OPTIMAL DESIGN OF ILC E-DRIVEN POSITRON SOURCE WITH MACHINE LEARNING

M. Kuriki *, S. Kuroguchi, T. Takahashi, Z. Liptak,
Graduate School for Advanced Science and Engineering, Hiroshima University
J. Urakawa, Japan Society for the Promotion of Science
Y. Enomoto, T. Omori, M. Fukuda, Y. Morikawa, K. Yokoya,
High Energy Accelerator Research Organization

Abstract

The International Linear Collider (ILC) is a next-generation electron-positron collider designed to operate in the center-of-mass energy range of 250 GeV to 1 TeV, allowing one to explore physics beyond the Standard Model. One of the most important components of ILC is the E-Driven positron source generating 2.0×10^{14} positron per second. The positron source should be composed of state-of-the-art technology and fully optimized operation and design. Conventional accelerator design methods involve sequential optimization, which is inefficient and difficult to achieve overall optimization. In this study, the Tree-structured Parzen Estimator (TPE) algorithm, one of the black-box optimization methods, was introduced to improve the design efficiency of the electron-driven positron source of the ILC. After optimization, positron yield defined as the number of the captured positron in DR acceptance normalized with the number of drive electron, of 1.48 was obtained, which is much higher than the 1.20 obtained by manual optimization. This significant improvement is expected to meet safety standards for target destruction with a larger margin. The optimization process was also expedited, reducing the time from about one week to about half a day. These results demonstrate the potential of machine learning techniques in accelerator design to provide a more comprehensive overall optimization by exploring a wider parameter space and avoiding local minima.

INTRODUCTION

ILC is an $e+e-$ linear collider with the center of mass energy 250 GeV - 1000 GeV [1]. It employs a super-conducting accelerator as the main accelerator. The beam is accelerated in a macro pulse with 1300 bunches by 5 Hz repetition. The bunch charge is 3.2 nC resulting in the average beam current 21 μA. This is a technical challenge, because the amount of positron per second is 40 times larger than that in SLC [2], which was the first linear collider.

The configuration of the positron source is schematically shown in Fig. 1. The positron generated by electron beam as the electromagnetic shower is captured and boosted up to 5 GeV by two linacs. In the E-Driven ILC positron source, the drive beam energy is 3.0 GeV and the target is 16 mm thick W-Re alloy rotating with 5 m/s tangential speed. FC (Flux Concentrator) generates a strong magnetic field along the beam axis to compensate the transverse momentum. 36 1.3 m L-band Standing Wave (SW) cavities with 0.5 Tesla solenoid field are placed for positron capture. This section is called as the capture linac. At the downstream, a chicane is placed to removes electrons. The positron booster is composed from 2.0 m L-and and 2.0 m S-band Traveling Wave (TW) cavities. ECS (Energy Compression Section) is composed from 3.0 m L-band TW cavities with chicane.

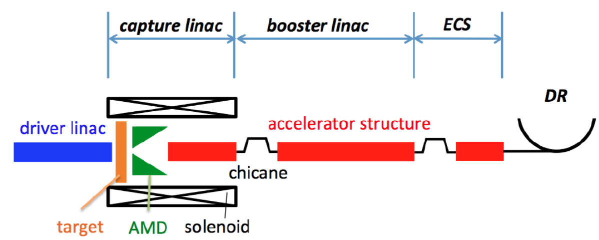

Figure 1: Configuration of E-Driven ILC positron source is schematically shown.

A first simulation was performed by T. Omori [3] only for the capture linac. A simulation with the tracking down to DR (Damping Ring) was made by Y. Seimiya [4], but no beam loading effect was accounted. A new simulation accounting the beam-loading effect was done by Kuriki and Nagoshi [5, 6]. For those simulations, the peak energy deposition density on the target is kept less than 35 J/g [7], which is considered to be a practical limit of the safety operation.

To obtain uniform intensity positrons over the pulse, the transient variation of the acceleration field by the beam loading has to be compensated so that positrons are accelerated uniformly. Compensation for the transient beam loading by Amplitude Modulation (AM) implemented by mixing of two inputs of klystron with Phase Modulation (PM) was proposed by Urakawa [8, 9]. The detail study of the compensation is discussed in Ref. [10, 11].

Figure 2 shows the pulse structure of the positron generation which has a 80 ns gap in the middle. A special treatment is needed to obtain a uniform intensity bunch train and it was stdied in Ref. [12].

