OPTIMIZATION OF THE FERMILAB BOOSTER USING A HYBRID BAYESIAN AND RL FRAMEWORK*

N. Kuklev†, J. Eldred, M. Balcewicz, R. Sharankova, J. St. John
Fermi National Accelerator Laboratory, Batavia, IL, USA

Abstract

The PIP-II project will raise the Fermilab Booster intensity and ramp rate. Beam losses will limit the maximum power and are hard to simulate. Presently, the Booster uses operator-guided empirical tuning - a challenging task due to high dimensionality, multiple objectives, critical safety constraints, and drifts. We developed a synergistic suite of Bayesian optimization (BO) and reinforcement learning (RL) tools to optimize and stabilize beam losses, including novel techniques for fast risk-aware Bayesian optimization and exploration. For the initial tune-up, we performed single and multi-objective tuning using scalarized objectives comprised of critical beam loss locations, achieving significant rebalancing of losses as well as an overall improvement in transmission efficiency. To build a data-driven surrogate, active learning was used to collect data while relying on risk-aware constraints to successfully avoid beam trips. A few thousand points were collected, and a GP surrogate validated for uncertainty-aware predictions. Several off-policy RL agent architectures were trained for long-term stabilization. In surrogate-based testing, SAC with BPM context and history embedding had the best performance with fast and robust convergence when subjected to energy and trajectory perturbations. Experimental testing is ongoing to enable operational use.

INTRODUCTION

The Fermilab Booster is a key component of the Fermilab accelerator complex providing 8 GeV protons to various users directly, as well as sending beam to the Recycler/Main Injector. For the future LBNF/DUNE program, the PIP-II linac upgrade will increase the power on target to 1.2 MW. It includes a new superconducting linac, a new BTL transfer line, and various improvements to the complex including the Booster. Key changes for the Booster are the increase of injection energy to 800 MeV, higher beam intensity (from 4.8×10^{12} to 6.7×10^{12}), a new painting scheme, new collimation and RF systems, and a faster ramp rate of 20 Hz. Given that the Booster is a 50-year-old machine with aging hardware, areas of concern have been identified in terms of machine reliability and capability to operate at PIP-II parameters. Significant beam dynamics challenges are expected due to extreme levels of space charge at injection, as well as new instabilities at transition crossing.

The maximum Booster power is determined primarily by administrative beam loss limits, with PIP-II requiring a two-fold reduction in relative losses as compared to the current levels. An extensive simulation campaign is ongoing to better understand and control beam dynamics, but so far it has proven difficult to obtain quantitative agreement with experimental data, especially for the local loss distribution. Thus, it is expected that an extensive experimental tune-up will be necessary to commission the Booster for PIP-II.

Driven by PIP-II requirements as well as challenges in the DOE GARD roadmap [1, 2], several AI/ML projects were started at Fermilab, focusing on the development of advanced optimization algorithms, fast differentiable simulations [3], and integrated facility-scale virtual accelerators (digital twins) [4]. Together, they enable both direct optimization using Bayesian methods, and a transition to more data-intensive but scalable methods like reinforcement learning (RL). In this paper, we describe our results in applying Bayesian and RL methods to experimentally control optics and injection properties for Booster loss reduction.

FERMILAB BOOSTER

The Fermilab Booster is a rapid cycling synchrotron with 24 sectors (short (S) and long (L) subsections). The main combined function magnets are connected to a shared resonant power supply. In addition, there are 48 independently powered correction elements with dipole, quadrupole, skew quadrupole, sextupole, and skew sextupole components [5]. The Booster diagram is shown in Fig. 1.

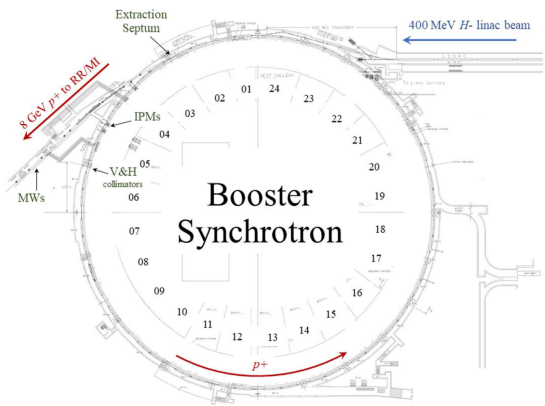

Figure 1: Booster synchrotron.

