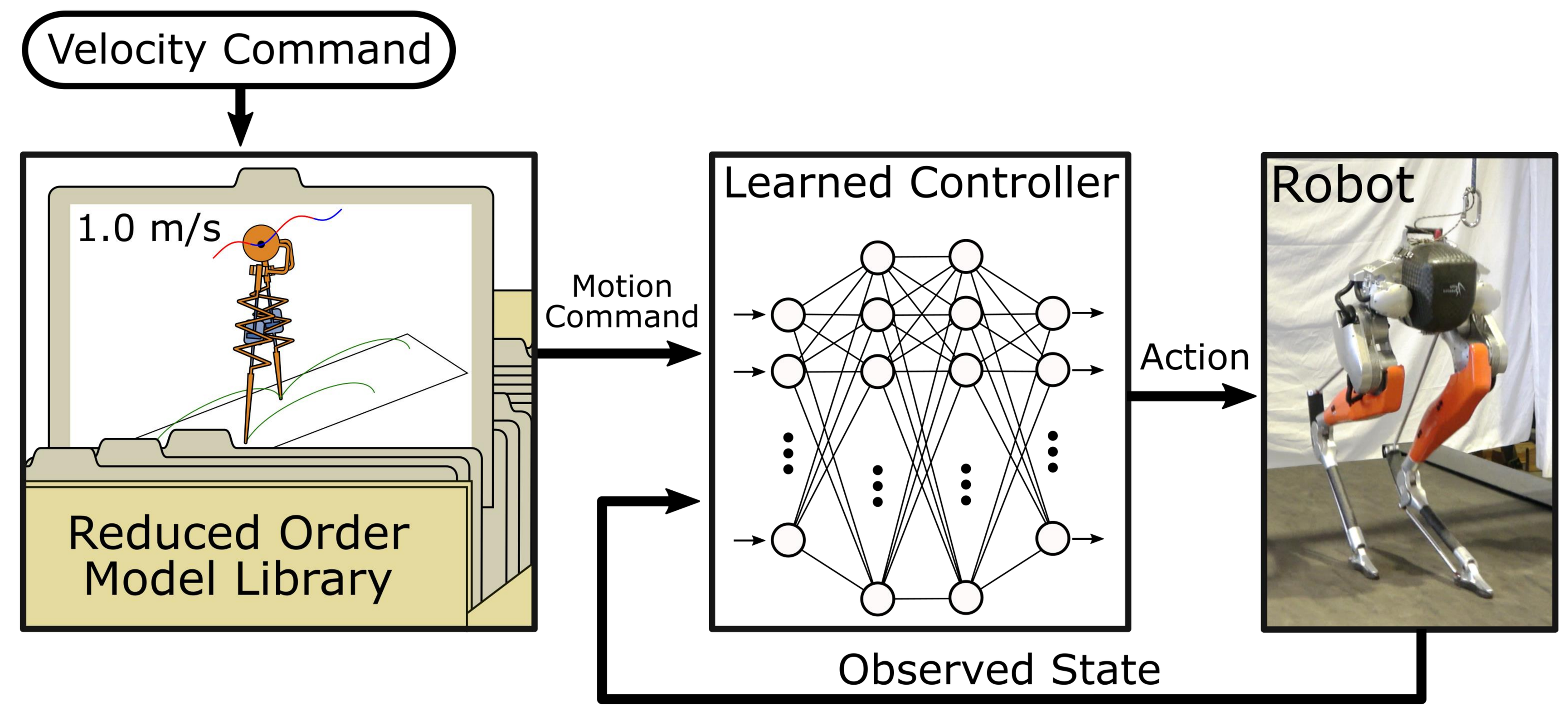
In this work, we describe an approach to achieving dynamic legged locomotion on physical robots which uses existing reduced order models of locomotion to guide reinforcement learning. Ultimately, our goal is that this work leads to a control hierarchy where the highest-level behaviors are planned through reduced-order models and lower level controllers utilize a learned policy that can bridge the gap between the idealistic model and the physical robot. Here we present the learned dynamic walking controller by showing that a range of walking motions from reduced-order models can be used as the command and primary training signal. Instead of a reactive planner, we use a library of optimized periodic walking motions from an actuated SLIP model. The resulting policies do not attempt to naively track the motion as a traditional trajectory tracking controller would, but instead balance immediate motion tracking with long term stability. The resulting controller is demonstrated on a physical Agility Robotics Cassie series bipedal robot at speeds up to 1.2 m/s. This work builds the foundation of a generic, dynamic learned walking controller that can be applied to many different tasks.

Actuated SLIP Reference Motion

We build a library of optimized steady state, double support walking gaits for the actuated spring loaded inverted pendulum (actuated SLIP) model. We solve for gaits from 0 to 2.0 m/s in 0.1 m/s increments. This library is using to command and train the learned controller.



The objective function for these gaits is to minimize integrated actuator acceleration while satisfying a series of constraints that ensure it is physically viable on Cassie.