The positron yield η is defined as the number of captured positrons in Damping Ring (DR) acceptance N_{e+} normalized with the number of electrons on the target N_{e-} as

* mkuriki@hiroshima-u.ac.jp

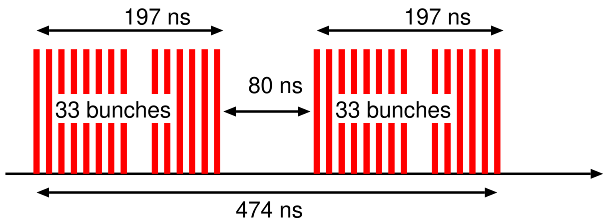

Figure 2: An example of macro-pulse structure in E-Driven ILC positron source. The macro pulse is composed from two mini-trains of 33 bunches with 6.15 ns bunch spacing. There is a gap of 80 ns.

$\eta = N_{e+}/N_{e-}$, where the DR acceptance is

$$\left(\frac{z}{35[mm]}\right)^2 + \left(\frac{\delta}{0.75[\%]}\right)^2 < 1$$
$$\gamma A_x + \gamma A_y < 70[mm], \tag{1}$$

in the longitudinal and transverse phase space.

The positron yield η has a large impact on the E-Driven positron source because the electron beam intensity on the target is inversely proportional to the positron yield η as $N_{e-} = N_{IP}/\eta$, where N_{IP} is the number of particles at the interaction point. N_{IP} is identical both for electron and positron. In the ILC design, it is 4.8 nC including 50% margin.

The primary concern in designing an ILC positron source is target destruction; from past experience with SLC [2], a safety threshold of 35 J/g is known for W-Re targets. The index is known as PEDD (Peak Energy Deposition Density) which is the maximum value of the energy normalized with the density. This safety threshold does not mean that the target will be destroyed if the value is exceeded, but rather that long-term stable operation is possible if the value is kept below the safety threshold. The goal of the design is to keep the energy load on the target below this value. 3. 0 GeV 3.75 nC electron bunch gives 33.6 J/g PEDD [6]. PEDD per electron bunch charge is 8.96 J/g.nC. It leads the following relation with η as

$$PEDD[J/g] = 8.96\frac{4.8}{\eta}. \tag{2}$$

Improving η gives more safety margin for the target operation.

BLACK BOX OPTIMIZATION

Black Box Optimization (BBO) is a mathematical optimization that aims to optimize (minimize or maximize) results in situations where only the output result can be obtained for a given input and no other information is given. The characteristics of this method can be understood by contrasting it with conventional optimization methods such as the Gradient descent method, which assumes that the differential coefficients are known. In the Gradient descent method, the objective function is optimized based on its differential coefficients, so there is always the possibility of falling into a pseudo-optimal solution (extreme value optimal solution), making optimizing a system with complex behavior difficult.

BBO is an appropriate method for optimizing systems with complex behavior. For example, if a system consists of many subsystems and the coupling between those subsystems and the objective function is not a simple linear combination, then most systems are complex. Accelerators are complex systems due to their large number of components, fundamentally nonlinear characteristics, and recursive nature, making BBO effective.

Direct search and Sequential Model-Based Global Optimization (SMBO) are typical approaches to BBO. The direct search method is a direct search type optimization algorithm that uses only the design variables and their objective function values. The direct search method does not use the gradient of the objective function, so it is difficult to fall into a local optimum, and thus has the feature of global optimization. Bayes optimization is a typical SMBO following three steps (1) generate a surrogate model from sample data based on a stochastic model, (2) construct an acquisition function corresponding to the stochastic model and find its optimum point, (3) sample the optimum point. Sampling is repeated at the optimum point.

In this study, we perform the system optimization with BBO by Tree-Structured Parzen Estimator (TPE) algorithm [13]. TPE is designed to optimize quantization hyperparameters and is an iterative process that uses history of evaluated hyperparameters to create a probabilistic model, which is used to suggest next set of hyperparameters to evaluate. The optimization is done in the following steps as,

1. Define a domain of hyperparameter search space. In our case, RF phase of accelerating cavity, momentum compaction of chicanes, etc.

2. Create an objective function that takes in hyperparameters and outputs a score. In our case, that is the number of positrons in DR acceptance. Note that the score in the algorithm is the inverse of the number because the score will be minimized.

