Berkeley Humanoid: A Research Platform for Learning-based Control

Qiayuan Liao, Bike Zhang, Xuanyu Huang, Xiaoyu Huang, Zhongyu Li, Koushil Sreenath

***Abstract*— We introduce Berkeley Humanoid, a reliable and low-cost mid-scale humanoid research platform for learning-based control. Our lightweight, in-house-built robot is designed specifically for learning algorithms with accurate simulation, low simulation complexity, anthropomorphic motion, and high reliability against falls. The narrow sim-to-real gap enables agile and robust locomotion across various terrains in outdoor environments, achieved with a simple reinforcement learning controller using light domain randomization. Furthermore, we demonstrate the robot traversing for hundreds of meters, walking on a steep unpaved trail, and hopping with single and double legs as a testimony to its high performance in dynamic walking. Capable of omnidirectional locomotion and withstanding large perturbations with a compact setup, our system aims for rapid sim-to-real deployment of learning-based humanoid systems. Please check our website `https://berkeley-humanoid.com/` and code `https://github.com/HybridRobotics/isaac_berkeley_humanoid/`.**

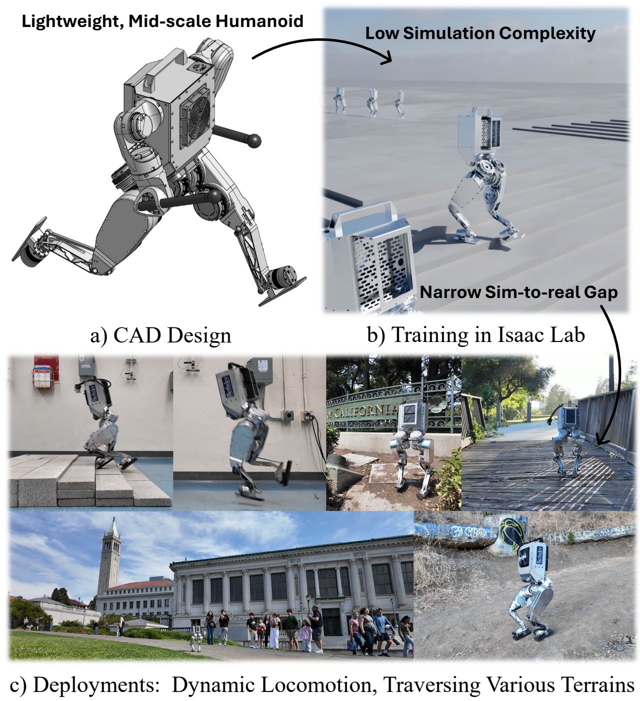

Fig. 1. Design, training, and sim-to-real deployment of our custom-built humanoid with learning-based control. With designs tailored for learning-based control, the robot is able to traverse a diverse set of outdoor terrains robustly, as well as climb stairs and perform single-leg hopping. Please check `https://youtu.be/8pR1HE-wMHw?si=1FNtxJhIay_Vdpur` for details.

I. Introduction

There is a strong need for mid-scale humanoid robots designed for effective deployment of learning-based policies, with the ability to perform highly dynamic motions, while being inexpensive and robust to falls and failures. Most current bipedal and humanoid robots [1]–[5] are larger, unsafe, and require a team of people to operate. Mid-sized robots [6]–[12], in comparison, are light-weight and easier to handle, requiring as few as one operator. They are more manageable to work with when experimenting with highly dynamic motions. Falls typically do not damage the environment or the robot, making the setup more forgiving. These robots can be deployed in cramped lab spaces, and creating rough terrain for testing is simple due to their limited ground clearance. Furthermore, their experiment-friendly nature and relatively low cost make it easier to scale up and conduct multi-agent research, especially in academia settings.

However, designing a mid-scale humanoid robot posts distinctive challenges. Mechanical design is more difficult due to limited space for housing components such as motors, sensors, and wiring, necessitating the use of compact power-dense actuators that are often very expensive or not available off-the-shelf [7], [13], [14]. Integrating all components in a compact volume without sacrificing performance or cost is difficult. Furthermore, as a more handy platform for testing highly dynamic tasks [9], it requires an even higher torque-to-weight ratio and greater impact reliability.

Control for mid-scale humanoids is also challenging, due to the instability from a low center of gravity, high natural frequency, and high sensitivity to external perturbations. Their lower mass and inertia make these robots more agile but also more sensitive: even small forces produce large motions. The shorter legs lead to a reduced stride length, requiring higher frequency leg movements to adjust foot placement rapidly, demanding precise coordination and control. Furthermore, learning-based policies face substantial sim-to-real gaps, particularly in such rapid and dynamic movements for controlling these robots. These characteristics mean that the actuation of the joints must be quick and accurate to support high-frequency motions, and the control policies need to be exceptionally precise and robust to match the short-time constants of the dynamics.

As shown in Figure 1, we propose to custom-build a mid-scale humanoid platform with a special emphasis on accommodating and facilitating learning-based control. To achieve accurate, robust, and agile control, we leverage a learning-based algorithm and focus on narrowing the sim-to-real gap with adequate hardware design. Learning-based algorithms enable us to leverage cheaper and noisier sensors to cut down costs. To optimize for simulation performance and accuracy while achieving high-performance actuation, we utilize custom modular actuators with simple, integrated transmission, hollow shafts, and EtherCAT field bus.

Our contributions are summarized as follows: (i) We present a reliable, low-cost, mid-scale humanoid research platform focusing on narrowing the sim-to-real gap with designs tailored for learning-based control. (ii) We demonstrate that our design choices facilitate us to be able to use a minimally composed control policy to perform dynamic and robust locomotion on complex terrain, notably the challenging task of walking on a steep, narrow, and unpaved trail. (iii) The codebase for policy training with the recent Isaac Lab release is open-sourced to support future humanoid research.

