

Whole-Body Model-Predictive Control of Legged Robots with MuJoCo

John Z. Zhang¹, Taylor A. Howell², Zeji Yi³, Chaoyi Pan³, Guanya Shi⁴, Guannan Qu³,
Tom Erez², Yuval Tassa², and Zachary Manchester¹

Abstract—We demonstrate the surprising real-world effectiveness of a very simple approach to whole-body model-predictive control (MPC) of quadruped and humanoid robots: the iterative linear-quadratic regulator (iLQR) algorithm with MuJoCo dynamics and finite-difference approximated derivatives. Building upon the previous success of model-based behavior synthesis and control of locomotion and manipulation tasks with MuJoCo in simulation, we show that these policies can easily generalize to the real world with few sim-to-real considerations. Our baseline method achieves real-time MPC while leveraging *whole-body* dynamics and collision detection on a variety of hardware experiments, including dynamic quadruped locomotion, a quadruped walking on two legs, and full-sized humanoid bipedal locomotion. Additionally, our GUI system enables users to interactively update robot behavior in real-time on the robot hardware, making task-specific parameter tuning easy and intuitive. Our code and experiment videos are available online at: https://johnzhang3.github.io/mujoco_ilqr.

I. INTRODUCTION

Enabling legged robots to achieve human and animal-level agility has been a decades-long challenge for robotics researchers. In addition to the challenges faced by other non-legged mobile robots (e.g., drones, autonomous vehicles, etc.), legged systems are generally high-dimensional and must effectively reason about making and breaking contact with the world. Advancements in model-based control [1], [2] and reinforcement learning (RL) [3], [4] methods have unlocked tremendous in-the-wild legged robot capabilities over the last 10-15 years. Over the same period, robotics simulation has seen significant growth in terms of physical accuracy [5]–[7], differentiability [7], [8], and parallelization performance [6], [8], [9]. Thanks to these advances in simulation technologies combined with incredible tools supported by the broader machine learning community such as PyTorch [10] and JAX [11], sim-to-real RL has enjoyed accelerated progress and become the standard approach for solving challenging problems like humanoid whole-body control. Interestingly, in comparison, model-based control researchers have generally favored novel, custom implementations of robot models [7], [12] and optimization solvers [13]–[15], in part due to the online computation requirements of model-predictive control (MPC) paradigm, making these

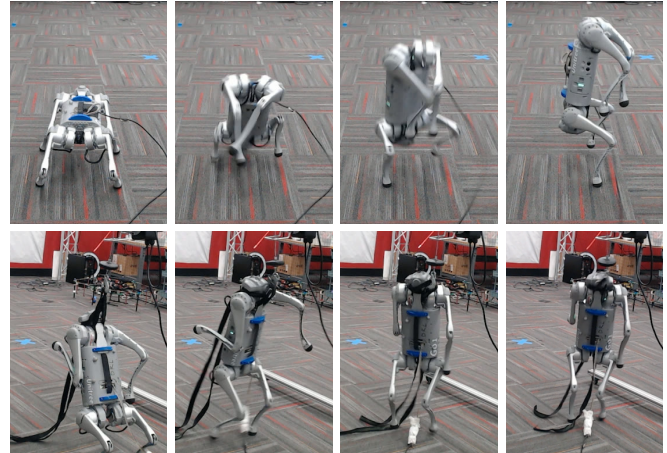


Fig. 1. A Unitree Go1 quadruped robot transitions from quadruped to handstand mode (top row) and walking on its hind legs (bottom row) using the MuJoCo iLQR policy.

works relatively more difficult to reproduce which, so far, has resulted in slower community adoption.

This paper aims to reduce this gap by providing an open-sourced baseline MPC algorithm and real-world legged robot implementation built on the MuJoCo physics engine [6], a standard, easy-to-use open-source robotics simulator. We show that a standard gradient-based MPC algorithm, in particular the iterative LQR (iLQR) algorithm, based on MuJoCo is surprisingly capable of solving a variety of challenging *real-world* tasks such as bipedal locomotion on a quadruped and full-sized humanoid robot in *real time*. By leveraging the efficient C implementation of MuJoCo as the backend for the forward model and derivative computations, our real-time MPC approach reasons about both whole-body dynamics and collision detection in the model, something difficult to achieve in previous open-source whole-body MPC algorithms [16], [17]. Additionally, we design an interactive GUI system that enables users to quickly modify key MPC parameters and observe *real-world* robot behaviors alongside a simulated twin. We hope that this effort lowers the barrier to entry for further model-based control research on legged robot hardware and eventually leads to accelerated research momentum.

Our specific contributions in this paper are:

- A simple-yet-surprisingly-effective baseline whole-body predictive control algorithm for real-world legged robot locomotion.
- An open-source interactive GUI system for real-world predictive control of legged robots.
- A set of hardware experiments demonstrating the effec-

¹ John Z. Zhang and Zachary Manchester are with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology.

² Taylor A. Howell, Tom Erez, and Yuval Tassa are with Google DeepMind.

³ Zeji Yi, Chaoyi Pan, and Guannan Qu are with the Department of Electrical and Computer Engineering, Carnegie Mellon University.

⁴ Guanya Shi is with the Robotics Institute, Carnegie Mellon University.

This work was completed at Carnegie Mellon University.

