DATA-DRIVEN MODEL PREDICTIVE CONTROL FOR AUTOMATED OPTIMIZATION OF INJECTION INTO THE SIS18 SYNCHROTRON

S. Hirlaender*, Paris Lodron Universität Salzburg, Austria
S. Appel†, N. Madysa, GSI Darmstadt, Germany

Abstract

In accelerator labs such as GSI / FAIR, automating complex systems is the key to maximize the time spent on physics experiments. This study explores the application of a data-driven model predictive control (MPC) to refine the multi-turn injection (MTI) process into the SIS18 synchrotron, departing from conventional numerical optimization methods. MPC is distinguished by its reduced number of optimization steps and its superior ability to control performance criteria, addressing issues like delayed outcomes and safety concerns – in this case septum protection. The study focuses on a highly sample-efficient MPC approach based on Gaussian processes, which lies at the intersection of model-based reinforcement learning and control theory. This approach merges the strengths of both fields, offering a unified and optimized solution and yielding a safe and fast state-based optimization approach beyond classical reinforcement learning and Bayesian optimization. Our study lays the groundwork for enabling safe online training for the SIS18 MTI issue, showing great potential to apply data-driven control in similar scenarios.

INTRODUCTION

Data-driven control theory and reinforcement learning (RL) hold significant potential for addressing control problems beyond the reach of classical control theory. These methods learn through direct interaction with the systems they control. However, RL faces challenges in accelerator control applications, including the need for large data sets for reliable performance and the trade-off between training stability and data efficiency. Enhancing reliability in particle accelerator control is crucial, particularly with the advent of new diagnostic tools and increasingly complex variable schedules. Standard algorithms often fall short, necessitating new strategies. This paper demonstrates the potential of data-driven model predictive control on the highly non-linear SIS18 injection simulation, achieving reliable performance within a feasible number of interactions suitable for real-world deployment.

Data-driven model predictive control

Model-Based Reinforcement Learning (MBRL) uses environment models to predict future states and rewards, significantly reducing the required amount of interaction with the real accelerator compared to model-free methods [1–4]. The accuracy and uncertainty of the model are crucial for the performance of MBRL algorithms.

Probabilistic models, specifically Gaussian processes (GPs), capture the uncertainty in the environment's dynamics and provide a measure of uncertainty in their predictions, which is essential for safe and efficient exploration. MPC is a control strategy that optimizes control inputs by solving a finite-horizon optimization problem at each time step based on predicted future states and rewards, considering system dynamics and constraints. GP-MPC [5] uses uncertainty information from the GP to make more informed decisions and to balance exploration and exploitation. This results in Probabilistic Model Predictive Control (P-MPC). This approach helps ensure safety and improve performance by avoiding regions with high uncertainty. The method requires fewer interactions with the environment to learn an effective policy, which is advantageous in scenarios where collecting data is expensive or time-consuming. It is applicable to a wide range of RL problems, especially those where data efficiency is critical. Examples beyond accelerator controls include robotics and autonomous driving. We propose a unified and optimized solution that yields safe and fast state-based optimization, situated at the intersection of MBRL and control theory. It demonstrates superior ability to control performance criteria and the ability to effectively address issues like delayed outcomes and safety concerns.

PROBLEM DEFINITION AND FORMULATION

FAIR, the Facility for Antiproton and Ion Research, will provide antiproton and ion beams of unprecedented intensity and quality, to drive the forefront of research on heavy-ion and antimatter [6]. Multi-turn injection (MTI) into SIS18 is one of the main bottlenecks to reach the FAIR intensity goals. An important limiting factor for intermediate charge-state ions is loss-induced vacuum degradation [7, 8]. Injection losses must be minimized to avoid a reduction in synchrotron performance due to loss-induced vacuum degradation [9].

