ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-TUPAB298

FIRST STEPS TOWARD AN AUTONOMOUS ACCELERATOR, A COMMON PROJECT BETWEEN DESY AND KIT*

Annika Eichler, Florian Burkart, Jan Kaiser, Willi Kuropka, Oliver Stein
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany
Erik Bründermann, Andrea Santamaria Garcia, Chenran Xu
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Abstract

Reinforcement learning algorithms have risen in popularity in the accelerator physics community in recent years, showing potential in beam control and in the optimization and automation of tasks in accelerator operation. The Helmholtz AI project "Machine Learning Toward Autonomous Accelerators" is a collaboration between DESY and KIT that works on investigating and developing reinforcement learning applications for the automatic start-up of electron linear accelerators. The work is carried out in parallel at two similar research accelerators: ARES at DESY and FLUTE at KIT, giving the unique opportunity of transfer learning between facilities. One of the first steps of this project is the establishment of a common interface between the simulations and the machine, in order to test and apply various optimization approaches interchangeably between the two accelerators. In this paper we present first results on the common interface and its application to beam focusing in ARES as well as the idea of laser shaping with spatial light modulators at FLUTE.

PROJECT GOAL

The authors in [1] have defined systems that can change their behavior in response to unanticipated events during operation as autonomous. While for self-driving cars different levels of autonomy have already been detailed in standards (SAE J3016), the definitions are less unique for the process industry, let alone particle accelerator operation. In [2] detailed levels for different kinds of process industry are given, where accelerator operation fits within the definition of control room operation. Given the autonomy levels there, one can conclude that accelerator operation varies from level 0 up to level 2, depending on the considered parts of the facility and whether it is a user or research facility. This means that in some cases only low-level automation (level 0) is available, while "automated system assisted start-up, transition, steady state and shutdown" as well as "manual fault correction supported by decision support systems" (level 2) is available in many user facilities. As a next step toward level 3, automated plant shutdown, start-up and transition on human request should be possible, which is the far-reaching goal of this project. The demands to achieve such levels will be even more pressing for future accelerators, which tend to increase in complexity to provide exceptional beams for research applications. Dedicated accelerators for very specific, medical and industrial applications may already be compliant with level 4 or 5 with minimal or no human supervision. A more flexible design from the start needs to employ advanced control methods to reach the same level of autonomy. However, this can be considered a sustainable investment to facilitate the reuse of such accelerators, if requirements for applications and products change.

Achieving this can be supported by tools from artificial intelligence (AI) and machine learning (ML), as it has been widely recognized by the accelerator community [3]. For this specific task, reinforcement learning (RL) appears to be a promising solution due to its ability to learn solution strategies looking into the future for complex problems, without the need for explicit labels or a well-defined model of the dynamics underlying the problem. In RL there exists a decision maker, called the *agent*. The agent can interact with the *environment*, the world in which the agent exists, by taking an *action*. The environment has a *state*, which is the immediate situation in which the agent finds itself. This state might however only be partially observed, with the part observed by the agent called an *observation*. By taking an action, the states are changed according to some transition dynamics, which are unknown to the agent. When an agent takes an action, he gets to know the next observation as well as the *reward*, which measures the immediate goodness of an agent's action in a given state. The strategy that the agent employs to take an action given the current state is called the *policy*. The goal is to find the policy that maximizes the *value function*, which is the sum of (discounted) future rewards over a time horizon or up to infinity expected under some policy. There are different types of RL algorithms: model-based ones that require or learn some model of the environment dynamics and model-free ones that do not. In model-free RL algorithms, one differentiates Q-learning approaches that learn the value function and policy gradient methods that try to learn the optimal policy directly. A good overview is provided in [4].

RL has been successfully used in accelerator operation to solve various tasks. In [5], the deep-deterministic-policy-gradient (DDPG) algorithm is used to optimize the booster current at BESSY II in Berlin. At FERMI FEL, RL is applied to optimize the laser alignment system and different algorithms are compared [6–8]. Furthermore, the output energy as well as the terahertz (THz) radiation is optimized in the beam-line by policy-gradient approaches [9]. Results at CERN for trajectory steering in AWAKE and LINAC4 are

* This work is funded by the Initiative and Networking Fund by the Helmholtz Association (Autonomous Accelerator, ZT-I-PF-5-6).

