

# Project Proposal

## Pin-Point Autonomous Lunar Landing

Silverio Manganaro / 1817504

Amine Ahardane / 2050689

November 25, 2023

### 1 Theory

We would like to replicate the reinforcement learning (RL) approach outlined in [1] using the identical simulation environment. We have read that this environment is built on pyBullet and is accessible online. However we have been unable to find it. We kindly request if it is possible to share it with us.

We plan to initiate our work by employing Advantage Actor-Critic (A2C) methodology as an initial framework. Subsequently, we intend to transition to the actual Soft Actor-Critic (SAC) algorithm [2], known for its ability to handle continuous action spaces and address various challenges in reinforcement learning.

However, our ultimate objective is to advance beyond traditional Soft Actor-Critic by incorporating some of the innovations presented in [3]. These modifications specifically focuses on enhancing the Experience Replay (ER) buffer of Soft Actor-Critic.

We are also considering the feasibility to implement an asynchronous version of the Actor-Critic [4].

### References

- [1] G. Ciabatti, D. Spiller, S. Daftry, R. Capobianco, and F. Curti, “Deep reinforcement learning for pin-point autonomous lunar landing: Trajectory recalculation for obstacle avoidance,” in *The Use of Artificial Intelligence for Space Applications*, C. Ieracitano, N. Mammone, M. Di Clemente, M. Mahmud, R. Furfaro, and F. C. Morabito, Eds. Cham: Springer Nature Switzerland, 2023, pp. 101–115.
- [2] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” 2018.
- [3] C. Banerjee, Z. Chen, and N. Noman, “Improved soft actor-critic: Mixing prioritized off-policy samples with on-policy experience,” 2021.
- [4] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, “Asynchronous methods for deep reinforcement learning,” 2016.