Corresponding email: jzhang3@mit.edu

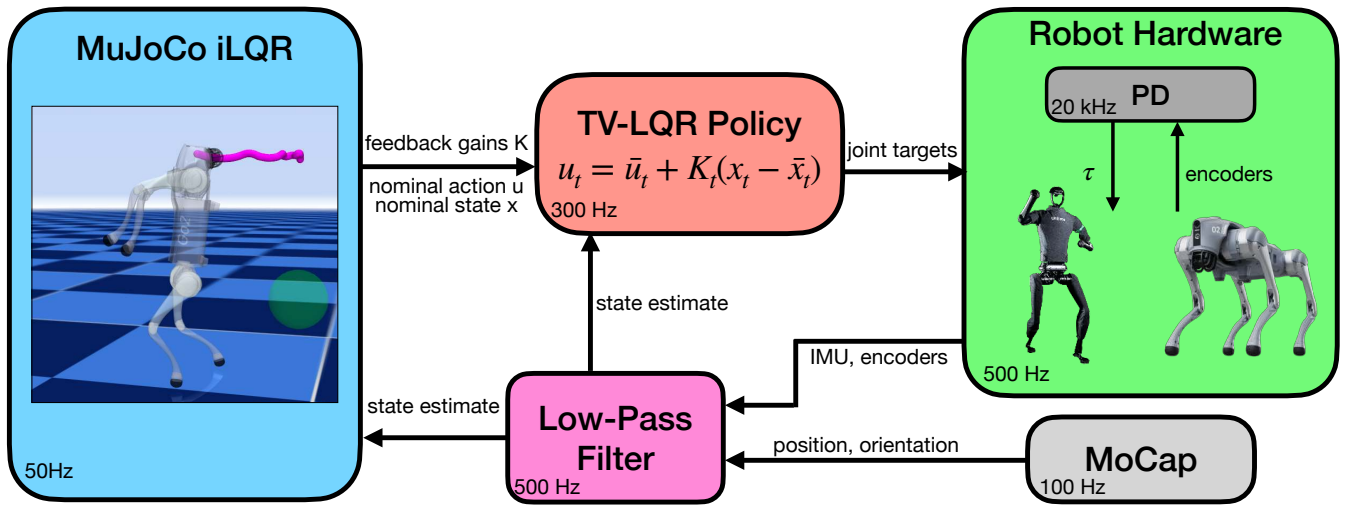


Fig. 2. System diagram for deploying the MuJoCo iLQR policy to the Unitree Quadruped and Humanoid robots. The iLQR algorithm provides control, state, and time-varying LQR (TV-LQR) feedback gain trajectories at 50 Hz. The TV-LQR feedback policy can then be updated at 300 Hz and passed to a joint-level PD controller. The robot’s state is estimated by fusing onboard joint encoders and motion capture data. Live state estimates are updated in the planner at 300 – 500 Hz and visualized in the MuJoCo MPC GUI. The cost categories are designed offline but the relative weights, goal locations, and iLQR hyperparameters can be adjusted by the user interactively in real time through the GUI.

tiveness of the baseline algorithm on both quadruped and humanoid robots across a variety of tasks.

The remainder of this paper is organized as follows: We begin by reviewing relevant literature on iLQR, whole-body MPC, and open-source efforts for model-based control in Sec. II. Next, we briefly introduce the MuJoCo contact model and iLQR in Sec. III, followed by key considerations and implementation details for transferring iLQR policies to hardware in IV. Then, we cover the hardware setup and experimental results in Sec. V. Finally, we conclude in Sec. VI by discussing current limitations of our system and directions for future work.

II. BACKGROUND AND RELATED WORKS

This section provides a brief review of the iLQR algorithm, a survey of relevant literature on MPC for legged robots, and other notable efforts for improving tooling for model-based control.

A. Iterative Linear-Quadratic Regulator

Differential Dynamic Programming (DDP), originally introduced in [18], solves the following nonlinear trajectory optimization problem:

$$\begin{aligned} \min_{u_{0:T-1}} \quad & \sum_{t=0}^{T-1} l(x_t, u_t) + l_f(x_T) \\ \text{subject to} \quad & x_{t+1} = f(x_t, u_t), \\ & x_0 = x(0), \end{aligned} \quad (1)$$

by iteratively solving the an locally approximated problem with Dynamic Programming. In addition to the nominal control sequence $u_{0:T-1}$, DDP also produces a time-varying linear feedback policy:

$$u_t = \bar{u}_t + K_t(x_t - \bar{x}_t) + \alpha k_t \quad (2)$$

where \bar{u} and \bar{x} are the current solution, K_t is the feedback gain matrix at time index t , k is an improvement to the current nominal control, and α is the line search step size. This single-shooting formulation only optimizes over the controls $u_{0:T-1}$ and recovers the corresponding states $x_{0:T}$ by rollingout the discrete-time dynamics $x_{t+1} = f(x_t, u_t)$. $l(x_t, u_t)$ and $l_f(x_T)$ are the running and terminal costs, respectively. In each iteration, derivatives of cost and dynamics w.r.t to the control sequence are computed around the current solution points in a process called linearization to form a subproblem with quadratic cost and linear constraints. The term iterative LQR [19], [20], or iLQR, generally refers to the Gauss-Newton approximations of the original DDP algorithm that is often more computationally efficient. The iLQR algorithm can also be modified to handle control limits [21], state constraints [13], [15], and contact [22].

As a single-shooting algorithm, iLQR maintains dynamically feasible state trajectories without convergence requirements, is amenable to warm starting, and naturally handles unstable systems, since the rollouts are performed with a feedback policy, all of these features make it appealing as an online controller. However, this method generally assumes the dynamics are smooth and differentiable. For robots with contact, the dynamics are non-smooth and the derivatives become nontrivial to compute [7], [23]. Our empirical results show that the combination of the MuJoCo soft contact model and its finite difference derivative approximation is sufficient for iLQR and, somewhat surprisingly, transfers well to robot hardware despite obvious model mismatch.

