# Leveraging Symmetry in RL-based Legged Locomotion Control

Zhi Su*, Xiaoyu Huang*, Daniel Ordoñez-Apraez, Yunfei Li, Zhongyu Li, Qiayuan Liao,
Giulio Turrisi, Massimiliano Pontil, Claudio Semini, Yi Wu, Koushil Sreenath

Project Webpage

IROS '24 ABU DHABI

## Morphological Symmetry

Robots are designed with morphological symmetry.
For example, reflection symmetry group, denoted as,

$$\mathbb{G} := \mathbb{C}_2 = \{e, g_s | g_s^2 = e\}$$

Symmetry transformation/action

Lagrangian function

Symmetry Group

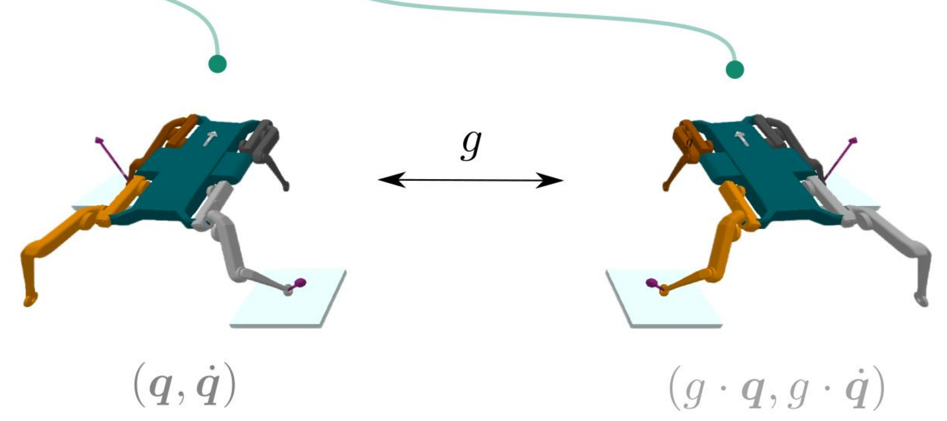
$$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \mathcal{L}(g \cdot \boldsymbol{q}, g \cdot \dot{\boldsymbol{q}}) \;|\forall g \in \mathcal{G}, \boldsymbol{q} \in \mathrm{Q}, \dot{\boldsymbol{q}} \in \mathrm{T}_{\boldsymbol{q}}\mathrm{Q}.$$

System state

Generalized coordinates

Symmetric state

Generalized velocities




## Why is Morphological Symmetry Important

An MDP is considered to have a symmetry group $\mathbb{C}_2$ if:

- The transition function is $\mathbb{C}_2$ -invariant
- The density of initial states is $\mathbb{C}_2$ -invariant
- The reward function is $\mathbb{C}_2$ -invariant

