ISBN: 978-3-95450-257-8 ISSN: 2222-5692 doi:10.18429/JACoW-ECRIS2024-MOP11

CONTINUOUS DATA-DRIVEN CONTROL OF THE GTS-LHC ION SOURCE AT CERN

V. Kain*, N. Bruchon, D. Küchler, B. Rodriguez Mateos, M. Schenk
CERN, Geneva, Switzerland
S. Hirlander, Paris Lodron University of Salzburg, Salzburg, Austria

Abstract

Recent advances with the CERN infrastructure for machine learning allow to deploy state-of-the-art data-driven control algorithms for stabilising and optimising particle accelerator systems. This contribution summarises the results of the first tests with different algorithms to optimise the intensity out of the CERN LINAC3 source. The task is particularly challenging due to the different latencies for the various control parameters that range from instantaneous to full response after only ~30 minutes. Next steps and vision towards full deployment and autonomous source control will also be discussed.

INTRODUCTION

The GTS-LHC 14.5 GHz Electron Cyclotron Resonance (ECR) ion source [1] at the CERN LINAC3 provides different heavy ion beams for the LHC, as well as the PS and SPS fixed target experiments. In the case of the main species of lead ions, the beams are produced by vaporisation of solid samples that are heated with an oven in the plasma chamber. Tuning the various parameters of the source, such as oven power, to maximise its intensity output as well as ensuring reproducible intensity during the pulse and on a shot-by-shot level is non-trivial and is frequently slow due to conditioning effects. For example, during commissioning or after a stop, the oven's power needs to be slowly ramped up until lead evaporation is initiated. It is then increased over two to four weeks to maintain a sufficiently high evaporation rate until the next oven refill. Figure 1 shows an example of the reconditioning of the source in May 2018 with the discussed slow ramp-up of the oven power and non-linear response of intensity over the course of about 11 h. Various other parameters need to be adjusted as well as part of this process that are not indicated in Fig. 1. All of this is usually done manually.

This paper summarises the first tests of deploying CERN's Generic Optimisation Framework [2] to automatically optimise the intensity out of the LINAC3 source with the final goal of making recovery after oven refills and commissioning more efficient and less dependent on singular experts. Algorithms to stabilise the performance after commissioning were also part of the investigation.

To date, only preliminary tests of various sample-efficient optimisation and stabilisation algorithms could be carried out. However, they were already sufficient to start addressing the challenging aspects of time-varying dynamics and control knobs that act at significantly different time scales. Another test is planned towards the end of the 2024 run where the lessons learned will be incorporated.

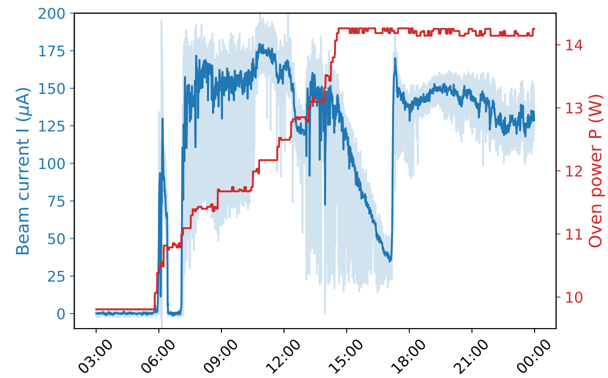

Figure 1: Evolution of the beam current measured by BCT.ITL05 at the end of the Low Energy Beam Transport in blue during lead ion beam setup in May 2018. The oven power (red) needs to be ramped up slowly. In this particular case this phase took roughly 11 h while tuning other parameters in addition.

GENERIC OPTIMISATION FRAMEWORK AND FRONTEND AT CERN

A significant step towards automating parameter optimisation and stabilisation was the implementation of the "Generic Optimisation Framework and Frontend" (GeOFF) in Python at CERN [2]. GeOFF standardises interfaces for optimisation tasks and provides adapters for various third-party packages such as SciPy, Stable Baselines 3, Scikit-Optimize, BoTorch. GeOFF tasks can scale to arbitrary complexity and depend on any Python package; they can use any controls system and even communicate with external simulation tools, as long as they have Python bindings. It comes with a GUI application, readily usable with the CERN control system in the various control rooms. It allows to add custom plotting in addition to a pre-defined set of plots that show the evolution of the objective function and the actors. It also allows to save the optimisation evolution in terms of objective function and actors, which was used to produce the plots in this paper.