B. Whole-body MPC for Legged Robots

Whole-body nonlinear MPC presents challenges to legged robots primarily because of the real-time requirements of reasoning over possible contact modes and computation of the high-degree-of-freedom dynamics and its derivatives.

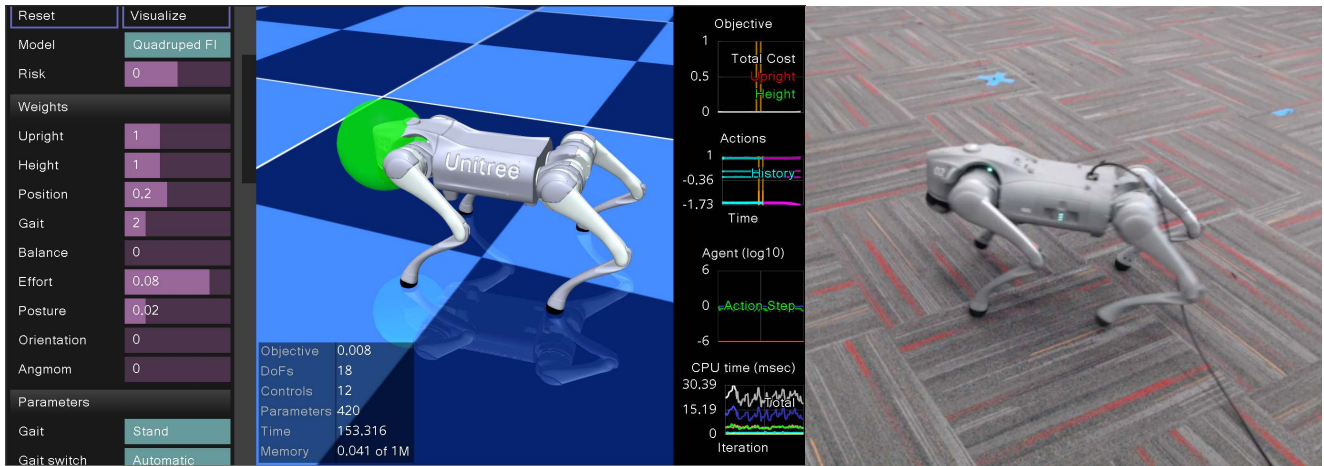


Fig. 3. The MuJoCo MPC GUI for deploying legged robots on hardware. This GUI enables the user to interactively control the real-world robot by changing the target position defined as the green sphere. Additionally, the user can update the planner agent parameters, and observe the simulated states and the real-world robot behaviors in real time.

Traditionally, real-time MPC has been achieved by simplifying models [14], [24] and heuristically choosing the contact modes [24], [25]. Older works at whole-body MPC have been deployed on humanoid robots [26], they generally fall short of the real-time requirements as an online controller, limiting their real-world capabilities. Thanks to advances in computer performance and increasingly mature implementations of dynamics libraries [27], real-time whole-body MPC has become much more computationally trackable on quadruped [28], [29] and humanoid robots [30] in recent years. Another interesting thread of research explores enabling whole-body MPC through GPU parallelization [31], but is so far limited to a single dynamics linearization of the model.

Still, these methods generally require custom modeling of the robot dynamics, non-trivial analytical derivatives of the discontinuous contact dynamics, and custom optimization solvers to run in real time, making these prior works difficult to reproduce and iterate upon. This paper shows that a much simpler and more straightforward approach, modeling the robot using an off-the-shelf simulator and approximating the derivatives via finite differencing, can also be very effective for quadruped and humanoid locomotion *without* specifying contact modes.

C. Open-Source Tooling for Model-Based Control

It is important to acknowledge previous open-source efforts to accelerate model-based control. In particular, MuJoCo MPC [32] implements a variety of derivative-based and derivative-free algorithms for predictive control and shows their effectiveness in simulation. [17], [33] successfully demonstrates the sim-to-real transfer of sampling-based MPC using the MuJoCo dynamics on open-loop stable tasks like quadruped walking. Similar to [17], this work builds on [32] with a focus on sim-to-real of the derivative-based iterative LQR (iLQR) algorithm. Unlike sampling-based MPC, we show that iLQR can successfully tackle inherently open-loop

unstable tasks such as bipedal walking.

Outside the MuJoCo ecosystem, Pinocchio [27] has become a popular toolbox in the community for efficiently computing rigid body dynamics and its derivatives. This dynamics toolbox has also enabled a variety of nonlinear trajectory optimization libraries designed for contact-rich robotics tasks like legged locomotion. For example, OCS2 [34] offers open-source software for quadruped locomotion with fixed contact modes, while Crocoddyle [16] and Aligator [35] offer more versatile control for quadruped and humanoid robots. However, these works do not leverage off-the-shelf simulators widely used in the robotics community and have a steep learning curve, which limits their reach and impact. On the other hand, Drake [5] provides advanced modeling and simulation features for accurate contact and friction dynamics for verification but is typically not fast enough for real-time control.

In comparison to prior projects (Table II-C), our approach bridges the current gap in tooling for model-based robotic control by enabling general controllers for both quadruped and humanoid robots using a popular, off-the-shelf, and fast robotics simulator. An additional advantage for using a mature simulator is the readily available collision detection algorithms that we can leverage during contact-rich planning and control. Additionally, we provide an interactive GUI for real-time control that enables rapid developments of robot behaviors in the real world.