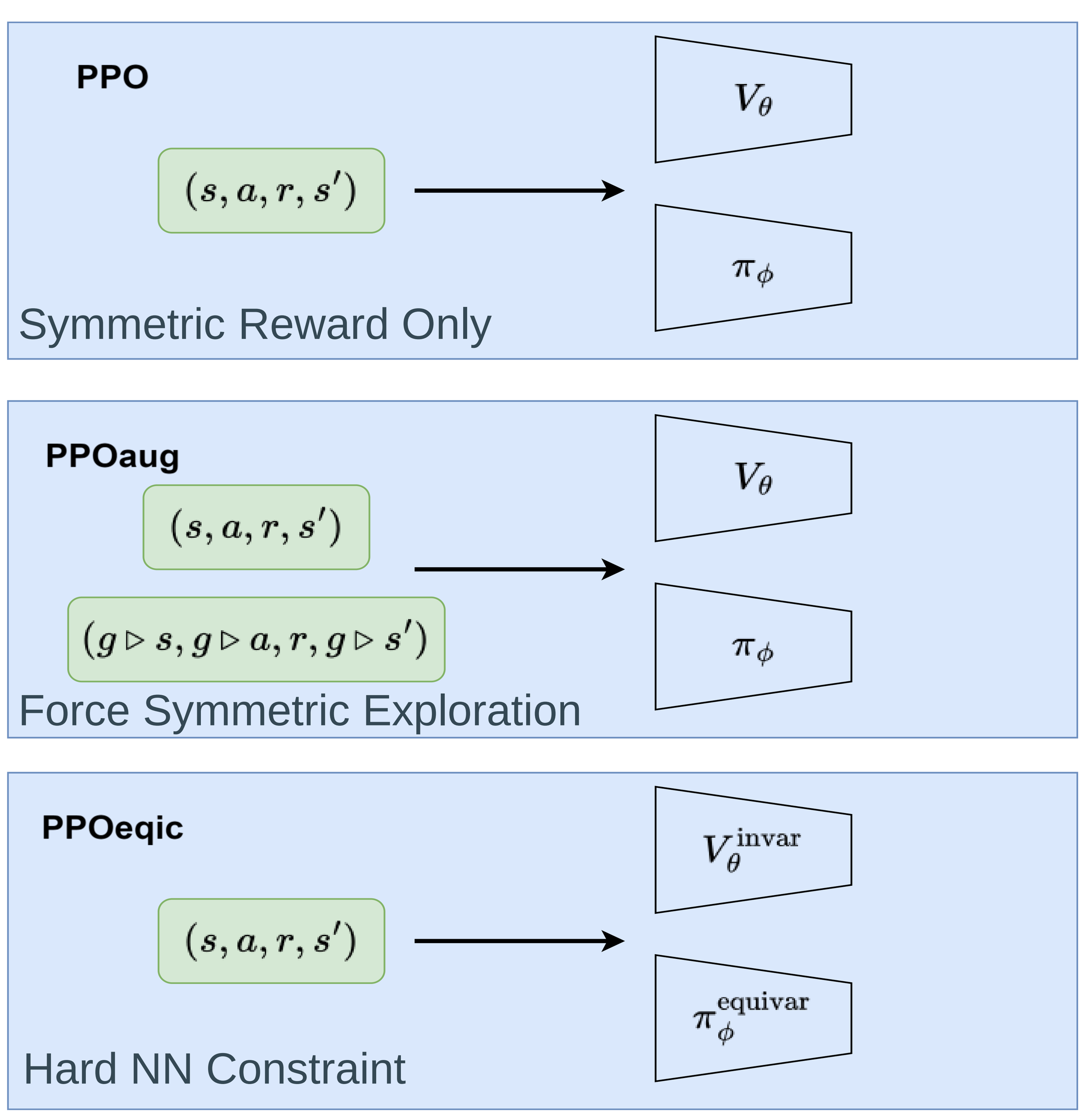
**Task with our symmetric robot can be modeled as symmetric MDP!**

Previous study [2] shows that symmetric MDPs possess $\mathbb{G}$ -equivariant optimal control policies and $\mathbb{G}$ -invariant value function.

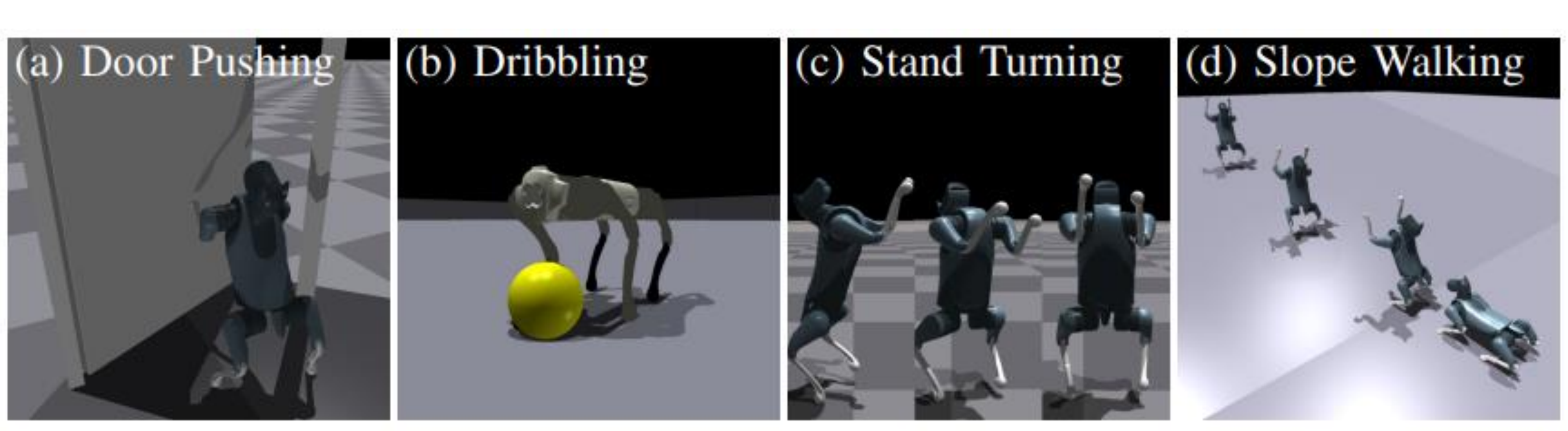
$$\pi(g_s \triangleright s) = g_s \triangleright \pi(s), \quad V(g_s \triangleright s) = V(s)$$

