14th International Particle Accelerator Conference,Venice, Italy JACoW Publishing
ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL028

BAYESIAN OPTIMIZATION FOR SASE TUNING AT THE EUROPEAN XFEL

C. Xu*, E. Bründermann, A.-S. Müller, A. Santamaria Garcia
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
S. Tomin, Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Abstract

Parameter tuning is a regular task and takes considerable time for daily operations at FEL facilities. In this contribution, we demonstrate SASE pulse energy optimization at the European XFEL with Bayesian optimization (BO) as an alternative approach to the widely used simplex method. Preliminary experimental results show that BO could reach a comparable performance as the simplex method, even with an out-of-the-box implementation. Compared to previous attempts, our version of BO does not require setting hyperparameters via additional measurements, thus effectively reducing the required effort for machine operators to use it during operation. On the other hand, BO has the potential to be further improved by introducing prior physical knowledge about the task and fine-tuning the algorithm to specific tasks. This makes BO a promising candidate for routine tuning tasks at particle accelerators in the future.

INTRODUCTION

Free electron lasers (FEL) are complex large-scale facilities and provide intense X-ray pulses for numerous user experiments. One particular challenge is the self-amplified spontaneous emission (SASE) process, which is highly sensitive to the beam condition and significantly affects the output FEL power. Therefore, it can be time-consuming to tune the FEL and maximize the pulse energy. At the European XFEL (EuXFEL) and many other particle accelerator facilities, automatic tuning is often used during the operation to aid human operators in the simultaneous tuning of multiple control parameters. These various methods have been developed and applied to such tuning tasks, including robust conjugate direction search [1], extremum seeking [2, 3], and the Nelder-Mead simplex method [4, 5]. Particularly, Bayesian optimization (BO) has gained a lot of attention recently as a sample-efficient method for black-box global optimization. It has been successfully applied to a wide range of tasks in accelerator physics, ranging from parameter optimization in simulation [6–8] to online tuning [9–15]. In early works, BO has also been applied for the FEL tuning task [16, 17]. In practice, however, it is not routinely used because applying BO often requires expert knowledge, for example, to set the hyperparameters appropriately. As a result, Nelder-Mead simplex remains to be the most used tool in daily operation at EuXFEL due to its simplicity and reliability.

Recent advances in software and design choices have significantly reduced the overhead of implementing BO, making it a viable plug-and-play tuning tool for various new tuning tasks. In this contribution, we revisit the potential of utilizing BO as an operation tool for routine tuning tasks at FELs. We benchmarked the performance of a BO implementation, without any fine-tuning, against the daily-used simplex tool for tuning the SASE pulse energy at EuXFEL.

OPTIMIZATION ALGORITHM

Bayesian optimization (BO) reduces the required optimization steps by utilizing a statistical surrogate model of the objective, mostly a *Gaussian process* [18] (GP), which is built with the observation data. A GP is defined by its mean and covariance function $\mathcal{G}(\mu, k)$. The covariance function, also known as the *kernel*, measures the similarity between two input points. In this contribution, we use the Matérn-5/2 kernel to construct the full covariance function

$$k(x,x') = \sigma^2_{\text{signal}} k_{\text{Matérn-5/2}}\left(\left\|\frac{x-x'}{l}\right\|\right) + \sigma^2_{\text{noise}}\delta(x,x'). \tag{1}$$

The signal variance σ^2_{signal} scales the covariance function and the extra white noise term σ^2_{noise} describes the measurement noise. The lengthscale l describes how sensitive the objective function is in each input parameter dimension.

The GP hyperparameters $\{\sigma^2_{\text{signal}}, l, \sigma^2_{\text{noise}}\}$ are key components of Bayesian optimization, as they characterize the underlying GP model. In the earlier BO implementation of the Ocelot optimizer [16, 17], the hyperparameters are determined beforehand by fitting the archived data. This requires additional effort when applying BO for a new tuning task. Instead of fixing the hyperparameters, we determine them dynamically using log-likelihood fits during the optimization. This data-driven approach is more robust against machine condition changes and makes BO applicable to new tuning tasks without available archived data.

