ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-TUPS50

ML-ENHANCED COMMISSIONING OF THE APS-U ACCELERATOR COMPLEX*

N. Kuklev†, L. Emery, M. Borland, H. Shang, V. Sajaev, Y. Sun,
J. Dooling, I. Lobach, K. Harkay, G. Fystro
Advanced Photon Source, Argonne National Laboratory, Lemont, IL, USA

Abstract

The Advanced Photon Source (APS) facility has just completed an upgrade to become one of the world's brightest storage-ring light sources. For the first time, machine learning (ML) methods have been developed and used as part of the baseline commissioning plan. One such method is Bayesian optimization (BO) – a versatile tool for efficient high-dimensional single and multi-objective tuning, as well as surrogate model construction and other purposes. In this paper we will present our development work on adapting BO to practical control room problems such as tuning linac and booster transmission efficiency, injection stabilization, enlarging storage ring dynamic and momentum apertures, and various other tasks. We will also show first experimental results of these efforts, including achieving initial beam capture in the APS-U storage ring. Given the success of BO methods at APS, we are working on tighter ML algorithm integration into the standard control room procedures through a dedicated graphical interface.

INTRODUCTION

New particle accelerator projects face increasing performance demands, resulting in tighter tolerances on accuracy and stability. With the recent successes of machine learning, there is immense interest in making use of smarter algorithms to implement generic tools to improve reliability, reduce expert workload, and provide higher performance to users.

A key application of ML for accelerators is in parameter optimization, whereby one or multiple objectives are tuned through an intelligent search of the parameter space. Conventional optimization methods previously applied at APS include simplex [1, 2], RCDS [3], and genetic algorithms [4]. Our recent efforts have focused on ML-based methods including Bayesian optimization (BO) [5], reinforcement learning [6], and others. BO is of special interest since it allows efficient black-box function optimization with few samples, taking advantage of any prior physics-model knowledge provided to the algorithm. This is especially valuable for new accelerators since simulation models might not yet be available or are not sufficiently well calibrated for offline training.

At APS, we developed methods to make BO applicable to a wide range of experimental systems, including those with high-dimensionality, time-dependent drift [7], or critical safety considerations [8], all without extensive expert tuning and automatically making use of high-fidelity archival data [9]. This paper presents applications of BO to several experimental tasks we performed as part of the APS-U commissioning.

* The work is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.
† nkuklev@anl.gov

BAYESIAN OPTIMIZATION

We first briefly review the math behind BO. The output being optimized is described by

$$\mathbf{y} = f(\mathbf{x}) + \varepsilon, \tag{1}$$

where $f(\mathbf{x})$ is the black-box function of interest and $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ the added noise. Using Gaussian Processes (GP) [10], a surrogate model for f can be parameterized as a multivariate normal distribution with a mean $m(\mathbf{x})$ and covariance kernel $k(\mathbf{x}, \mathbf{x}')$. The kernel is used to evaluate the similarity between values of f at $\mathbf{x}$ and $\mathbf{x}'$, and its' appropriate choice is critical for good GP convergence. In general, local kernels such as the square exponential (SE) are used, although more specialized ones are required for certain tasks [8, 11]:

$$k_{SE,i} = \sigma^2 \exp\left(\frac{-(x_i - x_i')^2}{2l^2}\right). \tag{2}$$

Kernel hyper-parameters are generally the output variance σ and lengthscale l.

GP surrogate model is used for acquisition function optimization in order to find the most promising next step(s). A typical choice is the upper confidence bound function that balances exploration and exploitation,

$$\mathrm{UCB}(\mathbf{x}) = \mu(\mathbf{x}) + \sqrt{\beta} * \sigma(\mathbf{x}) \tag{3}$$

where mean μ and variance σ are provided by the model and β is the tradeoff hyperparameter.

INJECTOR TUNING

In parallel with APS-U storage ring construction, injector commissioning activities were ongoing so as to validate various upgrades and prepare for robust delivery of high charge bunches. Details of these efforts are discussed in separate submissions for the Particle Accumulator Ring (PAR) [12] and the Booster [13].

An important part of this process was commissioning a set of new thermionic guns in the linac and re-optimizing the system for different charges/pulse counts. BO-based tuning was used extensively in these tasks, allowing for significant expansion of the scanned parameter spaces as compared to previous simplex-based tools. Figure 1 shows an example of such a ‘run’.

