Please indicate in what category you would like to participate in the 2024 RAPDASA-RobMech-PRASA-AMI Conference. Delete what is not applicable.

* Academic participation in the Conference with paper taken up in the proceedings of the conference: Extended abstract and full paper required. Choose the track with the best fit for your paper.

Lowering Reinforcement Learning Barriers for Quadruped Locomotion in Task Space Control

**L. L. Cooke1\* & C. Fisher2**

\* Department of Electrical and Electronic Engineering, University of Stellenbosch, South Africa, 22556303@sun.ac.za

2 Department of Electrical and Electronic Engineering, University of Stellenbosch, South Africa, cfisher@sun.ac.za

**ABSTRACT**

In contrast to traditional methods like model predictive control (MPC), deep reinforcement learning (DRL) offers a simpler and less modelling intensive option to develop quadruped locomotion policies. However, DRL presents a steep learning curve. Current research fails to include comprehensive design details, thereby increasing the barrier to entry for further research. This paper aims to facilitate entry into reinforcement learning simulations by illuminating design choices and offering comprehensive implementation details. Results demonstrate that training a quadruped robot in the task space yields natural locomotion and increased sample efficiency compared to conventional joint space frameworks.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\*Corresponding author

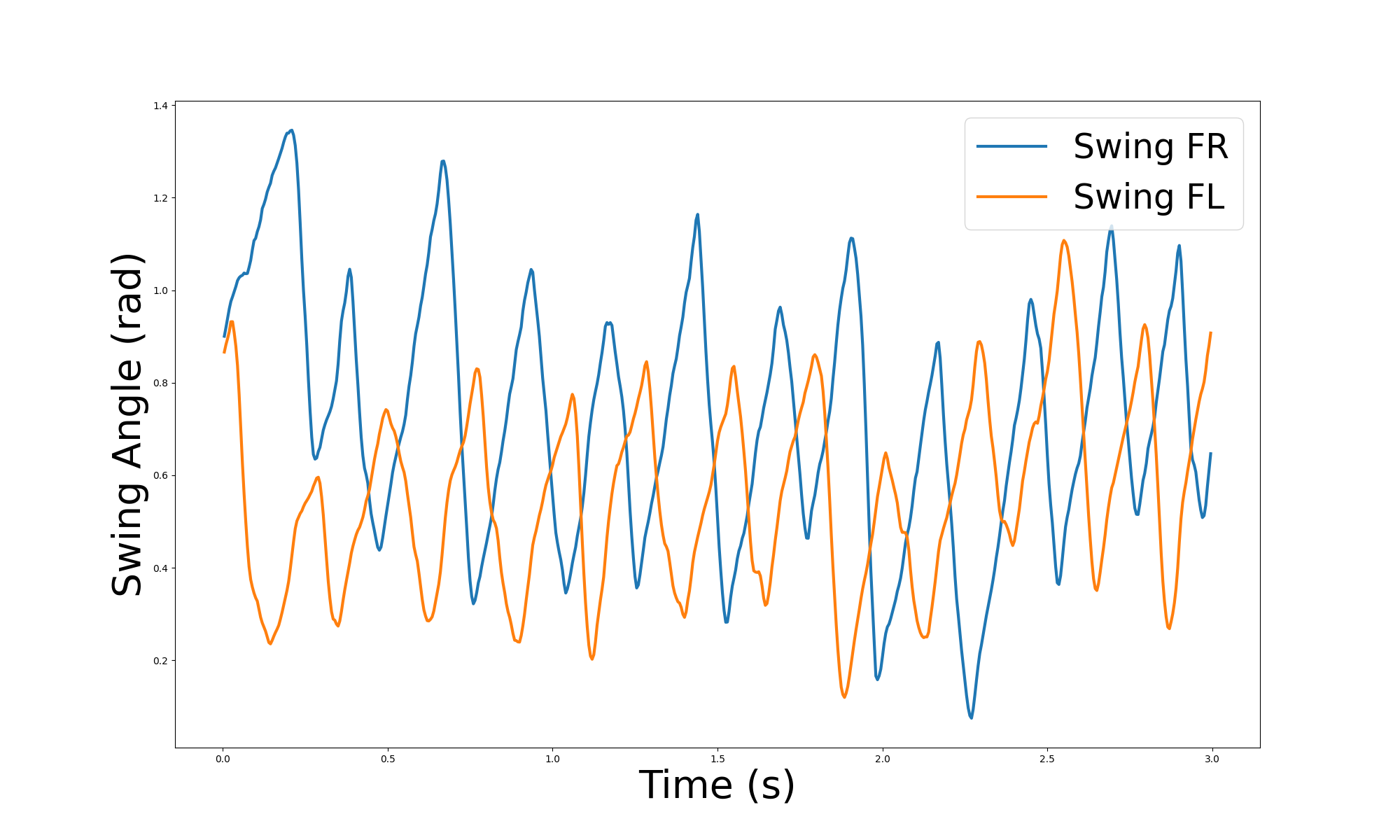
**INTRODUCTION**

Developing robust and adaptive locomotive control policies for legged robots is a well-established and challenging task. Traditional control methods, such as trajectory optimisation [1], state estimation [2], and model-predictive control (MPC) [3], have produced locomotion in varying environments. However, these techniques often require precise dynamic and kinematic modelling which can result in control policies that are arduous to develop. In contrast, deep reinforcement learning (DRL) has demonstrated the ability to create complex robotic control policies without the need for rigorous modelling [4], [5]. Despite reducing the complexity of traditional methods, DRL still presents a substantial entry barrier for new researchers [6]. This paper aims to reduce this barrier by motivating design choices, highlighting failure points, providing comprehensive implementation details, and advocating for the use of the task space as the choice of action space.

