ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL012

# BAYESIAN OPTIMIZATION CALIBRATION OF IONIZATION PROFILE MONITOR AT THE AGS COMPLEX*

Y. Gao†, K. Brown[1], J. Jamilkowski, J. Morris, L. Nguyen, V. Schoefer
C-AD Dept., Brookhaven National Laboratory, Upton, NY, USA
G. H. Hoffstaetter[2], W. Lin, CLASSE, Cornell University, Ithaca, NY, USA
[1]also at ECE Dept., Stony Brook University, Stony Brook, NY, USA
[2]also at Brookhaven National Laboratory, Upton, NY, USA



*Abstract*

The Ionization Profile Monitors (IPMs) are used to measure the transverse profiles of the beams accelerated at the Brookhaven National Laboratory (BNL) Alternating Gradient Synchrotron (AGS). These devices use Multi-Channel Plates (MCP) to collect electrons generated by ionization of the residual gas to get an image of the beam projection onto two transverse planes. The gain values of those channels are crucial for the accurate measurements of beam parameters. Various errors in the system can affect the channel gains - initial fabrication variations, channel aging, etc. Moreover, there are systematic errors associated with varying delays in the digitization paths for different groups of channels. In this work, we present a way of using Bayesian Optimization (BO) to calibrate the channel gains. Simulation results show that the BO approach can compensate for the errors quite well and enable better learning of the group sampling function.

## INTRODUCTION

Ionization Profile Monitors (IPMs) have been developed at Brookhaven National Laboratory (BNL) to measure transverse beam profiles in RHIC [1–3]. When the beam passes through the beamline, it ionizes the background gas and emits electrons. Those electrons are swept transversely from the beamline and collected by the Multi-Channel Plates (MCP) on 64 strip anodes oriented parallel to the beam axis. An IPM collects and measures the distribution of those electrons[1]. Ideally, the distributions should be independent of where the beam signal locates in the IPM. In other words, if the beam is moved across the channels, the IPM measurements from different locations should have identical shapes. One iteration of such channel scanning by moving beams across different channels is called a position scan.

The gain value of each channel is the dominant factor in determining what a final distribution looks like. There are various errors existing in the system that can affect the channel gains. Figure 1 illustrates the beam profiles from a position scan without calibration. We can see that the beam profile has a large variation when measured from different locations. The channel gain errors may result from initial channel-to-channel gain variations, depletion of channel gains due to aging, etc. There are also systematic errors associated with varying delays in the digitization paths for different groups of channels.

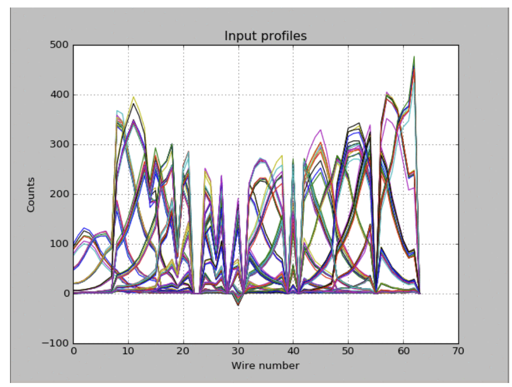


Figure 1: Beam profiles from a position scan.

Currently, the channel gains are calibrated manually. There are two ways to approach it. The first approach compares the beam profiles across different channels. It then tunes the channel gains so that the position scan produces almost equal profile amplitudes. The second approach compares the profiles on single channels. When the beam moves across different channels, each channel will record an envelope of the beam. The calibration is performed by setting gains which lead to equal envelopes.

Usually, the corrections for channels derived from those two methods agree well. The profiles after calibration are shown in Fig. 2. As we can see, the gain errors are mostly evened out and the beam profile measurements can be more useful now.

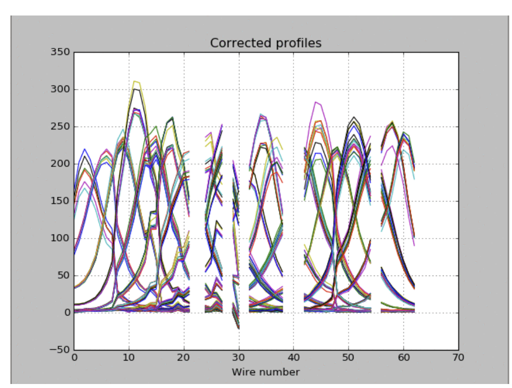


Figure 2: Beam profiles after calibration.