As MTI must fulfill Liouville's theorem, four bumper magnets create a closed orbit bump with a time variable so that the injection septum deflects the next incoming beamlet into an available horizontal phase space close to the formerly injected beamlets. During injection, loss can occur both on the septum and on the acceptance. If η characterizes the relationship between the lost and injected particles, the multiplication factor (i.e, the accumulated beamlets) follows

$$m = n(1 - \eta). \quad (1)$$

* simon.hirlaender@plus.ac.at
† s.appel@gsi.de

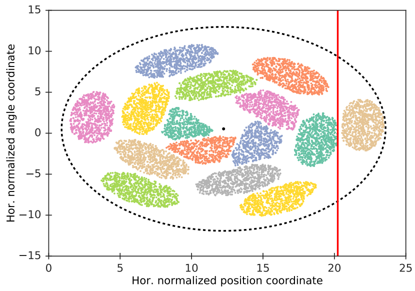

Figure 1: Snapshot of a MTI simulation. The red line indicates the septum, and the dashed line indicates acceptance.

n is the ratio between injection and revolution time. For loss-free injection, η is zero and the multiplication factor m is equal to the number of injected turns n. The center of the beam of the incoming beamlet x should be placed approximately so that the edge of the incoming beamlet touches the outside septum. The incoming beamlets will have a linear x and an angular x' displacement with respect to the closed orbit (x_c, x_c') and will therefore undergo betatron oscillations determined due to the horizontal tune Q_x. After one turn, the injected beamlets pass the injection point again without hitting the septum. If the orbit is not sufficiently fast, the beamlets will hit the inner side of the septum after the n_t revolution turns, depending on the betatron oscillation tune, and get lost. Additionally, the beamlets can also be lost at the beam pipe if the curvature of the incoming beamlet does not adapt to the ring acceptance curvature, depending on the mismatch between transfer line and SIS18.

Fig. 1 shows a snapshot of a MTI simulation. The loss areas, inside and outside the septum, as well as the acceptance, are visible. Inner beamlets lost particles in the septum earlier during the injection process and therefore did not overlap. The SIS18 MTI model has been implemented in the XSuite particle tracking code and was carefully validated against experiments [10–12]. For an ideal injection process without loss, the injected beam current will accumulate and will not be lost later, as shown by the red curve in Fig. 2. For poorly adjusted injection, during accumulation, particles will be lost and the accumulated beam current function diffs (black curve). The square of the root mean measures these differences and has been chosen for reward. The description of the state is given by the total loss after 50 turns, the loss at the septum, and the integral of the accumulated beam current divided by the point in time when no new particles are injected (see Fig. 1, 2). The actions are small Δ values for the six injection parameters $[x_c, x_c', x, x', \text{mismatch}, \Delta_{\text{reduction}}]$.

The Markov decision process (MDP)

The formulation of the problem as an episodic Markov decision process (MDP) for the injection problem is given as:

- State $= [\text{Reward}, \text{Loss}_{\text{Septum}}, \text{Integral}_1, \text{Integral}_2]$

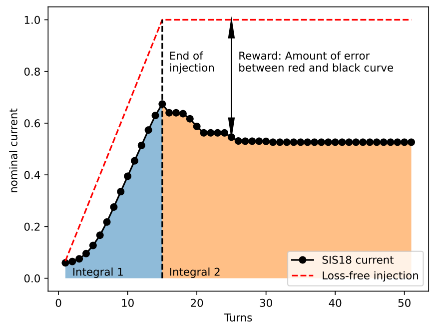

Figure 2: Nominal beam current (black) and the ideal current (red) during the injection process. The two integral values measure, in addition to the loss at the septum and the reward, the state.

- Reward $= \int (\text{Lossfree current}(t) - \text{SIS18 current}(t))\,\mathrm{d}t$
- Action $= [\Delta x_c, \Delta x_c', \Delta x, \Delta x', \Delta\text{mismatch}, \Delta_{\text{reduction}}]$
- Episodic design:
  - Episodes are initialized with initial values of the absolute actions uniformly sampled at random.
  - Only if a specific threshold is surpassed, the episode is reset (better reward than -1.9).
  - certain limits of the actions are exceeded to emulate hardware restrictions
  - if a specific step count of 25 termination has been reached the episode is truncated and reset.

The goal is that the agent learns to identify actions that swiftly move the state towards a reward within the specified threshold, thereby optimizing injection efficiency.

Simulation Results

Fig. 3 and Fig. 4 display the results of the highly-nonlinear injection problem. The experiment was simplified for this study due to the reward being concentrated in a small domain of the action space. The initial state was intentionally set close to the global optimum to enhance the likelihood of achieving non-constant, higher rewards. Fig. 4 shows the trajectories and actions for each episode within the environment. The top plot illustrates the state trajectories, with each colored line representing a different state and how it evolves over time. The bottom plot displays the action values taken at each step, with each action dimension represented by a different color. The x-axis represents cumulative steps, while the y-axis shows state and action values, respectively, with legends indicating the different states and actions for easy identification. Due to the complex nature of the problem, the policy begins near the optimum but initially overestimates the reward at the upper confidence bound, leading to oscillations around the solution. Once enough data is gathered,