In each BO step, the GP model is updated using the observed data to predict the posterior mean $\mu(x')$ and uncertainty $\sigma(x')$ of a new point x'. The GP model predictions are further used to calculate a so-called acquisition function α to efficiently guide the optimization. The next sample point is chosen by maximizing the acquisition function. Here we use the upper confidence bound (UCB)

$$\alpha_{\text{UCB}}(x) = \mu(x) + \sqrt{\beta}\sigma(x), \tag{2}$$

where the exploration and exploitation are explicitly weighted by a trade-off parameter β.

* chenran.xu@kit.edu

Content from this work may be used under the terms of the CC BY 4.0 licence (© 2022). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL028

A well-known problem of BO is its over-exploration behavior as a global optimization method. Due to the exploration-exploitation trade-off, it sometimes performs large steps toward unexplored regions of the parameter space. In practice, this can either lead to a sudden performance drop or even cause damage in the case of high-energy accelerators [13, 15]. Safety or step size constraints are often introduced to mitigate this issue. In this contribution, we employed a technique called *proximal biasing* [19]. Instead of setting hard limits on the optimization step sizes, the next sample point is chosen by maximizing the product of the acquisition function and a normal distribution $\mathcal{N}$ centered at the current settings $x^{(t)}$

$$x^{(t+1)} = \arg\max_{x} \alpha(x) \cdot \mathcal{N}(x^{(t)}, l_b^2). \tag{3}$$

This acts effectively as a soft limit, as the possibility to sample distant parameter settings is reduced depending on the biasing lengthscale l_b. We used the BoTorch [20] framework to implement the Bayesian optimization algorithm mentioned above.

SASE OPTIMIZATION AT EUXFEL

For the measurements at European XFEL, we focused on the FEL pulse energy optimization by adjusting the beam orbit in the undulator sections. The SASE1 beamline consists of 35 undulator cells, each equipped with two pairs of horizontal and vertical air coil correctors. In order to evaluate the performance of the optimization algorithms, we manually detuned the settings and reduced the X-ray pulse energy to about one order of magnitude lower than what is obtained in normal operation. This was achieved by disturbing the beam orbit at the beginning of the undulators section. The BO and simplex optimizer were tasked to restore the FEL performance by changing the air-coil corrector magnets downstream of the undulators. The FEL lasing process is highly non-linear and stochastic in nature, which leads to high noise in the measured X-ray pulse energies. The noisy signal further affects the performance of the optimizers and can lead to convergence in local optima, instead of the global one. To mitigate that, we average 30 pulses to obtain a cleaner signal, corresponding to 3 s per observation at 10 Hz repetition rate. The measured standard deviation of the signal depending on the pulse energy is shown in 1. Apart from the linear dependency for low pulse energies, most signal noises are centered at $\mu = 126\,\mu\text{J}$ and $\sigma = 100\,\mu\text{J}$ due to the stochastic beam condition. We first benchmarked BO and simplex methods using one pair of air coils with the same initial setting. The evolution of the two corrector values is shown in 2. Compared to the simplex behavior, where parameters were oscillating before converging, the proximal biased BO showed a much smoother convergence towards the optimum. Although both methods were able to maximize the pulse energy within 50 steps using air coils in this case, a smooth parameter change is expected to be more beneficial in situations where magnetic hysteresis is present.

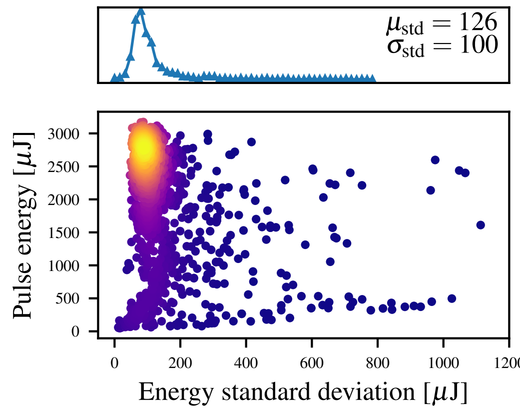

Figure 1: Measurement noise with respect to the photon pulse energy. The lower plot shows measured X-ray pulse energies averaged over 30 shots and the corresponding standard deviations, where the density of the measured points is color-coded. The upper plot shows the measurement noise integrated over all energies, giving an average noise of 126 µJ.