**A strong prior to guide RL exploration and behavior!**

## PPOaug & PPOeqic

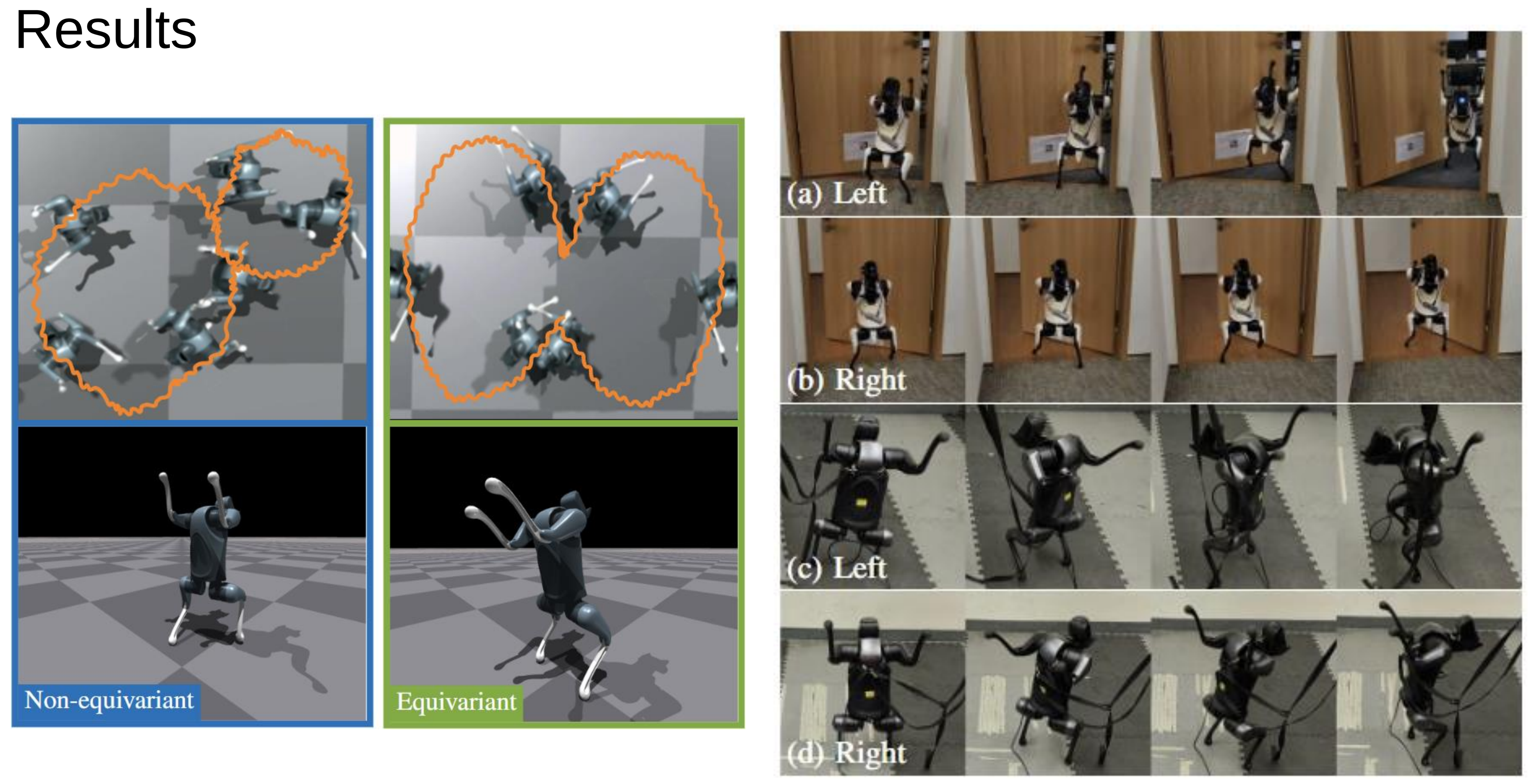



## Challenging Locomotion/Loco-manipulation Tasks

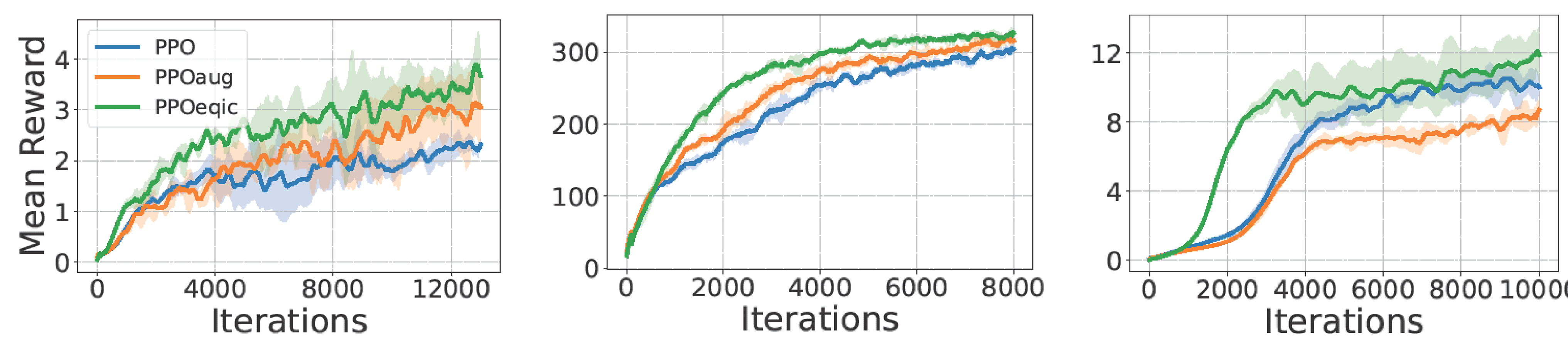



We investigate the efficacy of symmetry-incorporated policies on these four challenging tasks, including loco-manipulation with task-space symmetry and locomotion with motion-level symmetry.

