Photoinjector optimization using a derivative-free, model-based trust-region algorithm for the Argonne Wakefield Accelerator

N Neveu,[1,2] J Larson,[1] J G Power[1] and L Spentzouris[2]

[1] Argonne National Laboratory, Lemont, IL
[2] Illinois Institute of Technology, Chicago, IL

E-mail: `nneveu@hawk.iit.edu, jmlarson@anl.gov`

Abstract. Model-based, derivative-free, trust-region algorithms are increasingly popular for optimizing computationally expensive numerical simulations. A strength of such methods is their efficient use of function evaluations. In this paper, we use one such algorithm to optimize the beam dynamics in two cases of interest at the Argonne Wakefield Accelerator (AWA) facility. First, we minimize the emittance of a 1 nC electron bunch produced by the AWA rf photocathode gun by adjusting three parameters: rf gun phase, solenoid strength, and laser radius. The algorithm converges to a set of parameters that yield an emittance of 1.08 μm. Second, we expand the number of optimization parameters to model the complete AWA rf photoinjector (the gun and six accelerating cavities) at 40 nC. The optimization algorithm is used in a Pareto study that compares the trade-off between emittance and bunch length for the AWA 70 MeV photoinjector.

1. The AWA Facility

The 70 MeV rf photoinjector at the Argonne Wakefield Accelerator (AWA) facility [1] consists of an rf gun followed by six rf accelerating cavities, hereafter referred to as the linac. See Figure 1 for the beam line layout. The 1.5 cell rf gun operates at 1.3 GHz with three solenoids and a Cs_2Te photocathode excited by a 248 nm UV laser. Solenoid 1 (S_1) is used to buck the field at the cathode, while the other two solenoids (S_2 and S_3) are used for emittance compensation. The accelerating cavities, also operated at 1.3 GHz, are 7 cell standing-wave cavities [2] each with independently controllable phase. The cavities are labeled L_1-L_6 in Figure 1.

2. Optimization Algorithm

Model-based, derivative-free algorithms are frequently used to optimize computationally expensive simulations due to their judicious use of function evaluations. In cases specific to accelerator physics, beam properties at different operational parameters are observed; these methods then build models of the unknown function and minimize these models to identify candidate parameters to evaluate. BOBYQA [3] is one such method that is available via the NLopt [4] package and was used in this study. Given a candidate set of optimal parameters v^k, BOBYQA constructs a quadratic model using function values of points near v^k. This model is minimized in a neighborhood of v^k in order to produce a point $\hat{v}$. If $\hat{v}$ has a smaller objective function value than v^k, the estimate of the optimum is updated to $\hat{v}$, and a new model is
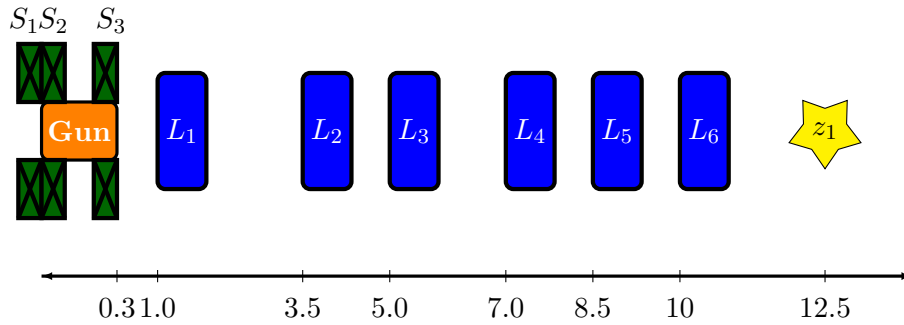
Figure 1. Layout of the AWA linac. The gun is enlarged to show solenoid detail. The physical length is $0.3\,\text{m}$. The cathode is located at $z = 0\,\text{m}$. Linac cavities are $0.85\,\text{m}$ long. Tick marks are located at the exit of the gun, entrance of each accelerating cavity, and location of optimization.

constructed. If $\hat{v}$ is not a sufficient improvement over v^k, the model around v^k is improved. For more information about derivative-free optimization, see [5].