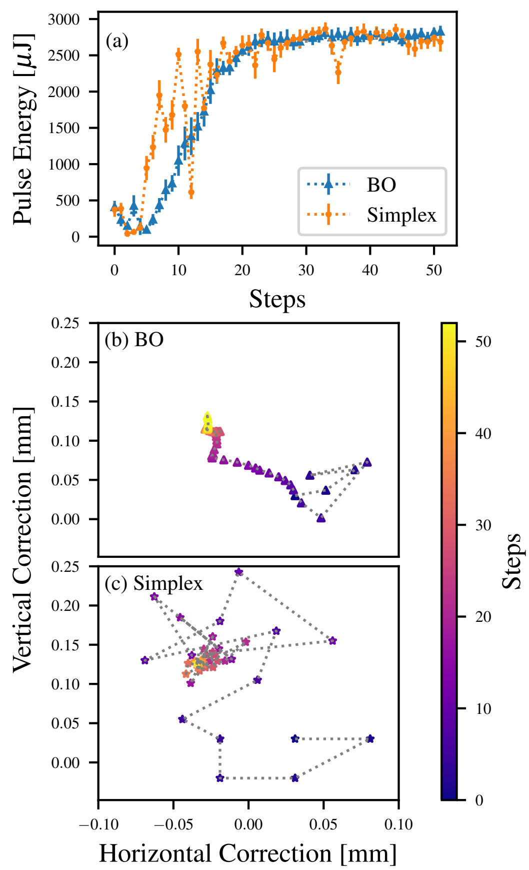

Figure 2: Progress of BO and simplex for the two-dimensional optimization task using one pair of horizontal and vertical correctors in an undulator. The measured X-ray pulse energies are shown in (a). The evolution of the corrector values during the optimization steps for BO and simplex are visualized in (b) and (c) respectively.

Content from this work may be used under the terms of the CC BY 4.0 licence (© 2022). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL028

Next, we repeated the optimization with an increasing number of randomly selected air coils. As can be seen in the 1, BO and simplex reached almost the same final pulse energies in most cases. The only exception is the four-dimensional case, where BO seemed to converge to a local optimum. Furthermore, we also calculated the steps to convergence for each method.

Table 1: Optimization results of BO and simplex with up to ten tuning parameters. The number of steps to convergence is defined as the step after which the variation of the objective values is smaller than the 226 µJ corresponding to the measured signal noise.

	Final pulse energy [µJ]		Convergence steps	
Dim.	BO	Simplex	BO	Simplex
2	2880	2864	22	36
4	2500	2900	27	28
6	2845	2852	34	45
8	3120	2944	102	105
10	3011	3049	84	78

Here we define convergence as the variation of the objective values is smaller than 226 µJ, corresponding to 1σ upper bound of the measurement noise. The required optimization steps to convergence increase with the number of tuning parameters, where BO reached a faster convergence than the simplex method in four out of five trials. Both methods were able to restore the FEL performance using 10 tuning parameters within 100 steps, corresponding to about 5 min of beamtime. Although one could increase the number of tuning parameters to include all the available correctors, it is important to note that using BO to tune a large number of parameters can significantly slow down the optimization process, as the search space grows exponentially. Therefore, it is often preferred in practice to run several optimizations with subsets of tuning parameters consecutively to find a good enough setting for its speed and robustness.

An important aspect of BO to be mentioned is that it provides more insight into the tuning task thanks to the GP model. For example, we visualized the task using four air coil correctors in 3. The predicted posterior mean function of each pair of input parameters is shown, where the other two parameters take the average values. It can be seen that the pulse energy is very sensitive with respect to both horizontal and vertical correctors in cell 3, whereas the condition is more relaxed for the horizontal corrector in cell 7. This could either assist the operator in monitoring the optimization process during operation or help gain a better understanding of the system's dynamics afterwards.

