

Fast longitudinal beam dynamics optimization in x-ray free electron laser linear accelerators

Ji Qiang*

Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA



(Received 3 July 2019; published 27 September 2019)

A high peak current, flat longitudinal phase space electron beam is desirable for efficient x-ray free electron laser (FEL) radiation in next generation light sources. To attain such a beam requires the extensive design of the linear accelerator (linac) including both linear and nonlinear effects. In this paper, we propose a lumped longitudinal beam dynamics model for fast optimization of the electron beam longitudinal phase space through the accelerator. This model is much faster than available tracking programs and also shows good agreement with the fully three-dimensional element-by-element multiparticle simulations. We applied this model in a parallel multiobjective differential evolution optimization program to an existing LCLS-II superconducting linac design and obtained an optimal solution with significantly higher core peak current than the original design.

DOI: [10.1103/PhysRevAccelBeams.22.094401](https://doi.org/10.1103/PhysRevAccelBeams.22.094401)

I. INTRODUCTION

High brightness, coherent x-ray radiation from a free electron laser (FEL) light source provides an important tool for scientific discovery in physics, chemistry, biology and other fields. To produce such a coherent radiation effectively requires the use of high brightness electron beam with high peak current, small energy spread, and small emittance inside the radiation undulator. The high brightness electron beam used in an x-ray FEL light source typically comes from a linear accelerator beam delivery system [1–6]. This system consists of a photoinjector to generate an initial high brightness electron beam, a rf linac to accelerate the beam to multiple GeV energy and to compress the beam to hundred or thousand Ampere peak current, and transport beam line to deliver the beam to different undulators.

To attain a high brightness electron beam at the entrance of undulator involves extensive design and optimization of the linear accelerator parameters. In previous studies, an analytical model was developed to choose the settings of linac accelerating gradients, phases, and magnetic chicane parameters [7,8]. This model gives a quick estimate of the rf cavity accelerating gradient amplitude, phase and bunch compressor parameter settings inside the linac but it does not include collective effects such as space-charge effects,

coherent synchrotron radiation (CSR), and structure or resistive wall wakefields inside the accelerator. It works well when the peak current is not very high and the collective effects are weak. It can be used as a starting point of longitudinal phase space optimization. However, when one pushes the limit of final core peak current, those collective effects are no longer negligible and can have major impact on electron beam longitudinal phase space distribution. This is especially true after the final stage of beam compression, the electron bunch length becomes as short as a few tens microns. Using the phase setting in the rf cavities to remove the correlated energy spread (also called chirp) of the electron beam becomes inefficient. Meanwhile, given the high peak current after compression, the collective effects such as wakefields can become dominant. For example, in the LCLS-II design, the final correlated energy spread of the electron beam is almost completely removed by the resistive wall wakefields before entering the undulator [9,10]. Including those collective effects in the longitudinal beam dynamics model is crucial for the high peak current accelerator design.

To include those effects, one normally resorts to detailed multiparticle element-by-element tracking simulations. In the past studies, a number of tracking codes such as LiTrack [11], Elegant [12], and IMPACT [13,14] were used for element-by-element electron linac simulations. Among those three codes, LiTrack handles only longitudinal beam dynamics and has the fastest computational speed. However, LiTrack does not include the CSR effect and the longitudinal space-charge effect in simulation. The other two codes include all those collective effects, but are much slower in computational speed due to the use of fully three-dimensional element-by-element multiparticle

*jqiang@lbl.gov

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

tracking. In this paper, we propose a lumped longitudinal beam dynamics model for fast longitudinal phase space optimization. This model is based on the assumption that out of the injector, inside the rf linac, when the electron beam energy is sufficiently high (e.g., ≥ 100 MeV), the transverse beam dynamics and the longitudinal beam dynamics can be reasonably decoupled. The transverse focusing elements such as quadrupoles are treated as drifts in the longitudinal beam dynamics model. The rf cavities in each section of the linac are lumped as one rf accelerating element. The magnetic bunch compression chicane is treated as a thin lens element except that the last dipole bending magnet is used for the CSR effect calculation. The longitudinal space-charge effect, the structure wakefields from the rf cavities, and the resistive wall wakefields from the transport beam line pipe are included in the model. Using the lumped elements significantly improves the speed of simulation while including all longitudinal collective effects. In the simulation, those collective effects are computed efficiently using an FFT based method. Using a weighted macroparticle method, given the initial current profile and correlated energy profile in longitudinal phase space, a small number of macroparticles (from a few hundreds to a thousand macroparticles) that corresponds to the longitudinal slice coordinates are needed in the simulation. All these make the longitudinal beam dynamics simulation very fast.