The parameters v^k are generated and supplied to the open source particle-in-cell code OPAL-t [6]. This code was chosen in part because it models the 3D space charge necessary to accurately simulate the linac. The optimization package NLopt and OPAL-t were used in combination with Python code written at Argonne National Laboratory to perform simulation evaluations and optimization. All the files needed to replicate the results in this paper are available at

www.mcs.anl.gov/∼jlarson/AWA

Interested parties are welcome to adapt the code to their needs and suggest improvements.

3. Optimization Parameters

When optimizing the gun, three parameters were varied: solenoid strength (S_3), gun phase (ϕ_g), and laser radius (R) of a uniform pulse. The minimized objective was emittance (ϵ_x). (The phase is defined as $0°$ at maximum energy gain.) When optimizing the entire linac, seven additional parameters were varied: the longitudinal laser full width at half maximum (T) and accelerating cavity phases (ϕ_L). The optimization parameters and bounds are given in table 1; we denote the set of ten optimization parameters as $v = [S_3, \phi_g, R, T, \phi_L]$, where $\phi_L = [\phi_{L_1}, \ldots, \phi_{L_6}]$ represents the phase of each linac cavity L_1-L_6.

Table 1. Parameter bounds for gun and linac optimization.

Variable	Range	Unit
Solenoid Strength	$0 \leq S_3 \leq 440$	amps
Phase of Gun	$-60 \leq \phi_g \leq 60$	degrees
Laser Radius	$0.1 \leq R \leq 30$	mm
Laser FWHM [1]	$2 \leq T \leq 10$	ps
Cavity Phase [1,2]	$-20 \leq \phi_L \leq 20$	degrees

[1] not varied during gun optimization

[2] $\phi_L = [\phi_{L_1}, \ldots, \phi_{L_6}]$

4. Gun Optimization

Much work has been done to optimize 1.5 cell rf guns at 1 nC [7]. This known solution was used as a baseline test of BOBYQA when applied to an accelerator application. An optimization of the single objective emittance (ϵ_x) was performed over a length of 5 m. All linacs were turned off and only gun parameters were varied. Non-varying parameters are listed in table 2; their values are based on work done at PITZ and AWA [7, 8]. Local optimization runs were started from five points with various distances from the optimum value. The optimization runs converged (in less than 100 function evaluations) to a parameter set ($S_3 = 269\,\text{A}$, $\phi_g = -3.0\,^\circ$, and $R = 0.6\,\text{mm}$) with an emittance of $1.08\,\mu\text{m}$. An exhaustive search of the parameter space was not done, and there may be other local minima that were not found. However, the results match expectations based on the literature.

Table 2. Non-varying parameters for gun optimization.

Parameter	Value
Charge	1 nC
Gradient	60 MV/m
Laser FWHM	20 ps
Laser Rise and Fall Time	6 ps
Kinetic Energy at Cathode	0.55 eV
S_1 and S_2	550 A

5. Linac Optimization

Next we performed a multiobjective optimization of the linac (Figure 1), by adjusting the ten parameters in table 1. The charge was set to 40 nC and was chosen for upcoming two-beam acceleration experiments [9]. Two objectives were considered: emittance, and bunch length, σ_z. The location of interest is $z_1 = 12.51\,\text{m}$, as this is the entrance of the first quadrupole magnet after the linac. We optimize ϵ_x instead of ϵ_{xy} because no asymmetric focusing elements were used in the linac. The non-varying parameters for all linac simulation runs are shown in table 3. The model used simulated emission from a Cs_2Te cathode using a laser with initial kinetic energy of 4 eV. These are typical operating conditions at AWA.

Table 3. Non-varying parameters for linac optimization.

