START-TO-END BEAM DYNAMICS OPTIMIZATION OF X-RAY FEL LIGHT SOURCE ACCELERATORS*

J. Qiang#, LBNL, Berkeley, CA, 94720, USA

Abstract

State-of-the-art tools have been developed that allow start-to-end modeling of the beam formation at the cathode, to its transport, acceleration, and delivery to the undulator. Algorithms are based on first principles, enabling the capture of detailed physics such as shot-noise driven micro-bunching instabilities. The most recent generation of the IMPACT code, using multi-level parallelization on massively parallel supercomputers, now enables multi-objective parametric optimization. This is facilitated by recent advances such as the unified differential evolution algorithm. The most recent developments will be described, together with applications to the modeling of LCLS-II.

INTRODUCTION

High brightness, coherent x-ray free electron laser (FEL) light sources provide an invaluable tool for scientific discovery in biology, chemistry, physics, and material science. Most of these FEL light sources use an accelerator beam delivery system to generate high quality electron beam needed for coherent x-ray radiation in an undulator. In order to achieve the desired electron beam quality within a reasonable cost, the accelerator system needs to be carefully designed. Typically, the accelerator system consists of a photo-injector as a front end to produce a high brightness electron beam, a linear accelerator (or equivalent accelerator) to accelerate the electron beam to the designed energy and to compress the beam to high peak current, and a final beam transport system to deliver the beam for different radiation undulator stations. As the quality of the electron beam such as emittance, peak current, energy spread plays a critical role in the production of the coherent x-ray radiation, it is important to optimize the electron beam quality during the accelerator design. In past studies, the accelerator design was typically divided into two sections, the injector section and the linear accelerator (linac) section. The injector was designed using the theory of space-charge emittance compensation and the multi-objective beam dynamics optimization [1-8]. After the injector optimization, an optimal solution from the injector output was selected as an input to the downstream linear accelerator. Using the electron beam information from the injector, the linear accelerator was then designed using analytical model, single pass tracking, and multi-objective optimization [9-12]. However, it turns out that final beam quality does not only depend on the linear accelerator settings, but also depend on the initial electron beam distribution. An optimal solution from the injector does not necessarily mean the best solution for the final beam quality. For example, the final electron beam longitudinal phase space distribution does not depend only on the peak current of the electron beam out of the injector, but also on its longitudinal phase space distribution. A highly nonlinear phase space distribution cannot be easily compressed in the linac to a high peak current at the end of the accelerator. In this paper, we report on start-to-end beam dynamics optimization using a multi-objective global optimization method and show an application to a future x-ray FEL light source LCLS-II accelerator design.

*Work supported by the Director of the Office of Science of the US Department of Energy under Contract no. DEAC02-05CH11231.

#jqiang@lbl.gov

PARALLEL MULTI-OBJECTIVE OPTIMIZATION METHOD

In accelerator community, multi-objective genetic algorithm (MOGA) such as NSGA-II has been widely used for beam dynamics optimization [13]. In this study, we developed a new multi-objective optimization algorithm based on a differential evolution method. The differential evolution method is a simple yet efficient population-based, stochastic, evolutionary algorithm for global parameter optimization [14]. In a number of studies, the differential evolution algorithm performed effectively in comparison to several stochastic optimization methods such as simulated annealing, controlled random search, evolutionary programming, the particle swarm method, and genetic algorithm [14-16].

The differential evolution algorithm starts with a population initialization. A group of NP solutions in the control parameter space is randomly generated to form the initial population. This initial population can be generated by sampling from a uniform distribution within the parameter space if no prior information about the optimal solution is available, or by sampling from a known distribution (e.g., Gaussian) if some prior information is available. After initialization, the differential evolution algorithm updates the population from one generation to the next generation until reaching a convergence condition or until the maximum number of function evaluations is reached. At each generation, the update step consists of three operations: mutation, crossover, and selection. The mutation and the crossover operations produce new candidates for the next generation population and the selection operation is used to select among these candidates the appropriate solutions to be included in the next generation. During the mutation operation stage, for each population member (target vector) x_i, $i = 1, 2, 3, \bullet\bullet\bullet, NP$ at generation G, a new mutant vector v_i is generated by following a mutation strategy. In past studies, multiple mutation strategies (>10) were proposed in the literature.

