3rd collaboration workshop on Reinforcement Learning for Autonomous Accelerators (RL4AA'25)

Wednesday, April 2, 2025 - Friday, April 4, 2025

DESY

Book of Abstracts

Contents

Talks / 1

Exploring Reinforcement Learning for Optimal Bunch Merge in the AGS

Author: Yuan Gao[None]

Co-authors: Armen Kasparian [1]; Daria Kuzovkova [2]; Eiad Hamwi [2]; Freddy F Severino [3]; Georg Hoffstaetter [2]; John Morris [3]; Jonathan Unger [2]; Keith Zeno [3]; Kevin Brown [3]; Levente Hajdu [3]; Malachi Schram [4]; Matthew Signorelli [2]; Vincent Schoefer [3]; Weijian Lin [3]; Xiaofeng Gu [3]; Yinan Wang [5]

[1] *JLab*

[2] *Cornell Univ.*

[3] *Brookhaven National Lab*

[4] *Jefferson Lab*

[5] *RPI*

Corresponding Authors: mgs255@cornell.edu, eh652@cornell.edu, wlin1@bnl.gov, georg.hoffstaetter@cornell.edu, ygao@bnl.gov, schram@jlab.org, brownk@bnl.gov, jeu8@cornell.edu, jtm@bnl.gov, severino@bnl.gov, lbhajdu@bnl.gov, armenk@jlab.org, zeno@bnl.gov, wangy88@rpi.edu, xgu@bnl.gov, dk638@cornell.edu, schoefer@bnl.gov

In BNL's Booster, the beam bunches can be split into two or three smaller bunches to reduce their space-charge forces. They are then merged back after acceleration in the Alternating Gradient Synchrotron (AGS). This acceleration with decreased space-charge forces can reduce the final emittance, increasing the luminosity in RHIC and improving proton polarization. Parts of this procedure have already been tested and are proposed for the Electron-Ion Collider (EIC). The success of this procedure relies on a series of RF gymnastics to merge individual source pulses into bunches of suitable intensity. In this work, we explore an RF control scheme using reinforcement learning (RL) to merge bunches, aiming to dynamically adjust RF parameters to achieve minimal longitudinal emittance growth and stable bunch profiles. Initial experimental results and ongoing system developments are presented and discussed.

Poster session / 2

Comparative Study of Machine Unlearning Techniques for Computer Vision and NLP Models

Author: Mahule Roy[1]

[1] *National Institute of Technology Karnataka Surathkal*

Corresponding Author: roymahule26@gmail.com

Machine unlearning is an emerging field in machine learning that focuses on efficiently removing the influence of specific data from a trained model. This capability is critical in scenarios requiring compliance with data privacy regulations or when erroneous data needs to be removed without retraining from scratch. In this study, I explore the importance of machine unlearning as a way to enhance privacy simultaneously not affecting the efficiency of machine learning models. Using the CIFAR- 10 and CIFAR-100 dataset, I implemented various unlearning methods like retraining on the retained set, instruction fine tuning a LLM model to forget biased sentences and distillation techniques. These methods allowed the models to forget specific contexts while not comprising on the model accuracy. My implementations yielded promising results in terms of unlearning effectiveness and I have used various unlearning metrics to compare with my implementation and the baseline performance. The outcomes demonstrate the potential these methods have to balance between privacy and model accuracy effectively.

Poster session / 3

Parallel Simulations for Faster Reinforcement Learning of Accelerator Control

Author: Finn O'Shea[1]

Co-authors: Auralee Edelen [2]; Christopher Hall [3]; Jonathan Edelen [3]

[1] *SLAC National Accelerator Laboratory*

[2] *SLAC*

[3] *RadiaSoft LLC*

Corresponding Authors: edelen@slac.stanford.edu, jedelen@radiasoft.net, chall@radiasoft.net, foshea@slac.stanford.edu

Reinforcement Learning methods typically require a large number of interactions with the environment to learn anything useful. This makes learning with sophisticated accelerator simulations difficult because of the total time required to train. On the other hand, learning with environments based on these accelerator codes is potentially very useful because they contain a lot of knowledge about accelerator systems. To ameliorate the problem of long wall-clock run times for these codes, we are using an Advantage Actor Critic (A2C) method to train an agent using the 2D accelerator code spiffe. Our end goal is to train an agent to control the FLEX accelerator, which is a candidate for FLASH radiation treatment. Herein we describe our progress, starting with learning at scale on simulations of the accelerator where we train an agent using one hundred simulations running in parallel.

Talks / 4

Geoff: Applications and Developments in 2024

Author: Penny Madysa[1]

Co-authors: Christoph Reinwald ; Daniel Kallendorf [1]; Erika Kazantseva [1]; Helmut Weick [1]; Martin Bajzek [1]; Sabrina Appel [2]; Stephane Pietri [1]

[1] *GSI Helmholtzzentrum für Schwerionenforschung*

[2] *GSI*

Corresponding Authors: h.weick@gsi.de, e.kazantseva@gsi.de, s.pietri@gsi.de, reinwald.christoph2019@outlook.de, m.bajzek@gsi.de, s.appel@gsi.de, n.madysa@gsi.de, daniel@kallendorf.com

The complexity of the GSI/FAIR accelerator facility demands a high level of automation in order to maximize time for physics experiments. Accelerator laboratories world-wide are exploring a variety of techniques to achieve this, from classical optimization to reinforcement learning.

Geoff, the Generic Optimization Framework & Frontend, is an open-source framework that harmonizes access to automation techniques and simplifies the transition towards them. It is maintained as part of the EURO-LABS project in cooperation between CERN and GSI.