Time-dependent current ramps with up to 64 points can be set, forming thousands of potential knobs. Such high dimensionality is infeasible so we identified two smaller scenarios. First, for Bayesian loss minimization, we chose a set of quadrupole ramp offsets either in 24D (short or long) or

* This manuscript has been authored by FermiForward Discovery Group, LLC under Contract No. 89243024CSC000002 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.
† nkuklev@fnal.gov

48D (combined). This was motivated by the experimentally observed increase in injection beam losses at lower tunes, likely due to space-charge detuning pushing particles onto the (Q_x/Q_y=6.5) resonance [6]. Changing quadrupole offsets enables both working point and low-order resonance strength variation. For the second set, we chose a coupled Linac-Booster setup consisting of the injection closed orbit bump, transfer line correctors, and cavity/debuncher phases (12D total). Energy and trajectory drift are typical during operation, and their compensation is a good use case for RL.

During studies, we collect thousands of data channels. Booster beam losses are measured by over 50 beam loss monitors (BLMs), which provide integrated proportional chamber signals at 12.5 kHz. Beam current is measured via a 12.5 kHz toroid and 20 kHz DCCT signals. A dedicated study event \$17 is used at a 2-6 s period. Python scripts are used to change settings, while monitoring the machine protection system to pause if the beam permit is lost.

SAFETY-CRITICAL BO

Bayesian optimization is uniquely suited to tasks where candidate evaluation is relatively expensive, and complications such as constraints and step sizes need to be satisfied. Numerous successful BO applications have been demonstrated in particle accelerators [7–9]. In BO, the output(s) are described by $\mathbf{y} = f(\mathbf{x}) + \varepsilon$ where $f(\mathbf{x})$ is the black-box function of interest and $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is 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 a kernel $k(\mathbf{x}, \mathbf{x}')$.

The fitted GP model is used for acquisition function optimization to find the best next step(s). For single-objective problems, a typical choice is the upper confidence bound function that balances (or strongly biases) exploration and exploitation. For multi-objective (MO) problems, there is no single best candidate. The goal is instead to increase the hypervolume - the volume relative to the reference point covered by the pareto front. We use a recently proposed noisy-EHVI acquisition function [11].

Our BO code is implemented in the `apsopt` toolkit [12], with a BoTorch/PyTorch computational backend [13].

Safety-Critical Operation

During initial testing, the standard BO algorithm had several large violations of efficiency constraints. This is expected due to how BO constraints are typically implemented. Because a single bad shot can exceed trip limits, additional safeguards were needed for parasitic data collection.

While methods like SafeOpt are well established [14], their performance at high dimensionality is unacceptable for online studies. Thus, we developed a comprehensive set of 'risk-aware' constraint terms including soft/hard risk thresholds, uncertainty limits, trust regions, and pessimistically biased model priors. By enabling the appropriate terms, we achieved a global preference for new locations with high model confidence with a heavy penalty for unknown, far away, or noisy regions. Importantly, no extensive hyperparameter tuning was necessary beyond establishing the risk tolerance criteria. A visualization of the 5 terms is shown in Fig. 2.

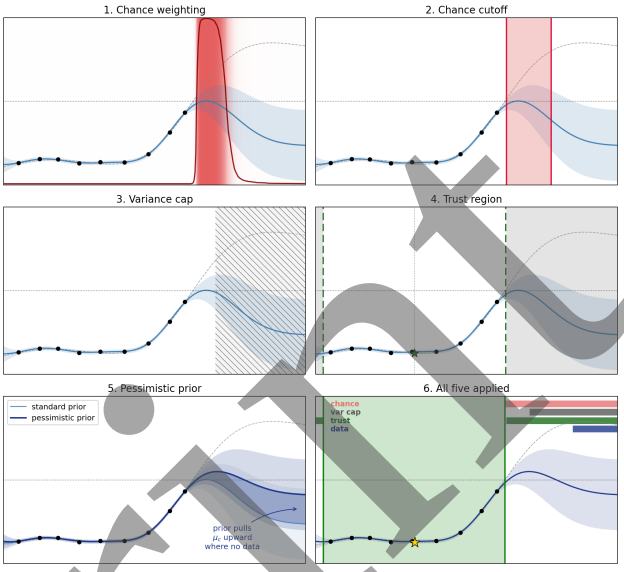

Figure 2: BO constraint types and their effect on the available acquisition function optimum search region.

To encourage better exploration with constraints, we also implemented a boundary bias term that pulls the optimizer closer to the threshold when model confidence is high, enabling an 'edge walking' behavior. We measured the number of algorithm violations and safe/unsafe region recall on a set of test problems showing no impact on safety (Fig. 3).