* Verena.Kain@cern.ch

ISBN: 978-3-95450-257-8 ISSN: 2222-5692 doi:10.18429/JACoW-ECRIS2024-MOP11

DESCRIPTION OF LINAC3 SOURCE OPTIMISATION PROBLEM AND FIRST TESTS

The tests described in this paper were carried out during a so-called Machine Development (MD) session as part of the LINAC3 lead ion commissioning phase between July 8-10, 2024. The objective was to optimise the intensity out of the source measured by the Beam Current Transformer (BCT) ITL.BCT05 as well as the rms intensity fluctuation σ over the pulse length, i.e. to maximise the intensity and minimise its rms. The objective function y that yielded the best results was

$$y = (-1) \cdot \bar{I} + 2 \cdot \sigma \,, \tag{1}$$

where $\bar{I}$ corresponds to the mean intensity measured at ITL.BCT05 and σ to the standard deviation of the slice of the intensity pulse of interest for injection into the downstream elements followed by the Low Energy Ion Ring (LEIR) [3]. The optimisation task was formulated with up to five degrees of freedom: the currents of the three solenoids that confine the plasma in the plasma chamber, the voltage of the bias disc electrode located at the entrance to the plasma chamber, and finally the oxygen gas injection regulated according to the voltage setting with a feedback controller. The used parameter ranges are summarised in Tab. 1. Whereas the effects of the solenoids and bias disc changes were immediately measurable with the BCT, any changes to the gas injection system would only stabilise after several minutes and a waiting time had to be applied before reading out the BCT. See Fig. 2 for the evolution and the "ringing" of the gas injection measured voltage during one of the optimisation runs. The waiting times when working with the gas injection system were either set to 30 s or 120 s. The plots shown in the following were all from the period with 30 s waiting time. While 30 s might not be long enough for full decay of the "ringing", especially for large settings changes, it was sufficient to allow for convergence for the overall optimisation together with the faster actors. A more detailed study on which waiting time to use will have to be carried out in a future test campaign.

To implement the optimisation task within GeOFF, while respecting the different delays in response for the various control knobs, the problem was split into two optimisation tasks running in parallel and independently with the same objective function. One task would adjust the "fast" degrees of freedom (solenoids and bias disc voltage) and another one would adjust the gas injection and only acquire intensities and rms after 120 s or 30 s, respectively. In a continuous control setup during e.g. gas injection adjustments, the "fast" actors were supposed to catch up with the new conditions of the source and optimise their settings, while the gas reading was stabilising. No other synchronisation between the two optimisation environments was built in during these first tests.

THE ALGORITHMS

In the following, the different algorithms tested during the MD will briefly be introduced. The most relevant results were achieved with Bayesian Optimisation [4] and Adaptive Bayesian Optimisation [5]. Note that the algorithms BOBYQA [6] for optimisation and Extremum Seeking (ES) [7] for stabilisation were also tested briefly – with success however only in the case of BOBYQA. However, we will not discuss BOBYQA and ES results further.

Bayesian Optimisation

Bayesian optimisation (BO) is a powerful black-box optimisation algorithm, which learns a probabilistic model of the objective function with Gaussian processes (GP) [4]. To make use of the model's uncertainty, the so-called *Acquisition Function* is optimised, rather than the mean $\mu(\mathbf{x})$ of the objective function. In our case, we used the *Upper Confidence Bound Acquisition Function* (UCB):

$$a(\mathbf{x}) = \mu(\mathbf{x}) + \sqrt{\beta}\,\sigma(\mathbf{x}) \,, \tag{2}$$

where $\mu(\mathbf{x})$ is the mean of the posterior GP and $\sigma^2(\mathbf{x})$ the variance. β is a hyperparameter that needs to be tuned for the specific application. It defines the balance between exploration and exploitation during the optimisation process. For the LINAC3 source tests, Bayesian Optimisation was implemented with BoTorch in a custom optimiser available in GeOFF only for the LINAC3 optimisation environments. To enforce "smooth" parameter optimisation, *proximal biasing* was applied [8].