We report on results that have been achieved with Geoff at GSI in 2024. The multi-turn injection of the SIS18 synchrotron has been optimized via multi-objective Bayesian optimization for the first time and a Pareto front has been built from real data. The existing optimization has also been analyzed in more detail. In addition, the use of a data-driven Gaussian Process Model Predictive Control (GP-MPC) framework has been studied in simulation.

We have also successfully used Geoff for beam centering and focusing at the GSI Fragment Separator (FRS). This task involved communication with multiple controls systems and between the different networks of the accelerator and the experiment complex. Algorithms as varied as track classification, distribution fitting and black-box optimization were used in tandem and demonstrate the flexibility of Geoff in the face of non-trivial user requirements.

In addition, Geoff has undergone a major update in 2024 that brings it in line with the latest developments of numerical and machine-learning software in the Python ecosystem.

Talks / 5

Autonomous Trajectory Steering of DC Beams at CERN's SPS Transfer Lines Using Reinforcement Learning

Authors: Adrian Menor de Onate[1]; Michael Schenk[1]; Niky Bruchon[None]; Verena Kain[1]

[1] *CERN*

Corresponding Authors: michael.schenk@cern.ch, br.niky@gmail.com, verena.kain@cern.ch, adrian.menor.de.onate@cern.ch

The slow extracted beams at the CERN Super Proton Synchrotron (SPS) are transported over several 100 m long transfer lines to three targets in the North Area Experimental Hall. The experiments require to eliminate intensity fluctuations over the roughly 5 s particle spill and hence to debunch the extracted beams. In this environment, secondary emission monitors (SEMs) have to replace the conventional beam position monitoring systems that rely on RF structure. Such monitors can be used to infer the intensity difference between two split foils, but do not readily provide position readings. Moreover, when the beam ends up on one of the foils, determining the appropriate corrector magnet settings remains a challenging task, as it is not possible to directly infer the beam deviation from the SEMs. In such scenarios, traditional trajectory control algorithms fail.
This paper summarises the application of reinforcement learning (RL) to successfully correct the beam trajectory using SEM readings. The RL policy is learnt offline in simulation, and can be successfully transferred to the real environment. Moreover, the RL policy is also tested in scenarios where the beam is lost in the line, and threading actions are needed. Results of the application of the RL policies in the real transfer line, and the different tests carried out in simulation are presented.

Talks / 6

Opportunities for Reinforcement Learning in Accelerator Automation

Authors: Jonathan Edelen[1]; Morgan Henderson[1]

[1] *RadiaSoft LLC*

Corresponding Authors: mhenderson@radiasoft.net, jedelen@radiasoft.net

For more than half a decade, RadiaSoft has developed machine learning (ML) solutions to problems of immediate, practical interest in particle accelerator operations. These solutions include machine vision through convolutional neural networks for automating neutron scattering experiments and several classes of autoencoder networks for de-noising signals from beam position monitors and low-level RF systems in the interest of improving and automating controls. As active deployments of our ML products have taken shape, one area which has become increasingly promising for future development is the use of agentic ML through reinforcement learning (RL). Leveraging our substantial suite of ML tools as a foundation, we have now begun to develop an RL framework for achieving higher degrees of automation for accelerator operations. Here we discuss our RL approaches for two areas of ongoing interest at RadiaSoft: total automation of sample alignment at neutron and x-ray beamlines, and automated targeting and dose delivery optimization for FLASH radiotherapy. We will provide an overview of both the ML and RL methods employed, as well as some of our early results and intended next steps.

Poster session / 7

Model Predictive Control with Gaussian Processes for Safe and Efficient RL: A Case Study at the CERN SPS

Author: Ferdinand Ferber[1]

[1] *CERN*

Corresponding Author: ferdinand.ferber@cern.ch

Classical, model-free Reinforcement Learning (RL) has achieved impressive results in areas where interactions with the environment are inexpensive, such as computer games or simulations. However, in many real-world applications, such as robotics or autonomous particle accelerators, interactions with the system are costly, which creates a need for sample-efficient RL algorithms. In addition, safety constraints must be respected even during the exploratory phases of the RL agent. To address both concerns, Reinforcement Learning based on Model Predictive Control using Gaussian Processes (GP-MPC) was proposed by Kamthe and Deisenroth in 2018. We present an implementation of this approach that is compatible with the Gymnasium API and test it on the problem of correcting power supply ripples (multiples of 50 Hz), which negatively affect the slow extraction spill from the CERN Super Proton Synchrotron (SPS) to the fixed-target physics experiments hosted in the North Experimental Area (NA) at CERN.

Talks / 8

A Benchmark for Deep Reinforcement Learning-Based Control of Liquid-Propellant Rocket Engines

Author: Kai Dresia[None]

Co-authors: Günther Waxenegger-Wilfing ; Jonas Dauer ; Simon Hirlaender ; Vincent Bareiß

Corresponding Authors: jonas.dauer@dlr.de, guenther.waxenegger@dlr.de, kai.dresia@dlr.de, bareiss.vincent@dlr.de, simon.hirlaender@plus.ac.at

Deep reinforcement learning (DRL) has demonstrated great potential for controlling and regulating complex real-world systems such as nuclear fusion reactors such as tokamaks and particle accelerators. Another promising application is the DRL-based control of liquid-propellant rocket engines (LPREs), which have been a focus of research at the German Aerospace Center (DLR) for the past six years. LPREs are safety-critical systems, where reliability and robustness are of utmost importance. A key difficulty in these systems is the discrepancy between simulation models and real-world behavior, combined with limited availability of real data. An ideal DRL-based controller should be capable of adapting to such errors to ensure robustness and reliability.

To address these challenges, we present a benchmark for LPRE control designed to evaluate DRL-based control strategies. This benchmark includes simulation software calibrated with experimental data, a dataset for fine-tuning, and the ability to simulate representative errors in both sensors and the system itself. The findings from this benchmark are expected to be transferable to particle accelerators and similar systems.