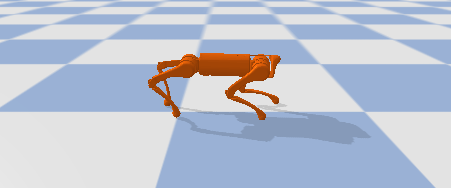
**Methodology and results**

The popular simulation environment, Pybullet [7], along with the Unitree A1 robot platform [8] were used to create a simulation environment for training, as indicated in Figure 1. To learn a policy, the Stable-Baselines3’s implementation of PPO [9] was used with minimal hyperparameter tuning. The RL agent selects actions to maximise a reward signal which consists of a forward velocity reward as well as an energy, orientation, and height penalty. The actions consist of foot positions in the robot’s leg frame i.e. the task space. These foot positions are then mapped to torques using kinematic relations and PD control. This framework was compared to an environment that utilised the joint space, as demonstrated in several other studies [4], [10]. In this environment the agent selects torques that are directly applied to the robot.

The task space RL framework achieved natural looking walking gaits, as shown in Figure 2. In comparison, the joint space agent produced locomotion that appeared less natural and inefficient, whilst requiring more samples. We observed that both reinforcement learning frameworks require considerable tuning of the control frequency, termination conditions and reward parameters to produce locomotion.

****

**Figure 2: Front leg swing angles for 3 seconds during a walk**

****

**Figure 1: Unitree A1 robot walking in simulation.**

**CONCLUSION**

The task space is an appropriate choice of action space for locomotion tasks, as it enables reduced reward shaping, and improved sample efficiency compared to typical joint space RL frameworks. The detailing of all design parameters can also expedite other research and enable further progress in more complex locomotion tasks. Furthermore, details such as termination conditions, policy frequency and action space choice can provide further research questions to advance the study of quadruped locomotion.

**REFERENCES**

**[1]** **Posa, M., Cantu, C., & Tedrake, R.** 2014. A direct method for trajectory optimization of rigid bodies through contact, *Int. J. Robot. Res.*, 33(1), pp. 69–81, doi: 10.1177/0278364913506757.

**[2]** **Erez, T., Lowrey, K., Tassa, Y., Kumar, V., Kolev, S., & Todorov, E.**, 2013. An integrated system for real-time model predictive control of humanoid robots, in *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, Atlanta, pp. 292–299. doi: 10.1109/HUMANOIDS.2013.7029990

**[3]** **Bloesch, M., *et al.*** 2013. State estimation for legged robots-consistent fusion of leg kinematics and IMU, *Robotics*, 17, pp. 17–24..

**[4]** **Fu, Z., Kumar, A., Malik, J., & Pathak, D.** 2021. Minimizing Energy Consumption Leads to the Emergence of Gaits in Legged Robots. Accessed: Oct. 23, 2023. [Online]. Available: http://arxiv.org/abs/2111.01674.

**[5]** **Bellegarda, G., Chen, Y., Liu, Z., & Nguyen, Q.** 2022. Robust High-Speed Running for Quadruped Robots via Deep Reinforcement Learning, *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Kyoto, Japan: IEEE, pp. 10364–10370. doi: 10.1109/IROS47612.2022.9982132.

**[6]** **Henderson, P., Islam, R., Bachman, P., Pineau, J., Precup, D., & Meger, D.** 2018. Deep Reinforcement Learning That Matters, *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), doi: 10.1609/aaai.v32i1.11694.

**[7]** **Coumans, E., & Bai, Y.** 2016 - 2023. PyBullet, a Python module for physics simulation for games, robotics and machine learning, [Online]. Available: http://pybullet.org.

**[8]** **Unitree**. A1 - highly integrated, pushing limits’. Accessed: Mar. 8, 2024. [Online]. Available: https://m.unitree.com.

**[9]** **Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., & Dormann, N.** 2021. Stable-Baselines3: Reliable Reinforcement Learning Implementations, *Journal of Machine Learning Research*, 22, pp. 1–8.

**[10]** **Haarnoja, T., Ha, S., Zhou, A., Tan, J., Tucker, G., & Levine, S.** 2019 Learning to Walk via Deep Reinforcement Learning. Accessed: Feb. 21, 2024. [Online]. Available: http://arxiv.org/abs/1812.11103.