Adaptive Bayesian Optimisation

To use Bayesian optimisation as a continuous control algorithm and make the algorithm adapt to changes – hence Adaptive Bayesian Optimisation (ABO), the objective function can be modelled as a function of the control parameters $\mathbf{x}$ and also as a function of time t. The kernel function, or prior covariance, of the GP is chosen such that it can represent the correlations in the data well. Following [5] the kernel that we use in ABO is a composite kernel with a *spectral mixture kernel* S for t and the *Matern kernel* M for $\mathbf{x}$:

$$k([t_1, \mathbf{x_1}], [t_2, \mathbf{x_2}]) = \theta_k \times S(t_1, t_2) \times M(\mathbf{x_1}, \mathbf{x_2}) \,, \tag{3}$$

where θ_k is the output scale.

ABO was implemented with BoTorch [9] and is again only available for the LINAC3 environments in GeOFF. To use ABO for continuous control, the data buffers for conditioning the GP models need to be truncated. These data buffers are also stored for subsequent warm-starts such that random (or any other policy) data collection is not necessary. As for BO, *proximal biasing* was used.

FIRST OBSERVATIONS

Figure 3 shows the evolution of the objective function and the normalised currents of three solenoids during optimisation with BO. *Proximal biasing* ensures that the solenoids

ISBN: 978-3-95450-257-8 ISSN: 2222-5692 doi:10.18429/JACoW-ECRIS2024-MOP11

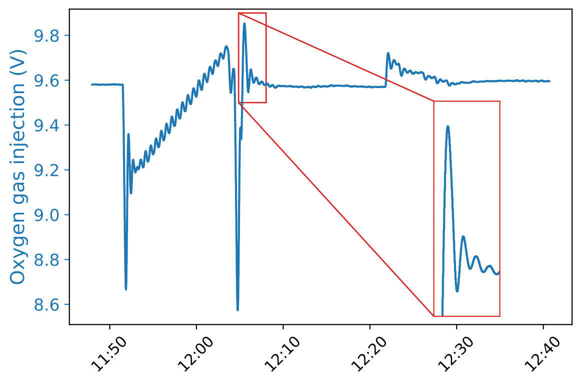

Figure 2: Evolution of the measured gas injection regulation system voltage during one of the tests on 9th of July described below. The larger the step size, the larger the excursions are around the set values. The oscillations settle eventually. See bottom plot of Fig. 5 for the set values during this phase.

smoothly arrive at the optimum settings avoiding big jumps, while the noisy objective function is minimised. The intensity after this optimisation was indeed higher than initially (improvement from 0.45 to 0.5 mA). A similar test with successful convergence was also carried out including the bias disc voltage, indicating that optimisation of the "fast" actors could be relatively easily automated. An important ingredient for good convergence was to not average over several acquisitions and let the GP learn the alleatoric uncertainty.

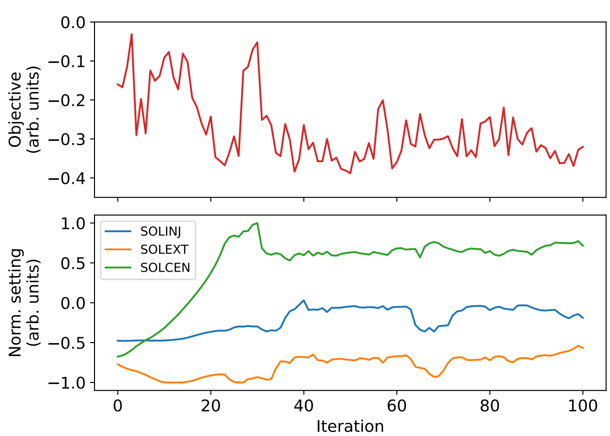

Figure 3: Bayesian Optimisation for the three solenoids around the LINAC3 source plasma chamber using proximal biasing and historical data to condition the GP at the start of the optimisation. The objective function is a combination of mean intensity and rms measured during the slice of interest of the ion pulse, see Eq. (1).