The use of multiple mutation strategies makes the differential evolution algorithm complicated to implement and use appropriately. In recent study, we developed a new adaptive unified differential evolution (AuDE) algorithm for global optimization [17]. The unified mutation strategy can be written as:

$$\vec{v}_i = \vec{x}_i + F_1(\vec{x}_b - \vec{x}_i) + F_2(\vec{x}_{r_1} - \vec{x}_i) + F_3(\vec{x}_{r_2} - \vec{x}_{r_3}) + F_4(\vec{x}_{r_4} - \vec{x}_{r_5})$$

This algorithm uses only a single mutation expression, but encompasses almost all commonly-used mutation strategies as special cases. It is mathematically simpler than the conventional algorithm with its multiple mutation strategies, and also provides users the flexibility to explore new combinations of conventional mutation strategies during optimization. This single unified mutation strategy is further tuned to improve the performance by using only three control parameters instead of four control parameters in the original unified algorithm [18]. Figure 1 shows the error of an objective function evolution with the newly tuned adaptive unified differential evolution algorithm (AuDE3) together with several conventionally used differential evolution algorithms using a noisy shifted Schwefel's problem 1.2 as a test function. The tuned unified differential evolution algorithm out performs the other three algorithms in this example.

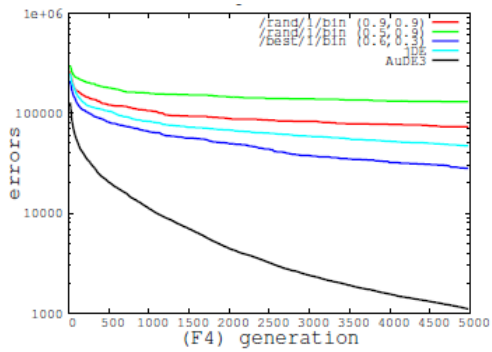

Fig. 1: The error of the objective as a function of number generation from the unified differential evolution algorithm (AuDE3), the conventional /rand/1/bin, /best/1/bin and jDE algorithms.

In many accelerator applications, one needs to optimize more than one objective function. The goal of multi-objective optimization is to find the Pareto front in the objective function solution space. The Pareto optimal front is a collection of all non-dominated solutions in the whole feasible solution space. Any other solution in the feasible solution space will be dominated by those solutions on the Pareto optimal front. In the multi-objective optimization, a solution A is said to dominate a solution B if all components of A are at least as good as those of B (with at least one component strictly better). The solution A is non-dominated if it is not dominated by any solution within the group. In this study, we have developed a new parallel multi-objective differential evolution algorithm with variable population size and external storage. The algorithm in each generation and external storage can be summarized in the following steps:

- Step 0: Define the minimum parent size, NPmin and the maximum size, NPmax of the parent population. Define the maximum size of the external storage, NPext.
- Step 1: An initial population of NPini parameter vectors is chosen randomly to cover the entire solution space.
- Step 2: Generate the offspring population using the adaptive unified differential evolutionary algorithm.
- Step 3: Check the new population against boundary conditions and constraints.
- Step 4: Combine the new population with the existing parent population from the external storage. Non-dominated solutions (Ndom) are found from this group of solutions and min(Ndom, NPext) of solutions are put back into external storage. Pruning is used if Ndom > NPext. NP parent solutions are selected from this group of solutions for next generation production. If NPmin ≤ Ndom ≤ NPmax, NP = Ndom. Otherwise, NP = NPmin if Ndom < NPmin and NP = NPmax if Ndom > NPmax. The elitism is emphasized through keeping the non-dominated solutions while the diversity is maintained by penalizing the over-crowded solutions through pruning.
- Step 5: If the stopping condition is met, stop. Otherwise, return to Step 2.