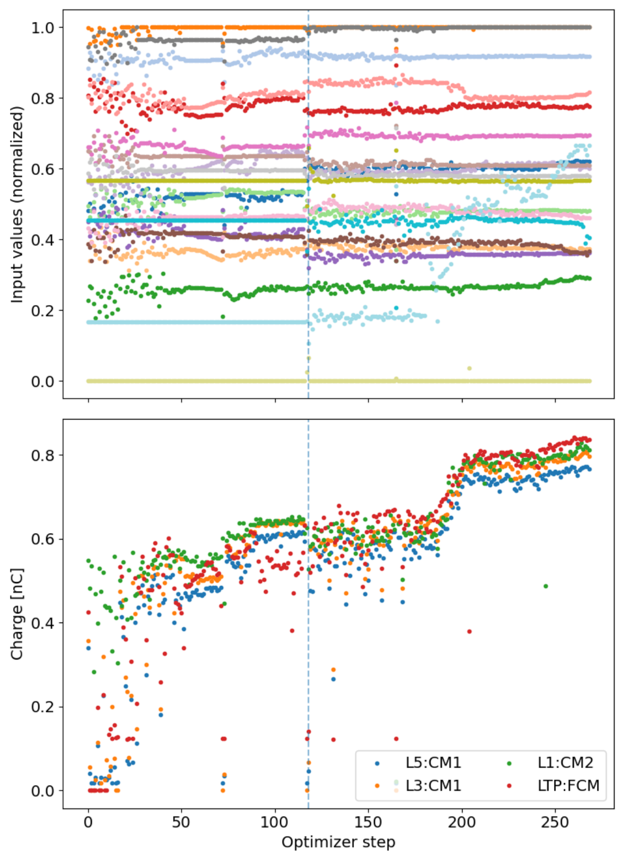

Figure 1: Experimental single-objective APS linac optimization with BO optimizer. Goal is to transmit highest possible charge to the end of the linac (denoted by LTP objective). Vertical line indicates when optimization parameter space was switched from 16D to 20D on the fly, causing some exploration jitter but quickly finding the new useful input parameter (lower pale blue trace).

Due to generally convex topology of the objective space for transmission efficiency and significant concerns about hysteresis, trust region [14] and step size limit features of our optimizer software APSopt (discussed in detail below) were used, allowing for a smooth improvement with minimal random exploration and small steps. We also performed dynamic objective and variable switching. Namely, for objectives, the single-objective goal was swapped to current monitors further downstream (L1 -> L3 -> L5 -> LTP) as transmission improved. For variables, parameter space was expanded on the fly based on expert intuition and beam trajectory. Such manipulations can be done while retaining already collected data and are unique to model-based methods like BO. We later extended BO procedure to measure and optimize PAR injection efficiency as the ultimate performance parameter, fully encompassing the linac and the transfer lines. An attempt to directly run the final 20D/LTP configuration with simplex did not appear to converge.

APS-U BEAM CAPTURE AND INJECTION EFFICIENCY

Initial steps of APS-U ring commissioning such as Booster transfer line tuning, first turn beam threading, and eventual beam capture have been extensively simulated and automated using standard linear optics methods [15]. However, we initially faced a number of challenges utilizing these tools due to magnet polarity errors, wrong setpoints, and instrumentation misconfigurations. While attempting to achieve beam capture, we decided to directly explore the parameter space of available injection parameters (transfer line corrector currents, septum and kicker voltages, RF phase, orbit bumps, etc.) through a high-dimensional BO configuration with a strong exploration bias in the acquisition function. As shown in Fig. 2, after several minutes of scanning with BPM 10 turn sum signal objective (from turn-by-turn acquisition), 8 uA beam current was noticed on the DCCT. This, to our knowledge, is the only time an ML tool has achieved initial beam capture in a particle accelerator.

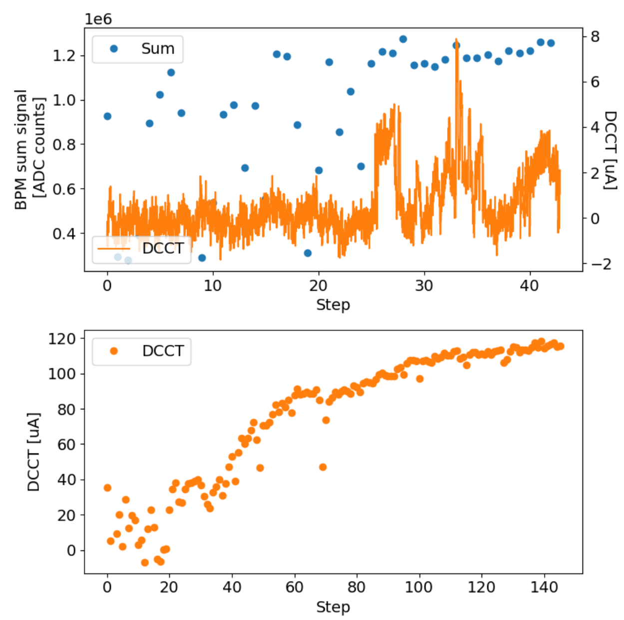

Figure 2: Top figure - first beam capture of the APS-U storage ring. BO with large exploration bias was used with BPM sum signal goal so as to provide a smoothly varying objective with gradient (as compared to mostly zero-valued DCCT current). Bottom picture - one of regular injection efficiency tuning runs with DCCT current as objective.

