Internalizing Geometric Stability for Learning Quadrupedal Recovery on Irregular Terrains

Boyuan Deng[1,2,*], Xu Yang[3], Yilin Mo[3], and Nikolaos Tsagarakis[1]

***Abstract*— Autonomous post-fall recovery on irregular terrain is a critical bottleneck for the long-duration deployment of legged robots. Prevailing methods often depend on flat-ground assumptions and heuristic rewards, which fail to generalize under irregular terrain and random initial states. This work addresses the role of geometry in a data-driven era by proposing that geometric stability criteria serve as a fundamental inductive bias for Reinforcement Learning (RL). We formalize the terrain-adapted optimal recovery posture as a multi-objective geometric optimization problem and introduce a unified Force-Angle (FA) stability margin to guide the policy search. We employ a Teacher-Student distillation scheme to enable an RL agent to learn recovery strategies that inherently satisfy these geometric requirements across a vast range of initial fallen configurations and terrain conditions. Our framework achieves a 90.26% recovery success rate on heterogeneous terrains, ensuring physically consistent and stable final postures that honor the underlying geometric boundaries. The supplementary material is available at `https://boyuandeng.github.io/TAFR-TeacherGuidedFramework/`**

I. INTRODUCTION

Autonomous legged systems are increasingly deployed in challenging environments for inspection and exploration [1], [2], [3]. However, unexpected falls caused by dynamic disturbances remain an inevitable interruption to autonomy, and current systems largely rely on manual intervention, severely limiting operational efficiency. Fundamentally, post-fall recovery is a complex motion planning problem requiring precise body adjustments through multiple discontinuous contacts [4].

While Deep Reinforcement Learning (DRL) shows promise in agile motor skills [5], [6], recovery on irregular terrain remains unsolved due to the lack of a principled definition of a ”stable posture” in non-Euclidean environments. Prevailing methods often depend on flat-ground assumptions and heuristic rewards that fail to generalize under uncertain topologies. Consequently, the randomness of initial configurations and contact conditions frequently lead to degraded performance or complete recovery failure. In this work, we argue that geometry serves as a Navigation Compass in the era of data-driven robotics. While DRL excels at navigating high-dimensional spaces for feasible trajectories , geometry defines the recovery targets derived from fundamental dynamic and kinematic principles. By anchoring a Teacher-Student PPO framework to the Force-Angle stability margin[7], we distill privileged geometric knowledge into deployable proprioceptive policies. Rather than a rigid constraint, geometry acts as an inductive bias that allows the agent to internalize strategies that inherently respect physical boundaries.

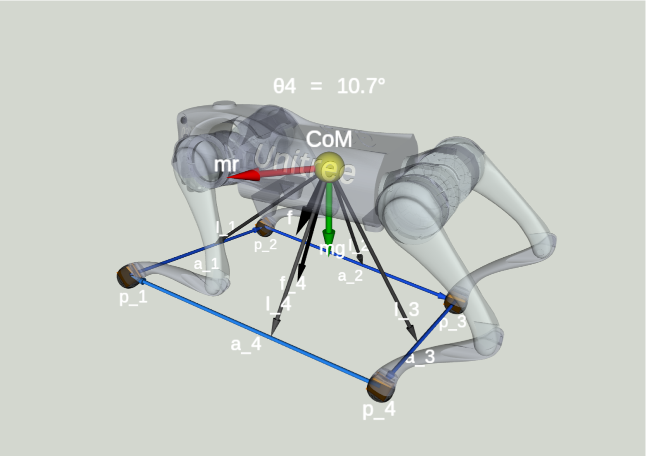

Fig. 1: Force–Angle Stability Measure for a quadrupedal robot. Foot contact points are denoted $\mathbf{p}_i$. Tip-over axes $\mathbf{a}_i$ are the ordered line segments connecting adjacent contacts; their associated normals l_i pass through the robot's CoM and intersect $\mathbf{a}_i$. The vector $\mathbf{f}$ is the sum of all forces and angular loads acting on the CoM, and $\mathbf{f}_i$ is its component associated with the i-th tip-over axis. The configuration shown attains the minimum stability angle $\beta = \theta_4$. All the vectors are represented in the inertial frame.

II. GEOMETRIC FORMULATION OF TERRAIN-ADAPTIVE STABILITY TARGETS

On irregular terrains, a robot's target pose cannot be a fixed joint configuration; instead, it must adapt to the underlying terrain topology.