Parameter	Value
Charge	40 nC
Laser Rise and Fall Time	1.0 ps
Gun Gradient	70 MV/m
S_1 and S_2	550 A
Cavity Gradient L_1–L_4	25 MV/m
Cavity Gradient L_5–L_6	27 MV/m

A 1,000 point sample of linac parameters were drawn from the domain in table 1 and simulated. Of these, 132 simulations completed without error, and the emittance and bunch

length at $z_1 = 12.51\,\mathrm{m}$ was recorded for each of these points. From the sample, the minimum and maximum values of emittance and bunch length were found (i.e: $\epsilon_{\min}$ and $\epsilon_{\max}$). The raw $\epsilon_x(v, z_1)$ and $\sigma_z(v, z_1)$ sample values were then shifted and scaled to produce $\bar{\epsilon}_x(v, z_1)$, and $\bar{\sigma}_z(v, z_1)$, which have a minimum value of 0 and a maximum value of 1 over the 132-point sample set. That is,

$$\bar{\epsilon}_x(v, z_1) = \frac{\epsilon_x(v, z_1) - \epsilon_{\min}}{\epsilon_{\max} - \epsilon_{\min}} \tag{1}$$

and $\bar{\sigma}_z(v, z_1)$ is defined similarly. This scaling is done in order to remove the difference in the units between emittance and bunch length when optimizing.

With the scaled values of $\bar{\epsilon}_x$ and $\bar{\sigma}_z$, a sequence of eleven optimization problems were solved by minimizing

$$f(v, w) = w\,\bar{\epsilon}_x(v, z_1) + (1 - w)\,\bar{\sigma}_z(v, z_1) \tag{2}$$

for $w \in \{0, 0.1, \ldots, 1\}$. For each weight w, BOBYQA was started from the sample point with the smallest value of $f(v, w)$. From the initial random sample, six unique starting points were chosen. (There were fewer starting points than weights because some samples had the smallest objective value for multiple weights. For example, the smallest values of $f(v, 0.4)$, $f(v, 0.5)$, and $f(v, 0.6)$ occurred when at the eighteenth sample point.) Some $f(v, w)$ values also had one or more linac phases near the initial $\pm 20°$ boundary. The ϕ_L boundary was expanded to $\pm 40°$ for those BOBYQA runs.

6. Pareto Front for AWA Linac
Since multiple objectives are under consideration in this case, a trade-off analysis is necessary. This can be aided by examining a Pareto front: the set of parameters for which no other point
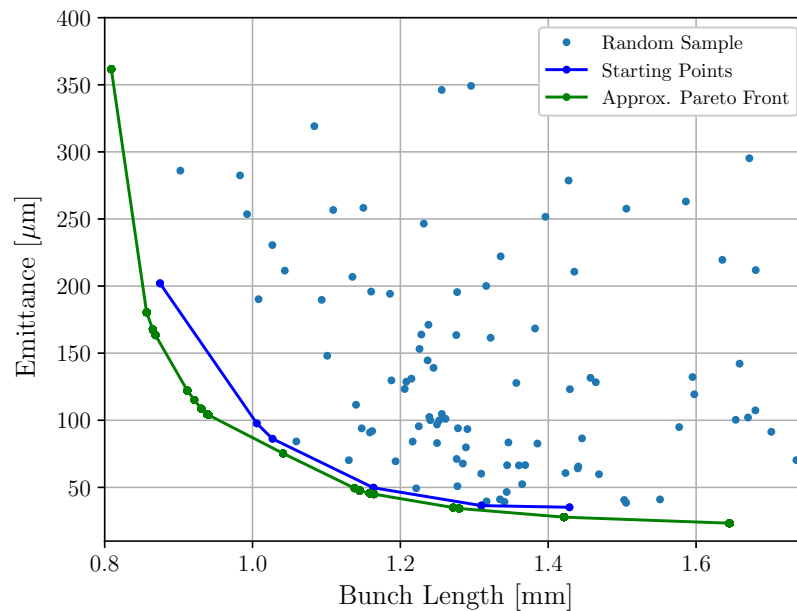

Figure 2. Random sample results, starting sample points, and resulting approximate Pareto front for the linac at $40\,\mathrm{nC}$. The Pareto front is the result of all BOBYQA evaluations.

exists that is better with respect to both objectives [10]. In Figure 2, blue dots show the

emittance and bunch length for the evaluated random sample. The sample points for which no other point has better emittance and bunch length are connected with a blue line. BOBYQA was started from these points, as described above, producing the green approximate Pareto front.