After establishing a good lattice configuration we continued to use BO repeatedly for injection efficiency tuning (see example in GUI section below), since APS-U utilizes a swap-out injection scheme and thus has high sensitivity to and complex interplay of optics matching, physical apertures, and dynamic acceptance. As commissioning proceeded, ring orbit and optics were adjusted multiple times and it typically took BO less than 10 minutes to recover efficiency. Work is ongoing to fully automate this process.

ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-TUPS50

BAYESIAN EXPLORATION AND MODELLING

Along with optimization tasks, BO can be used for explor-ing the parameter space and constructing surrogate models to better understand the objectives. This is performed by making acquisition function equal to uncertainty (i.e. very high β in Eq. (3)), sometimes referred to as Bayesian explo-ration (BE). One use-case we encountered was in mapping out the injection aperture limits to check for septum mis-alignments, since this determines the feasible closed orbit parameters through the injection straight. We scanned tra-jectory in 2D and 3D, the former case being shown in Fig. 3.

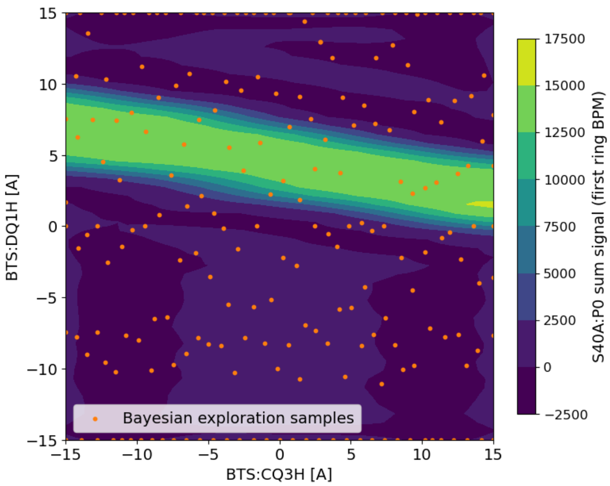

Figure 3: Bayesian exploration of the septum aperture using last 2 horizontal transfer line correctors as inputs and first ring BPM as output.

Our results were consistent with a grid scan and with simulations but required only 20% of the samples to generate. This sample efficiency will be even more prominent in higher-dimensional scans.

CONTROL ROOM INTEGRATION

In the control room, ML experts can run BO using Python notebooks and low level APIs for extra flexibility and debugging functionality. However, it is important to put ML tools into hands of other physicists and operators, which means abstracting away many of the details into presets and simple interfaces. To that end we have been designing a GUI front-end with a number of 'human-in-the-loop' features such as sampling suggestions but also single-click-to-run functionality. It is also fully compatible with standard APS file formats [16] and software deployment methods. A screenshot of current version after successful injection efficiency optimization is shown in Fig. 4.

The core algorithm execution engine has been well tested previously, and we are in the process of deploying the GUI for other physicists' use, iterating on feedback and incorporating various tutorials, failsafes, and other auxiliary features.

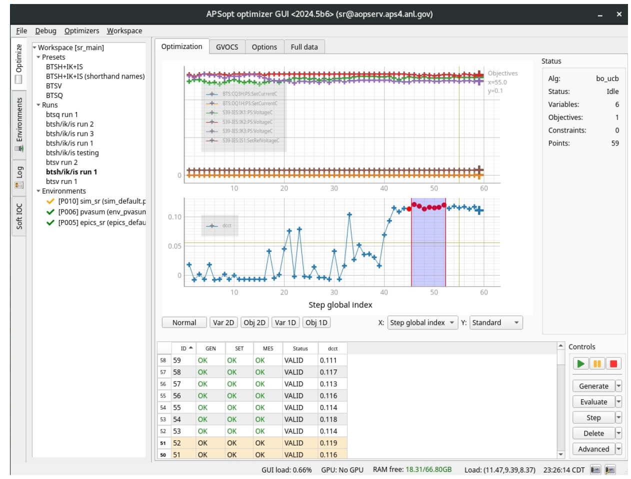

Figure 4: APSopt GUI being used for injection efficiency tuning after an orbit change.