The benchmark will be made freely accessible to the RL community, fostering further research in robust control applications. Our poster outlines the essential steps for deploying DRL controllers on real rocket engines and provides an overview of the benchmark's components and capabilities.

Talks / 9

Offline Reinforcement Learning-Based Control of LEIR Injection Efficiency via Data-Driven Surrogate Modeling

Author: Borja Rodriguez Mateos[1]

Co-authors: Michael Schenk [1]; Verena Kain [1]

[1] *CERN*

Corresponding Authors: verena.kain@cern.ch, borja.rodriguez.mateos@cern.ch, michael.schenk@cern.ch

Aging of the stripper foil and unexpected machine shutdowns are the primary causes for reduction of the injected intensity from CERN's Linac3 into the Low Energy Ion Ring (LEIR). As a result, the set of optimal control parameters that maximizes beam intensity in the ring tends to drift, requiring daily adjustments to the machine control settings. This paper explores the design of a Reinforcement Learning (RL) based auto-pilot that compensates for the drift of parameters and maintains a optimal beam intensity in the LEIR ring. It observes Time of Flight (ToF) measurements of the ion beam in the linac and Schottky signals in the ring to act on the relevant control knobs.

This autonomous agent is pre-trained on a data-driven surrogate model built from historical exploration of the high dimensional parameter space. A comparison of the performances of several RL algorithms will be done on different surrogate models designs to evaluate future design of the operational autopilot. This work holds promise for pre-training offline RL agents through data-driven surrogate models on tasks that are too complex or computationally expensive to simulate.

Poster session / 10

Leveraging Reinforcement Learning for Safe Navigation of Mobile Robots in Accelerator Tunnels

Author: André Dehne[1]

Co-authors: Hannes Voß [2]; Peer Stelldinger [2]

[1] *HAW Hamburg*

[2] *HAW-Hamburg*

Corresponding Authors: peer.stelldinger@haw-hamburg.de, hannes.voss@haw-hamburg.de, andre.dehne@haw-hamburg.de

The integration of mobile autonomous robots in accelerators introduces potential risks to the facility itself, including collisions with critical components, cables, and infrastructure. Such incidents could compromise the functionality and safety of the accelerator, necessitating robust solutions to mitigate these risks. This paper explores how Reinforcement Learning (RL) can be leveraged to enable autonomous vision-based navigation for mobile robots while incorporating self-supervised anomaly detection to enhance operational safety. Exemplified by the robotic platform "MARWIN" at the European XFEL, the findings provide an analysis of key factors influencing RL applicability, including adaptability, real-time performance, and training requirements. Concluding with practical implementation suggestions, this work aims to offer a roadmap for leveraging RL to develop safe and reliable robotic navigation systems tailored to the unique needs of accelerators.

Poster session / 12

Do you really think about consequences? Bridging Classical Control and Reinforcement Learning for Delayed Outcome Optimisation

Authors: Olga Mironova[1]; Simon Hirlaender[1]

Co-authors: Lorenz Fischl [1]; Thomas Gallien [2]

[1] *PLUS University Salzburg*

[2] *JOANNEUM RESEARCH*

Corresponding Authors: olga.mironova@plus.ac.at, lfischl@pm.me, simon.hirlaender@plus.ac.at, thomas.gallien@joanneum.at

This study explores advanced strategies for optimal control in systems with delayed consequences, using beam steering in the AWAKE electron line at CERN as a benchmark. We formulate the task as a constrained optimization problem within a continuous, primarily linear Markov Decision Process (MDP), incorporating measured system parameters and realistic termination criteria. A wide range of approaches is implemented and compared, including classical response matrix inversion, control-theoretic methods, reinforcement learning, and structured model-based techniques.

While classical methods like matrix inversion offer accurate convergence, they fail to account for delayed effects and are sensitive to noise. Control-theoretic approaches, such as Model Predictive Control (MPC), leverage known dynamics and handle delays effectively when models are available. Data-driven methods, including Proximal Policy Optimization (PPO), adapt to uncertainty and non-linearities but require large amounts of data. Structured GP-MPC bridges both paradigms by learning system dynamics using Gaussian Processes while respecting the problem's causal structure, significantly improving robustness and sample efficiency.

Our experiments highlight key performance differences, particularly in how each method handles delayed outcomes, noise, and structural assumptions. We find that exploiting the causal structure of the problem provides a notable advantage, and that method choice ultimately involves trade-offs between adaptability, data efficiency, and computational cost. These findings offer guidance for applying advanced control strategies in high-dimensional, partially structured environments.

Talks / 13

Generalizing LLMs for Accelerators with RL

Author: Antonin Sulc[1]

Co-author: Thorsten Hellert [1]

[1] *LBNL*

Corresponding Authors: thellert@lbl.gov, asulc@lbl.gov

Recent advances in fine-tuning large language models (LLMs) with reinforcement learning (RL) techniques have demonstrated their ability to generalize, unlike the often-used Supervised Fine-Tuning (SFT).
Many aspects of particle accelerators, such as beam parameters, have well-defined objectives, making them ideal candidates for RL-driven optimization.

In this work, we explore the capabilities of current open-source LLMs fine-tuned to understand the peculiarities of the ALS architecture.
We identify several optimization objectives that can be beneficial and create a small dataset to benchmark our hypothesis.
Our goal is to demonstrate how RL can enhance an LLM's ability to interpret accelerator states, optimize performance, and provide intelligent insights into beam dynamics.