II. Related Work

Humanoid Design. As shown in Table I, we categorize humanoid robots into three primary sizes: (a) full-scale, which corresponds to the size of an average adult, (b) mid-scale, comparable to the size of a child, and (c) miniature, which refers to tiny non-human-sized robots. Full-scale humanoid or biped research platforms typically have a large weight and use high gear ratio Harmonic Drive actuators [1]–[3]. These platforms are primarily enabling basic walking and arm manipulation. Some platforms utilize Cycloidal Drive Actuators for high-load joints, combined with spring and linkage designs [4], [16]. This setup simplifies the design of reduced-order, step-to-step model-based controllers. However, for more recent learning-based algorithms, these designs optimized for model-based control advertedly affect training and deployment. In comparison, more lightweight platforms featuring proprioceptive actuators [24] and primarily dummy arms have been recently developed, capable of performing more dynamic tasks [5], [15]. Besides full-scale humanoids, mid-scale or miniature humanoid research platforms have gained popularity over the recent years [10], [17]–[19], [25]. All of these platforms opt for proprioceptive actuators and are designed for better dynamic performance, but most of them lack fully articulated legs. On the other hand, new humanoid robots from some companies deviate from proprioceptive actuators: Tesla Optimus, for example, uses linear actuators and harmonic drives, some with load cells for force control, and features complex transmissions between joints and actuators [26]. Boston Dynamics' hydraulic Atlas [27] excels in highly dynamic tasks, and the newly released electric Atlas [27] showcases simplified joint designs with a large range of motion. The robots from companies are well-designed and well-tested, but unfortunately, most of them are not available for researchers in labs or do not provide access to modify or improve the low-level system.

TABLE I

COMPARISON OF EXISTING ELECTRIC HUMANOID LOCOMOTION RESEARCH PLATFORMS.

Robot	Size[a]	Avg. Leg [b] Len.(m)	Leg Dof	Weight (kg)	Price (USD)	Actuator[c] Type	Max HFE Tor.(Nm)	Max KFE Tor.(Nm)	Transmission Complexity	T/F Sensor
TORO [1]	F	∼0.4	6	76.4	-	H	100	130	++	Joint
LOLA [2]	F	∼0.44	6	68.2	-	H	370	390	+++	Feet
WALK-MAN [3]	F	∼0.38	6	132	-	H	270-400	270-400	++	Feet
Unitree H1 [5]	F	∼0.4	5	47	90K	P	270	360	+	✗
Digit [4]	F	∼0.5	6	50	250K	C, H	200	230	+++	✗
ARTEMIS [15]	F	∼0.38	5	37	-	P	250	250	+	Feet
Cassie [16]	F	∼0.5	5	35	250K	C, H	195	195	+++	✗
MIT [17]	M	∼0.28	5	24	-	P	72	144	+	✗
Unitree G1 [18]	M	∼0.3	6	35	16K	P	88	139	+	✗
HECTOR [19]	M	∼0.22	5	16	-	P	33.5	51.9	+	✗
iCub [20]	M	∼0.2	6	24	300K	H	40	40	++++	Feet
BRUCE [10]	T	∼0.17	5	3.3	6.5K	P	10.5	10.5	+	✗
NAO [21]	T	∼0.15	6	4.5	14K	S	1.61	1.61	+	Feet
DARwIn-OP [22]	T	∼0.09	6	2.8	-	S	2.35	2.35	+	Feet
Surena-Min [23]	T	∼0.085	6	3.3	-	S	3.1	7.3	+	✗
Ours	M	**∼0.2**	**6**	**16**[d]	**10K**[e]	**P**	**62.6**	**81.1**	**+**	✗

[a] F, M, and T represent full-scale, mid-scale, and miniature, respectively.
[b] Average length of thigh and calf.
[c] H, P, C, and S represent Harmonic Drive, Planetary, Cycloidal Drive, and Servo Motor with a high reduction ratio, respectively.
[d] Without arms. The estimated weight of two 4 DoF arms is 6kg, the total weight will be 22kg.
[e] Without arms. The estimated cost of two 4 DoF arms is 5K USD, the total non-profit cost will be 15K USD.

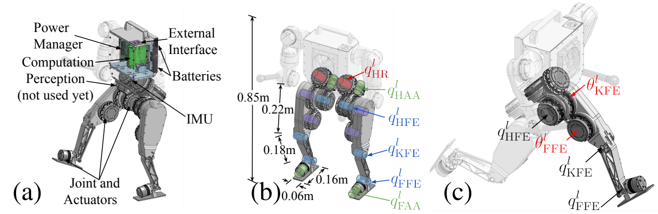

Fig. 2. Overview of design: (a) main components, (b) joints and key dimensions, (c) key actuators and joints of the left leg.

Humanoid Control. Humanoid control is a challenging problem in the robotics field. Utilizing control approaches ranging from heuristic-based methods to model-based control, humanoids have been equipped with stable movement abilities [28]–[32]. Recently, learning-based approaches demonstrate promising capabilities for humanoid robots, ranging from locomotion [33]–[37] to manipulation [38]–[41]. Dynamic humanoid locomotion has been demonstrated such as walking on rough terrain [42], [43], resisting large disturbances [44], running [45], and parkour [46]. These works often utilize complex neural networks and training pipelines for high expressiveness or require a history of state-action pairs for online adaptation, reducing the sim-to-real gap in deployment. In comparison, performing dynamic motions with a simple algorithm and architecture remains challenging. Furthermore, prior works often include wide distributions of domain randomization due to the higher robustness requirement to counteract the imprecise models with complex transmissions. However, excessive randomization may hinder successful policy learning or lead to exceedingly conservative policies [47]. Despite the progress in full-scale humanoid robots, learning control policies for smaller-scale humanoids pose different challenges due to the shorter-legged design as discussed in Sec. I. Prior works, such as teaching miniature humanoid robots to play soccer, address these challenges with large flat foot designs and servo motors [48], resulting in limited dynamic motion capabilities. In contrast, our design uses smaller flat feet and more powerful actuators, enabling more dynamic motions but presenting greater control challenges.

III. Design for Learning-based Control

We introduce our humanoid robot design by providing an overview of the system design and detailing specific design considerations tailored for learning-based control algorithms.