* Work supported by Brookhaven Science Associates, LLC under Contract No. DE-SC0012704 with the U.S. Department of Energy.
† ygao@bnl.gov
[1] An ideal beam profile should resemble a Gaussian shape.

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL012

However, the disadvantage of manual tuning is that the sample efficiency is very low, and it would not be practical to scan every possible setting in the real machine to select the best parameter. This work presents a way of using Bayesian Optimization (BO) to address this issue. After training, BO can efficiently find a solution that produces better calibration results than the existing method.

Bayesian Optimization (BO) is a powerful tool for optimizing an objective function $f$ with as few samples as possible [4, 5], and has been widely applied across various fields [6]. It is especially useful when the explicit expression of $f$ is unknown, and the evaluation of $f$ is expensive (e.g., any process involving human labor). Two essential components of BO are the surrogate model and acquisition function. In this work, we use a Gaussian Process (GP) with a Monte-Carlo (MC) based batch Expected Improvement (EI) acquisition function[7–9].

## SIMULATION RESULTS

We demonstrate the performance of the BO calibration by using a simulator. The simulator takes into account gain errors, group sampling function, and system white noise. It takes the gain calibrations as the inputs and outputs the corresponding signal strength for each channel.

### *Simulation Setups*

The gain errors used in the simulation are attained from a real machine scan in Apr. 2022, as shown in the top plots of Fig. 3.

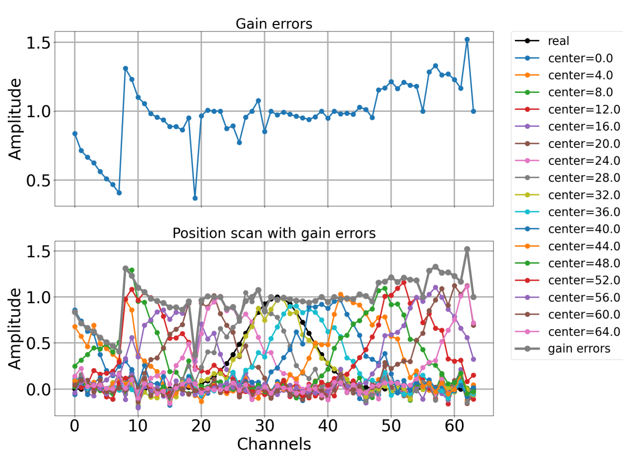


Figure 3: Position scan with gain errors acquired from the real system. It shows that the gain errors become the dominant factor in shaping the beam profiles.

A position scan under those gain errors is shown at the bottom of Fig. 3. The positions are selected across the 64 channels. The original beam distribution is plotted in black. As we can see, the gain errors (in grey) are the dominant factor in shaping the beam profiles, indicating the necessity of error calibration to obtain more useful beam profile measurements.

### *Gain Error Calibration using Bayesian Optimization*

We apply BO on both cross-channel and single-channel approaches. Due to the so-called curses of dimensionality [10], for the cross-channel approach, we organize the channels by groups and BO directly optimizes each group instead of individual channels.

Gaussian fit is performed on the beam profile measurements and the amplitude of the fit profile across all measured locations is used as the merit to guide the BO process. The goal is to make all the fitted amplitudes[2] close to 1. For the cross-channel approach, we use a group size of 4, 100 starting samples, and BO runs for 100 rounds. For the single-channel approach, 4 channels are optimized together with 100 starting samples and BO takes 100 samples. As a benchmark, a brute force method is also implemented where for each channel group, scanning is performed and the gain value that renders the highest objective is selected for that group.

Comparison results for the fitted amplitudes and sigmas are shown in Fig. 4. We can see that the calibration procedure significantly improves the beam profile quality by reducing their amplitude and sigma variations.