A. Geometric Optimization Objective

We formalize the optimal recovery posture by minimizing a multi-objective cost function $J(\tau)$ that incorporates contact geometry:

$$\min J = w_1 J_{force} + w_2 J_{effort} + w_3 J_{CoM} \tag{1}$$

where J_{force} penalizes the variance of vertical ground reaction forces to ensure uniform load sharing among legs. J_{effort} minimizes the norm of joint torques to reduce actuation demand, and J_{CoM} minimizes the distance between the Center of Mass (CoM) and the centroid of the support polygon, ensuring the robot resides in the center of its stability manifold—a

---

[1] Humanoids and Human-Centered Mechatronics (HHCM), Istituto Italiano di Tecnologia, Via Morego 30, Genoa, 16163, Italy

[2] Ph.D. Program of National Interest in Robotics and Intelligent Machines (DRIM), University of Genova, 16126 Genoa, Italy

[3] Department of Automation and BNRist, Tsinghua University, Beijing, China

[*] Corresponding author: boyuan.deng@iit.it

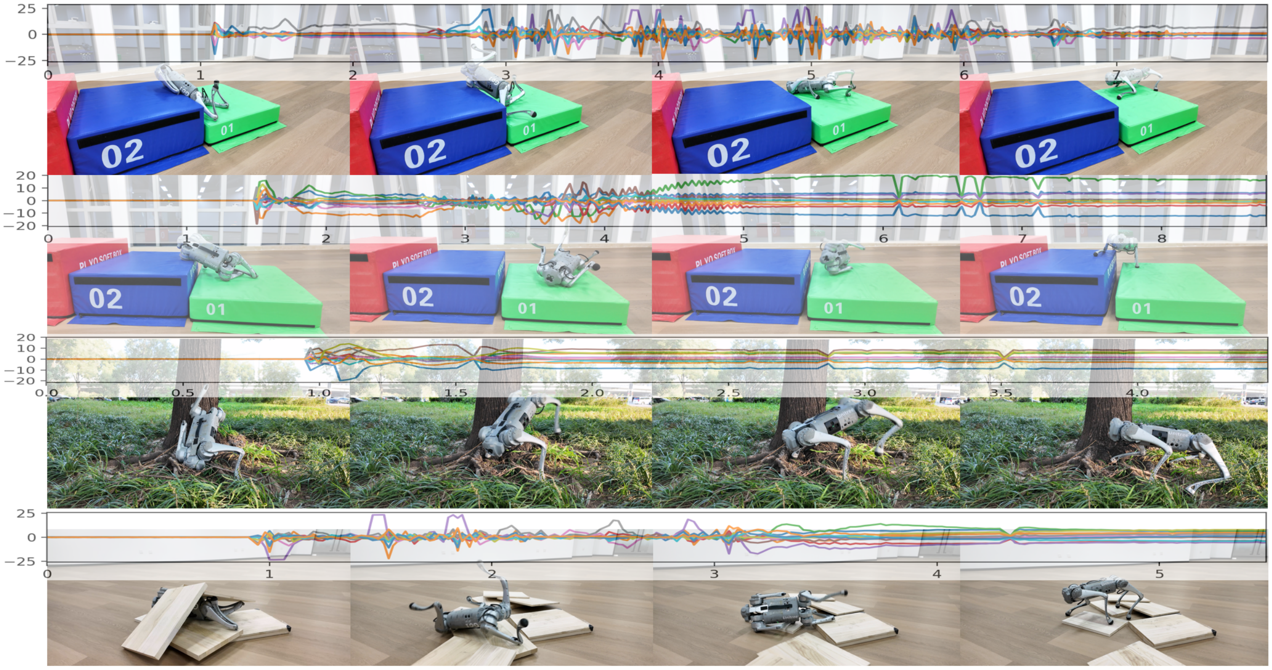

Fig. 2: Real-world experiments on challenging terrains (Flat, Foam, Stairs, Roots). Motor torques converge to similar values, indicating a near-optimal stable configuration.

geometric requirement for maximizing the stability margin on slopes [8], [9]

B. The Force-Angle Stability Margin

To evaluate recovery outcomes across diverse terrains, we adopt the Force–Angle Stability Measure β [7], [10]. Unlike planar support-polygon heuristics, β calculates the minimum angle between the resultant force vector and the tip-over axes formed by non-coplanar contacts. This geometric metric provides a differentiable proxy for the robot's physical limits, which we distill into the RL reward structure.