A. System Overview

Berkeley Humanoid is a 16 kg, fully electric drive mid-scale robot for humanoid research. The main component is shown in Figure 2(a). The robot has a torso and two 6 DoF legs, with a thigh length of 220 mm, a calf length of 180 mm, and a total height of 0.85 m in a nominal standing configuration, resembling a 5-year-old child in body size.

TABLE II

CUSTOM ACTUATOR SPECIFICATIONS.

Actuator	5013	8513	8518	10413
Mass (g)	251	756	856	1011
Gear Ratio	9:1	9:1	9:1	9:1
Hollow Shaft	✗	✓	✓	✓
Diameter × Thickness (mm)	54.6 × 53	104 × 50	104 × 55	123 × 50
Peak Torque (Nm)	9.7	45.3	62.6	81.1
Sustained Torque (Nm)	4.59	18.9	26.1	34.2
Max. Speed at 48V (rad/s)	83.7	40.7	29	27.9
Max. Power (W)	220	570	730	890
Rotor Inertia (kgm^2)	6.1e-6	6.9e-5	9.4e-5	1.5e-4
Joint	FAA	FFE, HR, HAA	HFE	KFE

Inside the torso, an Intel NUC, a power management board, and a cheap cellphone-level IMU sensor are installed along with 2 battery packs in a protected compartment. Each leg is equipped with 6 actuators, most of which are directly attached to the link and act as a joint. Two 4-DoF arms were designed but left out as we focus on the locomotion abilities in this work. To adapt to different torque requirements on each joint, we built 4 types of actuators and 2 types of motor drivers, as shown in Table II. These high-performance actuators allow our robot to perform highly dynamic maneuvers.

For the communication system, to achieve better accuracy with minimum latency, we opt for high-bandwidth EtherCAT protocol. We develop custom EtherCAT slaves for both motor drivers and the IMU. The onboard PC runs the custom EtherCAT master and communicates with the peripherals at frequencies ranging from 1 kHz to 4 kHz. USB and Ethernet connection are also supported for the perceptive sensors, such as RGBD cameras, lidar.

B. Design Considerations

Faster and More Precise Simulation. Learning-based locomotion algorithms mostly rely on model-free reinforcement learning with parallel simulation, making simulation cost-critical. In many instances of robot design for model-based control, transmission linkages with unilateral springs can reduce motor load and limb weight, absorb impacts, and be modeled with simplified models, but they also introduce complex dynamical equations that are difficult to simulate. To optimize efficiency, we eliminate such flexible components and closed kinematic chains, opting for direct actuator-joint transmissions. Leveraging cross-roller bearings, the actuators can be directly mounted at the joints, and their rotor inertia can be conveniently modeled by adjusting the joint mass matrix directly. One exception is the Foot Flexion/Extension (FFE) joint, where a linkage transmission is used to reduce leg inertia. However, this linkage is designed with a linear joint-actuator mapping, allowing us to treat it as a joint in the simulation similar to other joints. Furthermore, for better simulation, we opt for planetary gearboxes with a small gear ratio, which helps reduce overall friction and minimizes the quadratic amplification of cogging torque, friction, and reflected inertia errors. In addition, low latency (0.5-2 ms) with EtherCAT avoids the need to simulate latency. These optimizations enable over 90,000 simulation steps per second on an NVIDIA A4500 GPU.

TABLE III

COST OF EACH COMPONENT IN SMALL QUANTITY PRODUCTION.

Module	Actuator				Sensor	Misc		Off-the-shelf		Total
	5013	8513	8518	10413	IMU	Torso	Leg	PC	Battery	
Cost (USD)	422	570	639	676	50	410	974	347	153	9955
Quantity	2	6	2	2	1	1	2	1	2	-

Reliability and Low Cost. Our goal is to build a durable and low-cost robot. To achieve reliability, we use high-performance materials like 7075 and 6061 aluminum for the main structure and tool steel for the gearbox and linkage, ensuring the robot withstands heavy impacts while remaining lightweight, as opposed to [11], [49]–[52]. To improve cable durability, which is crucial for humanoids performing agile motions, we use hollow shaft designs, routing power, and communication cables through the joint axis to minimize tearing from movement. Reducing the number of sensors not only lowers the cost but also simplifies the structure and improves reliability. Eliminating the need for foot contact force sensors, our proprioceptive actuators enable us to leverage a generalized momentum observer [53] to accurately estimate foot contact forces without strain gauges as additional sensors. A significant advantage of learning algorithms is their robustness against sensor inaccuracies. Thus, we opt for cheaper sensors, such as a cellphone-level IMU ($1 for sensor IC, $50 with interface board) instead of the typical $1,000-priced units in prior robots [4], [7], [16], [54]. These optimizations lower the robot cost to around $10,000 (details in Table III), with further reductions possible through scaled production. Only the computers and batteries are commercially sourced for performance and safety.

Experiment-Friendly. Traditional full-sized humanoid robots are often heavier than adults and require multiple operators to handle, making experiments cumbersome and dangerous, especially with high-torque actuators. Unlike model-based methods, learning-based control especially requires extensive hardware testing to determine its effectiveness in real-world settings. To address this, we scale our robot down to a mid-scale humanoid and reduce its weight to merely 16 kg. This allows a single operator to conduct experiments indoors, and an additional cameraman outdoors, for all of the tasks including commanding, data collection, and resetting the robot. All experiments in this work are conducted with this setup.

Anthropomorphic Design. An important advantage of learning-based algorithms is their ability to directly imitate human motions without the need for tedious trajectory optimization; however, to fully leverage this capability, an anthropomorphic robot design is necessary. Our robot achieves this by incorporating 6 DoFs per leg, mirroring the dominant DoFs in human legs [55] and a common 6 DoF contact wrench in foot surface contact [56]. Compared to [5], [15]–[17], [19], the additional actuation in the ankle's roll direction not only improves stability in difficult poses, such as balancing on one foot or manipulating distant objects, but also enables learning from human motions directly. Additionally, joint limits are aligned with human physical constraints, protecting the hardware while providing sufficient range for natural motion imitation.

IV. Minimally Composed Learning Controller

With a humanoid platform designed for learning-based control, we are able to achieve robust and agile locomotion with a minimally composed RL controller. First, we outline the RL controller's design. Then, we explain how we reduce the sim-to-real gap based on the hardware platform.