TABLE I
A COMPARISON MODEL-BASED CONTROL TOOLS THROUGH CONTACT

	[5]	[16], [35]	[34]	Ours
Off-the-shelf simulator	✓	✗	✗	✓
Whole-body collision detection	✓	✗	✗	✓
No fixed contact mode	N/A	✗	✗	✓
Real-time control	✗	✓	✓	✓
Interactive GUI	✗	✗	✗	✓

III. ITERATIVE LQR WITH MUJoCo

This section provides a brief description of the MuJoCo soft contact model, the MuJoCo MPC toolbox, the details of the MuJoCo iLQR implementation, and sim-to-real considerations.

A. MuJoCo Soft Contact Model and Derivatives

The MuJoCo physics engine implements a soft contact model that is a convex approximation of the non-convex, discontinuous contact and friction models. While the interpenetration phenomena between objects (for example, the robot's foot and the floor) may be considered physically unrealistic, this convex formulation is fast, efficient, and provides a guaranteed solution, something difficult to do when solving non-convex problems. Additionally, in theory, the soft contact model offers smooth derivatives through contact. While the analytical derivatives are not yet provided, the finite different derivatives can be computed with little additional effort from the original time-stepping simulation problem.

To approximate the model derivatives, we use the forward difference method

$$f'(x) \approx \frac{f(x + \epsilon) - f(x)}{\epsilon} \quad (3)$$

as it requires only one additional simulation evaluation per dimension compared to two in centered difference, where f is an arbitrary function with input x and ϵ is the finite different tolerance.

B. Derivative Computation with MuJoCo

To solve a single iteration iLQR problem from Eq. 1, we take a second-order Taylor expansion and solve the resulting subproblem via Dynamic Programming [18], [20].

We define the cost function as the following:

$$l(x, u) = \sum_i^M w_i \cdot n_i(r_i(x, u)), \quad (4)$$

where r is a residual vector to be reduced when solving the problem, n is the norm function that returns a non-negative scalar, and w is non-negative scalar weight defining the importance of a residual term.

We compute the exact cost gradients:

$$\frac{\partial l}{\partial x} = \sum_i w_i \cdot \frac{\partial n_i}{\partial r} \frac{\partial r_i}{\partial x} \quad (5)$$

$$\frac{\partial l}{\partial u} = \sum_i w_i \cdot \frac{\partial n_i}{\partial r} \frac{\partial r_i}{\partial u} \quad (6)$$

and approximate the cost Hessians as:

$$\frac{\partial^2 l}{\partial x^2} \approx \sum_i w_i \cdot \left(\frac{\partial r_i}{\partial x} \right)^T \frac{\partial^2 n_i}{\partial r^2} \frac{\partial r_i}{\partial x} \quad (7)$$

$$\frac{\partial^2 l}{\partial u^2} \approx \sum_i w_i \cdot \left(\frac{\partial r_i}{\partial u} \right)^T \frac{\partial^2 n_i}{\partial r^2} \frac{\partial r_i}{\partial u} \quad (8)$$

$$\frac{\partial^2 l}{\partial x \partial u} \approx \sum_i w_i \cdot \left(\frac{\partial r_i}{\partial x} \right)^T \frac{\partial^2 n_i}{\partial r^2} \frac{\partial r_i}{\partial u} \quad (9)$$

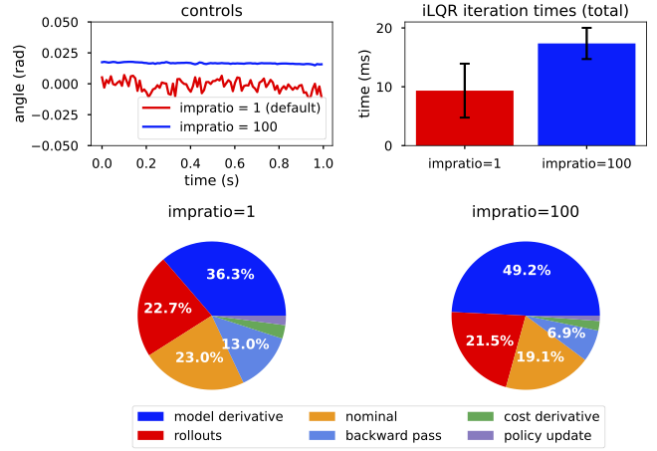


Fig. 4. Top left: thigh joint control trajectories of two $impratio$ contact settings on a quadruped robot standing in place. The default setting results in nonphysical foot slipping and jerky controls (red line) that are potentially dangerous on hardware. Top right: Increased $impratio$ prevents this issue (blue line) but increases compute times (top right). 1 standard deviation confidence interval for each bar is shown in black. Bottom: timing breakdowns of different iLQR components for each setting.

where the derivatives of the norm n (i.e. $\frac{\partial n}{\partial r}$, $\frac{\partial^2 n}{\partial r^2}$) are computed analytically and the Jacobians of the residual r (i.e. $\frac{\partial r}{\partial x}$, $\frac{\partial r}{\partial u}$) are computed via finite difference, Eq. 3.