The number of simulation evaluations needed to obtain convergence in the BOBYQA runs varied from a minimum of 107 evaluations to a maximum of 208 evaluations. In order to generate the Pareto front in Figure 2, a total of 2,492 simulation evaluations were completed. Most simulation evaluations took approximately 7 minutes, using 16 cores and 100,000 particles. Runs with phases near the domain boundary (table 1) took approximately 10 minutes. These numbers are driven by the amount of time OPAL-t needs to simulate the beam line.

The best-found objective value through each BOBYQA run is shown in Figure 3. Weight 0.0 and 1.0 are negative due to the scaling in Equation (1). While the weights, w, are always nonnegative, $\bar{\epsilon}_x(v, z_1)$ or $\bar{\sigma}_z(v, z_1)$ will be negative if BOBYQA finds a point with a value of $\epsilon_x(v, z_1)$ or $\sigma_z(v, z_1)$ that is less than $\epsilon_{\min}$ or $\sigma_{\min}$ from the initial sample.
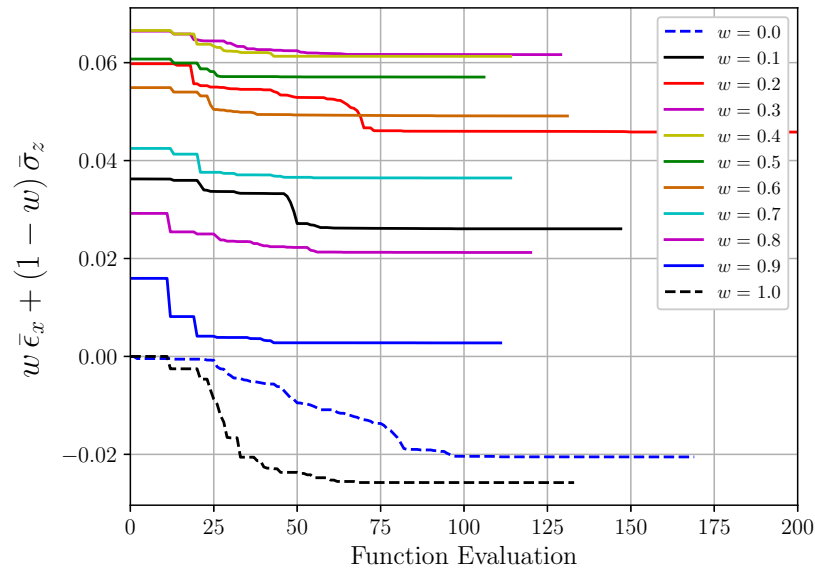

Figure 3. Minimum observed objective function values during eleven BOBYQA runs at 40 nC.

We note that seven of the BOBYQA runs converged to emittance values between $20 \, \mu$m and $50 \, \mu$m, which is shown in Figure 4 along with gun phase and bunch length for each of the 11 optimized points. We annotate Figure 4 with T because that parameter shows strong correlation with the gun phase. Other optimized parameters such as the laser radius were found to stay within a narrow range (10 mm–16 mm).