A. Reinforcement Learning Formulation

We formulate our tasks as Markov Decision Processes (MDPs) and leverage RL to solve them thanks to its promising performance in humanoid control. We create a minimally composed learning-based controller by formulating the MDP with minimal observations. Specifically, we only use immediate state feedback as the actor input, without formulating a short or long history [42], [57] or teacher-student training [58], [59] to estimate environment parameters. Similarly, we opt out of pre-defined phase signals [33] or reference motion [57] to reduce human biases. The immediate state feedback includes raw proprioceptive readings (base angular velocity, projected gravity vector, joint positions, velocities), base linear velocity $\mathbf{v}$ from a state estimator [60], [61], velocity commands $\mathbf{v}^c_{x,y}$ and ω^c_z, and the previous action. Likewise, the action is the desired joint positions, which are converted into torques by a PD controller on the motor driver. The reward function has four components: tracking desired velocities, smoothing by penalizing undesired velocities and actions, regularizing joint positions to avoid aggressive motions, and gait quality terms for longer air time and preventing slippage.

We design the architecture of the actor-critic with the most basic multilayer perceptron (MLP) networks. Specifically, each network has hidden sizes of $[512, 256, 128]$ neurons and ELU activation. The policy is optimized via PPO [62] and trained in Isaac Lab [63]. The RL policy executes at 50 Hz, the state estimator at 1 kHz, and the PD controller at 25 kHz.

This minimally-composed RL controller validates the efficacy of our hardware design for learning-based control. Without the ability to do online system identification or reference motion guidance, our policy relies solely on the synergy of the hardware and simulation setup to achieve a narrow sim-to-real gap, ensuring that the robust and agile locomotion performance in training can be fully demonstrated on the real-world robot.

B. Closing the Sim-to-Real Gap

Hardware Side. We focus on closing the sim-to-real gap through hardware design choices. The main factors of the sim-to-real gap, aside from sensor noise, are modeling errors and command execution rate, accuracy, and delay [64]–[66]. To reduce modeling errors, we choose the simulator-friendly design strategies as introduced in Sec. III. To improve command execution, we employ real-time Linux with EtherCAT, which leads to a precise execution rate and low communication latency; transparent proprioceptive actuator dynamics with high bandwidth torque control, so that the commanded torque is accurately tracked and has negligible actuator dynamics. All of these reduce the discrepancy between the hardware and the simulated dynamics.

Design-enabled Accurate Domain Randomization. While most of the learning controllers rely on domain randomization, extensive domain randomization slows down training and results in conservative policies [47]. To avoid this while still preserving a robust policy, in this work, we leverage a different approach aimed at providing accurate domain randomizations given the hardware design. For a humanoid robot performing locomotion tasks, we identify two sources of uncertainties: uncertainty in the robot physics property, e.g., the mass of each link, and that in performing tasks, e.g., contact with the environment.

For hardware uncertainty, our detailed design allows us to obtain a small and accurate range of parameter variations. Specifically, we use CAD to retrieve accurate mechanical parameters like rotor inertia and conduct simple experiments to characterize the friction of each actuator separately. This demonstrates the benefits of an in-house-built robot, as obtaining such detailed hardware parameters for commercial robots would be difficult.

For uncertainty in contact with the environment, we apply a wide range of domain randomization to cover as many real-world environment conditions as possible. This includes ground friction, restitution, and external perturbation forces from obstacles and unstable ground conditions.

Unlike previous work [57], [65], [67], we opt not to randomize properties that cannot be identified in these two categories, such as a general "motor strength" ratio or PD gains, which were often used as a "lazy approach" to approximate actuation uncertainties. However, because it is hard to accurately analyze the range of uncertainties with PD approximation, prior works rely on heuristics, which can lead to unnecessarily large ranges of domain randomization, which we aim to avoid.

As we will show later, with design-enabled accurate domain randomizations, we can achieve robust and agile locomotion skills when zero-shot transferring to robot hardware, even with a minimally-composed RL controller.

V. Experimental Validation

In our experiments, we aim to validate how our humanoid design facilitates learning locomotion control from three aspects: (1) The effectiveness of our minimally-composed RL controller in humanoid locomotion tasks. (2) The sim-to-real gap for the minimal RL policy with our adequate hardware design. (3) The hardware reliability of the robot.

A. Locomotion Performance

Compared to previous works leveraging advanced architectures, in this work, we emphasize how our adequate hardware design for learning-based algorithms facilitates us to achieve robust and agile locomotion performance with a basic RL controller introduced in Sec. IV.

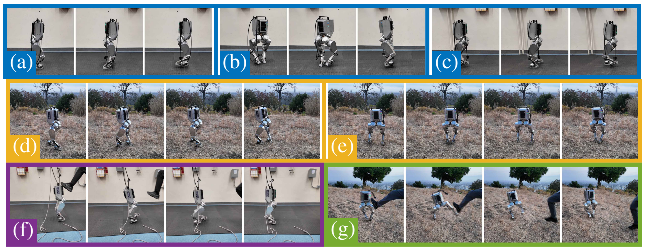

Fig. 3. Omnidirectional Walking and Disturbance Rejection. The robot is able to: (a-c) walk forward, turns in place, and walk backward in the lab environment. (d, e) walks forward and sideways in the wild, and recover from large external perturbations, such as being kicked (f) from behind while walking in the lab, and (g) from the side while walking in the wild.

Fig. 4. Walking on Various Terrains. The robot (a) walks on eight different types of outdoor terrains, (b) climbs a relatively steep and narrow unpaved hiking trail covered with dust and rocks, (c) traverses on an uneven pathway with slipping slabs, (d) makes a turn on rocky stairs.

Omnidirectional Walking. We train our robot to perform omnidirectional locomotion by following linear velocity commands in sagittal and lateral directions as well as angular velocity commands in yaw. In Figure 3, we show examples of walking forward, backward, and turning left and right.

