

IAF ASTRODYNAMICS SYMPOSIUM (C1)  
Attitude Dynamics (2) (2)

Author: Mr. Jan Loettgen

School of Engineering, University of Glasgow, United Kingdom, 2237859L@student.gla.ac.uk

Dr. Matteo Ceriotti

University of Glasgow, United Kingdom, matteo.ceriotti@glasgow.ac.uk

Dr. Gerardo Aragon-Camarasa

University of Glasgow, United Kingdom, gerardo.aragoncamarasa@glasgow.ac.uk

Dr. Kevin Worrall

University of Glasgow, United Kingdom, kevin.worrall@glasgow.ac.uk

APPLICATION OF DEEP REINFORCEMENT LEARNING FOR ATTITUDE CONTROL OF A  
SATELLITE IN THE PRESENCE OF UNCERTAINTIES**Abstract**

Classical attitude control strategies such as PID control require careful gain tuning for every specific satellite and manoeuvre. Modern Deep Reinforcement Learning (DRL) has shown to be a high-performance control strategy and a method that does not require parameter tuning for every specific satellite and manoeuvre. However, these results were obtained in simulation for highly idealised scenarios. This paper investigates the performance of DRL attitude controllers trained in idealised simulations when deployed into non-ideal, noisy simulations and explores whether training in noisy simulations improves performance. We also investigate how discrepancies in the inertia tensor between the training satellite and the final testing satellite affect the performance of the DRL attitude controller. Finally, the robustness against large impulse torques and small constant torques is investigated for both the cases where these torques are seen during training and where they are not.

The test case considered is a large-angle slew manoeuvre of a 6U-Cubesat, with reaction wheel actuation and sensors that directly measure the satellite angular velocity and attitude quaternion. The large-angle slew manoeuvre is formulated as a finite-horizon Markov Decision Process. The Proximal Policy Optimisation algorithm is used to train an attitude controller in simulation, to solve the Markov Decision Process. The idealised training satellite has noiseless sensors that measure the attitude quaternion and angular velocity exactly and actuators that produce the commanded control torques. The sensors on the non-idealised satellite have Gaussian white noise superimposed on their measurements. The non-idealised reaction wheels have Gaussian white noise superimposed onto their control torques.

Preliminary results were obtained by training an artificial neural network attitude controller using the Deep-Q-Network algorithm. Gaussian white noise with a standard deviation of 0.05 and a mean of 0 was applied to sensor measurements as described above. Preliminary findings for the controller trained on the ideal satellite show an increase in the mean pointing error from  $1.915^\circ$  to  $22.80^\circ$ , with the standard deviation increasing from  $0.84^\circ$  to  $11.59^\circ$ . When the controller is trained in a noisy environment, the mean pointing error was  $18.58^\circ$  with a standard deviation of  $11.33^\circ$ . This shows that the controller's performance improved when trained in a noisy environment compared to when trained in an ideal environment and then deployed into a noisy environment. We are investigating state of the art continuous action space DRL attitude controllers that achieve pointing accuracies of  $0.025^\circ$ , when trained in an idealised environment using a Proximal Policy Optimisation algorithm.