

# Fast control of plasma vertical displacement based on robust adversarial reinforcement learning

Thursday, 18 July 2024 09:30 (20 minutes)

Plasma with elongated configuration has the advantage of higher discharge parameters while at the cost of vertical displacement instability. Once the vertical displacement is out of control, it will inevitably lead to a major disruption, causing great damage to the device, which will have unacceptable consequences if it occurs on ITER. Therefore, active control of vertical displacement is necessary. The vertical displacement is affected by the passive structure, power supply delay, etc., which is a high-order system with complex response. As the system control ability is limited, when the perturbations are complex and diverse, the requirements for robustness of controlling are high. Deep learning has a strong learning capability, so we used a deep reinforcement learning approach to achieve fast control of plasma vertical displacement.

We first verified the feasibility of reinforcement learning to control plasma vertical displacement. We trained the vertical displacement controller using the Deep Deterministic Policy Gradient (DDPG) algorithm and tested its performance. After testing, we found that the dynamic response of the controller is better than the conventional PID control, but it is less resistant to PF coil current perturbations.

In order to increase the perturbation resistance of the model, we have adopted Robust Adversarial Reinforcement Learning (RARL). The strategy of RARL is to add an adversary who is also an agent, and the adversary will attack the weaknesses of the agent, so the agent needs to find the optimal strategy in the worst case scenario. It may be useful to refer to the DDPG-based RARL as DDPG-RARL. The traditional vertical displacement control cannot completely avoid the overcurrent of IC coil due to the perturbation of PF coil current. Therefore, in our work, the adversary attacks the controller by applying perturbations to the PF coil current based on the observations in the EAST.

We perform a comparative test of the model's resistance to perturbation by using an adversary to attack the DDPG-RARL-based controller, and then intercepting the attack pattern to attack the DDPG-based controller. We found that the training process yields adversaries with different characteristics. The adversaries can be categorized into two types, performing high-amplitude attacks and high-frequency attacks. We found that DDPG-RARL outperforms DDPG for both large amplitude attacks and high frequency attacks.

## Speaker's Affiliation

Institute of Plasma Physics, Chinese Academy of Sciences, Hefei

## Member State or IGO

China, People's Republic of

**Primary authors:** Prof. XIAO, Bingjia (Institute of Plasma Physics, Chinese Academy of Sciences); Dr LIU, Binnuo (Institute of Plasma Physics, Chinese Academy of Sciences); Prof. YUAN, Qiping (Institute of Plasma Physics, Chinese Academy of Sciences); Dr ZHANG, Ruirui (Institute of Plasma Physics, Chinese Academy of Sciences); Dr HU, Wenhui (Institute of Plasma Physics, Chinese Academy of Sciences); Dr HUAN, Yao (Institute of Plasma Physics, Chinese Academy of Sciences); Dr WANG, Yuehang (Institute of Plasma Physics, Chinese Academy of Sciences); Dr LUO, Zhengping (Institute of Plasma Physics, Chinese Academy of Sciences)

**Presenter:** Dr LIU, Binnuo (Institute of Plasma Physics, Chinese Academy of Sciences)

**Session Classification:** Machine Learning

**Track Classification:** Machine Learning