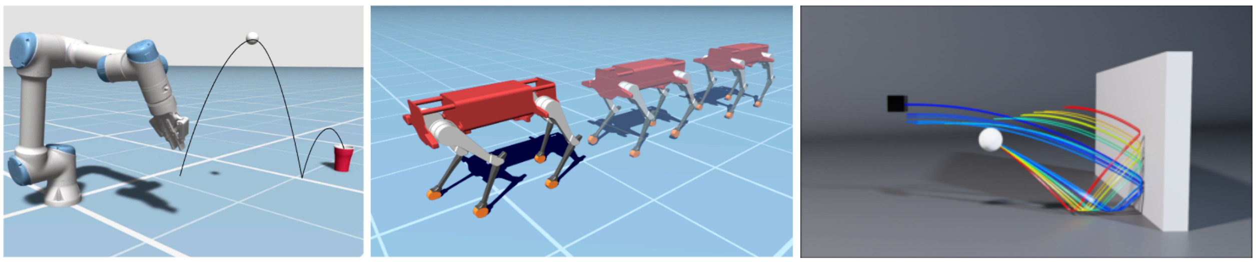
Applying differentiable simulation for policy optimization in robotics



Policy optimization through differentiable simulation in ADD [8] and Nvidia [Warp]

Challenges and Opportunities in Robotics Reinforcement Learning

Policy gradient algorithms are a cornerstone in robotics reinforcement learning (RL) for continuous control policy learning. However, the inherent complexity of robotic systems often results in prolonged computation times and impedes effective exploration. To address these challenges, two main strategies have emerged:

1. Leveraging Differentiable Simulators: Enhances sample efficiency [1, 2]
2. Utilizing Parallelizable Simulation: Offers significant speed improvements [3, 4]

Some advanced methods [5, 7] combine these approaches, aiming to harness the benefits of both. However, challenges persist:

- Training can be hampered by local minima and gradient instability (exploding/vanishing gradients)
- Need for specialized techniques to address these issues in contact-rich loco-manipulation environments

Critical Questions for Future Research

1. Necessity of Gradient Information:
   In an era of affordable and easily parallelizable simulations, is gradient information from simulations still crucial or beneficial for robotics RL?
2. Simulation Fidelity and Real-World Transfer:
   Are the assumptions made in differentiable simulators sufficiently realistic and transferable to real-world robotic applications?

Addressing these questions is vital for advancing the field of robotics RL and bridging the gap between simulation and real-world performance.

Possible roadmap:

1. Gradient-Based Method Implementation: Develop a method that leverages gradient information and has the potential for parallelization. Compare its performance against reinforcement learning (RL) baselines.
2. Algorithm Customization: Adapt the algorithm to utilize data from differentiable simulations, incorporating techniques to minimize variance in policy optimization.
3. Computational Trade-Off Analysis: Investigate the computational trade-offs associated with the proposed algorithm.
4. Sim-to-Real Transfer Testing: Evaluate the transferability of the optimized RL policies from simulation to real-world hardware.

Contact Person:
Gabriele Fadini (gabriele.fadini@infk.ethz.ch)

References:
[1] "PODS: Policy Optimization via Differentiable Simulation", Miguel Angel Zamora Mora, Momchil Peychev, Sehoon Ha, Martin Vechev, Stelian Coros, Proceedings of the 38th International Conference on Machine Learning
[2] "An End-to-End Differentiable Framework for Contact-Aware Robot Design", Jie Xu, Tao Chen, Lara Zlokapa, Michael Foshey, Wojciech Matusik, Shinjiro Sueda, Pulkit Agrawal Robotics: Science and Systems (RSS 2021)
[3] "Accelerated Methods for Deep Reinforcement Learning", Adam Stooke, Pieter Abbeel https://arxiv.org/abs/1803.02811
[4] "Isaac Gym: High-Performance GPU-Based Physics Simulation For Robot Learning", Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, Gavriel State, Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2), 2021
[5] "Accelerated Policy Learning with Parallel Differentiable Simulation", Jie Xu, Viktor Makoviychuk, Yashraj Narang, Fabio Ramos, Wojciech Matusik, Animesh Garg, Miles Macklin, ICRL 2022
[6] "Do Differentiable Simulators Give Better Policy Gradients?" H.J. Terry Suh, Max Simchowitz, Kaiqing Zhang, Russ Tedrake https://arxiv.org/abs/2202.00817

[7] “Adaptive Horizon Actor-Critic for Policy Learning in Contact-Rich Differentiable Simulation”, Ignat Georgiev, Krishnan Srinivasan, Jie Xu, Eric Heiden, Animesh Garg, https://arxiv.org/abs/2405.17784
[8] “ADD: Analytically differentiable dynamics for multi-body systems with frictional contact”. Moritz Geilinger et al. ACM Trans. Graph. 39.6, 2020.