III. LEARNING RECOVERY POLICIES UNDER GEOMETRIC BIAS

We utilize the Orbit/Isaac Lab simulation framework [11] to train a Teacher-Student architecture.

- **Geometric Teacher:** The teacher policy is granted privileged access to the terrain (elevation maps and contact forces). It learns to navigate the configuration space toward the optimal state defined in Section II.
- **Proprioceptive Student:** The student policy relies solely on proprioceptive histories. It does not "see" the terrain but is trained to imitate the teacher's geometric insights. Here, the neural network effectively learns an implicit mapping from temporal sensor data to the environment's geometric constraints.

IV. EXPERIMENTAL RESULTS

The effectiveness of the proposed framework was evaluated through extensive simulations and hardware deployments. We utilized the Isaac Lab framework [11] for massively parallel training of 2,048 agents with randomized initial states across six heterogeneous terrains. Experiments were conducted on the Unitree Go1 [12].

A. Recovery Success Rate

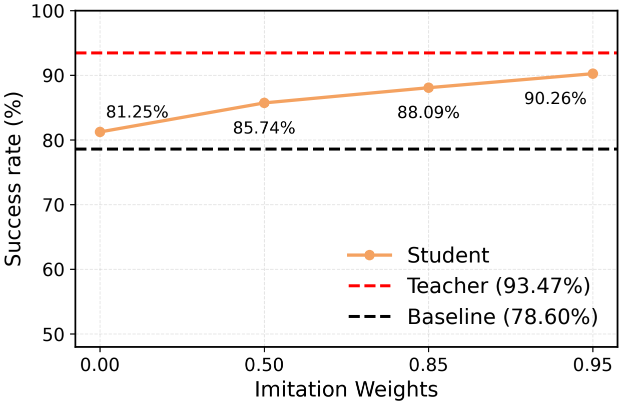

Fig. 3: Recovery success rates of the Student policy across different imitation weights compared to the Teacher and pure RL baseline [13].

A recovery trial is declared successful if, at the terminal timestep, the robot satisfies the following geometric and physical criteria: (i) the gravity vector projection onto the world z-axis is within a 0.1 error tolerance; (ii) the vertical ground-reaction force variance is below 0.5, ensuring stable four-leg contact; (iii) joint velocities are within safety limits; and (iv) the stability margin β is strictly positive.

As shown in Fig. 3, the **Geometric Teacher** policy achieves an upper-bound success rate of 93.47%. By internalizing these geometric insights through distillation, the **Proprioceptive Student** reaches 90.26% success. In contrast, the baseline RL method lacking explicit geometric guidance achieves only 78.6%. This significant performance gain validates that anchoring the learning process to formal geometric targets enables the acquisition of more robust recovery strategies than pure data-driven exploration.

B. *Tip-over Stability Margin Analysis*

To evaluate the quality of the final recovered posture, we analyzed the tip-over stability margin β for all successful trials across six different terrains. As illustrated in Fig. 4, the stability margins remain strictly positive in all cases, confirming the static reliability of the final configurations.

The mean values of β range from 12.06 to 14.78, with peak values reaching as high as 25.71. These results demonstrate that the learned policy successfully guides the robot into configurations that maximize stability relative to the specific terrain topology.

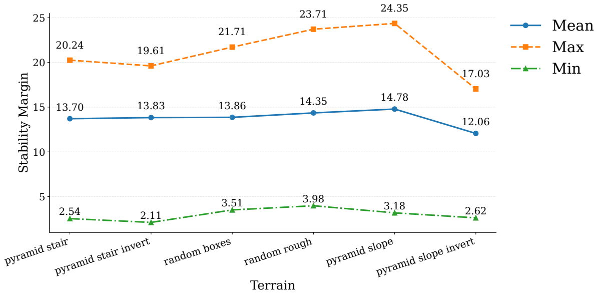

Fig. 4: Distribution of stability margins β across six irregular terrains. Mean values consistently remain positive, ensuring tipping resistance.