The fast longitudinal beam dynamics model can be used for the optimization of electron beam longitudinal phase space in a x-ray FEL linac. In the accelerator community, multiobjective beam dynamics optimizations were done using a tracking code such as Astra or Elegant [15–18]. These optimizations were all based on a multiobjective genetic algorithm (MOGA), NSGA-II [19]. The genetic algorithm is a useful but relatively old evolutionary method that was first proposed in 70s [20]. In recent years, in the evolutionary computation community, a new optimization method, differential evolution method, has been actively studied [21]. The differential evolution method is a simple yet efficient population-based stochastic evolutionary algorithm for global parameter optimization [22,23]. In a number of studies, this method showed better performance than several other stochastic optimization methods such as simulated annealing, controlled random search, evolutionary programming, particle swarm method, and genetic algorithm [22–25]. Recently, we developed a multiobjective differential evolution optimization method [26,27]. This method showed faster convergence in a comparison with the NSGA-II method [28]. In this study, we implemented the longitudinal beam dynamics model in this new multiobjective optimization program for fast longitudinal phase space optimization.

The organization of this paper is as follows: after the Introduction, we present the one-dimensional (1D) longitudinal beam dynamics model in Sec. II; We discuss about a

recently developed multiobjective differential evolution optimization algorithm in Sec. III; We apply the multiobjective longitudinal beam dynamics optimization to an existing LCLS-II design in Sec. IV; and draw conclusions in Sec. V.

II. LONGITUDINAL BEAM DYNAMICS MODEL

In this study, we focus only on the longitudinal beam dynamics, and neglect the transverse motion of electrons. Each electron macroparticle has longitudinal coordinates $(z, \Delta\gamma)$ with respect to the reference particle (s_0, γ_0) and charge weight w . Here, $z = s - s_0$ is the bunch length coordinate (z_{\max} corresponds to the bunch head and z_{\min} the bunch tail), $\Delta\gamma = \frac{E-E_0}{mc^2}$, E is the total energy of the particle, E_0 is the total energy of the reference particle, m is the rest mass of the particle, and c is the speed of light in vacuum. For the longitudinal beam dynamics study, we include only drifts, RF cavities, and magnetic compression chicanes as the beam line elements of the x-ray FEL linear accelerator. The other focusing elements such as quadrupoles are treated as drifts too.

For a macroparticle transporting through the lumped rf cavity element with total length L_{acc} , its longitudinal coordinates will be updated by the following equations using a leap-frog type of approximation:

$$z^+ = z_1 + \frac{L_{\text{acc}}}{2} \Delta\gamma_1 / (\gamma_{01} \beta_{01})^3 \quad (1)$$

$$\gamma_0^+ = \gamma_{01} + \frac{L_{\text{acc}}}{2} \frac{qV_{\text{acc}}}{mc^2} \cos(\phi_0) \quad (2)$$

$$\Delta\gamma_2 = \Delta\gamma_1 + L_{\text{acc}} \frac{qV_{\text{acc}}}{mc^2} (\cos(\phi_0 - kz^+) - \cos(\phi_0)) \quad (3)$$

$$z_2 = z^+ + \frac{L_{\text{acc}}}{2} \Delta\gamma_2 / (\gamma_0^+ \beta_0^+)^3 \quad (4)$$

$$\gamma_{02} = \gamma_0^+ + \frac{L_{\text{acc}}}{2} \frac{qV_{\text{acc}}}{mc^2} \cos(\phi_0) \quad (5)$$

where subscript 1 and 2 denote the quantity before and after the lumped cavity element respectively, $V_{\text{acc}} = qV_{rf}/L_{\text{acc}}$ is the accelerating gradient amplitude, k is the rf wave number, and ϕ_0 is the rf cavity design phase.

The magnetic bunch compression chicane is modeled as a thin lens element. The particle longitudinal position through the chicane is given by [29]:

$$z = z + R_{56} \frac{\Delta\gamma}{\gamma_0} + T_{566} \left(\frac{\Delta\gamma}{\gamma_0} \right)^2 + U_{5666} \left(\frac{\Delta\gamma}{\gamma_0} \right)^3 \quad (6)$$

where

$$R_{56} \approx 2\theta^2 \left(L_{db} + \frac{2}{3}L_b \right) \quad (7)$$

$$T_{566} \approx -\frac{3}{2}R_{56} \quad (8)$$

$$U_{5666} \approx 2R_{56} \quad (9)$$

where θ is the bending angle of one of dipole magnets (assuming that all four dipoles have the same bending angle amplitude), L_b is the length of the dipole magnet, and L_{db} is the distance between the first and the second (or between the third and fourth) dipole bending magnets. From our benchmark with fully 3D element-by-element tracking using 5th order transfer map for the dipole bending magnet, we need to increase the R_{56} by 0.5% in order to match the current profile after the chicane with that from the 3D model.