Disturbance Rejection. A crucial test of the robustness of the policy and the reliability of the hardware is the ability to recover from external perturbations. We exert instantaneous force randomly by kicking different parts of our robot while it is stepping in place. As shown in Figure 3, this perturbation causes a significant deviation from the nominal walking pose, making the robot almost fall over. Nevertheless, our robot is able to respond immediately, regain its stability from the perturbation within a few steps, and resume stepping.

In addition to the flat ground in the controlled lab environment, we repeat this test in outdoor environments, such as on uneven grass terrains. In these conditions, our robot is also able to recover from heavy external forces, as shown in Figure 3(g). This further showcases the robustness of our humanoid robot in real-world scenarios.

Walking on Various Terrains. Perhaps the best demonstration of the advanced performance of a humanoid is its capability to traverse various everyday environments robustly. As shown in Figure 4(a), our robot is able to walk robustly on diverse outdoor terrains, such as grass fields, brick sidewalks, unpaved trails, asphalt roads, bridges, concrete roads, running tracks, and tiled surfaces, as well as stairs and inclines.

Among these environments, we emphasize the two most challenging terrains. First, as shown in Figure 4(b) and the accompanying video, surprisingly, our robot is able to climb a relatively steep and narrow unpaved trail covered with dust and rocks, even though inclined surfaces were not in training distribution. This trail is a bit steep to climb even for adults, let alone our robot which resembles only a 5-year-old child in size. Specifically, the trail's average incline is 20 degrees, exceeding the ankle's upward pitch range, requiring the robot to step backward to maintain an upright torso. Despite this, our robot is able to walk stably, make turns, and recover from stepping on loose rocks.

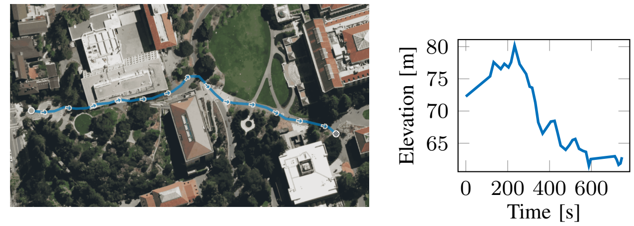

Fig. 5. Visualization of GPS data of a long-distance walk.

Second, as shown in Figure 4(c), our robot handles uneven pathways with large gaps and different heights between the slabs effectively. These gaps and slippery slabs require extra attention as this could cause a loss in balance, leading to the potential to fall over. On this challenging terrain, our robot is able to navigate both forward and backward inside the small pathway across changes in step heights and recover from stepping on slipping slabs.

In order to further demonstrate uneven terrain, we create a set of rocky stairs with step heights of 4 cm (10% of full leg length) and find that our robot is able to traverse the stairs smoothly and make turns on them, as seen in Figure 4(d).

Being able to handle these challenging terrains shows an advanced performance on locomotion control for our humanoid, even with such a basic RL controller. We attribute this to the careful adaptations for learning-based control algorithms in the hardware design.

Long Distance Walking. With the ability to traverse terrains and reject perturbations, the robot is able to perform relatively long-distance walking for several hundred meters over multiple terrains. As shown in Figure 5, the robot rambles freely on the campus of UC Berkeley for 10 minutes, traversing a total distance of 364 m with uphills and downhills. Furthermore, the robot is able to climb steadily along a rough hiking trail shown in Figure 4(b) for more than 5 minutes non-stop, covering 96 m in distance and an elevation gain of 10.5 m. The video of the campus walking can be found at `https://youtu.be/STbB12-oc_w` and the video of walking on the hiking trail at `https://youtu.be/Z2Bzs1mu7DA`.

B. *Evaluation of Sim-to-Real Transfer*

Because the majority of learning-based algorithms are trained entirely in simulation, the sim-to-real gap becomes a critical component of the performance of learning-based controllers in the real world. We demonstrate the small sim-

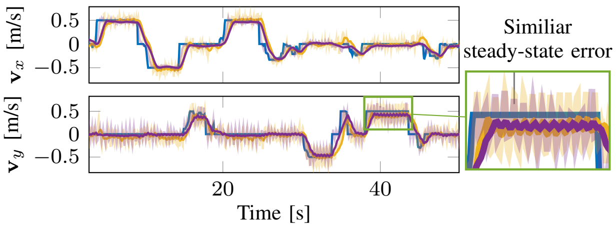

Fig. 6. Sim-to-real gap evaluation. We show trajectories for base linear velocities in command (blue), simulation (purple), and real-world (yellow). The values are smoothed by a moving average filter to better illustrate the steady-state error. We observe a closely aligned oscillation magnitude and frequency between sim and hardware, indicating a small sim-to-real gap.

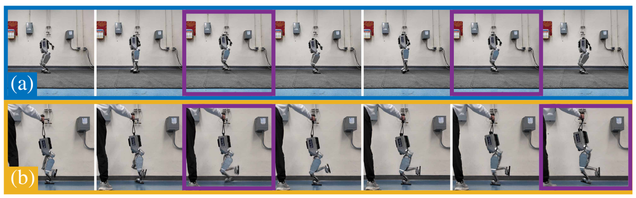

Fig. 7. Hopping with (a) both legs and (b) a single leg, with noticeable flight phases marked with purple frames.

to-real gap of our robot in two aspects: (i) A quantitative analysis of the locomotion task metrics. (ii) The ability to perform highly dynamic locomotion tasks.

First, we present a quantitative analysis of the sim-to-real transfer by plotting the tracking performance with random velocity commands given by the operator. As shown in Figure 6, our robot is able to follow the rapidly changing command closely in both lateral and sagittal directions with small steady-state errors. Over a 60-second trial, the average tracking error in the sagittal direction is 0.051 m/s in simulation and 0.058 m/s on hardware. In the lateral direction, the error is 0.086 m/s in simulation and 0.1156 m/s on hardware, respectively. Note that our RL controller is unable to perform online system identification or adaptation as it does not have access to the observation history during either training or deployment. Thus, these small differences in tracking errors indicate that the gap between the simulation MDPs during training and the MDPs of the real-world deployment is indeed small, which confirms the narrow sim-to-real gap for our hardware design.