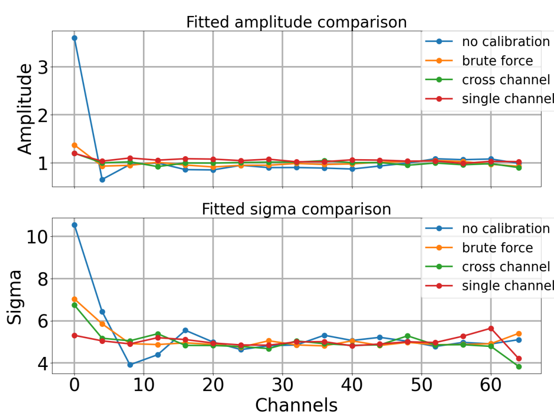


Figure 4: Comparison results of fitted amplitude and sigma from different schemes.

Table 1 summarizes the average and std values of the fitted amplitude and sigma with the best values underlined. We can see that both the cross-channel and single-channel methods generate less variational profiles than the greedy method, and are much better than the no calibration scenario. The single-channel method regulates both amplitude and sigma slightly better than the cross-channel method. The trade-off is that it takes longer for the single-channel method to get the optimal solution since at each time only one channel is optimized.

### *Learn the Group Sampling Function*

In addition to independent channel gains, the channels can have gain errors correlated by the instrument electron-

[2] Number 1 is used because it's the real signal's amplitude. It works for any other constant. The purpose is to equalize those fitted amplitudes.

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL012



Table 1: Average and std values of the fitted amplitude and sigma from different calibration schemes.

| Method | | Average | std |
|---|---|---|---|
| No calibration | amplitude | 1.09 | 0.63 |
| | sigma | 5.33 | 1.40 |
| Greedy | amplitude | 1.00 | 0.10 |
| | sigma | 5.11 | 0.55 |
| Cross-channel | amplitude | 1.00 | 0.06 |
| | sigma | 4.99 | 0.54 |
| Single-channel | amplitude | 1.06 | **0.04** |
| | sigma | 5.00 | **0.28** |

ics. The electrons impinging on the MCP create a pulse which is then sampled and digitized in groups of 8 sequentially in time. Since each group of 8 samples at a different time on the varying waveform, each group has an effective gain determined by its sampling time and the structure of the waveform. We call the time response function of the channels the group sampling function $\mathscr{F}$.

Hence, by controlling the starting time of the first group, the multipliers for an entire beam profile are settled. A typical scan of the starting time on the group sampling function $\mathscr{F}$ is shown in Fig. 5, where the starting time is increased from $-180$ to 180 with a step size of 20. We can see that under the same system condition, different starting times can manipulate the beam profiles on a noticeable scale.

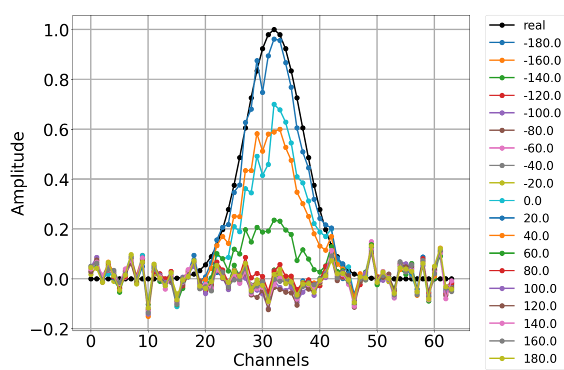


Figure 5: Beam profiles with different starting times from the group sampling function $\mathscr{F}$. The shapes can be manipulated by tuning the starting time.

To learn the group sampling function, we scan the starting time using the same fashion above. For each starting time, one signal in a group (usually the signal located at the center of a group) is selected as the group representative. All the maximum fitted amplitudes of the group representatives are recorded, and those amplitudes can be seen as a section of points from the function $\mathscr{F}$. Then a sketch of the function $\mathscr{F}$ can be plotted by connecting those sections.

Figure 6 shows the sketches of the function $\mathscr{F}$ before and after calibration[3]. We can see that after calibration, the learned function $\mathscr{F}$ has a better resolution.

[3] Here, we use the single-channel calibration.