Poster session / 14

Autonomous Optimization of Longitudinal Triple Splitting in the Proton Synchrotron Using Reinforcement Learning

Authors: Alexandre Lasheen[1]; Joel Axel Wulff[1]

[1] *CERN*

Corresponding Authors: joel.wulff@cern.ch, alexandre.lasheen@cern.ch

Achieving precise bunch spacing in the Large Hadron Collider (LHC) relies on advanced RF manipulations in the Proton Synchrotron (PS). Multiple RF systems covering a large range of revolution harmonics (7 to 21, 42, 84) allow performing bunch splitting manipulations. To minimize bunch-by-bunch variations in intensity, longitudinal emittance, and shape, precise tuning of relative RF amplitude and phase settings is essential. Since 2023, a machine-learning-based system has been deployed in the PS to automate the optimization of one such manipulation, the longitudinal triple splitting, utilizing deep Reinforcement Learning (RL) and Convolutional Neural Networks (CNN). As part of an effort to fully automate PS operations for LHC-beams, new longitudinal monitoring systems have enabled the transition from on-demand optimization to an autonomous, continuously monitoring system. Pre-trained RL models, with weights hosted centrally on the Machine Learning Platform (MLP), seamlessly interact with accelerator observation and control systems through the Unified Controls Acquisition and Processing (UCAP) framework. This work demonstrates a practical implementation for how to deploy ML-based models for autonomous control of complex accelerator processes, which is likely to play an essential part of future operational efficiency and stability in high-energy physics facilities.

Talks / 15

Reinforcement Learning for Laser Alignment

Author: Matthew Schwab[1]

[1] *FSU Jena*

Corresponding Author: matthew.schwab@uni-jena.de

Manual alignment of optical systems can be time consuming and the achieved performance of the system varies depending on the operator doing the alignment. A reinforcement learning approach using the PPO algorithm was used to train agents to align simple two-mirror optical setups, as well as a full regenerative laser amplifier. The goal is to produce agents that can reproducibly align the setup faster than a human and can correct long-term drifts in laser energy (time scale of approx. one hour) during operation. The work is still ongoing. Agents have been successfully implemented on hardware in the two-mirror setup, showing "super-human"performance in alignment time. The agents successfully "learn"to handle a significant amount of mechanical backlash in the used stepper motors and mirror mounts. Currently, the necessary hardware is being installed on a regenerative amplifier and agents are being further developed for this use case.

Poster session / 16

Robust Multi-Turn Injection at SIS18 via Gaussian Process Model Predictive Control

Authors: Benjamin Halilovic[None]; Sabrina Appel[1]; Simon Hirlaender[2]

[1] *GSI*

[2] *PLUS University Salzburg*

Corresponding Authors: s.appel@gsi.de, benjamin.halilovic@stud.plus.ac.at, simon.hirlaender@plus.ac.at

In advanced accelerator facilities like the heavy-ion synchrotron SIS18 at GSI in Darmstadt, ensuring stable and efficient multi-turn injection is crucial for achieving high-intensity beams. However, conventional control methods often lack the adaptability needed to handle rapidly changing beam dynamics, leading to suboptimal performance. To address this limitation, a data-driven Gaussian Process Model Predictive Control (GP-MPC) framework is employed, leveraging real-time updates to capture and predict complex injection behavior more accurately. We also systematically analyze the controller's behavior under variations in the incoming beam intensity and emittance, assessing its robustness against such perturbations.

Our simulation results demonstrate that GP-MPC can consistently manage multi-turn injection under realistic fluctuations of the incoming beam. By providing enhanced predictive accuracy through sample-efficient data-driven modeling, this work lays the groundwork for more robust and automated control strategies in such accelerator operations.

Poster session / 17

Towards Robust and Adaptive Decision-Making in Autonomous Robots: A Framework for Synchronized Digital Twins

Author: Hannes Voß[1]

Co-authors: Luca Jedelhauser [1]; André Dehne [2]; Stephan Pareigis [1]

[1] *University of Applied Sciences Hamburg*

[2] *HAW Hamburg*

Corresponding Authors: hannes.voss@haw-hamburg.de, stephan.pareigis@haw-hamburg.de, luca.jedelhauser@haw-hamburg.de, andre.dehne@haw-hamburg.de

The use of autonomous mobile robots in dynamic and uncertain environments requires adaptive and robust decision-making. Synchronized digital twins —real-time virtual counterparts of physical systems —offer a promising approach to improving planning, increasing robustness, and enhancing adaptability. However, developing such systems presents significant challenges, including balancing real-time synchronization with computational efficiency, ensuring the robustness of sensor fusion under domain shifts, and integrating learned models into real-world applications. This work identifies and investigates these challenges and explores how to address them. Therefore, existing approaches in synchronized digital twins for autonomous mobile robots are compared, identifying their strengths and limitations. Based on these insights, a conceptual framework that integrates simulation-driven reinforcement learning with real-time adaptation mechanisms under safety constraints is proposed. The framework is designed for efficient deployment across diverse robotic platforms and tasks. By contributing to bridging the gap between simulation and real-world execution, this research aims to advance the deployment of autonomous robots in safety-critical and unstructured environments.

Poster session / 18

Toward a Structured Reinforcement Learning Pipeline for Real-World Systems

Authors: Georg Schäfer[None]; Jakob Rehrl[1]; Simon Hirlaender[2]; Stefan Huber[1]

[1] *Salzburg University of Applied Sciences*

[2] *PLUS University Salzburg*

Corresponding Authors: stefan.huber@fh-salzburg.ac.at, georg.schaefer@fh-salzburg.ac.at, simon.hirlaender@plus.ac.at, jakob.rehrl@fh-salzburg.ac.at

Recent advances in reinforcement learning (RL) have shown great potential for managing complex systems in robotics, manufacturing, and beyond. However, translating RL successes from controlled experiments to real-world scenarios remains a significant challenge due to the absence of a standardized engineering pipeline that prioritizes thorough problem formulation. While data science and control engineering benefit from robust workflows, such as CRISP-DM and standard control-design processes, current RL practices often focus narrowly on hyperparameter tuning, overlooking foundational tasks like environmental design and reward specification.