CONCLUSION

ML tools have the potential to both be extremely powerful in certain types of tasks, but also problematic and detrimental if applied without careful setup and consideration of experimental realities. Through a concerted multi-year effort, APS has developed, tested, and now regularly uses ML optimization methods in a variety of operationally relevant hands-on tasks. In this paper we highlighted some of these in the injector and in the new APS-U storage ring, including key contributions to the commissioning process. Several further applications are slated to be explored soon, including nonlinear beam dynamics tuning, ion effect mitigation, orbit stabilization, and others. Our plan is to continue with ML method integration in a practical and data-driven fashion, while also incorporating new state-of-the-art techniques from ML community and pursuing high-fidelity modelling and digital twin environment for faster and more robust validation and deployment to the control room. We hope our experience encourages broader adoption of ML methods at accelerator facilities worldwide.

ACKNOWLEDGMENTS

Authors acknowledge ANL Laboratory Computing Resource Center (LCRC) for providing the computational resources, and APS operators for help in data collection.

REFERENCES

[1] L. Emery, M. Borland, and H. Shang, "Use of a General-Purpose Optimization Module in Accelerator Control," in *Proc. PAC'03*, Portland, OR, USA, May 2003. `https://jacow.org/p03/papers/WPPE002.pdf`

[2] X. Huang, "Robust simplex algorithm for online optimization," *Phys. Rev. Accel. Beams*, vol. 21, p. 104 601, 2018. `doi:10.1103/PhysRevAccelBeams.21.104601`

ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-TUPS50

[3] X. Huang, J. Corbett, J. Safranek, and J. Wu, "An algorithm for online optimization of accelerators," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 726, 2013. doi:10.1016/j.nima.2013.05.046

[4] M. Borland, V. Sajaev, L. Emery, and A. Xiao, "Multi-objective direct optimization of dynamic acceptance and lifetime for potential upgrades of the Advanced Photon Source," ANL/APS/LS-319, 986303, 2010, ANL/APS/LS–319, 986 303. doi:10.2172/986303

[5] M. Balandat *et al.*, "BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization," in *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 21 524–21 538. https://proceedings.neurips.cc/paper/2020/hash/f5b1b89d98b7286673128a5fb112cb9a-Abstract.html

[6] V. Francois-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An Introduction to Deep Reinforcement Learning," *Foundations and Trends® in Machine Learning*, vol. 11, no. 3-4, pp. 219–354, 2018. doi:10.1561/2200000071

[7] N. Kuklev, M. Borland, G. I. Fystro, H. Shang, and Y. Sun, "Online Accelerator Tuning with Adaptive Bayesian Optimization," in *Proc. NAPAC'22*, Albuquerque, NM, USA, 2022, pp. 842–845. doi:10.18429/JACoW-NAPAC2022-THXD4

[8] N. Kuklev, M. Borland, G. Fystro, H. Shang, and Y. Sun, "Robust adaptive bayesian optimization," in *Proc. IPAC'23*, Venice, Italy, 2023, pp. 4428–4431. doi:10.18429/JACoW-IPAC2023-THPL007

[9] N. Kuklev, I. Lobach, H. Shang, and R. Soliday, "Upgrades to logging and ml analytics architecture at APS," in *Proc. IPAC'23*, Venice, Italy, 2023, pp. 4432–4435. doi:10.18429/JACoW-IPAC2023-THPL008

[10] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. 23, 2005. doi:10.7551/mitpress/3206.001.0001

[11] A. G. Wilson, Z. Hu, R. Salakhutdinov, and E. P. Xing, *Deep Kernel Learning*, 6, 2015. doi:10.48550/arXiv.1511.02222

[12] K. Harkay *et al.*, "Particle accumulator ring restart and readiness for Advanced Photon Source upgrade commissioning," presented at IPAC'24, Nashville, TN, USA, May 2024, paper TUPG06, this conference.

[13] J. Calvey *et al.*, "APS upgrade booster commissioning," presented at IPAC'24, Nashville, TN, USA, May 2024, paper TUPG07, this conference.

[14] D. Eriksson, M. Pearce, J. Gardner, R. D. Turner, and M. Poloczek, "Scalable global optimization via local Bayesian optimization," in *Advances in Neural Information Processing Systems*, 2019, pp. 5496–5507. http://papers.nips.cc/paper/8788-scalable-global-optimization-via-local-bayesian-optimization.pdf

[15] V. Sajaev, "Commissioning simulations for the argonne advanced photon source upgrade lattice," *Phys. Rev. Accel. Beams*, vol. 22, p. 040 102, 2019. doi:10.1103/PhysRevAccelBeams.22.040102

[16] M. Borland, "A Self-Describing File Protocol for Simulation Integration and Shared Post-Processors," in *Proc. PAC'95*, Dallas, TX, USA, May 1995.