We take the original nonlinear discrete-time dynamics:

$$x_{t+1} = f(x_t, u_t) \quad (10)$$

and linearize as:

$$\delta x_{k+1} = \frac{\partial f}{\partial x} \delta x_k + \frac{\partial f}{\partial u} \delta u_k \quad (11)$$

where dynamics Jacobians $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial u}$ are computed via finite difference, Eq. 3. Note that because residuals in Eq. 4 are implemented as MuJoCo sensors, we can efficiently compute all the Jacobians $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial u}$, $\frac{\partial r}{\partial x}$, and $\frac{\partial r}{\partial u}$ via a *single* call to the MuJoCo finite difference utilities function [6].

Once all derivatives are computed, we solve the resulting problem using the Riccati-recursion that produces the updated nominal controls u and a time-varying linear feedback policy K , Eq. 2. We perform this update once before returning the current-best nominal trajectories and feedback policy without convergence checks. We then use the previous solution to warm start a new iLQR iteration with the latest state estimation, Fig. 2. Additionally, our interactive GUI allows the user to update the residual terms such as target height, goal positions, etc and adjust the weights assigned to each residual term in real-time on the robot, Fig. 3.

IV. IMPLEMENTATION DETAILS

This section documents several implementation details for successfully deploying real-time MuJoCo iLQR on hardware.

A. Contact Modeling

While the MuJoCo default contact parameters are fast and physically reasonable, they also often lead to contact slipping due to the solver's inability to enforce friction

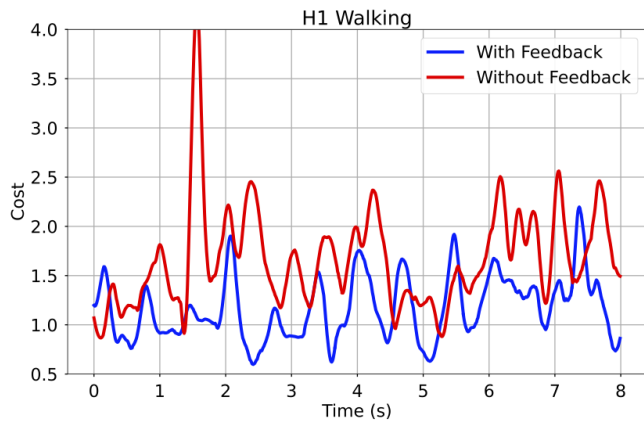


Fig. 5. Cost comparison between the iLQR policy with (blue) and without (red) TV-LQR feedback gains applied to the nominal control sequence on an H1 humanoid robot trotting task on hardware. The TV-LQR policy improves task performance by $\sim 30\%$.

constraints [36]. This slipping dynamics from the planner model results in jerky control trajectories (Fig. 4 top left) that are difficult to execute on the hardware. We address this issue by increasing the `impratio` value, which roughly corresponds to the solver’s ability to trade off sliding versus penetration, from the default 1 to 100. This change also has the added effect of increasing the solve times of the simulation problem and, as a result, the iLQR iteration times. For our quadruped system, the time for a single iLQR iteration increases from ~ 10 ms to ~ 20 ms on a 12th-gen Intel i7 CPU. We do not find the added compute time to be an issue during real-world deployment. Interestingly, the moderate ground penetration from the softness of the contact model does not cause issues in the sim-to-real transfer in our experience.

B. State and Action Representation

We represent the state x of the robot in the standard way, with a floating base position and attitude quaternion, followed by joint angles and corresponding linear, angular, and joint-angle velocities. We include a low-level joint-space PD controller in the dynamics model of the robot such that the inputs to the model that are optimized by iLQR are joint angle references. As a result, our approach achieves direct *whole-body control* of the robot without the model hierarchies commonly seen in traditional model-based MPC algorithms [24]. By leveraging a fast off-the-shelf simulator, our method is simpler and more accessible than methods relying reduced-order models and hierarchical control approaches.

C. Time-Varying Linear Feedback

We compare the real-world performance of the iLQR policy with and without TV-LQR feedback on an H1 humanoid robot trotting-in-place task. The TV-LQR policy improves task performance but only *marginally* compared to directly executing the nominal open-loop control sequence on the robot, resulting in an average of 30.1% improved tracking

performance over an 8 second window, Fig. 5. Note that the cost spike around 1.5s in the red line (without feedback) in Fig. 5 indicates a temporary policy failure recovered with gantry support. The trial with feedback policy does not fail during the same window. Our empirical results are aligned with prior work [37] which showed linear feedback on smoothed dynamics seems unsatisfactory for stabilizing contact-rich plans.

D. Dynamics Derivative Approximation

Computing the dynamics derivatives at every knot point is often the most computationally expensive step in iLQR, Fig. 4. We implement an optional heuristic by skipping dynamics derivative evaluations at some knot points along the planning horizon and instead interpolate from nearby evaluations. This heuristic is motivated by the fact that robot dynamics are approximately linear locally and, therefore, do not change much between nearby knot points [38]. This is implemented as `skip_deriv` in the GUI, where the integer value corresponds to the number of knot points to skip before computing the next derivative. In the tasks we consider in this paper, we find the iLQR update frequency to be sufficient when paired with TV-LQR policy. However, this option can still be beneficial for enabling iLQR in real time on more articulated systems or when compute is limited.

Note that while we choose to use forward difference approximation for computation reasons, we find that the centered difference method does not work on locomotion tasks, even in simulation without real-time requirements. We hypothesize that the derivative information into the contact (this can happen when perturbing the robot state in the negative height direction) is not informative during locomotion but leave further investigation for future work.