In this work, we introduce an *RL engineering pipeline* that bridges existing gaps by integrating several key components. First, our approach begins with systematic problem identification and the formalization of the Markov Decision Process - defining states, actions, and rewards in alignment with physical constraints. Next, we ensure a careful setup of the environment, optimization objectives, and training procedures that respect the inherent nature of the optimization problem. These steps are followed by iterative agent training, hyperparameter optimization, evaluation, and eventual deployment. By adapting best practices from both CRISP-DM and classical control, our methodology enhances reproducibility and efficiency in RL development.

We validate our framework through a case study on a 1-degree-of-freedom (1-DoF) helicopter testbed. Our experiments indicate that targeted modifications, such as normalizing observations, randomizing initial conditions, extending episode horizons, and incorporating action penalties, yield measurable improvements in sample efficiency and training stability in both simulation and hardware. Moreover, our results suggest that this pipeline can be generalized to more complex systems, paving the way for more robust real-world RL applications.

Overall, our framework not only clarifies design decisions in RL but also offers a promising pathway to overcome long-standing challenges in deploying RL solutions outside of controlled experimental settings.

Talks / 20

Reinforcement Learning-Driven Optimal Control of Perfect Entangling Two-Qubit Gates

Author: Leander Grech[1]

Co-author: Simon Hirlaender [2]

[1] *University of Malta*

[2] *PLUS University Salzburg*

Corresponding Authors: simon.hirlaender@plus.ac.at, leander.grech@um.edu.mt

Noisy intermediate-scale quantum (NISQ) computers promise a new paradigm for what is possible in information processing, with the ability to tackle complex and otherwise intractable computational challenges, by harnessing the massive intrinsic parallelism of qubits. Central to realising the potential of quantum computing are perfect entangling (PE) two-qubit gates, which serve as a critical building block for universal quantum computation. Quantum optimal control, which involves shaping electromagnetic pulses that drive quantum gates, aims at optimising the use of NISQ computers. In this work, Reinforcement Learning (RL) was leveraged to optimise for pulses which can generate PE gates. We train a range of RL agents within robust simulation environments to identify pulse shapes that push gate operation towards their theoretical limits. In particular, a trained RL agent can generate a 12 ns pulse, with a 50 ps sampling time, that yields near-maximal entanglement for the two-qubit gate, while maintaining unitarity error on the order of 10-4. Selected agents are then validated on higher-fidelity simulations, proving how RL-based methods can reduce overhead in hardware calibration and accelerate experimentation. Moreover, this approach is hardware agnostic, enabling broad applicability across different quantum platforms. Ultimately, this work shows how RL can be used to expedite optimisation and refine control processes in the presence of vast and intricate parameter spaces.

Poster session / 21

Towards Few-Shot Reinforcement Learning in Particle Accelerator Control

Author: Simon Hirlaender[1]

Co-authors: Andrea Santamaria Garcia [2]; Annika Eichler [3]; Chenran Xu [4]; Jan Kaiser [3]; Verena Kain [5]

[1] *PLUS University Salzburg*

[2] *University of Liverpool*

[3] *DESY*

[4] *IBPT*

[5] *CERN*

Corresponding Authors: annika.eichler@desy.de, verena.kain@cern.ch, ansantam@liverpool.ac.uk, jan.kaiser@desy.de, simon.hirlaender@plus.ac.at, chenran.xu@kit.edu

This paper investigates the automation of particle accelerator control using few-shot reinforcement learning (RL), a promising approach to rapidly adapt control strategies with minimal training data. With the advent of advanced diagnostic tools and increasingly complex accelerator schedules, ensuring reliable performance has become critical. We focus on the physics simulation of the AWAKE electron line—a representative platform that allows us to disentangle the intrinsic control challenges from the performance of various algorithmic approaches.

In our study, we address two primary challenges: the scarcity of high-fidelity simulations and the presence of partially observable environments. To overcome these hurdles, we apply well-established methods in a novel context. Specifically, we leverage meta-reinforcement learning to pre-train agents on simulated environments with variable uncertainties, enabling them to quickly adapt to real-world scenarios with only a few additional training episodes. Moreover, when suitable simulations are unavailable or prohibitively costly, we employ a model-based strategy using Gaussian processes to facilitate efficient few-shot direct training.

While the underlying methods are not new, their application in the particle accelerator domain represents a significant step forward. Our results demonstrate that few-shot RL can markedly enhance control efficiency and adaptability, paving the way for more robust and automated accelerator operations. This work thus opens new avenues for integrating advanced RL techniques into complex physical systems, bridging the gap between simulation and real-world deployment.

Talks / 22

Multi-Agent Reinforcement Learning to achieve reliable wireless communication

Authors: Peter Dorfinger[1]; Roland Kwitt[2]; Sabrina Pochaba[None]; Simon Hirlaender[2]

[1] *Salzburg Research Forschungsgesellschaft*

[2] *PLUS University Salzburg*

Corresponding Authors: simon.hirlaender@plus.ac.at, sabrina.pochaba@salzburgresearch.at

Reinforcement learning (RL) is gaining more and more importance in the field of machine learning (ML). One subfield of RL is Multi-Agent RL (MARL). Here, several agents learn to solve a problem simultaneously rather than a single agent. For this reason, this approach is suitable for many real-world problems.
Since learning in a multiple agent scenario is highly complex, further conflicts can arise in addition to the difficulties in single-agent RL. These are, for example, scalability problems, non-stationarity or ambiguous learning goals.