As a next step ABO was tested for the "fast" actors in combination with optimisation of the slow gas injection system. The best configuration was achieved with $\beta = 1.5$ in the UCB acquisition function and $\beta_{prox} = 0.5$ for proximal biasing. Figure 4 shows the evolution of the objective function as well as the three solenoid settings for this case. The GP of the algorithm was conditioned with a previously recorded dataset, thus the solenoids start with good settings and maintain those for the first roughly 115 iterations. The sharp increase of the objective function away from the optimum is caused by the start of the optimisation of the gas injection system, which collected its initial data by a linear ramp of its settings between predefined bounds. See Fig. 2 for the gas injection voltage acquisition and Fig. 5 for the perspective of the controller during the optimisation of this slow actor, respectively. The linear ramp instead of the usual random policy for initial data collection was chosen to minimise the excursions of the gas injection regulation system. The solenoids do not manage to establish the original performance for most of the linear ramp of the gas injection system and stay at the bounds of their allowed ranges (the bounds correspond to $(-1, 1)$ in the plots as GeOFF works with normalised settings). Together the slow and fast tasks converge however eventually to an objective value that is not as good as the initial one, but very close to the optimum obtained during the optimisation phase. Note that the starting gas injection setting and corresponding objective function were not used for building the model, see Fig. 5. This is believed to be the reason for not getting back to the initial optimum. The result is promising and indicates that adequate optimisation of the system gas injection system and solenoids (plus bias disc) is in principle possible with the used techniques. If the optimisation routines were synchronised, where instead of running ABO for the fast actors in the shadow, an optimisation of the fast actors is triggered after each step of the gas injection system, the global optimum should be within reach. This would also allow for easier tuning of the adequate waiting time after changing the settings on the gas injection system.

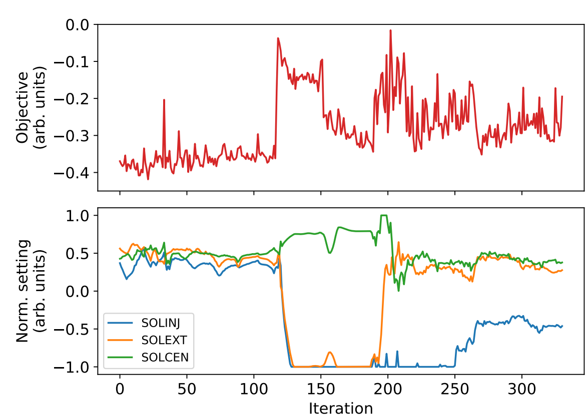

Figure 4: Evolution of objective function (upper plot) and three solenoids (lower plot, normalised settings) during Adaptive Bayesian Optimisation. At iteration ~120, ABO is launched on the gas injection system, see Fig. 5. The three solenoids eventually converge to a new optimum given the changed setting of the gas injection system.

Table 1: Ranges of the Various Parameters used for LINAC3 Source Optimisation

Parameter	Parameter Range	Unit
solenoid inj	[1100, 1265]	A
solenoid cen	[1100, 1140]	A
solenoid ext	[0, 400]	A
bias disc	[0, 500]	V
gas injection	[9, 9.6]	V

NEXT STEPS

Towards the end of the 2024 run another test to control the LINAC3 source is foreseen and the lessons learned so far will be incorporated. Given the results obtained and the experience of the expert, the optimisation task for the next step will be implemented as *one* GeOFF optimisation problem. The different parameters will be optimised/stabilised in a hierarchical and sequential manner: the outer loop will run ABO to optimise the gas injection system; at each iteration of this outer loop, first the solenoids will be optimised with BO (conditioned on previous data) and then the bias disc. The waiting times for the outer loop for the slow gas injection system will be established dynamically based on the acquisition of the gas injection regulation voltage together with the acquisition of the current of the extraction power supply as additional information.