As a test of above algorithm, we used the following two objective functions, which are to be minimized:

$$f_1(\mathbf{x}) = x_1$$
$$f_2(\mathbf{x}) = g(\mathbf{x})\left[1 - (x_1/g(\mathbf{x}))^2\right]$$
$$g(\mathbf{x}) = 1 + 9\left(\sum_{i=2}^{n} x_i\right)/(n-1)$$

The optimal Pareto front for these two objectives is:

$$f_2 = 1 - {f_1}^2$$
$$x_1 \in [0,1]$$
$$x_i = 0,$$
$$i = 2, \ldots, n$$

The final optimal Pareto front from both the numerical solution and the analytical solution is shown in Fig. 2. Figure 3 also shows the distance to the Pareto front as a function of the number of objective function evaluations from the above algorithm and the widely used genetic algorithm NSGA-II. It is seen that new proposed algorithm (called "variation population with external storage differential evolution algorithm" or VPES) can be significantly faster than the widely used genetic algorithm.

One advantage of the evolutionary based optimization algorithm is that it is very easy to be parallelized on multi-processor high performance computer. Here, we have used a multi-group method in which the population is uniformly distributed among the total number of groups. Each single group of processors is used for one objective function evaluation using the parallel beam dynamics simulation. This two-level parallelization method makes the start-to-end simulation scalable on a supercomputer with a large number of processors.

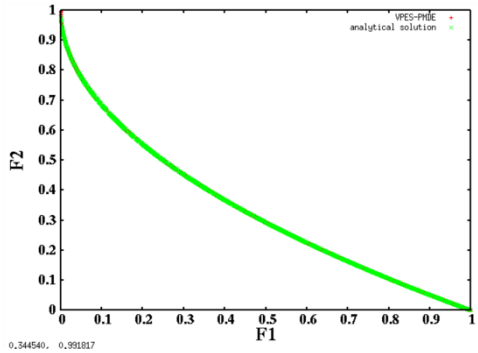

Fig. 2: The Pareto optimal front from the VPES and analytical solution.

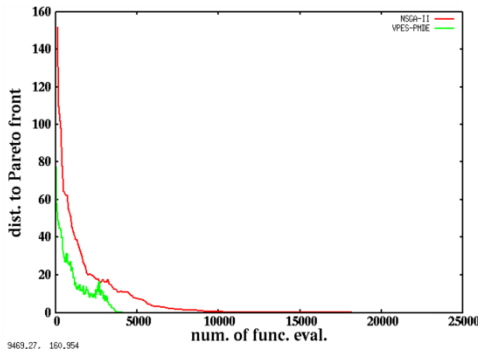

Fig. 3: Distance to the analytical Pareto front from the VPES and the NGSA-II algorithm.

START-TO-END BEAM DYNAMICS OPTIMIZATION

The start-to-end beam dynamics simulations presented in this study were done using a 3D parallel beam dynamics simulation framework IMPACT [19-21]. It includes a time-dependent 3D space-charge code module IMPACT-T for injector modeling and a position-dependent 3D space-charge code module for linac and beam transport system model. The simulation starts from the generation of photo-electrons at the photo-cathode following the initial laser pulse distribution and the given initial thermal emittance. The electron macroparticles out of the cathode will be subject to both the external fields from a DC/RF gun and solenoid, and the space-charge/image charge fields from the Coulomb interaction of the particles among themselves. After exiting from the injector, the electron macroparticle will transport through a linear accelerator and beam transport system that includes laser heater, bunch compressors, accelerating RF cavities, harmonic linearizer, and magnetic focusing elements. Besides the 3D space-charge effects, the simulation also includes coherent synchrotron radiation (CSR) effects through a bending magnet, incoherent synchrotron radiation inside the bending magnet, RF cavity structure wakefield, and resistive wall wakefield.