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-TUPAB298

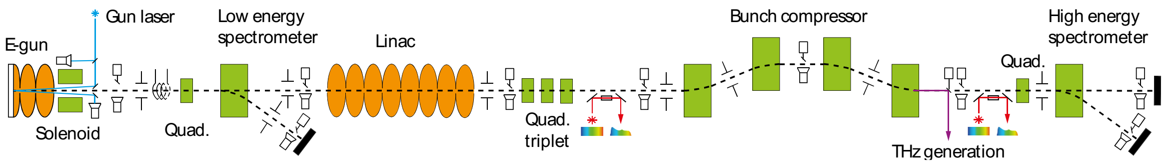


Figure 1: Schematic overview of the FLUTE layout showing the diagnostics and magnets [10].

presented using the Q-learning approach NAF. Simulation results with a high-fidelity GPU-accelerated online multi-particle beam dynamics simulator for the A3C algorithm are presented in [11]. Simulation studies on controlling the micro-bunching instability with RL are presented in [12]. To satisfy low-latency requirements given by the longitudinal beam dynamics, the RL controller was implemented on hardware [13].

While autonomous start-up is the far reaching goal of this project, a realistic goal is the control of the bunch profile for two accelerators within the project (see Table 1) that are very similar to each other. This will allow us to address the interesting questions of transferability of such agents. The results will therefore have an immediate impact on other accelerators as well.

In order to address the transferability of agents from simulation to experiment as well as from one facility to the other, a shared code base is required in order to easily interchange RL agents and use existing algorithms. Standard interfaces such as the one defined by OpenAI Gym [14] are widely used for this purpose in the RL community. Furthermore, to enable the transfer from simulation to experiment or from one facility to another, a standardized machine interface to the experiment and simulation will be needed.

Table 1: Comparison of the Final Electron Beam parameters for Two Linear Accelerators Considered in the Project

	FLUTE	ARES
Final energy [MeV]	40 - 50	100 - 155
Bunch charge [pC]	1 - 3000	0.5 - 30
Bunch length [fs]	1 - 300	0.2 - 10
Pulse repetition rate [Hz]	1 - 10	10 - 50

ACCELERATORS

ARES and FLUTE are two test facilities with accelerators of similar characteristics, as shown in Table 1, dedicated to research. They provide a unique environment to easily explore algorithm transferability among facilities, helping to determine how much facility-specific adaption is necessary. This experience will be of major importance for the long-term goal of automated start-up and the application of algorithms to more complex user facilities such as the European XFEL, which employs similar subsystems, e.g. lasers for photoinjection, magnets for steering, etc.

FLUTE (Ferninfrarot Linac- Und Test-Experiment) is a test facility with a new versatile electron linear accelerator for accelerator physics, as well as a source of intense broadband THz pulses for photon science [15]. Research topics include the development of single-shot femtosecond (fs) diagnostics [16], synchronization on a fs level, systematic bunch compression, and electron beam instability studies. The layout of FLUTE is shown in Fig. 1, where the low energy section up to the first spectrometer is in operation. Instead of the linear accelerator (linac) module, a Faraday cup is currently mounted on the straight beam pipe for charge measurements.

ARES (Accelerator Research Experiment at SINBAD) is an S-band radio frequency linac at the DESY Hamburg site equipped with a photoinjector and two independently driven traveling wave accelerating structures [17]. The main research focus is the generation and characterization of sub-femtosecond electron bunches at relativistic particle energy. The generation of short electron bunches is of high interest for radiation generation, i.e. by free electron lasers. The limits of particle beam diagnostic technologies and novel electron acceleration methods will also be investigated at the site. The linac is currently upgraded with a magnetic chicane and a particle beam diagnostic line including a PolariX X-band transverse deflecting structure [18]. The final layout of ARES is shown in Fig. 2.

FIRST STEPS

Using Spatial Light Modulators for Electron Beam Shaping at KIT

The properties of an electron bunch created at an RF photoinjector heavily depend on the RF-cavity design and photocathode properties, as well as on the characteristics of the driving laser. We propose to use spatial light modulators (SLMs) to achieve laser pulse shaping and modulation [19]. However, the SLM cannot work with the UV laser required for the photocathode and must be placed before the third harmonic generation stage. The optical propagation and other non-linear optical transformations can distort the modulated laser and degrade the pulse shaping quality. To mitigate this effect, we train a convolutional neural network to learn the inverse process of the optical propagation. Compared to the classical Gerchberg-Saxton method [20], this ML-based approach produces better transverse laser shaping results on a proof-of-principle setup. A more detailed discussion is given in [21]. In the future, the 3D distribution of the laser

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-TUPAB298

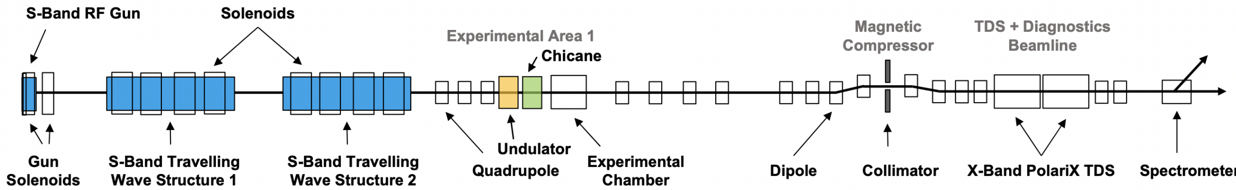


Figure 2: Schematic overview of the ARES layout.

pulse can be used as input for an RL algorithm to provide adaptive control.