V. EXPERIMENTS AND RESULTS

This section presents the interactive GUI setup on robot hardware platforms and a variety of hardware experiments. We start by demonstrating our system on basic quadrupedal locomotion on Go1 and Go2 robots in Sec. V-B. Next, we show that MuJoCo iLQR naturally extends to open-loop unstable tasks like quadruped walking on two legs in Sec. V-C. Finally, we deploy our system on a human-sized Unitree H1 humanoid robot in Sec. V-D. Our open-source software and experiment videos are available at:

Omitted for Anonymous Review

A. Interactive GUI for Real-World Legged Robots

We interface with the MuJoCo MPC GUI and iLQR planner via Python for robot hardware deployment. The interface takes the latest state estimation and receives the current control. The state estimation is computed by fusing Optitrack MoCap [39] position and attitude measurements at 100Hz via ROS and the robot joint angle position and velocity measurements at 500Hz. The floating base linear and angular velocities in the body frame are calculated by applying a low-pass filter to finite difference values of position and attitude measurements. The planner’s actions are communicated to

the robot via Unitree SDK for the Go1 robot and Unitree SDK 2 for the Go2 and H1 robots. All robot controls are represented as joint targets and tracked using Unitree internal low-level PD controllers. We compute the MPC policies on a desktop equipped with a 13th generation Intel i9 CPU chip with 20 cores, a different CPU from earlier in the paper. We update the iLQR policy at $\sim 50\text{Hz}$ and use the TV-LQR policy to stabilize the robots between solves at $\sim 300\text{ Hz}$. Unless otherwise noted, we use a default prediction horizon of 0.35s and discretize the dynamics at 100Hz in the iLQR planner. Note that, while we use a desktop-level CPU in our experiments, we see similar policy update frequencies from recent top-end laptop Arm-based CPUs such as the M-series chips from Apple, but leave deployment for future work.

Since the MPC planners are implemented in C++ and update asynchronously, using the Python interface does not affect the planning frequency. While we observe a $\sim 3\text{ms}$ message passing overhead from the Python API, the planner is robust to this unmodeled delay for the tasks presented in this paper.

B. Quadruped Locomotion

We first validate our system using a simple quadruped locomotion task. The task is designed for the robot to follow a reference gait and walk to the target locations. We successfully deploy the policy to the Unitree Go1 and Go2 robots with 12 degrees of freedom (DoF) joint actuators. The GUI allows the user to move the target interactively and the real-world robot can follow the virtual target, as illustrated in Fig. 3. Note that while we keep the whole-body collision geometry in the planner model, generally only the 4 spherical contact points are active during the locomotion tasks considered.

The residual terms for the locomotion task include tracking position and orientation while keeping balance and maintaining a desired torso height. Control efforts are penalized to minimize energy usage. We also include a nominal gait pattern in the residual, but note that this is not implemented as a constraint. As a result, the iLQR solver is free to discover new contact modes if they reduce the overall cost. More detailed descriptions of the residual terms can be found in Tab. V-C. Bipedal locomotion and humanoid tasks follow a similar structure.

C. Bipedal Locomotion for Quadruped Robots

Next, we show that the iLQR policies naturally handles open-loop unstable tasks, such as a quadruped walking on two legs (Fig. 1). In comparison, MPPI [17] can tackle locomotion tasks that are stable, but fails under unstable dynamics such as during bipedal walking. In Fig. 1, we successfully enable a quadruped robot to walk gracefully on its back legs alone while using front legs to main balance and getting up to a hand stand pose from an initial quadruped configuration with the MuJoCo iLQR policy.

D. Humanoid Locomotion

Finally, we deploy the iLQR policy on an H1 full-sized humanoid robot to track a periodic trotting gait, as shown

TABLE II
TASK RESIDUAL FOR QUADRUPEDAL AND HUMANOID LOCOMOTION

Term	Description
Upright	Keeps torso orientation upright (z-axis pointing up)
Height	Controls torso height relative to average foot position
Position	Tracks head position to target location (x, y, z)
Gait	Controls foot lifting patterns during gait cycles (one per foot)
Balance	Keeps capture point within support polygon (x, y components)
Effort	Penalizes actuator forces to minimize energy consumption
Posture	Keeps joints near home configuration
Yaw	Controls heading direction (x, y components of heading vector)
Angular	Controls rotational dynamics (x, y, z components)

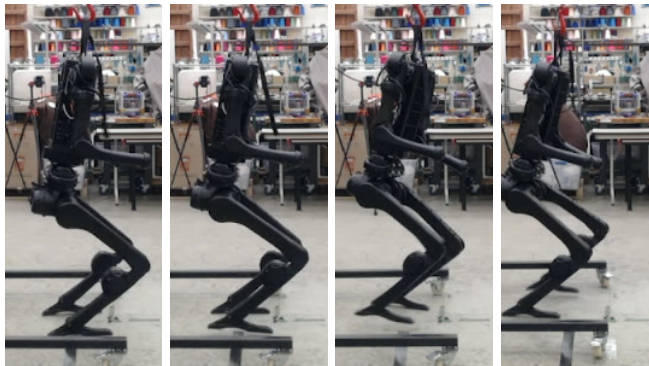


Fig. 6. A Unitree H1 robot trotting in place with MuJoCo iLQR.

in Fig. 6. Similar to the quadruped locomotion task in Fig. 3, we ask the robot to walk to a position defined as the green sphere. To achieve real-time control on a more complex system, we disable collision checking in the robot body other than two spherical contact points on each foot. Additionally, we disable the DoFs in the robot upper body equipped with only relative encoders that provide unreliable joint angle estimates since they are not critical for locomotion. While this reduces the iLQR computation time, we do not believe it to be necessary for the success of the task. Overall, the system has 10 DoF joint actuators and 4 contact points. We use a prediction horizon of 0.5s.