To explore the difficulties of MARL, we have implemented the environment for a wireless network communication problem. There we look at the assignment of frequency resources to guarantee a reliable communication. Thus, different devices in a given area should learn on their own, which frequency they can use without disturbing the communication of other devices. As an overlap of frequencies can lead to a lack of communication, it is important that all devices select their own frequencies.

To accomplish this task, all devices that need to communicate become agents. The given area in which their communication takes place, the so-called communication cell, is the environment of the MARL problem. At each time step, every device chooses a communication channel (a frequency band) that it wants to use for its communication.
To ensure that the problem can be solved reliably, each agent receives the following information in its state: the communication channel used in the previous step, the own Quality of Service~(QoS) achieved by the last action, a vector of all neighbouring devices and the communication channels the neighbouring devices used in their last action.
After all agents have selected a communication channel, they receive their next state and a reward. We choose the reward to be the sum of the achieved QoS of all agents, since a shared reward avoids adversarial behavior and leads to cooperation between the agents.

We train this MARL task using a Q-Learning algorithm. We train our agents as well with a NashQ algorithm, which is adapted for Multi-Agent learning from Game Theory. The results show that the agents learn to communicate in a reliable way. However, the number of agents influences the training, since the number of possible state combinations increases exponentially with the number of agents. By comparison of the two different algorithms, the NashQ algorithm needs slightly less episodes to converge to an optimal policy than the Q-Learning algorithm.

Talks / **23**

Reinforcement Learning for In-Spill Optimization of the Mu2e Resonant Extraction: Compensating Non-Stationarity

Author: Jason St. John[1]

Co-authors: Aakaash Narayanan [1]; Andrew Whitbeck [1]; Ben Hawks [1]; Kyle Hazelwood [2]; Maira Khan [1]

[1] *Fermilab*

[2] *Fermilan*

Corresponding Authors: bhawks@fnal.gov, maira@fnal.gov, awhitbe1@fnal.gov, aakaashn@fnal.gov, stjohn@fnal.gov, kjh@fnal.gov

We present design considerations and challenges for the fast machine learning component of a third-order resonant beam extraction regulation system being commissioned to deliver steady beam rates to the mu2e experiment at Fermilab. Dedicated quadrupoles drive the tune toward the 29/3 resonance each spill, extracting beam at kV multiwire septa. The overall Spill Regulation System consists of (1) a "slow"process using ~100-spill averages to adjust the base quad ramp infrequently, (2) a feedforward harmonic content compensator, and (3) the "fast"ML agent reacting during each ongoing spill with on-the-fly additive corrections to the sum of (1) and (2).

We have demonstrated improved beam-rate steadying for a fast ML agent compared to a PID controller using a quasi-physical spill simulation, and demonstrated distillation of that simulation into a predictive surrogate model. Current work includes a data-and-training pipeline to generate data-aware surrogates with real-world dynamics, even as the dynamics shift unpredictably. The surrogates are to act as RL environments against which to train our fast ML control agents before deploying them on FPGA in the live system. Further current efforts focus on modeling and controlling beam loss around the storage ring, understanding additional available hardware inputs to the model, and the interplay of these with beam-steadying performance.

Poster session / 27

Studies on Virtual Platform for the HALF Beamline

Authors: Jiahui Zhang[None]; Xueting Wu[1]

Co-authors: Dadi Zhang ; Gongfa Liu ; Liuguo Chen

[1] *USTC*

Corresponding Author: xueting.wu@ustc.edu.cn

The autonomous alignment and optimization of synchrotron beamlines pose significant challenges. Traditionally, manual alignment is time-consuming and experience-dependent process, often requiring extensive diagnostic efforts and data collection. With the construction of the Hefei Advanced Light Facility (HALF) underway, the development of a virtual platform for beamlines will be an invaluable tool for beamline scientists and users. This platform will enable software testing and improve the prediction of optical element parameters in advance. In this paper, we present the development and comprehensive study of a virtual platform representing beamline BL10 at HALF. Additionally, we explore the integration of an AI-driven control system for optical elements control of synchrotron radiation beamlines within the virtual platform.

Talks / 31

RL for beam phase space distribution tuning at LCLS

Authors: Auralee Edelen[1]; Zihan Zhu[1]

[1] *SLAC*

Corresponding Author: edelen@slac.stanford.edu

Beams at LCLS require precise shaping in position-momentum phase space to meet the needs of different users. In particular, the shape of the longitudinal phase space needs to be customized, while ensuring the transverse phase space meets the requirements for Free Electron Laser (FEL) lasing. We present results of using RL for longitudinal phase space shaping, and compare these with approaches using Bayesian optimization. We also present results on combining longitudinal phase space shaping with transverse phase space control.

Poster session / 32

High-Speed Differentiable Simulations as an Enabler for Reinforcement Learning-based FEL Tuning Agents

Author: Jan Kaiser[1]

Co-authors: Ryan Roussel [2]; Zihan Zhu [2]; Jenny Morgan [2]; Kishansingh Rajput [3]; Malachi Schram [3]; Auralee Edelen [2]; Daniel Ratner [2]; Annika Eichler [1]

[1] *DESY*

[2] *SLAC*

[3] *Jefferson Lab*

Corresponding Authors: jan.kaiser@desy.de, annika.eichler@desy.de, dratner@slac.stanford.edu, edelen@slac.stanford.edu

The photon pulse intensity is one of the key performance metrics of Free Electron Laser (FEL) facilities and has a direct impact on their experimental yield. To date, FEL intensity tuning is a time-consuming manual task that requires expert human operators to have significant skill and experience. Autonomous tuning methods have been demonstrated to reduce setup times and improve the attained working points on other tuning tasks, but existing numerical optimisation algorithms struggle with the high dimensionality and complexity of FEL intensity maximisation. Reinforcement learning-trained optimisers (RLOs) promise the capabilities to successfully be applied to even such complex tuning tasks, but their requirement for training interactions makes online training infeasible. Offline training in simulation has previously been proposed as an effective solution, but FEL simulations are prohibitively complex and slow for use in RLO training. We propose the coupling of novel high-speed differentiable beam dynamics simulations with modular neural network surrogates to address this issue at the example of tuning the FEL intensity at LCLS. Not only have similar approaches been shown to reduce the wall-time required for training by orders of magnitude, making training for FEL tuning feasible, but the availability of gradients also promises an avenue for profoundly reducing the number of required training samples through gradient-based reinforcement learning.