The start-to-end beam dynamics simulation is integrated with the parallel multi-objective optimization program described above. Figure 4 shows a schematic plot of the global optimization including both the injector control parameters and the linac control parameters in the start-to-end beam dynamics optimization. Here, the start-to-end simulation is treated as an objective function in the parallel multi-objective optimizer. The parallel optimizer will call the IMPACT simulation by passing the injector control parameters and the linac control parameters into the objective function. The injector control parameters normally include laser pulse transverse size and length, RF gun amplitude and phase, solenoid strength, buncher and boosting cavity amplitudes and phases. The linac control parameters include linac section 1 cavity amplitude and phase, harmonic linearizer amplitude and phase, bunch compressor 1 bending angle, linac section 2 cavity amplitude and phase, bunch compressor 2 bending angle, and so on. There can be more than 20 total number of control parameters for the global optimization. Instead of starting with direct global optimization in the entire control parameter space, we start the optimization with reduced control parameter space that contains only the injector control parameters. The two objective functions, final project transverse emittance and rms bunch length (directly related to peak current) at the exit of the injector are optimized subject to a number of constraints. These constraints are final electron beam energy, beam energy chirp, longitudinal phase space nonlinearity, and so on. After a Pareto optimal front is found for these two objective functions at the exit of the injector, these optimal injector control parameters are combined with some randomly sampled control parameter solutions in the linac. Using the optimal injector control parameters as a partial initial component in the global control parameter solution significantly saves the computational time and speeds up the convergence of the final global solution. During the global beam dynamics optimization, one of the objective (transverse emittance) from the original injector optimization becomes a constraint to the new objective functions. Those solutions at the exit of the injector that can not satisfy this constraint for final start-to-end optimization will be automatically excluded at the beginning of the global optimization. For the global beam dynamics optimization, in the linac section, we decouple the electron beam dynamics in the transverse direction from that in the longitudinal direction. This is because at this energy, the longitudinal bunch length is mostly frozen, the final longitudinal phase space is primarily determined by the linac RF settings, bunch compressor settings, and longitudinal collective effects such as longitudinal space-charge effects, longitudinal wakefields, and coherent synchrotron radiation. The transverse beam dynamics is primarily determined by the lattice matching and transverse space charge effect, and CSR. In the transverse direction, we would like to minimize the emittance growth in the linac. In the longitudinal direction, we would like to attain a higher peak current with flatter longitudinal phase space. Two objective functions are defined for the global longitudinal beam dynamics optimization. These two functions are fraction of charge and rms energy spread inside a given

longitudinal window. The output from the injector such as energy, emittance, and energy spread are used as constraints for the global optimization. Besides the constraint at the exit of the injector, we also put constraints at the final linac output such as energy, peak current etc. After the global longitudinal beam dynamics optimization is done, the transverse emittance growth in the linac is minimized through transverse beam dynamics optimization by retuning a number of quadrupole settings along the linac including the transverse space-charge and CSR effects.

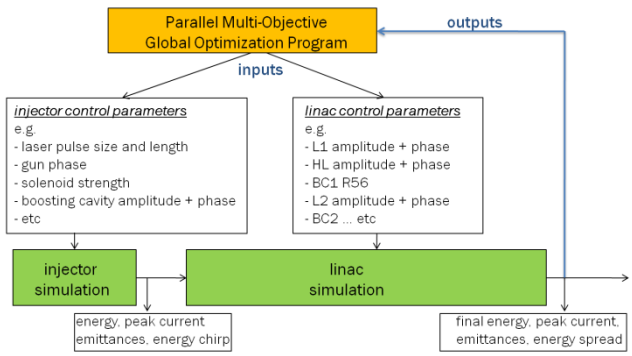

Fig. 4: A schematic diagram of the global beam dynamics optimization.

