Learning More With Less: Sample-Efficient Model-Based RL for Loco-Manipulation

Benjamin Hoffman[1], Jin Cheng[1], Chenhao Li[2], Stelian Coros[1]

[1]Computational Robotics Lab, ETH Zurich

[2]ETH AI Center, ETH Zurich

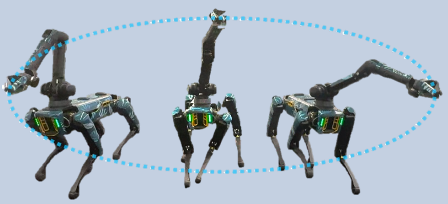
Paper

Website

1 Introduction

Loco-manipulation platforms face **complex dynamics** and the black-box nature of state-of-the-art commercial platforms, making them **hard to accurately model** and **control** in practice, especially when **data** and **simulation** access is **limited**.

We address this challenge by combining a **hand-crafted kinematic model** with **BNN–based MBRL** to efficiently learn the robot's dynamics from **limited data** and derive **accurate control policies**, validated on the **Boston Dynamics Spot**.



2 Approach

1. Formulate MDP: $(S, A, f, r, \gamma, \mathbf{s}_0)$

2. Derive kinematic model from first principles:

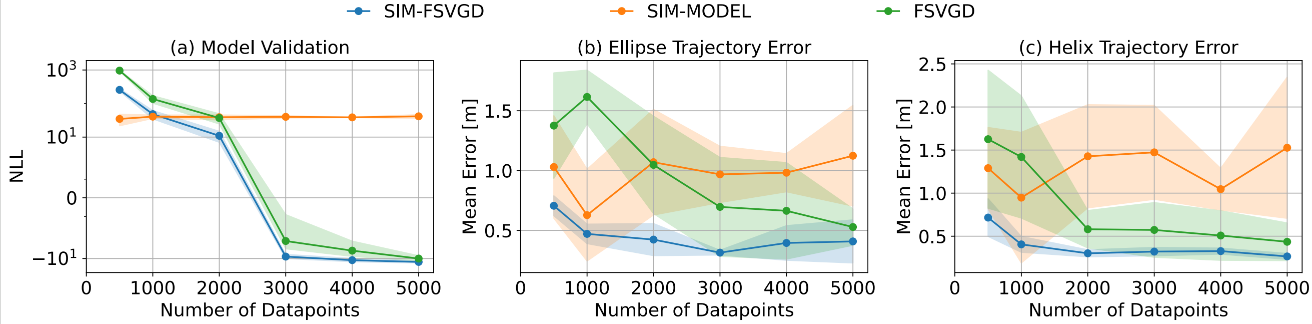
$$\mathbf{v}_{\text{base},t+1} = \alpha_{\text{base}} \cdot \mathbf{v}_{\text{base},t} + (1-\alpha_{\text{base}}) \cdot \mathbf{u}_{\text{base},t} + \boldsymbol{\beta}_{\text{vel, base}}$$
$$\mathbf{p}_{\text{base},t+1} = \mathbf{p}_{\text{base},t} + \Delta t \cdot \mathbf{v}_{\text{base},t+1} \cdot \gamma_{\text{base}} + \boldsymbol{\beta}_{\text{pos, base}}$$
$$\mathbf{v}_{\text{ee},t+1} = \alpha_{\text{ee}} \cdot \mathbf{v}_{\text{ee},t} + (1-\alpha_{\text{ee}}) \cdot \mathbf{u}_{\text{ee},t} + \mathbf{v}_{\text{ee, rot.-ind.},t} + \mathbf{v}_{\text{base},t+1} + \boldsymbol{\beta}_{\text{vel,ee}}$$
$$\mathbf{p}_{\text{ee},t+1} = \mathbf{p}_{\text{ee},t} + \Delta t \cdot \mathbf{v}_{\text{ee},t+1} \cdot \gamma_{\text{ee}} + \boldsymbol{\beta}_{\text{pos,ee}}$$

3. Use kinematic model to efficiently learn a model $\hat{f}_\theta(s_t, u_t)$ of the system dynamics via SIM-FSVGD [1]

4. Use learned model of system dynamics to derive feedback control policies π via RL

3 Results

We demonstrate the effectiveness of our approach by accurately tracking two trajectories with the Spot, consistently outperforming our SIM-MODEL *(no BNN)* and FSVGD [2] *(no kinematic model)* baselines.



4 Conclusion

By deriving a hand-crafted kinematic model and using BNN–based dynamics learning via SIM-FSVGD, our approach enables efficient policy learning for loco-manipulation on a commercial quadruped with a manipulator.

Future work could explore including the end-effector orientation into our action- and state space to enable more complex tasks. Further, fully exploiting Spot's dynamic capabilities by tracking trajectories our performing tasks that require even faster motion of the base, such as catching a ball, could be investigated.

References

[1] J. Rothfuss, B. Sukhija, L. Treven, F. Dörfler, S. Coros, and A. Krause. Bridging the Sim-to-Real Gap with Bayesian Inference. Proc. of IROS, 2024.

[2] Z. Wang, T. Ren, J. Zhu, and B. Zhang. Function Space Particle Optimization for Bayesian Neural Networks. In ICLR, 2019.