Second, we showcase the ability to perform highly dynamic motions by demonstrating a hopping controller trained with the same settings as in Sec. IV except for the rewards. As shown in Figure 7(a), our robot can perform omnidirectional hops, accelerate, and decelerate while maintaining balance. Notably, the robot further demonstrates exceptional agility by being able to perform hops using only one leg in Figure 7(b), a highly challenging feat. Although a safety rope is used and minor balance assistance is needed during single-leg hopping experiments, the rope is mostly slack, and the robot is able to maintain its balance on its own. Compared to complex algorithm designs in prior works, this further shows that the hardware design facilitates us to perform agile motions with simple algorithmic design.

C. Hardware Reliability

Validating the robot's reliability is challenging. A notable systematic perturbation analysis is presented in [68], which applies perturbation forces but does not push the robot to the point of falling and focuses primarily on controller reliability rather than hardware reliability. We attempt to demonstrate the hardware's reliability more intuitively by recording instances of the robot falling. We documented a total of 38 times our robot falling over on various terrains including concrete pavements and unpaved roads. Note that these falls resulted from debugging or systematic errors in low-level systems (e.g., actuator controller, IMU filtering) or from testing the model-based controller, rather than the final policy described in the paper. Thanks to the reliable and lightweight design, we did not experience any damage to the hardware itself except for two failures caused by loose screws and glue. In most fallovers, we are able to reset the robot and resume the control policy within 3 to 5 seconds. The ability to reset easily and rapidly not only relieves the burden of experiments but more importantly, is necessary for the ultimate goal of scalable real-world deployment.

VI. Conclusion and Future Works

In this work, we have presented Berkeley Humanoid, a reliable and low-cost research platform for learning-based bipedal locomotion control with a narrow sim-to-real gap. Our in-house-built humanoid robot is designed specifically for learning-based control, featuring accurate simulation, low simulation complexity, easier real-world experiments, and anthropomorphic ranges of motion. As a result, our minimally-designed policy, without history or phase signal as input, is able to withstand large, random external perturbations and perform omnidirectional locomotion over challenging terrains, demonstrating the efficacy of these learning-oriented design features and the resulting narrow sim-to-real gap. Notably, it demonstrates the ability to walk long distances on campus, climb steadily along steep and narrow unpaved hiking trails, and hop with a single leg, a highly dynamic feat. Furthermore, the Berkeley Humanoid has served as an open research platform for various other locomotion studies across the globe [69], [70].

Our future plan is to conduct experiments with two pre-designed arms, enabling future loco-manipulation research. As a reliable, low-cost, and high-performance research platform, the ultimate goal is to deploy scalably for learning in the real world.

Acknowledgment

This work was supported in part by The Robotics and AI Institute. We want to thank Jiaze Cai for the generous help with the experiments, and Yufeng Chi for suggesting the name of the robot. We'd also like to express our gratitude to Prof. Wei Zhang and Pan Motor for their valuable discussions and assistance with the actuators. K. Sreenath has financial interests in The Robotics and AI Institute. He and the company may benefit from the commercialization of the results of this research.

References

[1] J. Englsberger, A. Werner, C. Ott, B. Henze, M. A. Roa, G. Garofalo, R. Burger, A. Beyer, O. Eiberger, K. Schmid, *et al.*, "Overview of the torque-controlled humanoid robot toro," in *2014 IEEE-RAS International Conference on Humanoid Robots*, pp. 916–923, IEEE, 2014.

[2] P. Seiwald, S.-C. Wu, F. Sygulla, T. F. Berninger, N.-S. Staufenberg, M. F. Sattler, N. Neuburger, D. Rixen, and F. Tombari, "Lola v1. 1–an upgrade in hardware and software design for dynamic multi-contact locomotion," in *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*, pp. 9–16, IEEE, 2021.

[3] N. G. Tsagarakis, D. G. Caldwell, F. Negrello, W. Choi, L. Baccelliere, V.-G. Loc, J. Noorden, L. Muratore, A. Margan, A. Cardellino, *et al.*, "Walk-man: A high-performance humanoid platform for realistic environments," *Journal of Field Robotics*, vol. 34, no. 7, pp. 1225–1259, 2017.

[4] Agility Robotics, "Meet digit: The newest robot from agility robotics," 2024.

[5] Unitree Robotics, "Unitree h1," 2024.

[6] J. Ramos, B. Katz, M. Y. M. Chuah, and S. Kim, "Facilitating model-based control through software-hardware co-design," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 566–572, IEEE, 2018.

[7] B. G. Katz, *A low cost modular actuator for dynamic robots*. PhD thesis, Massachusetts Institute of Technology, 2018.

[8] M. Chignoli, D. Kim, E. Stanger-Jones, and S. Kim, "The MIT humanoid robot: Design, motion planning, and control for acrobatic behaviors," in *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*, pp. 1–8, IEEE, 2021.

[9] B. Katz, J. Di Carlo, and S. Kim, "Mini cheetah: A platform for pushing the limits of dynamic quadruped control," in *2019 international conference on robotics and automation (ICRA)*, pp. 6295–6301, IEEE, 2019.

[10] Y. Liu, J. Shen, J. Zhang, X. Zhang, T. Zhu, and D. Hong, "Design and control of a miniature bipedal robot with proprioceptive actuation for dynamic behaviors," in *2022 International Conference on Robotics and Automation (ICRA)*, pp. 8547–8553, IEEE, 2022.

[11] F. Grimminger, A. Meduri, M. Khadiv, J. Viereck, M. Wüthrich, M. Naveau, V. Berenz, S. Heim, F. Widmaier, T. Flayols, *et al.*, "An open torque-controlled modular robot architecture for legged locomotion research," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3650–3657, 2020.

[12] A. B. Ghansah, J. Kim, K. Li, and A. D. Ames, "Dynamic walking on highly underactuated point foot humanoids: Closing the loop between hzd and hlip," *arXiv preprint arXiv:2406.13115*, 2024.

[13] M. Hutter, C. Gehring, D. Jud, A. Lauber, C. D. Bellicoso, V. Tsounis, J. Hwangbo, K. Bodie, P. Fankhauser, M. Bloesch, *et al.*, "Anymal-a highly mobile and dynamic quadrupedal robot," in *2016 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 38–44, IEEE, 2016.