3. Get observations (score) using randomly selected sets of hyperparameters,

4. Sort the collected observations by score and divide them into two groups based on some quantile. The first group (x_1) contains observations that gave the best scores and the second one (x_2) - all other observations,

5. Two densities $l(x_1)$ and $g(x_2)$ are modeled using Parzen Estimators (also known as kernel density estimators) which are a simple average of kernels centered on existing data points,

6. Draw sample hyperparameters from $l(x_1)$, evaluating them in terms of $l(x_1)/g(x_2)$, and returning the set

that yields the minimum value under $l(x_1)/g(x_1)$ corresponding to the greatest expected improvement. These hyperparameters are then evaluated on the objective function.

7. Update the observation list from step 3

8. Repeat step 4-7 with a fixed number of trials or until time limit is reached.

We choose TPE algorithm because it was stable even with a large number of parameters and trials. Time complexity is integral to the computation power for optimization. The time complexity of TPE algorithm is $O(dn \ln n)$, where d is the number of hyperparameters and n is the number of trials. It should be compared with that of Gaussian process regression which is typically $O(n^3)$. Optimization with TPE is lighter and faster than the Bayesian optimization with the Gaussian process regression.

ACCELERATOR OPTIMIZATION WITH BBO

The design of the E-Driven positron source for ILC is performed with several simulation codes. GEANT4 is employed to simulate the positron generation in the target with the 3 GeV electron beam. Positrons from the target is focused by Flux Concentrator which is a pulsed magnet followed by the capture linac composed from APS (Alternate Periodic Structure) standing wave cavity. This part is simulated with GPT (General Particle Tracer). After the capture linac, the positron is accelerated up to 5 GeV by the booster linac, which is composed from L-band and S-band traveling wave cavity. Before the booster linac, there is a chicane to remove electrons. After the booster, ECS (Energy Compression Section) is placed which is composed from chicane section and RF section. From the first chicane to ECS including the booster linac, SAD is employed. In this study, we apply BBO with the SAD simulation part.

The input parameters are

- The design momentum at the first chicane,
- The bending angle of the first chicane,
- K-value of Q triplets at the first chicane. There are two triplets at the upstream and downstream of the chicane.
- Booster linac RF phase,
- Booster linac RF voltage,
- The bending angle of the ECS chicane,
- RF Phase of ECS RF section,
- RF voltage of ECS RF section,
- Z-center of the Damping Ring RF.

The output function is the number of captured positrons in DR acceptance (dynamic aperture). As the TPE algorithm, we used TPESampler of Optuna. The simulation is performed with the single bunch, but the RF voltage is determined by considering the beam loading compensation with the amplitude and phase modulations [12].

The optimization is performed with a generic desktop PC and 15 out of 16 threads were used for the optimization. 10-fold speedup compared with a single thread was achieved. After the first optimization, the range of each parameter was narrowed and performed another round of optimization. We call the first optimization as the global parameter search and the second optimization as the local parameter search.

Figure 3 shows the results of the global optmization. The horizontal axis shows the number of trials, and the vertical axis is the objective value, the number of the captured positron. Points are the value of each trial and the line gives the best value in the progress. In the global optimization, the objective function (number of the captured positrons) is improved to 1414, which corresponds to 1.41 positron yield.

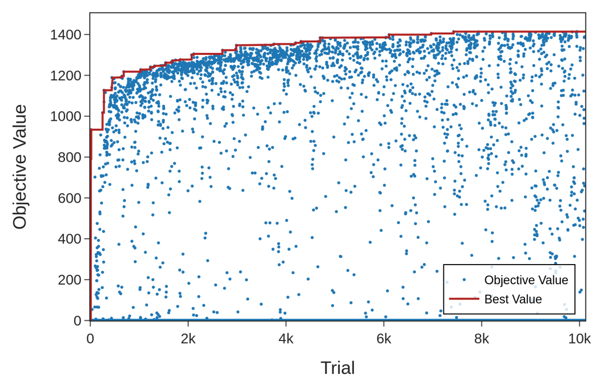

Figure 3: The progress of the trials of the global optimization. The horizontal axis shows the number of trials, and the vertical axis is the objective value, the number of the captured positron. Points are the value of each trial and the line gives the best value.

Figure 4 shows the results of the local optimization. The horizontal axis shows the number of trials, and the vertical axis is the objective value, the number of the captured positron. Points are the value of each trial and the line gives the best value. In the local optimization, the objective function (number of the captured positrons) is improved up to 1483.