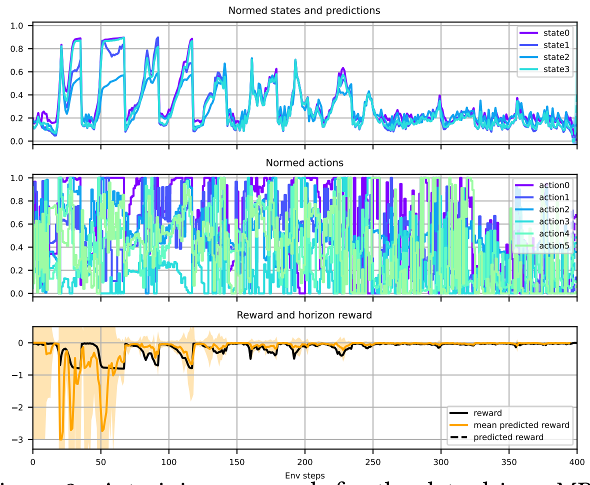

Figure 3: A training approach for the data-driven MPC involves resetting the episodes, which appear as jumps in the graphs.

the episodes consistently conclude successfully after approximately 300 interactions. Figure 3 (lower plot) distinctly illustrates the decease in the uncertainty in the expected reward (orange-shaded) during the training and the true reward (black).

Model-free Deep RL algorithms

Model-free off-the-shelf RL algorithms were also evaluated. Soft Actor-Critic (SAC) [13] and Proximal Policy Optimization [14] were successfully tested but required a considerable number of interactions with the system, making experiments on the real machine infeasible without prior tuning on a simulation. Fig. 5 shows an experiment employing the SAC algorithm. The training shown, achieves rewards not as high as an good value found by numerical optimization using BOBYQA within 5000 interactions. Additional numerical optimization techniques, such as BOBYQA optimization, have been successfully implemented but typically lack state information and do not develop a model incrementally [15, 16]. Despite their ease of use, these methods are likely to be replaced by more adaptive solutions over time, as demonstrated in this study.

SUMMARY AND OUTLOOK

This paper discusses the application of data-driven MPC integrated with GPs to enhance the MTI process at the SIS18 synchrotron within the GSI facility. This approach has demonstrated the ability to reduce the optimization steps required and improve the efficiency of the MTI process. Additionally, we evaluated the limitations of traditional reinforcement learning methods in terms of their high demand for interactions, which complicates their application without extensive prior simulation adjustments. Looking forward, the study paves the way for further development of data-driven control strategies in particle accelerator operations and similar complex systems. The next step is to facilitate real-time applications in the operational environment in several additional scenarios such as incorporating the prior knowledge from the simulation. The integration of advanced machine learning methods with traditional control systems holds significant promise for revolutionizing the operational capabilities of research facilities like GSI/FAIR, moving towards fully automated, highly efficient systems.

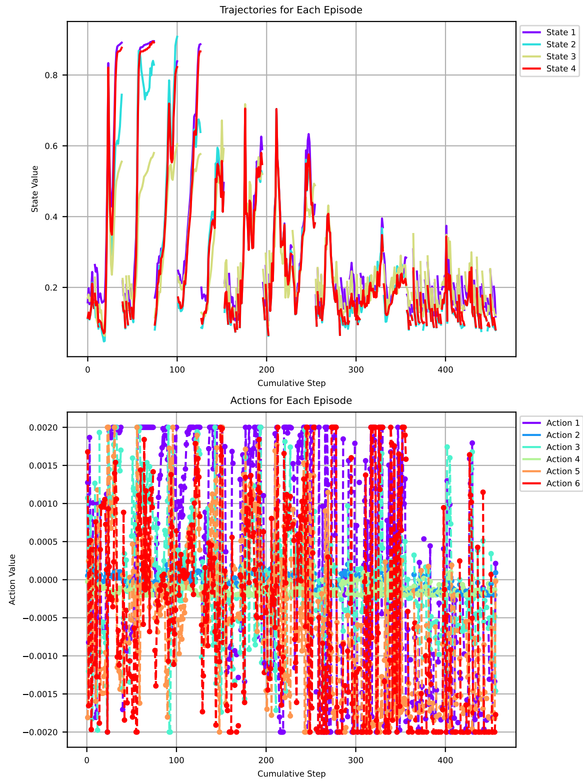

Figure 4: An experiment showing several episodes during the learning process.