Reinforcement Learning for Beam Parameter Optimization at DESY

Focusing and centering the electron beam on diagnostic screens along the beam line are frequently performed and time-consuming tasks during start-up and working point changes. They are also well defined tasks suitable as proof of concept applications for RL on particle accelerators. We consider focusing and centering the beam using a quadrupole triplet and corrector magnets in the ARES experimental area (EA) as a particular instance of this type of task. We propose an RL formulation of this task that enables an RL agent to focus and center the beam in just a few iterations and thereby significantly reduce the time required for optimizing the beam parameters. Such an RL agent observes the beam parameters on the screen and current magnet settings in order to propose a set of actions for changing the magnets' settings, improving the beam's focusing and centering. A flow chart of this setup is shown in Fig. 3.

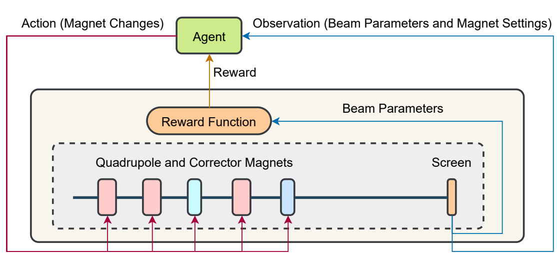
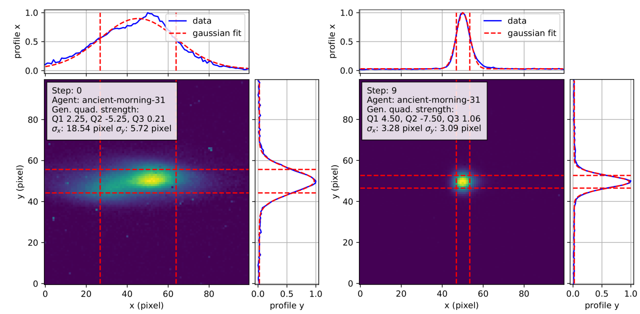


Figure 3: RL environment for beam optimisation in the ARES EA. The RL environment is shown in beige with the accelerator (or simulation) shown in grey. Quadrupoles are denoted in red, cyan and blue denote vertical and horizontal correctors respectively.

The DDPG [22] and TD3 [23] algorithms were tested, and further investigations of policy gradient methods such as PPO [24] are planned. Initial results on simulations using DDPG and TD3 are promising and demonstrated the algorithm's ability to solve the beam optimization task. A measurement campaign at the ARES accelerator has been carried out as well. Preliminary agents demonstrated their ability to transfer well from simulation to the real machine. Figure 4 shows a screen image taken at ARES before and after focusing by an RL agent.



(a) ARES step 0. (b) ARES step 9.

Figure 4: (a) Before and (b) after running a trained DDPG agent at ARES. The beam focus is clearly improved.

Transferability

A quadrupole triplet with correctors similar to the one in the ARES EA can be found at FLUTE. While similar in structure, the exact specifications of these two sections differ between the two accelerators. In order to test the transferability of the agents trained on ARES, an environment representing the appropriate section of FLUTE was created. First tests have shown that the RL agents trained on ARES were able to optimize the beam parameters on the FLUTE accelerator as well.

For the purpose of this initial transferabilty test, we developed a shared interface built around OpenAI Gym. As a result both facilities are compatible with this important RL standard and able to transfer agents between machines, as well as from simulation to machine without any code modifications. This interface is general enough to be easily adapted to other accelerator-related problems.

SUMMARY AND OUTLOOK

"Machine Learning Toward Autonomous Accelerators" is a cooperation between DESY and KIT. The goal of this project is to develop novel, AI-inspired methods that aid in the start-up of the accelerator in particular. Initial results in the areas of electron beam shaping at FLUTE and beam centering and focusing at ARES have demonstrated that machine learning methods such as reinforcement learning have the potential to substantially speed up accelerator start-up through autonomous optimization and thereby maximize the beam time available for experiments

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-TUPAB298

REFERENCES

[1] D. P. Watson and D. H. Scheidt, "Autonomous systems", *Johns Hopkins APL technical digest*, vol. 26, no. 4, pp. 368–376, 2005.

[2] T. Gamer, M. Hoernicke, B. Kloepper, R. Bauer, and A. J. Isaksson, "The autonomous industrial plant -future of process engineering, operations and maintenance", *IFAC-PapersOnLine*, vol. 52, pp. 454–460, 2019. doi:10.1016/j.ifacol.2019.06.104

[3] A. Edelen *et al.*, "Opportunities in machine learning for particle accelerators", 2018. arXiv: 1811.03172.