We performed the optimization process three times. In each optimization, no previous knowledge is used. In each optimization, the global search was performed followed by the local search. The best objective values (number of the caputured positions) for each optimization were 1480, 1483, and 1495. The value of 1495 may have been over-optimized for statistical fluctuations; 1480 and 1483 are reasonable value.

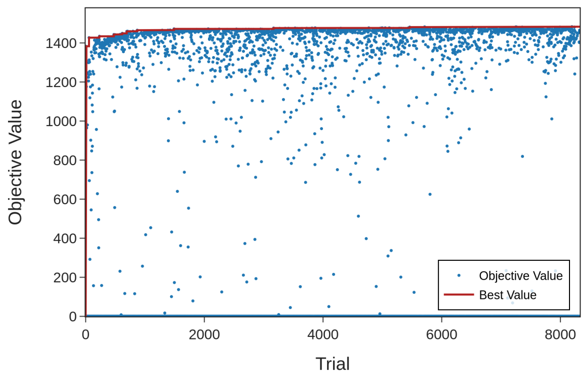

Figure 4: The progress of the trials of the local optimization. The horizontal axis shows the number of trials, and the vertical axis is the objective value, the number of the captured positron. Points are value of each trials and the line gives the best value.

In the simulation, we inject 1000 electrons on the target. The positron yield which is the number of the captured positrons normalized with the number of electrons on the target is obtained divided with 1000. The positron yield is evaluated as 1.48 ± 0.04, where we consider only the statical error.

The highest positron yield before the BBO study was 1.20 [6], which was obtained optimization by hand. The yield was improved by 23%

Figure 5 shows the objective value (Number of the captured positron) as a function of each parameter are shown. The color scale of the points shows the progress of the trials. The parameters are a) Booster RF phase, b) Booster L-band cavity voltage, c) Bending angle of the first chicane, d) Z-position of Damping Ring, e) Bending angle of the ECS chicane, f) RF phase of ECS RF, g) Peak voltage of ECS RF, h) The design momentum of the first chicane, i) K-value of the first FODO, j) K-value of the second FODO.

The design momentum at the first chicane is optimized to much lower than the actual average momentum, ~ 240 MeV/c. In the downstream of the capture linac, positrons are widely spread out in momentum space. Optimizing the optical design for the lower-momentum positrons probably improves the overall transmission.

A NEW DESIGN OF E-DRIVEN ILC POSITRON SOURCE

As a result of the optimization, the electron bunch charge on the target becomes 3.24 nC giving 28.7 J/g PEDD. These numbers should be compared with 3.75 nC and 33. 6 J/g by H. Nagoshi [6]. The electron bunch charge and PEDD is decreased 15 % improving the safety margin of the target greatly. In addition, the operations of the capture linac and the booster linac are also improved because there is a heavy beam loading effect by a large beam current up to 2A. Especially, the beam loading in the capture linac is very serious because not only positrons but also electrons contribute to the heavy beam loading effect. The actual accelerating electric field is determined from conditions that suppress accelerating voltage fluctuations due to transient beam loading, which is highly dependent on the beam current [14]. In the capture linac, the average acceleration field is only 5 MV/m with 2 A beam current. If the beam current is reduced by 15%, i.e. 1.7 A, the average acceleration field becomes 8 MV/m, i.e. 60% improved. An improvement by the same mechanism is expected also in the booster linac, but the effect is smaller because the beam loading is lighter than that in the capture linac.

The evolution of the longitudinal phase space, z in the horizontal axis and δ in the vertical axis is shown in Fig. 6, 7, 8, and 9, where δ is defined as the energy spread normalized with the average. Figure 6 shows the distribution after the capture linac, where the positrons are captured in a RF bucket, but distributed over the wide RF phase space.

As shown in Fig. 7, the phase space distribution after the chicane is ”skewed” due to the effect of the momentum compaction. The momentum compaction (chicane angle) is one of the optimized parameters. It indicates that the average z-directional spread is suppressed by optimizing the momentum compaction of the chicane.

Figure 8 shows the longitudinal phase space after the booster. It can be seen that the positrons are distributed along the curve of the accelerating RF. From this point, it can be seen that in order to suppress the energy spread after acceleration in the booster, it is important to optimize the momentum compaction of the chicane before the booster to minimize the z spread.

Figure 9 shows the longitudinal phase space distribution after ECS. ECS is composed from chicanes generating momentum compaction and RF section. ECS rotates the distribution by 90 degree and the energy spread is compressed. The chicane angle, RF phase, and the RF amplitude are the optimized parameter. By optimizing those parameters, the matching condition corresponding to adjustment of the rotating angle to be 90 degree is met and the nonlinearity is controlled. The elliptical circle shows the longitudinal DR acceptance. The positron distribution is well optimized to the DR acceptance.