Poster session / 33

Towards Lattice-Agnostic Reinforcement Learning Agents for Transverse Beam Tuning

Authors: Jan Kaiser[1]; Chenran Xu[2]; Andrea Santamaria Garcia[3]; Annika Eichler[1]

[1] *DESY*

[2] *IBPT*

[3] *University of Liverpool*

Corresponding Authors: chenran.xu@kit.edu, annika.eichler@desy.de, ansantam@liverpool.ac.uk, jan.kaiser@desy.de

Reinforcement learning (RL) has been successfully applied to various online tuning tasks, often outperforming traditional optimization methods. However, model-free RL algorithms typically require a high number of samples, with training processes often involving millions of interactions. As this time-consuming process needs to be repeated to train RL-based controllers for each new task, it poses a significant barrier to their broader application in online tuning tasks. In this work, we address this challenge by extending domain randomization to train general lattice-agnostic policies. We focus on a common task in linear accelerators: tuning the transverse positions and sizes of electron bunches by controlling the strengths of quadrupole and corrector magnets. During training, the agent interacts with environments where the magnet positions are randomized, enhancing the robustness of the trained policy. Preliminary results demonstrate that this approach enables policies to generalize and solve the task on different lattice sections without the need for additional training, indicating the potential for developing transferrable RL agents. This study represents an initial step toward rapid RL deployment and the creation of lattice-agnostic RL controllers for accelerator systems.

Poster session / 34

Cheetah –A High-speed Differentiable Beam Dynamics Simulation for Machine Learning Applications

Authors: Jan Kaiser[1]; Chenran Xu[2]; Andrea Santamaria Garcia[3]; Annika Eichler[1]

[1] *DESY*

[2] *IBPT*

[3] *University of Liverpool*

Corresponding Authors: jan.kaiser@desy.de, chenran.xu@kit.edu, ansantam@liverpool.ac.uk, annika.eichler@desy.de

Machine learning has emerged as a powerful solution to the modern challenges in accelerator physics. However, the limited availability of beam time and the high computational cost of simulation codes pose significant hurdles in generating the necessary data for training state-of-the-art machine learning models. Furthermore, optimisation methods can be used to tune accelerators and perform complex system identification tasks. However, they too require large numbers of samples of expensive-to-compute objective functions in order to achieve state-of-the-art performance. In this work, we introduce Cheetah, a PyTorch-based high-speed differentiable linear-beam dynamics code that enables fast collection of large datasets and sample-efficient gradient-based optimisation, while being easy to use, straightforward to extend and integrating seamlessly with widely adopted machine learning tools. Ultimately, we believe that Cheetah will simplify the development of machine learning-based methods for particle accelerators and fast-track their integration into everyday operations of accelerator facilities.

Poster session / 35

Advanced Optimization Strategies for the ESS-Bilbao Proton Injector, LEBT and RFQ Tuning

Authors: Ibon Bustinduy[1]; Juan Luis Muñoz[2]; Konrad Altenmüller[2]

[1] *ESS-BILBAO*

[2] *ESS-Bilbao*

Corresponding Authors: kaltenmueller@essbilbao.org, jlmunoz@essbilbao.org, ibustinduy@essbilbao.org

The ESS-Bilbao injector is a multipurpose machine that will accelerate protons up to 3 MeV. It will be used to produce neutrons by means of a Beryllium target. The first part of the injector has been running smoothly for more than a decade. This is formed by a proton source of the Electron Cyclotron Resonance (ECR) type that posseses unique characteristics. The subsequent Low Energy Transport section (LEBT) makes use of two solenoid magnets to transport and optimize the beam parameters for the Radio Frequency Quadrupole (RFQ). To enhance LEBT transmission and address the continuous plasma impedance variations during operation, we employ numerical optimization techniques, specifically Bayesian optimization.

On the other hand, we have used genetic algorithms to tune the RFQ cavity. In the RFQ resonant cavity (which is a 3 meters long linear accelerator) there are deviations from the desgiend uniform voltage profile, caused by local frequency variations that are a consequence of the modulation, the assembly and other assimetries. This deviations are corrected by adjusting the penetration of 62 plunger tuners. The combined action of the tuners modifies the field profile in order to achieve an uniform voltage (field profile). The adequate combination of tuners penetration is usually selected by a combination of an experimental technique (bead-pull measurements of field profiles, and their change for all the tuner actions) together with an algebraic approach for inverting the transfer matrix of the tuning system. This is usually done by a SVD psudo-inerse or equivalent algorithm. We propose an alternative method that instead of using the SVD approach selects the optimum tuner configuration by employing a genetic algorithm. The results of the complete assembly yield promising results, and are compared to the actualt measurements of the ESS-Bilbao RFQ.