## Results




Incorporating symmetry yields task-level symmetric trajectories, better sim-to-real performance, and more robust behavior.




(a) Door Pushing (b) Dribbling (c) Slope Walking

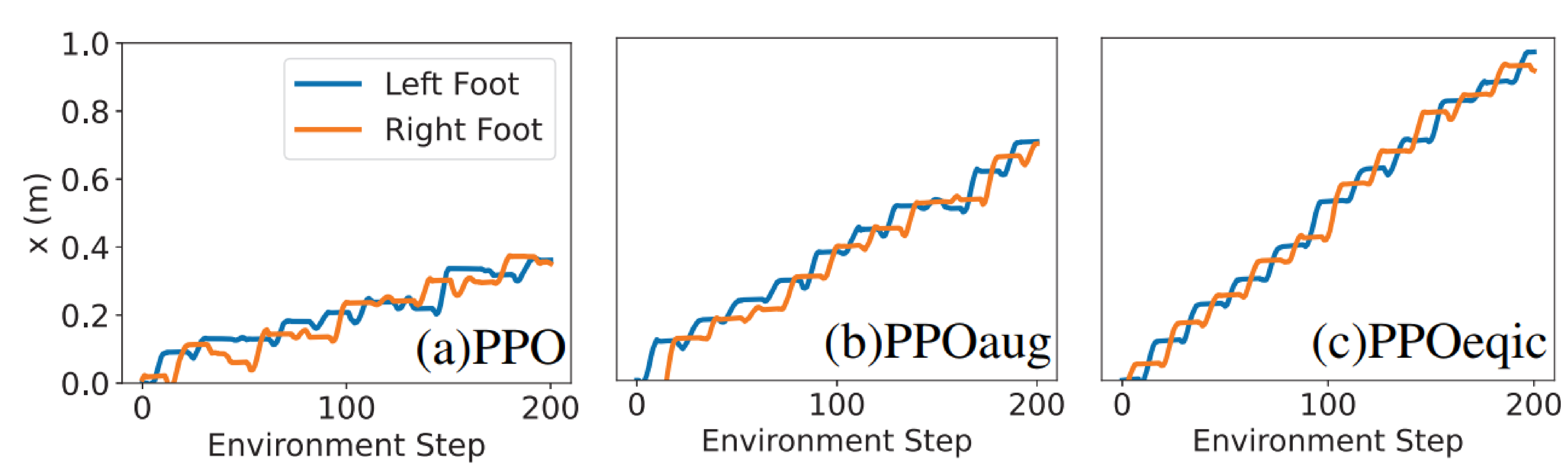
Comparison of training curves of PPO, PPOaug, PPOeqic. PPOeqic consistently demonstrates the highest training returns and sample efficiency in all tasks.