In this study, the simulations of the capture linac and the booster linac are performed with the old parameters, lower acceleration field. A further improvement is possible, especially on the capture linac by setting the correct acceleration field with the 3.24 nC bunch charge. A consistent study by a start-to-end simulation is the next issue.

SUMMARY

We performed a design optimization as BBO with TPE algorithm for the E-Driven ILC positron source. 1.48 positron yield η is obtained as a result of BBO. This number is 23% improved compared with the previous studies. The electron bunch on the target is 3.24 nC, decreased by 15% compared with 3.75 nC. PEDD is also decreased 28.7 J/g compared

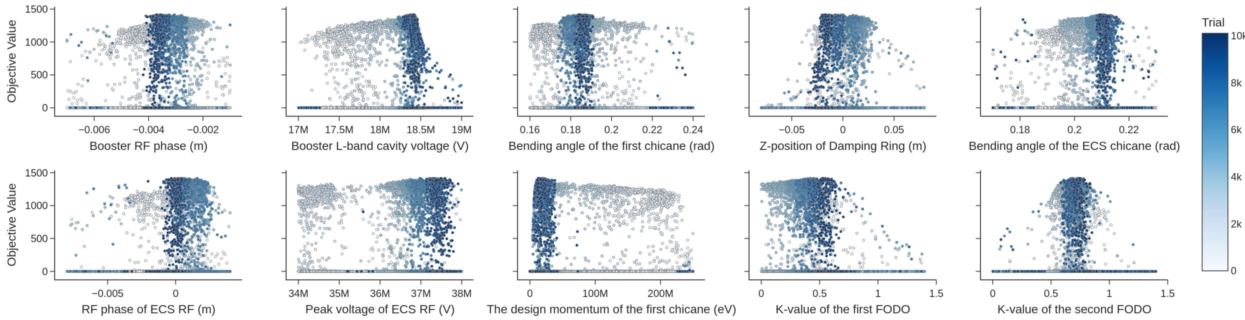

Figure 5: The objective value (Number of the captured positron) as a function of each parameters are shown. The color scale of the points shows the progress of the trials. The parameters are a) Booster RF phase, b) Booster L-band cavity voltage, c) Bending angle of the first chicane, d) Z-position of Damping Ring, e) Bending angle of the ECS chicane, f) RF phase of ECS RF, g) Peak voltage of ECS RF, h) The design momentum of the first chicane, i) K-value of the first FODO, j) K-value of the second FODO.

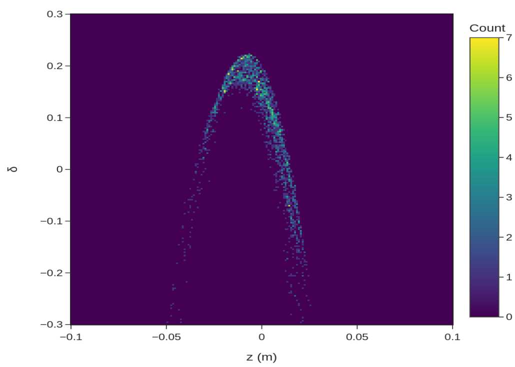

Figure 6: The longitudinal phase space after the capture linac is shown.

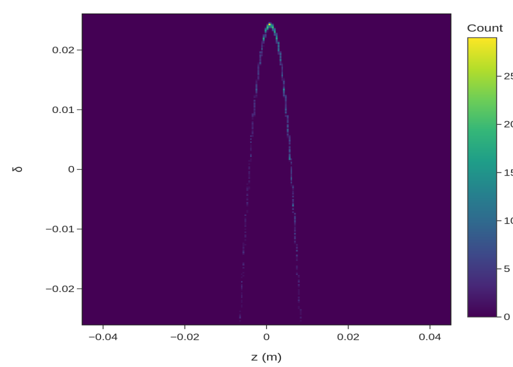

Figure 8: The longitudinal phase space after the booster linac is shown.