The optimized linac phases maximize energy gain while minimizing the energy spread. The energy spread of the beam exiting the gun depends strongly on ϕ_g. There are three distinct regions where the optimized points had similar gun phases (less than $\pm 10°$) which resulted in nearly identical linac phases. For example: all six phases, ϕ_L, varied by less than $10°$ for $w \in \{0, 0.1\}$, less than $5°$ for $w \in \{0.4, 0.5, 0.6\}$, and less than $10°$ for $w \in \{0.7, 0.8, 0.9\}$. This indicates optimized linac phases may benefit a range of gun settings during operation. AWA will conduct experimental measurements in order to verify these results.
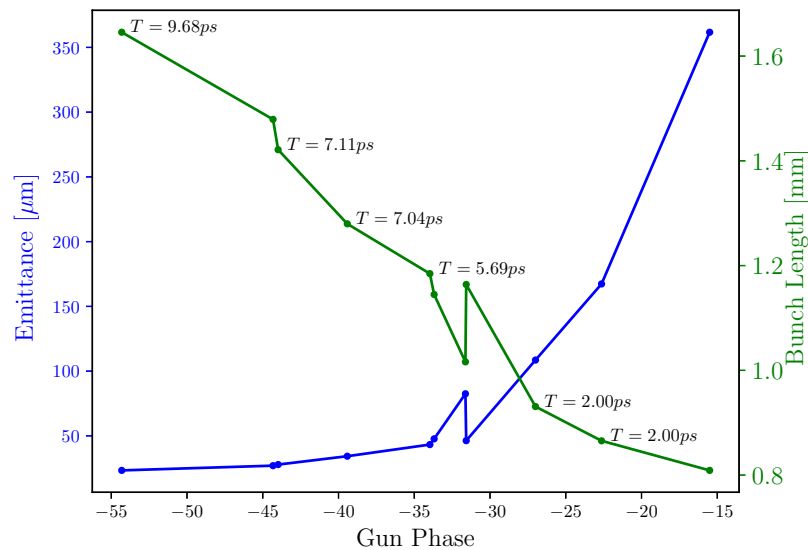
Figure 4. Bunch length and emittance vs. gun phase at each optimized point in the Pareto front for the linac at 40 nC. The phase of the maximum energy gain is 0°.

7. Conclusion

Using an AWA beam line as the simulation model, we used the BOBYQA algorithm to optimize the emittance produced by the gun at 1 nC. Using the same algorithm, we performed a multiobjective analysis for the linac at 40 nC. A Pareto front comparing the trade-off between bunch length and emittance was generated for the linac. This analysis will be used to decide future operating parameters at the AWA during high-charge experiments. In total, only 2,492 simulation evaluations were needed to produce the approximate Pareto front. Future work will include a refinement of results using 3D field maps for all cavities, experimental measurements to verify the Pareto front, and a comparison with evolutionary algorithms.

Acknowledgments

References

[1] Conde M *et al.* 2017 Research program and recent results at the Argonne Wakefield Accelerator facility (AWA) *Proc. of IPAC'17* (Copenhagen, Denmark) p WEPAB132
[2] Power J, Conde M, Gai W, Li Z and Mihalcea D 2010 Upgrade of the drive LINAC for the AWA facility dielectric two-beam accelerator *Proc. of IPAC'10* (Kyoto, Japan) p THPD016
[3] Powell M 2009 *The BOBYQA algorithm for bound constrained optimization without derivatives* University of Cambridge report NA2009/06 University of Cambridge
[4] Johnson S G 2017 *The NLopt nonlinear-optimization package* `http://ab-initio.mit.edu/nlopt`
[5] Conn A R, Scheinberg K and Vicente L N 2009 *Introduction to Derivative-Free Optimization* MPS/SIAM Series on Optimization (SIAM)

[6] Adelmann A *et al.* 2008-2017 *The OPAL (Object Oriented Parallel Accelerator Library) Framework* Tech.
      Rep. PSI-PR-08-02 Paul Scherrer Institut URL `https://amas.psi.ch/OPAL`
[7] Stephan F *et al.* 2010 *Phys. Rev. ST Accel. Beams* **13** 020704
[8] Neveu N, Spentzouris L, Power J G, Piot P, Metzger-Kraus C, Russell S J, Adelmann A and Ha G 2016
      Benchmark of rf photoinjector and dipole using ASTRA, GPT, and OPAL *Proc. of NAPAC'16* (Chicago,
      IL)
[9] Jing C *et al.* 2015 The two beam acceleration staging experiment at Argonne Wake-
      field Accelerator facility *Proc. of IPAC'15* (Richmond, VA, USA) p WEPJE020 URL
      `http://accelconf.web.cern.ch/AccelConf/IPAC2015/papers/wepje020.pdf`
[10] Ehrgott M 2005 *Multicriteria Optimization* (Springer Science & Business Media)