[4] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. Cambridge, NY, USA: MIT press, 2018.

[5] L. V. Ramirez, T. Mertens, R. Mueller, J. Viefhaus, and G. Hartmann, "Adding machine learning to the analysis and optimization toolsets at the light source BESSY II", in *Proc ICALEPCS'19*, New York, NY, USA, Oct. 2019. doi:10.18429/JACoW-ICALEPCS2019-TUCPL01

[6] N. Bruchon *et al.*, "Toward the application of reinforcement learning to the intensity control of a seeded free-electron laser", in *Proc. ICMT'19*, IEEE, Salerno, Italy, Oct. 2019, pp. 1–6. doi:10.1109/ICMECT.2019.8932150

[7] N. Bruchon *et al.*, "Basic reinforcement learning techniques to control the intensity of a seeded free-electron laser", *Electronics*, vol. 9, no. 5, p. 781, 2020. doi:10.3390/electronics9050781

[8] S. Hirlaender and N. Bruchon, "Model-free and bayesian ensembling model-based deep reinforcement learning for particle accelerator control demonstrated on the fermi fel", 2020. arXiv: 2012.09737.

[9] F. O'Shea, N. Bruchon, and G. Gaio, "Policy gradient methods for free-electron laser and terahertz source optimization and stabilization at the fermi free-electron laser at elettra", *Phys. Rev. Accel. Beams*, vol. 23, no. 12, p. 122 802, 2020. doi:10.1103/PhysRevAccelBeams.23.122802

[10] M. Nasse *et al.*, "First electron beam at the linear accelerator FLUTE at KIT", in *Proc. IPAC'19*, Melbourne, Australia, May 2019, pp. 882–884. doi:10.18429/JACoW-IPAC2019-MOPTS018

[11] X. Pang, S. Thulasidasan, and L. Rybarcyk, "Autonomous control of a particle accelerator using deep reinforcement learning", 2020. arXiv: 2010.08141.

[12] T. Boltz *et al.*, "Feedback design for control of the microbunching instability based on reinforcement learning", in *Proc. IPAC'19*, Melbourne, Australia, May 2019, pp. 104–107. doi:10.18429/JACoW-IPAC2019-MOPGW017

[13] W. Wang *et al.*, "Accelerated deep reinforcement learning for fast feedback of beam dynamics at KARA", *IEEE Trans. Nucl. Sci.*, in print, 2021. doi:10.1109/TNS.2020.00771

[14] G. Brockman *et al.*, "Openai Gym", 2016. arXiv: 1606.01540

[15] M. J. Nasse *et al.*, "FLUTE: A versatile linac-based THz source", *Rev. Sci. Instrum.*, vol. 84, no. 2, p. 022 705, 2013. doi:10.1063/1.4790431

[16] M. Yan *et al.*, "Design of a time-resolved electron diagnostics using THz fields excited in a split ring resonator at FLUTE", in *Proc. IBIC'17*, Barcelona, Spain, Sep. 2017, pp. 475–478. doi:10.18429/JACoW-IBIC2016-TUPG56

[17] E. Panofski *et al.*, "Status Report of the SINBAD-ARES RF Photoinjector and LINAC Commissioning", in *Proc. IPAC'19*, Melbourne, Australia, May 2019, p. 5. doi:10.18429/JACOW-IPAC2019-MOPTS026

[18] B. Marchetti *et al.*, "Experimental Demonstration of Novel Beam Characterization using a Polarizable X-Band Transverse Deflection Structure", *Scien. Rep.*, vol. 11, no. 1, p. 3560, 2021. doi:10.1038/s41598-021-82687-2

[19] G. Loisch *et al.*, "Photocathode laser based bunch shaping for high transformer ratio plasma wakefield acceleration", *Nucl. Instrum. Meth. A*, vol. 909, pp. 107–110, 2018. doi:10.1016/j.nima.2018.02.043

[20] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures", *Optik*, vol. 35, pp. 237–246, 1972.

[21] C. Xu *et al.*, "Machine Learning Based Spatial Light Modulator Control for the Photoinjector Laser at FLUTE", presented at the 12th Int. Particle Accelerator Conf. (IPAC'21), Campinas, Brazil, May 2021, paper WEPAB289, this conference.

[22] T. P. Lillicrap *et al.*, "Continuous control with deep reinforcement learning", 2015. arXiv: 1509.02971

[23] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods", in *Proc .35th Int. Conf. Machine Learning*, Stockholm, Sweden, Jul. 2018, pp. 1582–1591.

[24] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms", 2017. arXiv: 1707.06347