| Method | | Mean SR (%) | Max SR (%) | RSI | OOD Mean SR (%) | OOD Max SR (%) | OOD RSI |
|---|---|---|---|---|---|---|---|
| PPO | trained on 1 side | 43.40 ± 1.73 | 44.47 | 199.96 | 27.46 ± 2.34 | 30.04 | 199.99 |
| | trained on 2 sides | 61.18 ± 7.56 | 69.63 | 12.25 | 42.98 ± 2.21 | 45.51 | 3.45 |
| PPOaug | trained on 1 side | 54.39 ± 32.56 | 86.98 | 3.02 | 38.24 ± 19.17 | 52.85 | 1.40 |
| | trained on 2 sides | 50.24 ± 36.52 | 74.98 | 3.77 | 36.46 ± 25.98 | 53.78 | **0.17** |
| PPOeqic | trained on 1 side | **65.65** ± 23.16 | **87.96** | **0.98** | **44.15** ± 9.39 | 50.63 | 0.88 |
| | trained on 2 sides | 59.92 ± 17.39 | 74.74 | 1.65 | 38.56 ± 15.54 | **55.95** | 0.32 |

Comparison of performance on Door Pushing Task. PPOeqic demonstrates both higher success rate and better symmetry index in both cases, indicating a better task-level symmetric policy.

| Method | | Mean SR (%) | Max SR (%) | RSI | OOD Mean SR (%) | OOD Max SR (%) | OOD RSI |
|---|---|---|---|---|---|---|---|
| PPO | trained on 1 side | 43.40 ± 1.73 | 44.47 | 199.96 | 27.46 ± 2.34 | 30.04 | 199.99 |
| | trained on 2 sides | 61.18 ± 7.56 | 69.63 | 12.25 | 42.98 ± 2.21 | 45.51 | 3.45 |
| PPOaug | trained on 1 side | 54.39 ± 32.56 | 86.98 | 3.02 | 38.24 ± 19.17 | 52.85 | 1.40 |
| | trained on 2 sides | 50.24 ± 36.52 | 74.98 | 3.77 | 36.46 ± 25.98 | 53.78 | **0.17** |
| PPOeqic | trained on 1 side | **65.65** ± 23.16 | **87.96** | **0.98** | **44.15** ± 9.39 | 50.63 | 0.88 |
| | trained on 2 sides | 59.92 ± 17.39 | 74.74 | 1.65 | 38.56 ± 15.54 | **55.95** | 0.32 |

Comparison of performance on Stand Turning Task. PPOeqic demonstrates less error and energyconsumption, indicating a more optimal policy.




Comparison of performance on Slope Walking Task. Plots of the feet positions in the desired walk direction. PPOeqic presents the most symmetric interweaving gait pattern and walks at the desired speed.

## Conclusions

- This work incorporates morphological symmetry as a strong prior in RL to enhance robot behavior and training efficiency.
- We introduce data augmentation and equivariant neural network methods to form equivariant policies under symmetric MDPs.
- Both approaches demonstrate improved performance and robustness in challenging bipedal locomotion and loco-manipulation tasks.

## References

[1] D. Ordoñez-Apraez, M. Martin, A. Agudo, and F. Moreno-Noguer, "On discrete symmetries of robotics systems: A group-theoretic and data-driven analysis," in *Robotics Science and System (RSS)*, 2023.
[2] M. Zinkevich and T. Balch, "Symmetry in markov decision processes and its implications for single agent and multi agent learning." Citeseer, 2001.