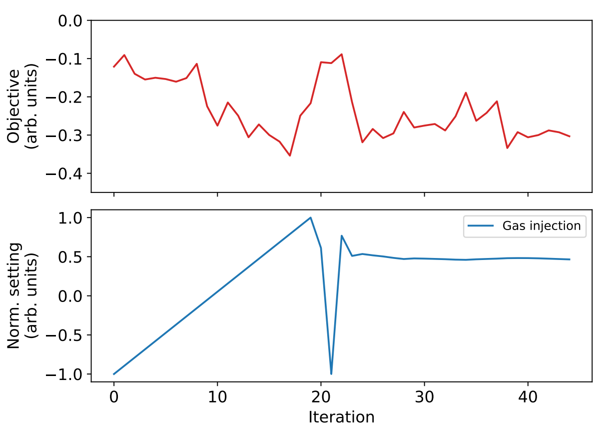

Figure 5: Evolution of objective function (upper plot) and gas injection setting (lower plot, normalised settings) during Adaptive Bayesian Optimisation. The linear ramp is used to avoid large excursions of the gas injection system regulator and "confuse" the "fast" actors that are running ABO in parallel.

SUMMARY

The LINAC3 ECR ion source at CERN provides heavy ion beams for the LHC and the PS and SPS fixed target experiments. Lead ion beams are produced through vaporisation of solid samples in a plasma chamber. Optimising the lead beam intensity out of this source is time consuming, non-trivial and is usually done manually relying on year-long experience of a few experts. This paper summarises the first steps towards automating the source commissioning as well as the performance stabilisation thereafter with algorithms based on Bayesian Optimisation. The CERN Generic Optimisation Framework and Frontend (GeOFF) was used as a platform for implementation and test execution with its ready-made GUI and various features. Particularly challenging for controlling the LINAC3 source are the different latencies in response involved with the various control knobs. This was addressed by running several optimisation/stabilisation tasks in parallel, separating the "slow" and "fast" actors. The first results were promising and showed that Bayesian Optimisation and Adaptive Bayesian Optimisation are sample-efficient enough and can deal well with the noisy environment. For guaranteed convergence it was however proposed for the next test to modify the setup of optimisation tasks and trigger for each iteration of the slow system an optimisation of the fast actors and synchronise the tasks in this manner. Given the results obtained already, this should allow convergence to the global optimum in a robust manner and to track changes adequately.

REFERENCES

[1] C. E. Hill *et al.*, "GTS-LHC: A New Source For The LHC Ion Injector Chain", in *AIP Conference Proceedings*, ECRIS'04, Berkeley, California (USA), March 2005, p. 127–130. doi:10.1063/1.1893381

[2] N. Madysa *et al.*, "Generic Optimisation Framework and Frontend (GeOFF)". doi:10.5281/zenodo.8434512

[3] M. Chanel, "LEIR: the low energy ion ring at CERN,"*Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 532, no. 1–2, pp. 137–143, Oct. 2004. doi:10.1016/j.nima.2004.06.040

[4] R. Roussel *et al.*, "Bayesian Optimisation Algorithms for accelerator physics", 2023. doi:10.48550/arXiv.2312.05667

[5] A.G. Wilson and R.P. Adams, "Gaussian Process Kernels for Pattern Discovery and Extrapolation", 2013. doi:10.48550/arXiv.1302.4245

[6] M.J.D Powell, "The BOBYQA algorithm for bound constrained optimization without derivatives", DAMTP 2009/NA06, Cambridge, England, 2009. https://ml4physicalsciences.github.io/2021/files/NeurIPS_ML4PS_2021_82.pdf

[7] A. Scheinker *et al.*, "Extremum Seeking Optimal Controls of Unknown Systems", 2018. doi:10.48550/arXiv.1808.05181

[8] R. Roussel *et al.*, "Proximal Biasing for Bayesian Optimization and Characterization of Physical Systems", in *NeurIPS'21*. doi:10.18429/JACoW-ICALEPCS2021-MOPV039

[9] M. Balandat *et al.*, "BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization", Advances in Neural Information Processing Systems 33. doi:10.48550/arXiv.1910.06403