V. DISCUSSION AND CONCLUSION

Returning to the central theme of this workshop, our work demonstrates that the rise of data-driven methods does not diminish the significance of geometric principles; rather, it redefines their utility. In the context of post-fall recovery, geometry serves as the "Navigation Compass" by providing formal definitions of physical stability that simple heuristic rewards cannot capture. Traditional geometric optimization for multi-contact recovery on irregular terrains requires solving highly non-linear and non-convex constraints in real-time, a task that remains computationally prohibitive for embedded systems [7], [10]. The core insight of our framework is that deep reinforcement learning can be leveraged as an efficient offline optimizer. By training on a wide distribution of randomized initial states and terrains, the neural network learns to internalize these complex geometric requirements, essentially acting as a high-dimensional non-linear approximator for the optimal recovery manifold.

Empirical evidence from both extensive simulations and real-world deployments confirms that the learned student policy successfully internalizes the prescribed geometric objectives, effectively reaching the target configurations with high reliability. By unifying theory-backed geometric targets with a teacher-guided learning paradigm, this study offers a general recipe for reliable fall recovery on complex terrains. However, an inherent challenge persists in the paradigm of data-driven robotics: while reinforcement learning can navigate high-dimensional, non-convex landscapes where traditional optimization remains computationally prohibitive or prone to local minima, providing a formal proof that the learned strategy represents the global optimum remains elusive. Furthermore, the stochastic nature of the learning process itself offers no theoretical guarantees regarding the stability of the training trajectory or the optimality of the converged policy. Future research will focus on incorporating formal methods and lyapunov-based stability certificates into the learning loop to ensure both the robustness of the training process and the certified optimality of the resulting recovery behaviors.

REFERENCES

[1] H. Christensen, N. Amato, H. Yanco, M. Mataric, H. Choset, A. Drobnis, K. Goldberg, J. Grizzle, G. Hager, J. Hollerbach *et al.*, "A roadmap for us robotics–from internet to robotics 2020 edition," *Foundations and Trends® in Robotics*, vol. 8, no. 4, pp. 307–424, 2021.

[2] M. Tranzatto, F. Mascarich, L. Bernreiter, C. Godinho, M. Camurri, S. Khattak, T. Dang, V. Reijgwart, J. Loeje, D. Wisth *et al.*, "Cerberus: Autonomous legged and aerial robotic exploration in the tunnel and urban circuits of the darpa subterranean challenge," *arXiv preprint arXiv:2201.07067*, vol. 3, 2022.

[3] A. Bouman, M. F. Ginting, N. Alatur, M. Palieri, D. D. Fan, T. Touma, T. Pailevanian, S.-K. Kim, K. Otsu, J. Burdick *et al.*, "Autonomous spot: Long-range autonomous exploration of extreme environments with legged locomotion," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 2518–2525.

[4] C. D. Bellicoso, M. Bjelonic, L. Wellhausen, K. Holtmann, F. Günther, M. Tranzatto, P. Fankhauser, and M. Hutter, "Advances in real-world applications for legged robots," *Journal of Field Robotics*, vol. 35, no. 8, pp. 1311–1326, 2018.

[5] 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.

[6] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, "Learning to walk in minutes using massively parallel deep reinforcement learning," in *Conference on Robot Learning*. PMLR, 2022, pp. 91–100.

[7] E. Papadopoulos and D. A. Rey, "The force-angle measure of tipover stability margin for mobile manipulators," *Vehicle System Dynamics*, vol. 33, no. 1, pp. 29–48, 2000.

[8] D. V. Lee, "Effects of grade and mass distribution on the mechanics of trotting in dogs," *Journal of Experimental Biology*, vol. 214, no. 3, pp. 402–411, 2011.

[9] C. Gehring, C. D. Bellicoso, S. Coros, M. Bloesch, P. Fankhauser, M. Hutter, and R. Siegwart, "Dynamic trotting on slopes for quadrupedal robots," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015, pp. 5129–5135.

[10] V. S. Medeiros, E. Jelavic, M. Bjelonic, R. Siegwart, M. A. Meggiolaro, and M. Hutter, "Trajectory optimization for wheeled-legged quadrupedal robots driving in challenging terrain," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4172–4179, 2020.

[11] 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.

[12] U. Robotics, "Unitree go1," 2024. [Online]. Available: https://github.com/unitreerobotics/unitree_ros/tree/master/robots/go1_description

[13] B. Deng, L. Rossini, J. Wang, W. Wang, and N. Tsagarakis, "Learning to recover: Dynamic reward shaping with wheel-leg coordination for fallen robots," *arXiv preprint arXiv:2506.05516*, 2025.