VI. CONCLUSIONS AND FUTURE WORK

We present a very simple but surprisingly capable approach to whole-body MPC of quadrupeds and humanoid robots using MuJoCo. We hope that the successful hardware validation of this baseline method by leveraging a widely adopted physics engine can encourage researchers to leverage existing tooling for model-based control research. Despite the promising results, several key limitations remain for future work.

First, reliable state estimation remains a key challenge for model-based planning and control of legged robots. For this reason, many researchers prefer the RL paradigm which easily supports learning control policies directly from a history of sensor measurements. Our current system relies

on marker-based motion capture to obtain good robot position measurements. Future work should develop easy-to-use tooling for full-state estimation from the robot's onboard sensors alone to enable our robots to walk outside controlled laboratory environments. Second, the community also needs additional tooling for rigorous system identification of the robot's joint actuation and contact dynamics to overcome the significant sim-to-real gap.

There are also several fundamental limitations of the iLQR algorithm. First, iLQR struggles with contact mode exploration. As a second-order derivative-based local planner, iLQR excels when the reference contact model schedule is provided in the task specification (cost function), which is typically the case for locomotion. However, for more contact-rich whole-body loco-manipulation tasks, derivative-free sampling-based methods have shown much more promise at discovering useful contact modes without prespecification [17]. Furthermore, iLQR rollouts and backward passes are both fundamentally serial operations. As computers become more parallelizable given the rise of multi-core CPUs and massively parallel GPUs, research on MPC algorithms that can leverage this computation paradigm becomes increasingly important. Finally, iLQR suffers from other issues tied to its single-shooting nature, such as sensitivity to the initial guess, numerical instability, and poor convergence over longer horizons. Future work should extend the current MPC libraries to make multiple-shooting and collocation methods more accessible to the community.

ACKNOWLEDGMENT

The authors thank Arun Bishop, Swaminathan Gurusamy, and Shuo Yang for insightful discussions and feedback. This work is supported by a Google award.

REFERENCES

- [1] R. Deits, S. Kuindersma, M. P. Kelly, T. Koolen, Y. Abe, and B. Stephens, "Robot movement and online trajectory optimization," U.S. patent 11 833 680, 12 5, 2023, uS Patent 11,833,680. [Online]. Available: <https://patents.google.com/patent/US11833680B2/>
- [2] R. Grandia, F. Jenelten, S. Yang, F. Farshidian, and M. Hutter, "Perceptive locomotion through nonlinear model-predictive control," *IEEE Transactions on Robotics*, vol. 39, no. 5, pp. 3402–3421, 2023.
- [3] J. Lee, M. Bjelonic, A. Reske, L. Wellhausen, T. Miki, and M. Hutter, "Learning robust autonomous navigation and locomotion for wheeled-legged robots," *Science Robotics*, vol. 9, no. 89, p. eadi9641, 4 2024.
- [4] D. Hoeller, N. Rudin, D. Sako, and M. Hutter, "Anymal parkour: Learning agile navigation for quadrupedal robots," *Science Robotics*, vol. 9, no. 88, p. eadi7566, 3 2024.
- [5] R. Tedrake and the Drake Development Team, "Drake: Model-based design and verification for robotics," 2019. [Online]. Available: <https://drake.mit.edu>
- [6] E. Todorov, T. Erez, and Y. Tassa, "Mujoco: A physics engine for model-based control," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 5026–5033.
- [7] T. Howell, S. Le Cleac'h, J. Bruedigam, Z. Kolter, M. Schwager, and Z. Manchester, "Dojo: A differentiable physics engine for robotics," *arXiv preprint arXiv:2203.00806*, 2022. [Online]. Available: <https://arxiv.org/abs/2203.00806>
- [8] M. Macklin, "Warp: A high-performance python framework for gpu simulation and graphics," <https://github.com/nvidia/warp>, March 2022, nVIDIA GPU Technology Conference (GTC).
- [9] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, "Learning to walk in minutes using massively parallel deep reinforcement learning," 2022. [Online]. Available: <https://arxiv.org/abs/2109.11978>
- [10] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems* 32, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds. Curran Associates, Inc., 2019, pp. 8024–8035.
- [11] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, "JAX: composable transformations of Python+NumPy programs," 2018. [Online]. Available: <http://github.com/google/jax>
- [12] J. Carpentier, L. Montaut, and Q. L. Lidec, "From compliant to rigid contact simulation: a unified and efficient approach," 2024. [Online]. Available: <https://arxiv.org/abs/2405.17020>
- [13] T. A. Howell, B. E. Jackson, and Z. Manchester, "Altro: A fast solver for constrained trajectory optimization," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.
- [14] S. Le Cleac'h, T. A. Howell, S. Yang, C.-Y. Lee, J. Zhang, A. Bishop, M. Schwager, and Z. Manchester, "Fast contact-implicit model predictive control," *IEEE Transactions on Robotics*, vol. 40, pp. 1617–1629, 2024.
- [15] W. Jallet, A. Bambade, E. Arlaud, S. El-Kazdadi, N. Mansard, and J. Carpentier, "Proxddp: Proximal constrained trajectory optimization," *arXiv preprint*, 2023, preprint submitted on 8 Dec 2023.
- [16] C. Mastalli, R. Budhiraja, W. Merkt, G. Saurel, B. Hammoud, M. Naveau, J. Carpentier, L. Righetti, S. Vijayakumar, and N. Mansard, "Crocodyl: An Efficient and Versatile Framework for Multi-Contact Optimal Control," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020.
- [17] J. Alvarez-Padilla, J. Z. Zhang, S. Kwok, J. M. Dolan, and Z. Manchester, "Real-time whole-body control of legged robots with model-predictive path integral control," *arXiv preprint arXiv:2409.10469*, 2024, available at: <https://arxiv.org/abs/2409.10469>.
- [18] D. H. Jacobson and D. Q. Mayne, *Differential Dynamic Programming*. Elsevier, 1970.
- [19] Y. Tassa, T. Erez, and E. Todorov, "Synthesis and stabilization of complex behaviors through online trajectory optimization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 4906–4913.
- [20] W. Li and E. Todorov, "Iterative linear quadratic regulator design for nonlinear biological movement systems," in *Proceedings of the 1st International Conference on Informatics in Control, Automation and Robotics*, 2004.
- [21] Y. Tassa, N. Mansard, and E. Todorov, "Control-limited differential dynamic programming," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 1168–1175.
- [22] N. J. Kong, C. Li, G. Council, and A. M. Johnson, "Hybrid ilqr model predictive control for contact implicit stabilization on legged robots," *IEEE Transactions on Robotics*, vol. 39, no. 6, pp. 4712–4727, 2023.
- [23] N. J. Kong, J. Joe Payne, J. Zhu, and A. M. Johnson, "Saltation matrices: The essential tool for linearizing hybrid dynamical systems," *Proceedings of the IEEE*, vol. 112, no. 6, pp. 585–608, 2024.
- [24] G. Bledt, M. J. Powell, B. Katz, J. Di Carlo, P. M. Wensing, and S. Kim, "Mit cheetah 3: Design and control of a robust, dynamic quadruped robot," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 2245–2252.
- [25] M. H. Raibert and H. B. Brown, "Dynamically stable legged locomotion," 1983.
- [26] J. Koenemann, A. Del Prete, Y. Tassa, E. Todorov, O. Stasse, M. Bennewitz, and N. Mansard, "Whole-body model-predictive control applied to the hrp-2 humanoid," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 3346–3351.
- [27] J. Carpentier, G. Saurel, G. Buondonno, J. Mirabel, F. Lamiroux, O. Stasse, and N. Mansard, "The pinocchio c++ library – a fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives," in *IEEE International Symposium on System Integrations (SII)*, 2019.
- [28] G. Kim, D. Kang, J.-H. Kim, S. Hong, and H.-W. Park, "Contact-implicit model predictive control: Controlling diverse quadruped motions without pre-planned contact modes or trajectories," *The*