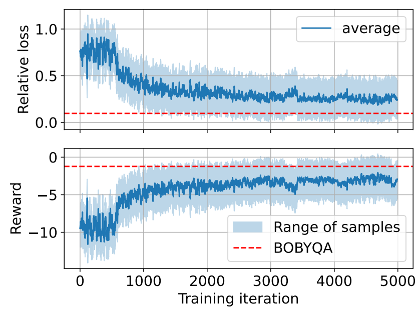

Figure 5: Training with Soft Actor Critic (SAC). The red line indicates the best result with the BOBYQA algorithms.

ACKNOWLEDGMENTS

S. Hirlaender gratefully acknowledge the support of the WISS 2025 project 'IDA-lab Salzburg' (20204-WISS/225/197-2019 and 20102-F1901166-KZP) and the other authors the EURO-LABS project funded by the European Union's Horizon Europe Research and Innovation program under Grant Agreement no. 101057511.

REFERENCES

[1] S. Hirlaender, L. Lamminger, G. Zevi Della Porta, V. Kain, Ultra fast reinforcement learning demonstrated at CERN AWAKE, JACoW IPAC 2023 (2023) THPL038. doi:10.18429/JACoW-IPAC2023-THPL038.
URL https://cds.cern.ch/record/2886522

[2] S. Hirlaender, N. Bruchon, Model-free and Bayesian Ensembling Model-based Deep Reinforcement Learning for Particle Accelerator Control Demonstrated on the FERMI FEL (2022). arXiv:2012.09737.

[3] S. Hirlaender, et al., Towards few-shot reinforcement learning in particle accelerator control, in: presented at the 14th International Particle Accelerator Conf. (IPAC'24), Nashville, Tennessee, USA, May 2024 paper TUPS60, this conference.

[4] V. Kain, et al., Test of Machine Learning at the CERN LINAC4, JACoW HB 2021 (2022) 181–185. doi:10.18429/JACoW-HB2021-TUEC4.
URL https://cds.cern.ch/record/2841807

[5] S. Kamthe, M. P. Deisenroth, Data-efficient reinforcement learning with probabilistic model predictive control (2018). arXiv:1706.06491.

[6] P. Spiller, et al., FAIR Technical Design Report December 2008, Tech. rep. (2008).

[7] E. Mustafin, et al., A theory of the beam loss-induced vacuum instability applied to the heavy-ion synchrotron SIS18, Nucl. Instr. and Meth. A 510 (3) (2003) 199–205. doi:10.1016/S0168-9002(03)01811-4.

[8] C. Omet, P. Spiller, J. Stadlmann, D. H. H. Hoffmann, Charge change-induced beam losses under dynamic vacuum conditions in ring accelerators, New J. Phys. 8 (11) (2006) 284–284. doi:10.1088/1367-2630/8/11/284.

[9] S. Appel, O. Boine-Frankenheim, F. Petrov, Injection optimization in a heavy-ion synchrotron using genetic algorithms, Nucl. Instr. and Meth. A 852 (2017) 73–79. doi:10.1016/j.nima.2016.11.069.

[10] S. Appel, O. Boine-Frankenheim, Multi-turn injection into a heavy-ion synchrotron in the presence of space charge, arXiv 1403.5972v (2014).

[11] S. Appel, et al., Injection optimization through generation of flat ion beams, Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip. 866 (April) (2017) 36–39. doi:10.1016/j.nima.2017.05.041.
URL http://linkinghub.elsevier.com/retrieve/pii/S0168900217305946

[12] G. Iadarola, et al., Xsuite: An Integrated Beam Physics Simulation Framework, JACoW HB2023 (2024) TUA2I1. arXiv:2310.00317, doi:10.18429/JACoW-HB2023-TUA2I1.

[13] T. Haarnoja, A. Zhou, P. Abbeel, S. Levine, Soft actor-critic algorithms and applications, arXiv preprint arXiv:1812.05905 (2018).

[14] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms, arXiv preprint arXiv:1707.06347 (2017).

[15] S. Appel, N. Madysa, SIS18 Machine Studies: Beam-Based Optimization of the Multi-Turn Injection, Report (to be published.

[16] S. Appel, et al., presented at the 14th International Particle Accelerator Conf. (IPAC'24), Nashville, Tennessee, USA, May 2024 paper MOPS68, this conference.