# Domain Randomization for Robust Sampling-Based Model Predictive Control on Bipedal Wheeled Robots

## Motivation:

Reinforcement learning (RL) with domain randomization (DR) enables robust sim-to-real transfer for legged robots [1, 2], but produces fixed policies that cannot adapt online. GPU-accelerated sampling-based MPC [3, 4, 5, 6] offers online adaptability and interpretable objectives, yet suffers from **brittle sim-to-real transfer** requiring extensive manual tuning [4]. Traditional robust MPC (tube MPC, H-infinity) struggles with contact-rich dynamics [7, 8], and DR has not been systematically studied for sampling-based MPC.

We target the LIMX Tron1, a bipedal wheeled robot (8 DoF: 3 per leg + 1 per wheel) that is complex enough to stress-test robustness yet compact enough for thorough ablation. Crucially, sampling-based MPC *already* evaluates thousands of rollouts in parallel; randomizing dynamics parameters across these rollouts adds minimal cost, transforming each MPC step into an implicit distributionally robust optimization. This can be further accelerated with learned surrogate dynamics [9, 10]. If DR can close the robustness gap, it would combine RL's sim-to-real reliability with MPC's online adaptability.

## Goal:

We aim to integrate domain randomization directly into the sampling-based MPC loop for robust sim-to-real transfer on the LIMX Tron1, without requiring online parameter estimation or per-robot tuning. Concretely:

- Build a DR-MPC framework that randomizes dynamics parameters (friction, mass, actuator gains, contact properties) across GPU-parallelized rollouts using physics simulators or learned surrogates [11, 12]
- Ablate key design choices: parameter selection, randomization range, and cost aggregation (mean, worst-case, CVaR [13, 14])
- Show that DR-MPC achieves sim-to-real robustness competitive with RL while retaining MPC's online adaptability

## Approach:

We propose two complementary instantiations:

**(1) DR-MPC with Physics Simulation.** Building on DIAL-MPC [4] or Hydrax [5] with MuJoCo MJX [15], we randomize dynamics parameters across parallel rollouts and select actions via MPPI/CEM [16, 17] that minimize cost across the parameter distribution. Cost aggregation ranges from mean (risk-neutral) to CVaR for tunable risk sensitivity [18].

**(2) DR-MPC with Learned Surrogate Dynamics.** We train a neural dynamics model conditioned on physical parameters (mass, friction, stiffness, latency) on domain-randomized data [9, 10, 19], replacing full physics simulation with cheap forward passes. Lipschitz regularization ensures smooth rollouts, and GPU-parallel execution leverages CusADi [20] or GATO [21]. The Tron1's compact 8-DoF state space makes surrogate learning feasible; a hierarchical architecture [22] can further decompose planning across timescales.

**Ablation axes:** parameter selection (friction, mass, actuator gains, damping), randomization range (5%–30%), cost aggregation (mean, minimax, CVaR), and sample budget split between action and parameter diversity [23].

**Evaluation.** Tasks include velocity tracking, push recovery, slope traversal, and turning on the LIMX Tron1. Baselines: nominal MPC [4], DR-trained RL [1, 24, 25], and analytical MPC [7]. Metrics: robustness under perturbations, tracking accuracy, energy, planning frequency (at least 50 Hz), and online adaptability. Hardware validation is planned pending access.

## General Details:

The student should bring along the following attributes:

1. Proficiency in Python and machine learning (familiar with PyTorch/JAX).
2. Good knowledge of model predictive control and/or reinforcement learning.
3. Experience with physics simulation (MuJoCo, Isaac Gym) is preferred.
4. Familiarity with GPU-accelerated computing is a plus.

## Interested?

Reach out to Jin Cheng (jin.cheng@inf.ethz.ch) with your CV and transcripts.

# References


[1] Xue Bin Peng et al. Sim-to-real transfer of robotic control with dynamics randomization. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2018.

[2] Fabio Muratore et al. Robot learning from randomized simulations: A review. *Frontiers in Robotics and AI*, 2022.

[3] Viktor Makoviychuk et al. Isaac Gym: High performance GPU-based physics simulation for robot learning. *arXiv preprint*, 2021.

[4] Le Li et al. DIAL-MPC: Diffusion-inspired annealing for sampling-based model predictive control. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2025.

[5] Vince Kurtz. Hydrax: GPU-accelerated sampling-based MPC. *arXiv preprint*, 2024.

[6] Yiyu Zhang et al. Whole-body MPC of legged robots with MuJoCo. *arXiv preprint*, 2025.

[7] Farbod Farshidian et al. Real-time motion planning of legged robots: A model predictive control approach. In *IEEE-RAS International Conference on Humanoid Robots*, 2017.

[8] Ruben Grandia et al. Feedback MPC for torque-controlled legged robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.

[9] Samuel A. Moore et al. Learning legged MPC with smooth neural surrogates. *arXiv preprint*, 2026.

[10] Kurtland Chua et al. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

[11] Nicklas Hansen et al. Temporal difference learning for model predictive control. In *International Conference on Machine Learning (ICML)*, 2022.

[12] Bo Ai et al. A review of learning-based dynamics models for robotic manipulation. *Science Robotics*, 2025.

[13] Lukas Schneider et al. Learning risk-aware quadrupedal locomotion using distributional reinforcement learning. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[14] Jiyuan Shi et al. Robust quadrupedal locomotion via risk-averse policy learning. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[15] Google DeepMind. MuJoCo playground. *arXiv preprint arXiv:2502.08844*, 2025.

[16] Zeji Yi et al. CoVO-MPC: Theoretical analysis of sampling-based MPC and optimal covariance design. *arXiv preprint*, 2024.

[17] Chaoyi Pan et al. Sampling-based methods for optimal control: Theory, algorithms, and applications. *ICRA Tutorial*, 2025.

[18] Carlos E. Luis et al. Model-based epistemic variance of values for risk-aware policy optimization. *arXiv preprint*, 2023.

[19] Bhavya Sukhija et al. Gradient-based trajectory optimization with learned dynamics. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2023.

[20] Se Hwan Jeon et al. CusADi: A GPU parallelization framework for symbolic expressions and optimal control. *arXiv preprint*, 2024.

[21] Aiden Du et al. GATO: GPU-accelerated and batched trajectory optimization for scalable edge MPC. *arXiv preprint*, 2025.

[22] Kai Ishihara et al. Hierarchical learning framework for whole-body MPC of a real humanoid robot. *arXiv preprint*, 2024.

[23] Cristina Pinneri et al. Sample-efficient cross-entropy method for real-time planning. In *Conference on Robot Learning (CoRL)*, 2021.

[24] Ashish Kumar et al. RMA: Rapid motor adaptation for legged robots. In *Robotics: Science and Systems (RSS)*, 2021.

[25] Ashish Kumar et al. Adapting rapid motor adaptation for bipedal robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022.