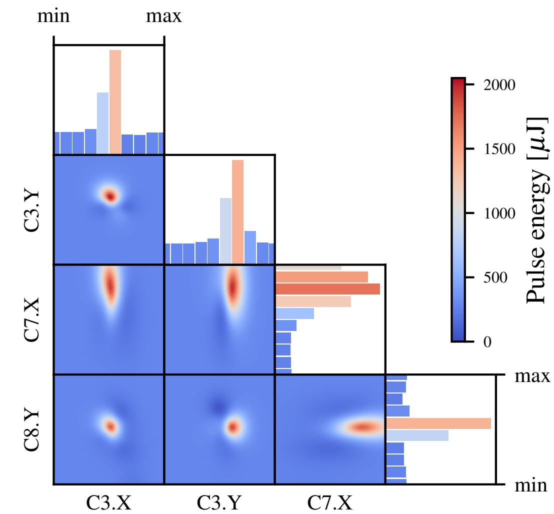

Figure 3: GP Model visualization of the tuning task with four correctors located in undulator cells 3, 7, and 8, where X and Y denote horizontal and vertical respectively. The two-dimensional subspace plots show the posterior mean predicted by the GP model with different combinations of correctors, with the other two taking the averaged values. The one-dimensional histograms show the dependencies of the pulse energy on the individual corrector.

CONCLUSION AND OUTLOOK

We successfully applied BO to the SASE pulse energy tuning task at the European XFEL. Even without fine-tuning it to the tasks, BO achieved comparable final pulse energies and slightly faster convergence speed than the existing simplex method in the Ocelot optimizer, which is used in day-to-day operations. With the help of advanced software implementations, one no longer needs dedicated beamtime or relies on archived data for hyperparameter tuning, which significantly reduces the overhead of applying BO for a new tuning task. In addition, the fact that no expert knowledge is needed during the application makes BO a promising operator-friendly tool in the accelerator control room. Domain knowledge could be incorporated into the BO to further increase its performance, for example by a physics-informed kernel [11] or neural network priors [21, 22] for the GP model. This extends BO's applicability to higher-dimensional tuning tasks or systems with more complex dynamics.

The Bayesian optimizer used in this paper is implemented as a general-purpose tool, which could be easily applied to various tuning tasks at different accelerators. In the future, it is foreseen to integrate BO with the existing optimization frameworks like Ocelot optimizer and Badger [23] and use it as a routine tuning tool. The code package and data used in this paper are available at the GitHub repository [24].

ACKNOWLEDGEMENTS

The authors thank the EuXFEL operation group and Zihan Zhu for their support for live tests at the European XFEL. C. Xu acknowledges the support by the DFG- funded Doctoral School "Karlsruhe School of Elementary and Astroparticle Physics: Science and Technology".

Content from this work may be used under the terms of the CC BY 4.0 licence (© 2022). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-THPL028

Content from this work may be used under the terms of the CC BY 4.0 licence (© 2022). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

REFERENCES

[1] X. Huang, J. Corbett, J. Safranek, and J. Wu, "An algorithm for online optimization of accelerators," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 726, pp. 77–83, 2013. doi:10.1016/j.nima.2013.05.046

[2] N. Bruchon, G. Fenu, G. Gaio, M. Lonza, F. A. Pellegrino, and L. Saule, "Free-electron laser spectrum evaluation and automatic optimization," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 871, pp. 20–29, 2017. doi:10.1016/J.NIMA.2017.07.048

[3] A. Scheinker *et al.*, "Model-independent tuning for maximizing free electron laser pulse energy," *Physical Review Accelerators and Beams*, vol. 22, 8 2019. doi:10.1103/PhysRevAccelBeams.22.082802

[4] S. Tomin *et al.*, "Progress in automatic software-based optimization of accelerator performance," 2016, pp. 3064–3066.

[5] I. Agapov, G. Geloni, S. Tomin, and I. Zagorodnov, *Automatic tuning of free electron lasers*, 2017.