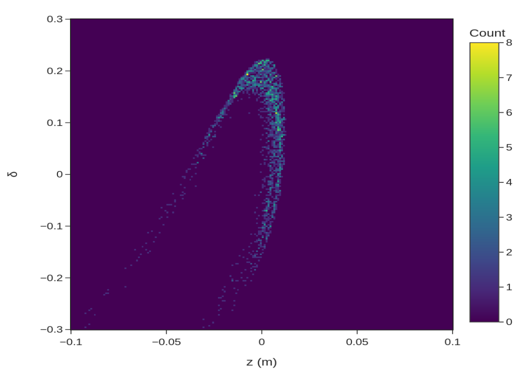

Figure 7: The longitudinal phase space after the first chicane in the upstream of the booster linac is shown.

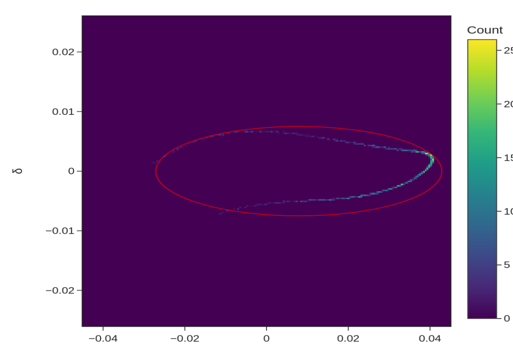

Figure 9: The longitudinal phase space after ECS is shown. The ellipse circle shows the longitudinal DR acceptance.

with 33.6 J/g. Large improvements on the capture linac and the booster linac are expected because of the heavy beam loading, especially on the capture linac. The study was performed with the old parameters on those linacs and a further improvement is expected.

ACKNOWLEDGEMENTS

This work is partly supported by Grant-in-Aid for Scientific Research (B) and US-Japan Science and Technology Cooperation Program in High Energy Physics.

REFERENCES

[1] ILC Technical Design Report, Rep. KEK-Report, 2013.

[2] SLC Design Report, Rep. SLAC-R-714, 1984.

[3] Omori T *et al.* A conventional positron source for international linear collider, *Nucl. Instr. and Meth. A*, vol. 672, no. 52, 2012. doi:10.1016/j.nima.2011.12.032

[4] Seimiya Y *et al.*, Positron capture simulation for the ILC electron-driven positron source, *Prog. Theor. Exp. Phys.*, vol. 2015, no. 10, Oct. 2015. doi:10.1093/ptep/ptv136

[5] Kuriki M *et al.*, "Electron Driven ILC Positron Source with a Low Gradient Capture Linac", in *Proc. LINAC'16*, East Lansing, MI, USA, Sep. 2016, pp. 430–433. doi:10.18429/JACoW-LINAC2016-TUPRC008

[6] Nagoshi H, Kuriki M *et al.*, A design of an electron driven positron source for the international linear collider, *Nucl. Instr. and Meth. A* vol. 953, p. 163134, 2020. doi:10.1016/j.nima.2019.163134

[7] Takahashi T, Radiation effect on the target and capture devices, in *Proc. of LCWS2018*, Arlington, Texas, USA, Oct. 2018.

[8] Kuriki M *et al.*, "Beam Loading Compensation of APS Cavity with Off-Crest Acceleration in ILC e-Driven Positron Source", in *Proc. IPAC'21*, Campinas, Brazil, May 2021, pp. 1368–1371. doi:10.18429/JACoW-IPAC2021-TUPAB015

[9] Kuriki M *et al.*, in *Proc. of PASJ2021*, Takasaki, Japan, Aug. 2021, paper MOP031.

[10] Kuriki M, "Energy Spread Compensation in Arbitrary Format Multi-Bunch Acceleration With Standing Wave and Traveling Wave Accelerators", in *Proc. IPAC'18*, Vancouver, Canada, Apr.-May 2018, pp. 307–310. doi:10.18429/JACoW-IPAC2018-MOPMF076

[11] Kuriki M *et al.*, "Beam Loading Compensation of Standing Wave Linac With Off-Crest Acceleration", presented at the IPAC'22, Bangkok, Thailand, Jun. 2022, paper WEPOPT024.

[12] Kuriki M, "Energy Spread Compensation in Arbitrary Format Multi-Bunch Acceleration With Standing Wave and Traveling Wave Accelerators", presented at the IPAC'18, Vancouver, Canada, May 2018, paper MOPMF076.

[13] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for Hyper-Parameter Optimization", in Advances in Neural Information Processing Systems 24, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2011, pp. 2546–2554.

[14] Kuriki M *et al.*, *Journal of Physics: Conference Series* **2687** (2024) 022012. doi:10.1088/1742-6596/2687/2/022012