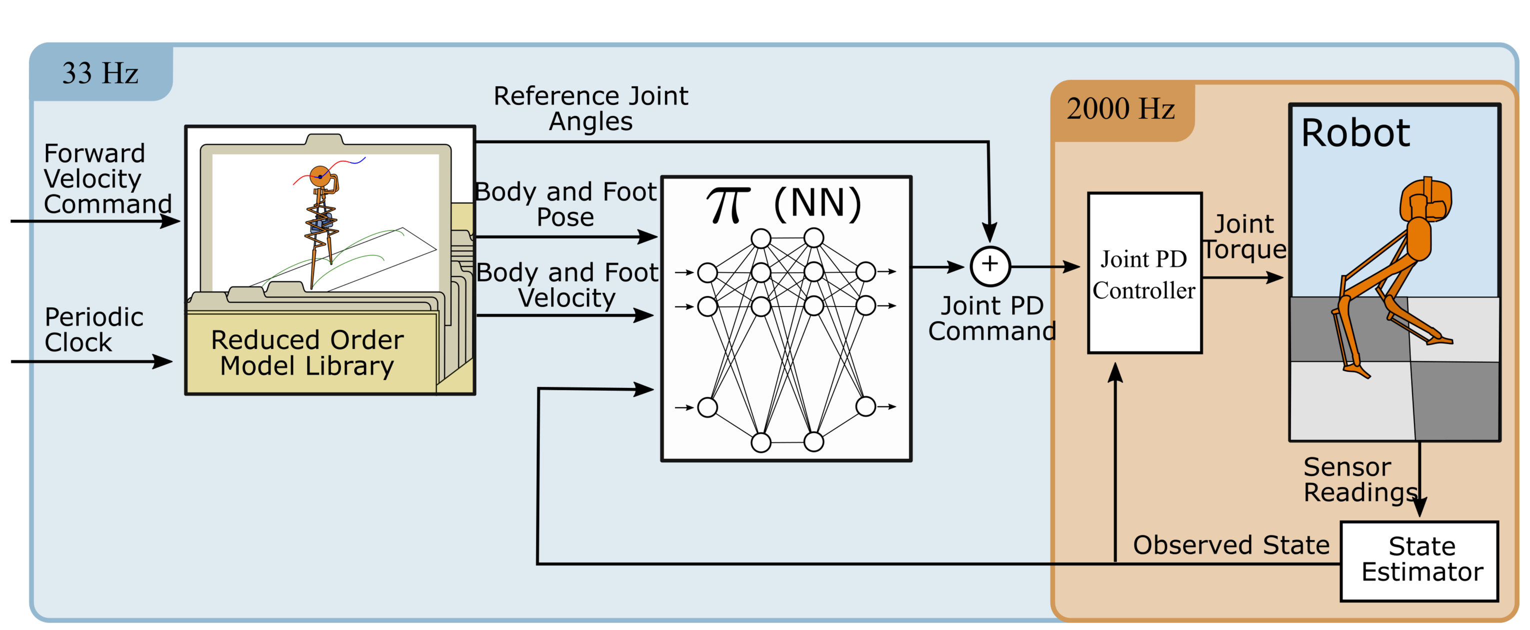
Learned Controller

Our learned controller uses the library of ASLIP reference trajectories as a learning signal to capture the most important features of the reduced-order model's motion.

Policy Input (64D)	Policy Output (10D)
Robot State (from state estimator) ROM Desired Task Space Pose ROM Desired Task Space Velocity	Motor PD Targets

We learn control policies using a simulated model of Cassie in the MuJoCo physics simulator and PPO. **The majority of the reward function consists of matching the reference body velocity and foot position**

Reward Term	% of Max Reward
ROM Task Space Pose Matching	30
ROM Task Space Vel Matching	30
Smooth Action Commands	10
Foot Orientation	20
Body Drift	10