[14] A. Hattori, *Design of a high torque density modular actuator for dynamic robots*. PhD thesis, Massachusetts Institute of Technology, 2020.

[15] T. Zhu, *Design of a highly dynamic humanoid robot*. University of California, Los Angeles, 2023.

[16] Agility Robotics, "Cassie sets a guinness world record," 2022.

[17] A. SaLoutos, E. Stanger-Joncs, Y. Ding, M. Chignoli, and S. Kim, "Design and development of the mit humanoid: A dynamic and robust research platform," in *2023 IEEE-RAS 22nd International Conference on Humanoid Robots (Humanoids)*, pp. 1–8, IEEE, 2023.

[18] Unitree Robotics, "Unitree g1," 2024.

[19] J. Li, J. Ma, O. Kolt, M. Shah, and Q. Nguyen, "Dynamic loco-manipulation on hector: Humanoid for enhanced control and open-source research," *arXiv preprint arXiv:2312.11868*, 2023.

[20] A. Parmiggiani, M. Maggiali, L. Natale, F. Nori, A. Schmitz, N. Tsagarakis, J. S. Victor, F. Becchi, G. Sandini, and G. Metta, "The design of the icub humanoid robot," *International journal of humanoid robotics*, vol. 9, no. 04, p. 1250027, 2012.

[21] D. Gouaillier, V. Hugel, P. Blazevic, C. Kilner, J. Monceaux, P. Lafourcade, B. Marnier, J. Serre, and B. Maisonnier, "Mechatronic design of nao humanoid," in *2009 IEEE international conference on robotics and automation*, pp. 769–774, IEEE, 2009.

[22] I. Ha, Y. Tamura, H. Asama, J. Han, and D. W. Hong, "Development of open humanoid platform darwin-op," in *SICE annual conference 2011*, pp. 2178–2181, IEEE, 2011.

[23] A. Nikkhah, A. Yousefi-Koma, R. Mirjalili, and H. M. Farimani, "Design and implementation of small-sized 3d printed surena-mini humanoid platform," in *2017 5th RSI International Conference on Robotics and Mechatronics (ICRoM)*, pp. 132–137, IEEE, 2017.

[24] P. M. Wensing, A. Wang, S. Seok, D. Otten, J. Lang, and S. Kim, "Proprioceptive actuator design in the mit cheetah: Impact mitigation and high-bandwidth physical interaction for dynamic legged robots," *Ieee transactions on robotics*, vol. 33, no. 3, pp. 509–522, 2017.

[25] A. Wang, J. Ramos, J. Mayo, W. Ubellacker, J. Cheung, and S. Kim, "The hermes humanoid system: A platform for full-body teleoperation with balance feedback," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pp. 730–737, IEEE, 2015.

[26] H. Khan, R. Featherstone, D. G. Caldwell, and C. Semini, "Bio-inspired knee joint mechanism for a hydraulic quadruped robot," in *2015 6th International Conference on Automation, Robotics and Applications (ICARA)*, pp. 325–331, IEEE, 2015.

[27] Boston Dynamics, "Atlas," 2024.

[28] M. H. Raibert, *Legged robots that balance*. MIT press, 1986.

[29] S. Kajita, F. Kanehiro, K. Kaneko, K. Yokoi, and H. Hirukawa, "The 3d linear inverted pendulum mode: A simple modeling for a biped walking pattern generation," in *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No. 01CH37180)*, vol. 1, pp. 239–246, IEEE, 2001.

[30] S. Kuindersma, R. Deits, M. Fallon, A. Valenzuela, H. Dai, F. Permenter, T. Koolen, P. Marion, and R. Tedrake, "Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot," *Autonomous robots*, vol. 40, pp. 429–455, 2016.

[31] Y. Ding, C. Khazoom, M. Chignoli, and S. Kim, "Orientation-aware model predictive control with footstep adaptation for dynamic humanoid walking," in *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, pp. 299–305, IEEE, 2022.

[32] 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*, 2024.

[33] J. Siekmann, Y. Godse, A. Fern, and J. Hurst, "Sim-to-real learning of all common bipedal gaits via periodic reward composition," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7309–7315, IEEE, 2021.

[34] I. Radosavovic, T. Xiao, B. Zhang, T. Darrell, J. Malik, and K. Sreenath, "Real-world humanoid locomotion with reinforcement learning," *Science Robotics*, vol. 9, no. 89, p. eadi9579, 2024.

[35] R. P. Singh, Z. Xie, P. Gergondet, and F. Kanehiro, "Learning bipedal walking for humanoids with current feedback," *IEEE Access*, 2023.

[36] A. Tang, T. Hiraoka, N. Hiraoka, F. Shi, K. Kawaharazuka, K. Kojima, K. Okada, and M. Inaba, "Humanmimic: Learning natural locomotion and transitions for humanoid robot via wasserstein adversarial imitation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 13107–13114, IEEE, 2024.

[37] I. Radosavovic, B. Zhang, B. Shi, J. Rajasegaran, S. Kamat, T. Darrell, K. Sreenath, and J. Malik, "Humanoid locomotion as next token prediction," in *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

[38] X. Cheng, Y. Ji, J. Chen, R. Yang, G. Yang, and X. Wang, "Expressive whole-body control for humanoid robots," *arXiv preprint arXiv:2402.16796*, 2024.

[39] T. He, Z. Luo, W. Xiao, C. Zhang, K. Kitani, C. Liu, and G. Shi, "Learning human-to-humanoid real-time whole-body teleoperation," *arXiv preprint arXiv:2403.04436*, 2024.

[40] J. Dao, H. Duan, and A. Fern, "Sim-to-real learning for humanoid box loco-manipulation," *arXiv preprint arXiv:2310.03191*, 2023.

[41] Z. Fu, Q. Zhao, Q. Wu, G. Wetzstein, and C. Finn, "Humanplus: Humanoid shadowing and imitation from humans," *arXiv preprint arXiv:2406.10454*, 2024.

