

5th ICFA Beam Dynamics Mini-Workshop on Machine Learning for Particle Accelerators



Contribution ID: 75

Type: **Invited talks**

Explainable physics-based constraints on reinforcement learning for accelerator controls-15'+5'

Wednesday 9 April 2025 11:40 (20 minutes)

We present a reinforcement learning (RL) framework for controlling particle accelerated experiments that builds explainable physics-based constraints on agent behavior. The goal is to increase transparency and trust by letting users verify that the agent's decision-making process incorporates suitable physics. Our algorithm uses a learnable surrogate function for physical observables, such as energy, and uses them to fine-tune how actions are chosen. This surrogate can be represented by a neural network or by a sparse dictionary model. We test our algorithm on a range of particle accelerator controls environments designed to emulate the Continuous Electron Beam Accelerator Facility (CEBAF) at Jefferson Lab. By examining the mathematical form of the learned constraint function, we are able to confirm the agent has learned to use the established physics of each environment. In addition, we find that the introduction of a physics-based surrogate enables our reinforcement learning algorithms to reliably converge for difficult high-dimensional accelerator controls environments.

Authors: COLEN, Jonathan (Old Dominion University); SCHRAM, Malachi; RAJPUT, Kishansingh (Jefferson Lab); KASPARIAN, Armen (JLab)

Presenter: COLEN, Jonathan (Old Dominion University)

Session Classification: Optimisation and Control

Track Classification: Optimisation and Control