APPLICATION TO AN X-RAY FEL ACCELEATOR DESIGN OPTIMIZATION*

As an application, we applied the above global multi-objective beam dynamics optimization tool to an LCLS-II design optimization with a 20 pC charge. The LCLS-II is a high repetition rate (1 MHz) x-ray FEL that will deliver photons of energy between 200 eV and 5 keV [22-23]. It consists of a high repetition rate photo-injector to generate and accelerate the electron beam to about 100 MeV, a laser heater (LH) to suppress microbunching instability, a section of superconducting linac L1 to accelerate the beam to 250 MeV, a bunch compressor BC1, a second section of superconducting linac L2 to accelerate the beam to 1.6 GeV, a bunch compressor BC2, and a third section of superconducting linac L3 to accelerate the beam to 4 GeV, a long bypass transport line, and a magnetic kicker to spread the electron beam to a soft x-ray transport beam line and to a hard x-ray transport beam line. The superconduting linacs in all three sections are made of 1.3 GHz 9 cell superconducting cavities except the two cryomodules of 3.9 GHz third harmonic cavities right before the BC1 to linearize longitudinal phase space.

For the global longitudinal beam dynamics optimization of this accelerator, we have defined 22 control parameters: 12 in the injector, 10 in the linac. The 12 control parameters in the injector are laser transverse size, laser pulse flat-top length, VHF gun RF phase, buncher cavity amplitude and phase, two solenoid strengths, the 1st boosting cavity amplitude and phase, and the 2nd boosting cavity amplitude and phase and the last cavity phase. The 10 control parameters in the linac are the linac section one amplitude and phase, 3rd harmonic cavity amplitude and phase, bending angle in bunch compressor one, linac section two amplitude and phase, bending angle in bunch compressor two, and linac amplitude and phase.

We first did two objective optimization of the injector design with the 12 control parameters. These two objectives are final rms emittance and rms bunch length. Figure 5 shows the Parato front of these two objectives from the injector optimization. Here, we have set a final peak current to be lower than 20 A, final rms emittance to be less than 1 um, final electron beam energy to be greater than 85 MeV, final rms energy spread to be less than 100 keV. It is seen that the rms emittance approaches to 0.1 um with the rms bunch length close to 1 mm. The simulation was done using 10,000 macroparticles with 32x32x64 grid points.

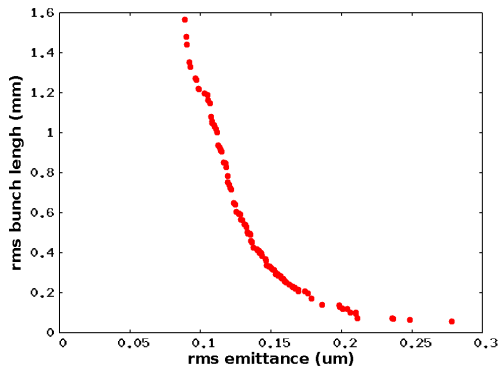

Fig. 5: The Pareto front of the injector beam dynamics optimization.

In most previous studies, a solution is selected from this Pareto front as the solution of the injector. Using this solution from the injector as an initial distribution, the linac optimization is carried out to find the best solution at the end of the accelerator. In this study, instead of using a single solution from the above Pareto front, we combined the 12 control parameters of these solutions together with the 10 linac control parameters that randomly sampled from the allowed solution space as the first generation of the global start-to-end multi-objective longitudinal beam dynamics optimization. Figure 6 shows the Pareto front of the two objective functions from the global optimization. These two objective functions are the negative fraction of charge inside and the rms energy spread inside a window between -7 and 9 um. In this plot, we also show the Pareto front from only the linac optimization using the solution from the injector as an initial distribution. It is seen that the Pareto front from the global is significantly better than that from the linac only optimization. For the same amount of charge inside the window, the global solution has 40% less energy spread in some region. For the same level of the final rms energy spread, the global solution has 15% larger amount of charge. In this simulation,

*We would like to thank the LCLS-II physics design team for the LCLS-II application study.

besides those constraints for the beam at the exit of the injector, we also put constraints on the final beam energy to be greater than 3.9 GeV, final rms energy spread to be less than 2.5 MeV, fraction of charge inside the window between 0.3 and 0.9.