- International Journal of Robotics Research*, Oct. 2024. [Online]. Available: <http://dx.doi.org/10.1177/02783649241273645>
- [29] G. Lunardi, T. Corbères, C. Mastalli, N. Mansard, T. Flayols, S. Tonneau, and A. D. Prete, "Reference-free model predictive control for quadrupedal locomotion," *IEEE Access*, vol. 12, pp. 689–698, 2024.
- [30] C. Khazoom, S. Hong, M. Chignoli, E. Stanger-Jones, and S. Kim, "Tailoring solution accuracy for fast whole-body model predictive control of legged robots," *IEEE Robotics and Automation Letters*, vol. 9, no. 12, pp. 11 074–11 081, 2024.
- [31] A. L. Bishop, J. Z. Zhang, S. Gurumurthy, K. Tracy, and Z. Manchester, "Relu-qp: A gpu-accelerated quadratic programming solver for model-predictive control," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 13 285–13 292.
- [32] T. Howell, N. Gileadi, S. Tunyasuvunakool, K. Zakka, T. Erez, and Y. Tassa, "Predictive Sampling: Real-time Behaviour Synthesis with MuJoCo," dec 2022. [Online]. Available: <https://arxiv.org/abs/2212.00541>
- [33] H. Xue, C. Pan, Z. Yi, G. Qu, and G. Shi, "Full-order sampling-based mpc for torque-level locomotion control via diffusion-style annealing," 2024. [Online]. Available: <https://arxiv.org/abs/2409.15610>
- [34] F. Farshidian *et al.*, "OCS2: An open source library for optimal control of switched systems," [Online]. Available: <https://github.com/leggedrobotics/ocs2>.
- [35] W. Jallet, A. Bambade, S. El Kazdadi, C. Justin, and M. Nicolas, "aligator." [Online]. Available: <https://github.com/Simple-Robotics/aligator>
- [36] MuJoCo Team, "Preventing slip - MuJoCo documentation," <https://mujoco.readthedocs.io/en/3.2.7/modeling.html#preventing-slip>, 2023, accessed: 2025-02-25.
- [37] Y. Shirai, T. Zhao, H. J. T. Suh, H. Zhu, X. Ni, J. Wang, M. Simchowitz, and T. Pang, "Is linear feedback on smoothed dynamics sufficient for stabilizing contact-rich plans?" 2024. [Online]. Available: <https://arxiv.org/abs/2411.06542>
- [38] D. Han, H. Eom, J. Noh, and J. S. Shin, "Data-guided model predictive control based on smoothed contact dynamics," *Computer Graphics Forum*, vol. 35, no. 2, pp. 533–542, May 2016, (formerly Sung Yong Shin).
- [39] "Optitrack motion capture system," 2025, available at: <https://www.optitrack.com>.