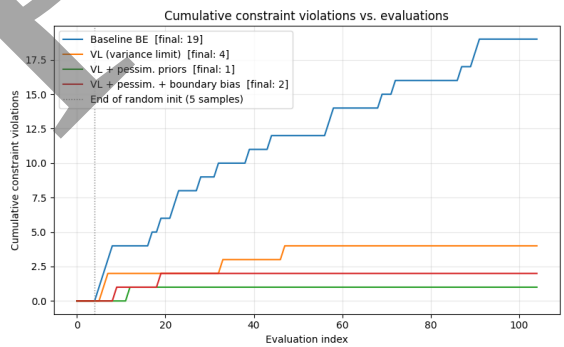

Figure 3: Violation counts for various risk terms. All variants have standard (chance weight) constraints enabled.

RESULTS

BO With Quad Offsets

Quadrupole offset tuning was initially done in 24D (QL) while ramping the beam intensity from 50 % to 75 % to 100 % of the operational one. No beam aborts or other issues were encountered. A full optimization run was performed in 48D (QL+QS), demonstrating an improvement of 0.4 % to the transmission efficiency. Due to the already high efficiency this is a notable current loss reduction.

Using the above data, we reviewed the loss/efficiency correlations - some areas showed a pareto front with useful tradeoffs. Based on the MOBO algorithm (with efficiency as the first objective), we developed a way to model losses as independent GP outputs but then use a scalar objective to combine them, which scales linearly with the number of outputs. We selected several key locations (L15,L22,L01), and successfully demonstrated a transfer of losses to the S06 collimation region without efficiency degradation. The best observed margin improvement of 25 % is shown in Fig. 4.

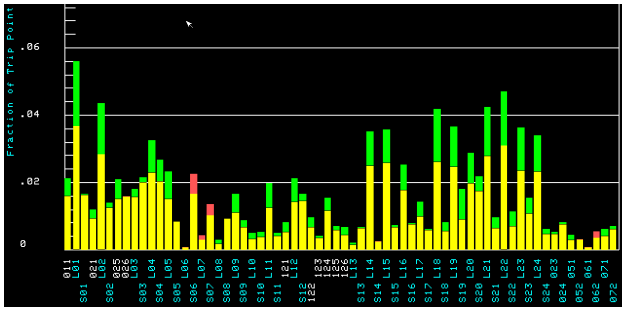

Figure 4: BLM loss rebalancing into S06 collimation region. Green/red indicates lower/higher vs reference values.

RL for Injection Efficiency

For long-term stabilization, due to the low maximum data collection rate (≤1 Hz), only off-policy RL methods are feasible. Training them still requires an initial surrogate model, followed by various online transfer learning techniques. We used Bayesian exploration to collect a dataset covering all safe areas of the injection parameter space through first doing active learning, and then fake 'RL rollouts' with random walks within a safe area, shown on the right half of Fig. 5.

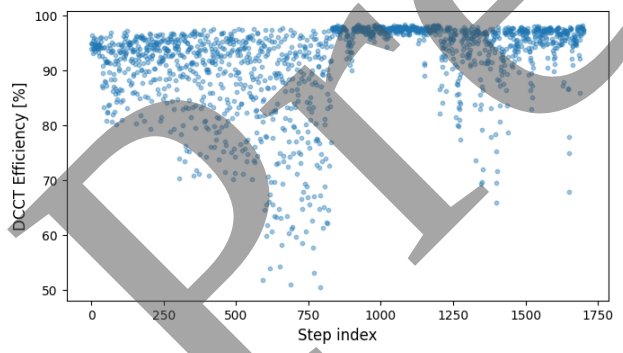

Figure 5: RL surrogate data collection. Note the stepwise drops in efficiency, enabled by relaxed constraints.

Using the full ~2k point dataset, we generated a GP surrogate model and validated the statistical properties (distribution of residuals, lengthscale convergence, etc.). Then, a SAC policy (Stable Baselines 3 implementation) was trained with a 5 % (of full range) delta action space of correctors/phases. We experimented with the naive version, but quickly found it useful to add a state vector of BPM data and history as context. After hyperparameter tuning, promising variants were quickly found. Full efficiency recovery and stability were demonstrated from many random initializations, as shown in Figs. 6 and 7. Experimental tests are in progress to confirm these encouraging simulation results.

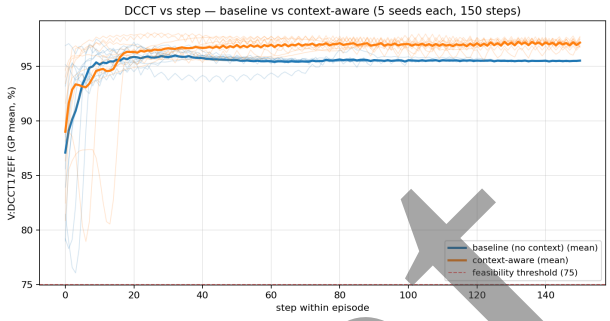

Figure 6: DCCT efficiency traces of several RL rollouts.