CONCLUSIONS

In this paper, we reported on a global start-to-end beam dynamics multi-objective optimization method for x-ray FEL accelerator design. Using the recently developed parallel variable population with external storage multi-objective differential evolution algorithm and together with the start-to-end parallel beam dynamics simulation, we have shown in an LCLS-II design example that the global optimization including control parameters for both the injector and the linac can result in a better optimal solution than the linac only optimization using an optimal solution from the injector.

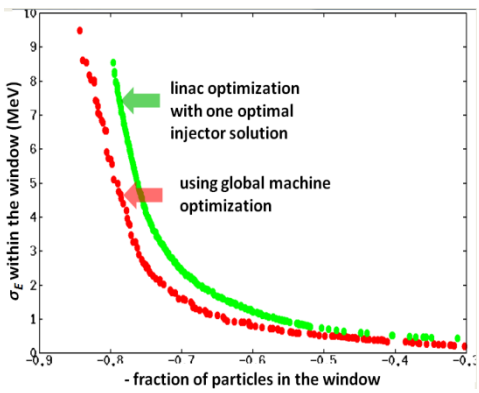

Fig. 6: The Pareto front from the global beam dynamics optimization and from the linac only optimization using one optimal injector solution.

REFERENCES

[1] B. E. Carlsten, *Nucl. Instrum. Methods Phys. Res.*, Sect. A285, 313, 1989.
[2] L. Serafini and J. Rosenzweig, *Phys. Rev. E* 55, 7565, 1997.
[3] I. V. Bazarov and C. K. Sinclair, *Phys.Rev.ST Accel. Beams* 8, 034202, 2005.
[4] C. F. Papadopoulos *et al.*, in *Proc. of FEL2010*, Malmo, Sweden, p. 479.
[5] J. Qiang *et al.*, in *Proc. IPAC2013*, Shanghai. China, p. 1031, 2013.
[6] J. Qiang and C. Mitchell, in *Proc. FEL2014*, THP020, 2014.
[7] C. Mitchell *et al.*, in *Proc. IPAC2016*, Busan, Korea, 2016, p. 1699.
[8] H. Qian *et al.*, in *Proc. IPAC2016*, Busan, Korea, 2016, p. 3979.
[9] J. Arthur *et al.*, LCLS Conceptual Design Report, SLAC-R-593, 2002.
[10] I. Zagorodnov and M. Dohlus, *Phys. Rev. ST – Accel. Beams* 14, 014403, 2011.
[11] S. DiMitri *et al.,* Physics Reports 539, p.1, 2014.
[12] L. Wang *et al.*, in *Proc. FEL2014*, Basel, Switzerland, 2014, p. 763.
[13] K. Deb *et al.*, IEEE Trans. Evol. Comp, Vol. 6, p. 182, (2002).
[14] R. Storn and K. Price, *Journal of Global Optimizational,* 341-359, 1997.
[15] K. Price *et al.*, Differential Evolution - A Practical Approach to Global Optimization, Springer, Berlin, 2005.
[16] M. M. Ali and A. Torn, Computers and Operations Research, Elsevier, no. 31, p. 1703, 2004
[17] J. Qiang and C. Mitchell, Lawrence Berkeley National Laboratory Report LBNL-6853e, 2014.
[18] J. Qiang *et al.*, to appear in *Proc. CEC2016*, Vancouver, 2016.
[19] J. Qiang *et al.*, *PRST-AB* 9, 044204, 2006.
[20] J. Qiang *et al.*, *J. of Comp. Phys.*, 163, 434 2000.
[21] J. Qiang et al., PRST-AB 12, 100702 (2009).
[22] T. O. Raubenheimer, "Technical challenges of the LCLS-II CW X-ray FEL," in *Proc. IPAC2015*, Richmond, VA, USA.
[23] P. Emma *et al.*, in *Proc. FEL2014*, Basel, Switzerland, THP025. 2014.