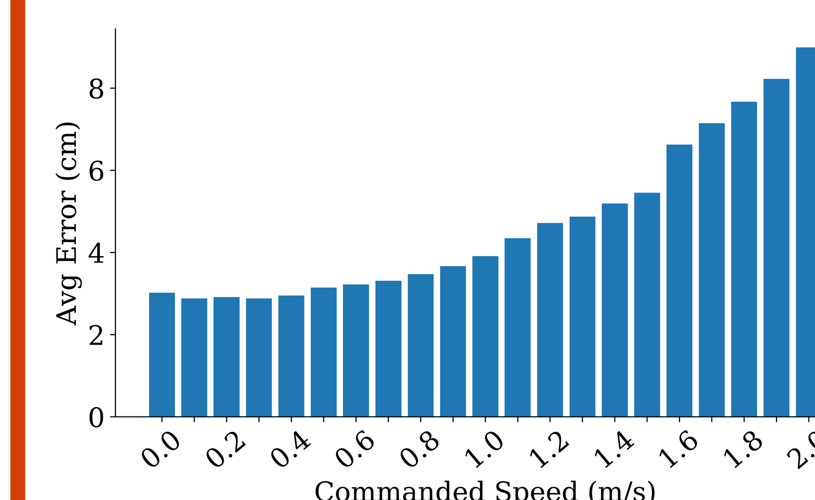


Video Here of Hardware Demonstration
https://youtu.be/QsWrjXI5hsg

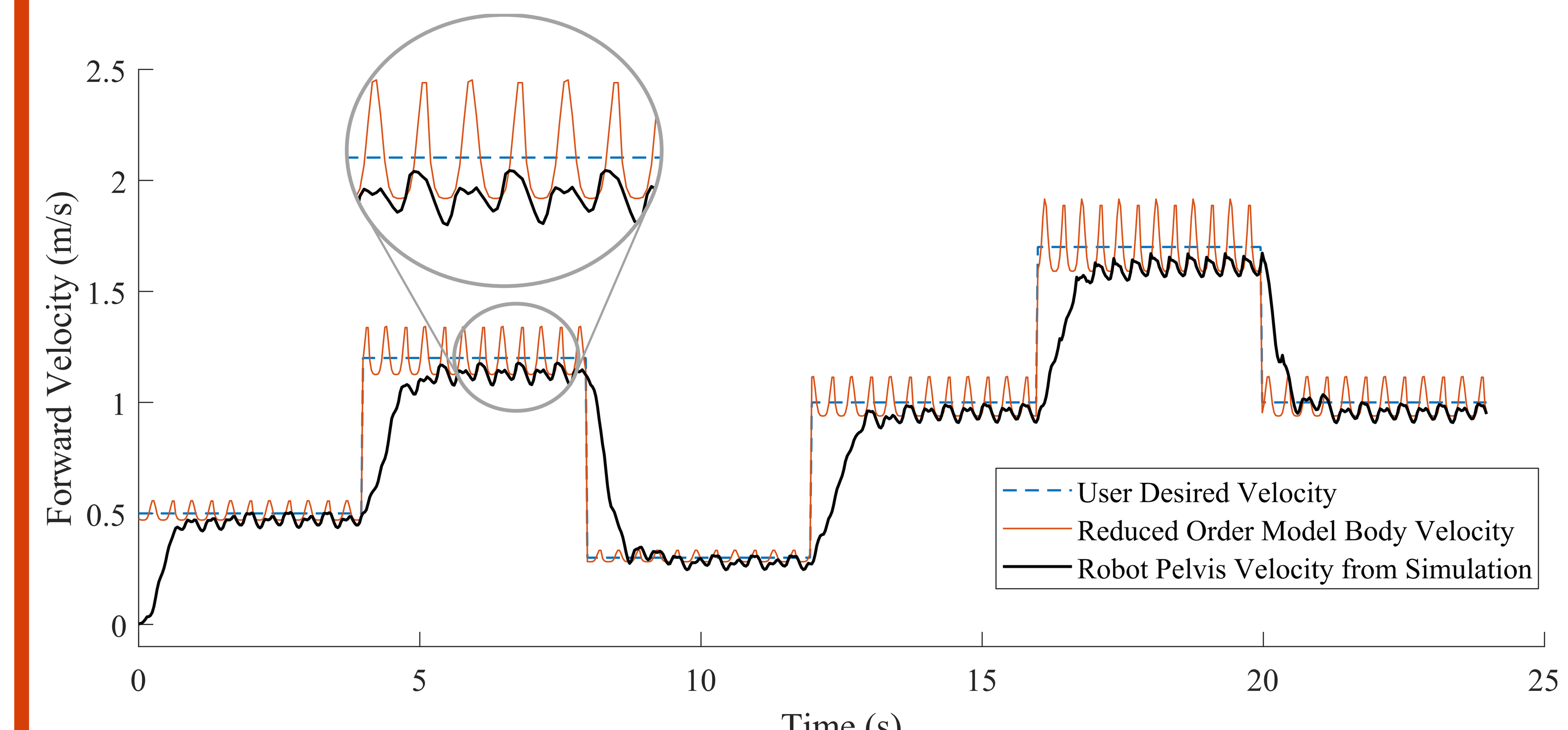
We directly transfer policies trained in simulation to hardware and observe that **the produced motion resembles the underlying spring-mass motion**.

Results

Policies learn to step in place after about 25 million timesteps and can track the library reference motion after about 175 million timesteps.



Especially at lower speeds, **the learned controller is capable of closely matching the reference motion's foot locations.** This is a critical feature for any high-level footstep planning that will use this control pipeline.



In simulation the controller tracks the changing velocity commands well and handles speed changes effectively.

Acknowledgements

This work was supported by NSF grant DARPA contract W911NF-16-1-0002, NSF Grant No. CMMI-1653220 and NSF Grant No. DGE-1314109. We thank John Warila and Dylan Albertazzi for their assistance rendering videos, Helei Duan, Jonah Siekmann, Lorzeno Bermillo, and Pedro Morais for productive discussions and feedback and Stephen Offer and Intel Labs for their computing resources and support.