[6] A. F. Pousa *et al.*, "Bayesian optimization of laser-plasma accelerators assisted by reduced physical models," 2022. doi:10.48550/arxiv.2212.12551

[7] F. Irshad, S. Karsch, and A. Döpp, "Multi-objective and multi-fidelity bayesian optimization of laser-plasma acceleration," *Physical Review Research*, vol. 5, p. 013 063, 1 2023. doi:10.1103/PHYSREVRESEARCH.5.013063

[8] C. Xu, E. Bründermann, A.-S. Müller, A. S. Garcia, M. Schwarz, and J. Schäfer, "Optimization Studies of Simulated THz Radiation at FLUTE," in *Proc. IPAC'22*, Bangkok, Thailand, 2022, paper WEPOMS023, pp. 2292–2295. doi:10.18429/JACoW-IPAC2022-WEPOMS023

[9] R. J. Shalloo *et al.*, "Automation and control of laser wakefield accelerators using bayesian optimization," *Nature Communications*, vol. 11, pp. 1–8, 1 2020. doi:10.1038/s41467-020-20245-6

[10] S. Jalas *et al.*, "Bayesian optimization of a laser-plasma accelerator," *Physical Review Letters*, vol. 126, 10 2021. doi:10.1103/PhysRevLett.126.104801

[11] A. Hanuka *et al.*, "Physics model-informed gaussian process for online optimization of particle accelerators," *Phys. Rev. Accel. Beams*, vol. 24, p. 072 802, 7 2021. doi:10.1103/PhysRevAccelBeams.24.072802

[12] R. Roussel, A. Hanuka, and A. Edelen, "Multiobjective bayesian optimization for online accelerator tuning," *Physical Review Accelerators and Beams*, vol. 24, p. 062 801, 6 2021. doi:10.1103/PHYSREVACCELBEAMS.24.062801

[13] J. Kirschner, M. Mutný, A. Krause, J. Coello de Portugal, N. Hiller, and J. Snuverink, "Tuning particle accelerators with safety constraints using bayesian optimization," *Phys. Rev. Accel. Beams*, vol. 25, p. 062 802, 6 2022. doi:10.1103/PhysRevAccelBeams.25.062802

[14] Y. Gao *et al.*, "Bayesian optimization experiment for trajectory alignment at the low energy rhic electron cooling system," *Physical Review Accelerators and Beams*, vol. 25, 2022. doi:10.1103/PhysRevAccelBeams.25.014601

[15] C. Xu, T. Boltz, A. Mochihashi, A. Santamaria Garcia, M. Schuh, and A.-S. Müller, "Bayesian optimization of the beam injection process into a storage ring," *Phys. Rev. Accel. Beams*, vol. 26, p. 034 601, 3 2023. doi:10.1103/PhysRevAccelBeams.26.034601

[16] M. McIntire, T. Cope, S. Ermon, and D. Ratner, "Bayesian Optimization of FEL Performance at LCLS," in *Proc. of International Particle Accelerator Conference (IPAC'16), Busan, Korea, May 8-13, 2016*, Busan, Korea, 2016, pp. 2972–2975. doi:doi:10.18429/JACoW-IPAC2016-WEPOW055

[17] J. Duris *et al.*, "Bayesian optimization of a free-electron laser," *Physical Review Letters*, vol. 124, 12 2020. doi:10.1103/PhysRevLett.124.124801

[18] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2005.

[19] R. Roussel *et al.*, "Turn-key constrained parameter space exploration for particle accelerators using bayesian active learning," *Nature Communications 2021 12:1*, vol. 12, pp. 1–7, 1 2021. doi:10.1038/s41467-021-25757-3

[20] M. Balandat *et al.*, "Botorch: A framework for efficient monte-carlo bayesian optimization," *Advances in Neural Information Processing Systems*, vol. 2020-December, 2019. https://arxiv.org/abs/1910.06403v3

[21] K. Hwang *et al.*, *Prior-mean-assisted bayesian optimization application on frib front-end tunning*, 2022.

[22] C. Xu, R. Roussel, and A. Edelen, *Neural network prior mean for particle accelerator injector tuning*, 2022.

[23] Z. Zhang *et al.*, "Badger: The Missing Optimizer in ACR," in *Proc. IPAC'22*, Bangkok, Thailand, 2022, pp. 999–1002. doi:10.18429/JACoW-IPAC2022-TUPOST058

[24] C. Xu, *Bayesian Optimization Attempts for European XFEL Tuning*. https://github.com/cr-xu/bo-4-euxfel