Poster session / 36

Efficient Accelerator Operation with Artificial Intelligence Based Optimization Methods

Author: Evangelos Matzoukas[None]

Co-authors: Anke-Susanne Müller [1]; Chenran Xu [2]; Giovanni De Carne [3]

[1] *KIT*

[2] *IBPT*

[3] *Karlsruhe Institute of Technology*

Corresponding Authors: chenran.xu@kit.edu, giovanni.carne@kit.edu, anke-susanne.mueller@kit.edu, evangelos.matzoukas@kit.edu

Tuning injectors is a challenging task for the operation of accelerator facilities and synchrotron light sources,
particularly during the commissioning phase. Efficient tuning of the transfer line is essential for ensuring
optimal beam transport and injection efficiency. This process is further complicated by challenges such as
beam misalignment in quadrupole magnets, which can degrade beam quality and disrupt operations. Traditional
tuning methods are often time-consuming and insufficient for addressing the complexities of high-dimensional
parameter spaces.
In this work, we explore the use of advanced AI methods, including Bayesian optimization, to automate and
improve the tuning process. Initial results, demonstrated on the transfer line of KARA (Karlsruhe Research
Accelerator) at KIT (Karlsruhe Institute of Technology), show promising improvements in beam alignment
and transport efficiency, representing first steps toward more efficient and reliable accelerator operation.
This study is part of the RF2.0 project, funded by the Horizon Europe program of the European Commission,
which focuses on advancing energy-efficient solutions for particle accelerators.

Poster session / 37

Efficient data-driven model predictive control for online accelerator tuning

Authors: Chenran Xu[1]; Simon Hirlaender[2]

Co-authors: Andrea Santamaria Garcia [3]; Annika Eichler [4]; Borja Rodriguez Mateos [5]; Christian Hespe [4]; Jan Kaiser [4]

[1] *IBPT*

[2] *PLUS University Salzburg*

[3] *University of Liverpool*

[4] *DESY*

[5] *CERN*

Corresponding Authors: jan.kaiser@desy.de, simon.hirlaender@plus.ac.at, ansantam@liverpool.ac.uk, annika.eichler@desy.de, chenran.xu@kit.edu, borja.rodriguez.mateos@cern.ch, christian.hespe@desy.de

Reinforcement learning (RL) is a promising approach for the online control of complex, real-world systems, with recent success demonstrated in applications such as particle accelerator control. However, model-free RL algorithms often suffer from sample inefficiency, making training infeasible without access to high-fidelity simulations or extensive measurement data. This limitation poses a significant challenge for efficient real-world deployment. In this work, we explore data-driven model-predictive control (MPC) as a solution. Specifically, we employ Gaussian processes (GPs) to model the unknown transition functions in the real-world system, enabling safe exploration in the training process. We apply the GP-MPC framework to the transverse beam tuning task at the ARES accelerator, demonstrating its potential for efficient online training. This study showcases the feasibility of data-driven control strategies for accelerator applications, paving the way for more efficient and effective solutions in real-world scenarios.

Poster session / 38

Feedback Optimization on the EuXFEL Electron Dump Beamline

Authors: Christian Hespe[1]; Jan Kaiser[1]; Jannis Lübsen[2]; Frank Mayet[1]; Matthias Scholz[1]; Annika Eichler[1]

[1] *DESY*

[2] *TUHH/DESY*

Corresponding Authors: jannis.luebsen@tuhh.de, jan.kaiser@desy.de, matthias.scholz@desy.de, christian.hespe@desy.de, annika.eichler@desy.de, frank.mayet@desy.de

At the European XFEL, the main beam dump serves to absorb all electron bunches that are not required for the downstream scientific experiments. Due to the large beam power of the accelerator, controlling the dump temperature is a crucial component in its operation. Currently, this is done in an open-loop feed-forward manner. However, due to unforeseen drifts and changes in the setup of the accelerator, sporadic manual interventions by the machine operators are necessary to maintain regular operation, binding attention that could be spend better on other tasks.

For this reason, we present *feedback optimization* as a powerful and flexible solutions to automatically control the beam position along the main EuXFEL electron beam dump line. Two variants are investigated, model-based relying on the *Cheetah* particle accelerator optics simulation and model-free learning the model response over time, and the controller has been evaluated on the accelerator successfully.

39

Structured Doctoral Studies in Interdisciplinary Data Science

Authors: Christiane Ehrt[1]; Heike Hufnagel Martinez[2]

[1] *CFEL*

[2] *DESY*

Corresponding Authors: heike.hufnagel.martinez@desy.de, christiane.ehrt@cfel.de

The increasing automatization and the surging number and resolution of sensors in scientific experiments result in large, heterogeneous, and complex data collections. Data Science is, therefore, a key technology in modern natural sciences and materials science. Data-intensive research at the Science City Hamburg Bahrenfeld that centers around several large-scale user facilities for research in the structure of matter requires beyond off-the-shelf software solutions for data processing and analysis. To develop these novel methods, scientists must be highly skilled in the corresponding natural science field and computer science or applied mathematics alike. DASHH addresses this need by linking the competencies of more than 120 scientists with a high international reputation in research for the structure of matter, computer science, and applied mathematics in Hamburg.
Data Science in Hamburg –Helmholtz Graduate School for the Structure of Matter (DASHH) offers interdisciplinary doctoral projects with data challenges from application fields such as structural biology, particle physics, accelerator science, materials science, and photon science.

Welcome session / 40

Organisational Information

Welcome session / 41

Introduction to DESY

Welcome session / 42

Introduction to DMA

Welcome session / 43

Introduction to DASHH

44

6D Phase Space Beam Modelling Using Point Cloud Approach

Author: Hayg Guler[1]

[1] *IJCLab*

Corresponding Author: hayg.guler@ijclab.in2p3.fr

6D Phase Space Beam Modelling Using Point Cloud Approach

Talks / 45

6D Phase Space Beam Modelling Using Point Cloud Approach

Author: Hayg Guler[1]

[1] *IJCLab*

Course / 46

Introduction to Reinforcement Learning

Corresponding Author: ansantam@liverpool.ac.uk

Welcome session / 47

Opening and Organisational Information