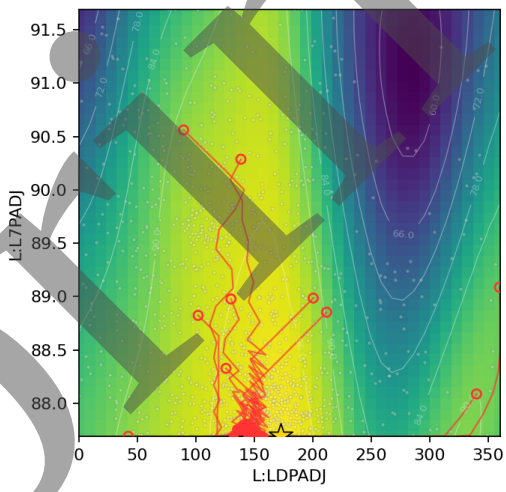

Figure 7: Rapid RL convergence in 2D projection along the two energy knobs of the optimization.

CONCLUSION

Improving Booster losses is critical to achieving PIP-II design parameters. In this paper we highlighted our work on data-driven loss modelling and optimization. We developed techniques for safe and fast Bayesian exploration, and used them to both rebalance and reduce the losses directly, but also to safely generate high quality surrogate models. Based on the collected data, RL policy architectures were explored and promising candidates identified that were able to robustly recover performance after energy and trajectory drift. Our future work will aim to experimentally measure the RL performance and robustness, as well as fully automate the data collection and policy training process such that it can be continuously rerun by operators.

REFERENCES

[1] Accelerator beam physics research roadmap, `https://science.osti.gov/hep/-/media/hep/pdf/2022/ABP_Roadmap_2023_final.pdf`

[2] S. Nagaitsev *et al.*, "Accelerator and beam physics research goals and opportunities", *arXiv*, 2021. doi:10.48550/arXiv.2101.04107

[3] N. Kuklev *et al.*, "Facility-scale differentiable virtual accelerator at fermilab", in *Proc. NAPAC'25*, Sacramento, CA, USA, Aug. 2025, pp. 106–109. doi:10.18429/JACoW-NAPAC2025-MOP028

[4] N. Kuklev, M. Borland, L. Emery, Y. Hidaka, H. Shang, and Y. Sun, "Unified differentiable digital twin for the IOTA/FAST facility", in *Proc. IPAC'25*, Taipei, Taiwan, Jun. 2025, pp. 2901–2904. doi:10.18429/JACoW-IPAC25-THPM101

[5] E. J. Prebys *et al.*, "New corrector system for the Fermilab booster", *Conf. Proc. C*, vol. 070625, p. 467, 2007. doi:10.1109/PAC.2007.4440247

[6] J. Eldred, V. Lebedev, K. Seiya, and V. Shiltsev, "Beam intensity effects in Fermilab booster synchrotron", *Phys. Rev. Accel. Beams*, vol. 24, no. 4, p. 044001, Apr. 2021. doi:10.1103/PhysRevAccelBeams.24.044001

[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, Aug. 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, May 2023, pp. 4428–4431. doi:10.18429/JACoW-IPAC2023-THPL007

[9] R. Roussel *et al.*, "Bayesian optimization algorithms for accelerator physics", *Phys. Rev. Accel. Beams*, vol. 27, no. 8, p. 084801, Aug. 2024. doi:10.1103/PhysRevAccelBeams.27.084801

[10] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, Nov. 2005. doi:10.7551/mitpress/3206.001.0001

[11] S. Daulton, M. Balandat, and E. Bakshy, "Parallel Bayesian optimization of multiple noisy objectives with expected hypervolume improvement", in *Proc. NeurIPS'21*, virtual, Dec. 2021, pp. 2187–2200.

[12] N. Kuklev *et al.*, "High efficiency multi-objective bayesian algorithm for aps-u nonlinear dynamics tuning", in *Proc. IPAC'25*, Taipei, Taiwan, Jun. 2025, pp. 2905–2908. doi:10.18429/JACoW-IPAC2025-THPM102

[13] M. Balandat *et al.*, "BoTorch: a framework for efficient Monte-Carlo Bayesian optimization", in *Proc. NeurIPS'20*, virtual, Dec. 2020, pp. 21524–21538.

[14] F. Berkenkamp, A. P. Schoellig, and A. Krause, "Safe controller optimization for quadrotors with Gaussian processes", in *Proc. ICRA'16*, Stockholm, Sweden, May 2016, pp. 491–496. doi:10.1109/icra.2016.7487170