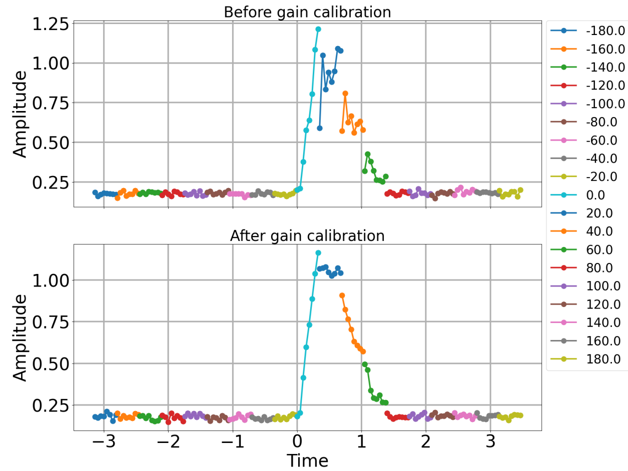


Figure 6: Comparisons of the group sampling function sketches before and after gain calibration. We can see that calibration improves the resolution of the learned function $\mathscr{F}$.

Furthermore, when the gain errors are systematic, calibration is necessary to learn a sensible group sampling function, as shown in Fig. 7.

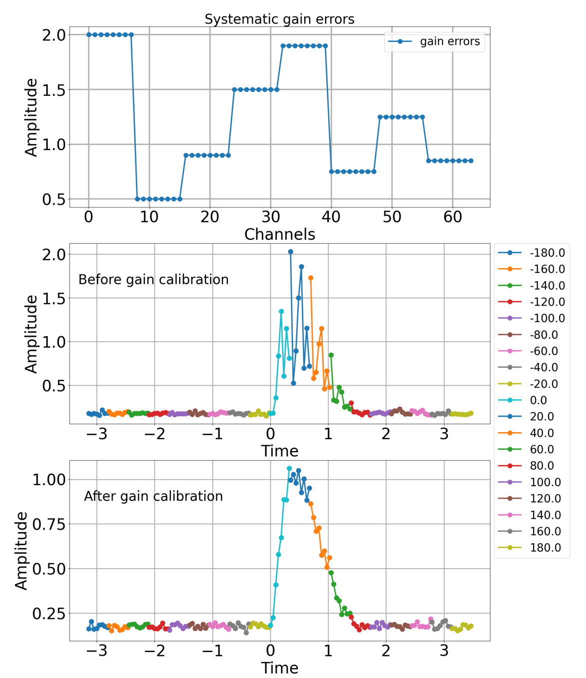


Figure 7: With systematic gain errors (top), calibration is necessary to learn a sensible group sampling function.

## CONCLUSION

In this work, we explore the possibility of using Bayesian optimization to calibrate the channel gains in the AGS IPMs. The simulation results show that BO can effectively correct the gain errors and facilitate the learning of the group sampling function.

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL012

## REFERENCES

[1] R. Connolly, C. Dawson, J. Fite, *et al.*, "A RHIC-Style IPM in the Brookhaven AGS,"

[2] R. Connolly, P. Cameron, R. Michnoff, and S. Tepikian, "Performance of the RHIC IPM," in *Proc. PAC'01*, Chicago, IL, USA, Jun. 2001. https://jacow.org/p01/papers/TPAH033.pdf

[3] R. Connolly, P. Cameron, W. Ryan, T. J. Shea, R. Sikora, and N. Tsoupas, "A Prototype Ionization Profile Monitor for RHIC," in *Proc. PAC'97*, Vancouver, Canada, May 1997.

[4] E. Brochu, V. M. Cora, and N. de Freitas, *A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning*, 2010.

[5] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

[6] B. Shahriari, K. Swersky, Z. Wang, *et al.*, "Taking the human out of the loop: A review of bayesian optimization," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016. doi:10.1109/JPROC.2015.2494218

[7] D. J. Rezende, S. Mohamed, and D. Wierstra, "Stochastic backpropagation and approximate inference in deep generative models," in *Proceedings of the 31st International Conference on Machine Learning*, 2014.

[8] J. T. Wilson, R. Moriconi, F. Hutter, *et al.*, "The reparameterization trick for acquisition functions," *ArXiv*, vol. abs/1712.00424, 2017.

[9] M. Balandat, B. Karrer, D. Jiang, *et al.*, "Botorch: A framework for efficient monte-carlo bayesian optimization," in *Advances in Neural Information Processing Systems*, vol. 33, 2020.

[10] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, 2007.