[42] J. Siekmann, K. Green, J. Warila, A. Fern, and J. Hurst, "Blind bipedal stair traversal via sim-to-real reinforcement learning," in *Proceedings of Robotics: Science and Systems*, 2021.

[43] X. Gu, Y.-J. Wang, X. Zhu, C. Shi, Y. Guo, Y. Liu, and J. Chen, "Advancing Humanoid Locomotion: Mastering Challenging Terrains with Denoising World Model Learning," in *Proceedings of Robotics: Science and Systems*, 2024.

[44] B. van Marum, A. Shrestha, H. Duan, P. Dugar, J. Dao, and A. Fern, "Revisiting reward design and evaluation for robust humanoid standing and walking," *arXiv preprint arXiv:2404.19173*, 2024.

[45] D. Crowley, J. Dao, H. Duan, K. Green, J. Hurst, and A. Fern, "Optimizing bipedal locomotion for the 100m dash with comparison to human running," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 12205–12211, IEEE, 2023.

[46] Z. Zhuang, S. Yao, and H. Zhao, "Humanoid parkour learning," *arXiv preprint arXiv:2406.10759*, 2024.

[47] Y. Chebotar, A. Handa, V. Makoviychuk, M. Macklin, J. Issac, N. Ratliff, and D. Fox, "Closing the sim-to-real loop: Adapting simulation randomization with real world experience," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 8973–8979, IEEE, 2019.

[48] T. Haarnoja, B. Moran, G. Lever, S. H. Huang, D. Tirumala, J. Humplik, M. Wulfmeier, S. Tunyasuvunakool, N. Y. Siegel, R. Hafner, *et al.*, "Learning agile soccer skills for a bipedal robot with deep reinforcement learning," *Science Robotics*, vol. 9, no. 89, p. eadi8022, 2024.

[49] K.-S. Labs, "K-scale," 2024.

[50] K. Urs, C. E. Adu, E. J. Rouse, and T. Y. Moore, "Design and characterization of 3d printed, open-source actuators for legged locomotion," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1957–1964, IEEE, 2022.

[51] T.-G. Song, Y.-H. Shin, S. Hong, H. C. Choi, J.-H. Kim, and H.-W. Park, "Drpd, dual reduction ratio planetary drive for articulated robot actuators," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 443–450, IEEE, 2022.

[52] A. J. Fuge, C. W. Herron, B. C. Beiter, B. Kalita, and A. Leonessa, "Design, development, and analysis of the lower body of next-generation 3d-printed humanoid research platform: Pandora," *Robotica*, vol. 41, no. 7, pp. 2177–2206, 2023.

[53] S. Haddadin, A. Albu-Schaffer, A. De Luca, and G. Hirzinger, "Collision detection and reaction: A contribution to safe physical human-robot interaction," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3356–3363, IEEE, 2008.

[54] Boston Dynamics, "Spot - the agile mobile robot," 2024.

[55] S. Kudo, M. Fujimoto, T. Sato, and A. Nagano, "Optimal degrees of freedom of the lower extremities for human walking and running," *Scientific Reports*, vol. 13, no. 1, p. 16164, 2023.

[56] S. Caron, Q.-C. Pham, and Y. Nakamura, "Stability of surface contacts for humanoid robots: Closed-form formulae of the contact wrench cone for rectangular support areas," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5107–5112, IEEE, 2015.

[57] Z. Li, X. B. Peng, P. Abbeel, S. Levine, G. Berseth, and K. Sreenath, "Reinforcement learning for versatile, dynamic, and robust bipedal locomotion control," *arXiv preprint arXiv:2401.16889*, 2024.

[58] A. Kumar, Z. Fu, D. Pathak, and J. Malik, "Rma: Rapid motor adaptation for legged robots," in *Robotics: Science and Systems*, 2021.

[59] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, "Learning quadrupedal locomotion over challenging terrain," *Science robotics*, vol. 5, no. 47, p. eabc5986, 2020.

[60] T. Flayols, A. Del Prete, P. Wensing, A. Mifsud, M. Benallegue, and O. Stasse, "Experimental evaluation of simple estimators for humanoid robots," in *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*, pp. 889–895, IEEE, 2017.

[61] J. Carpentier, G. Saurel, G. Buondonno, J. Mirabel, F. Lamiraux, 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.

[62] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[63] M. Mittal, C. Yu, Q. Yu, J. Liu, N. Rudin, D. Hoeller, J. L. Yuan, R. Singh, Y. Guo, H. Mazhar, A. Mandlekar, B. Babich, G. State, M. Hutter, and A. Garg, "Orbit: A unified simulation framework for interactive robot learning environments," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3740–3747, 2023.

[64] Z. Xie, P. Clary, J. Dao, P. Morais, J. Hurst, and M. Panne, "Learning locomotion skills for cassie: Iterative design and sim-to-real," in *Conference on Robot Learning*, pp. 317–329, PMLR, 2020.

[65] J. Tan, T. Zhang, E. Coumans, A. Iscen, Y. Bai, D. Hafner, S. Bohez, and V. Vanhoucke, "Sim-to-real: Learning agile locomotion for quadruped robots," in *Proceedings of Robotics: Science and Systems*, 2018.

[66] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso, V. Tsounis, V. Koltun, and M. Hutter, "Learning agile and dynamic motor skills for legged robots," *Science Robotics*, vol. 4, no. 26, p. eaau5872, 2019.

[67] X. B. Peng, E. Coumans, T. Zhang, T.-W. Lee, J. Tan, and S. Levine, "Learning agile robotic locomotion skills by imitating animals," in *Proceedings of Robotics: Science and Systems*, 2020.

[68] B. Weng, G. A. Castillo, Y.-S. Kang, and A. Hereid, "Towards standardized disturbance rejection testing of legged robot locomotion with linear impactor: A preliminary study, observations, and implications," *arXiv preprint arXiv:2308.14636*, 2023.

[69] Ryu Kanghyun, "Curricullm: Automatic curricula design for learning complex robot skills using large language models," 2024.

[70] Zixuan Chen, "Learning smooth humanoid locomotion through lipschitz-constrained policies," 2024.