Collective effects such as longitudinal space-charge effect, structure and resistive wall wakefields, and coherent synchrotron radiation play an important role in the longitudinal beam dynamics and are included in this model. For the longitudinal space-charge effect, instead of using the space-charge impedance model in the frequency domain [14], we assume that the electron beam is a round cylinder with separable uniform transverse density distribution and longitudinal density distribution. The longitudinal space-charge field on the axis is given as:

$$E_z^{sc}(0, 0, z) = \frac{1}{4\pi\epsilon_0} \frac{2}{a^2} \int \frac{\gamma_0(z-z')\rho(z')}{(\gamma_0^2(z-z')^2 + r'^2)^{3/2}} r' dz' dr'. \quad (10)$$

After integrating with respect to the transverse radial dimension, the longitudinal space-charge field on the axis can be written as:

$$E_z^{sc}(z) = \frac{1}{4\pi\epsilon_0} \frac{2}{a^2} \left(\int_{z_{\min}}^z \rho(z') dz' - \int_z^{z_{\max}} \rho(z') dz' - \int_{z_{\min}}^{z_{\max}} \frac{\gamma_0(z-z')\rho(z')}{\sqrt{\gamma_0^2(z-z')^2 + a^2}} dz' \right) \quad (11)$$

where a is the radius of the cylinder, z_{\min} and z_{\max} denote the minimum and the maximum longitudinal bunch length positions, and ρ is the electron beam longitudinal charge density distribution. The above convolution can be computed efficiently using an FFT based method [30,31].

The longitudinal wakefields from both the structure wakefields of rf cavities and the resistive wall wakefields are included in the model. The longitudinal field from the wakefields are calculated from the following convolution:

$$E_z^{wk}(z) = \int_z^{z_{\max}} W_L(z-z')\rho(z') dz' \quad (12)$$

where $W_L(s)$ is the longitudinal wake function. For example, for a TESLA type superconducting cavity, the short-range longitudinal wake function is given as [32]:

$$W_L(s) = 41.5 \times 10^{12} \exp\left(-\sqrt{\frac{s}{0.00174}}\right). \quad (13)$$

This convolution can also be computed efficiently using the FFT based method [14].

The coherent synchrotron radiation through a bending magnet can be calculated from the following integral:

$$E_z^{csr}(z) = \int_{z_{\min}}^z W_{csr}(z, z')\rho(z') dz' \quad (14)$$

where $W_{csr}(s)$ is the longitudinal CSR wake function that includes both transient and steady-state radiations through a bending magnet following Saldin *et al.* (Case A-D) [33]. The CSR wake function has a very sharp variation around the origin of bunch length coordinate. To avoid the use of large number of numerical grid points to resolve the sharp variation, instead of using the original CSR wake function and calculating the above integral directly, we divide this integral into the summation of a number of small interval integrals and rewrite the integral for the CSR wakefield at location z_k as:

$$E_z^{csr}(z_k) = \sum_{k'=1}^k \int_{z_{k'}-h/2}^{z_{k'}+h/2} W_{csr}(z_k, z')\rho(z') dz'. \quad (15)$$

If we assume that the longitudinal density $\rho(z')$ is constant within that small interval, the above integral at k slice can be approximated as [34,35]:

$$E_z^{csr}(z_k) = \sum_{k'=1}^k \rho(z_{k'}) w_{k,k'} \quad (16)$$

where

$$w_{k,k'} = \int_{z_{k'}-h/2}^{z_{k'}+h/2} W_{csr}(z_k, z') dz' \quad (17)$$

$$= I_{csr}(z_k, z'_k - h/2) - I_{csr}(z_k, z'_k + h/2) \quad (18)$$

where

$$I_{csr}(z, z_t) = - \int_{z_{\min}}^{z_t} W_{csr}(z, z') dz'. \quad (19)$$

In case A (transient at entrance) where radiation source electron is in front of bending dipole magnet while the observer electron is inside the magnet, the integrated CSR wake function is given as:



FIG. 1. A schematic of LCLS-II linac.

$$I_{csr} = \frac{\gamma r_c m c^2}{R} \left(\frac{1}{\zeta} - \frac{2(\hat{\phi} + \hat{y}) + \hat{\phi}^3}{(\hat{\phi} + \hat{y})^2 + \hat{\phi}^4/4} \right) \quad (20)$$

where $\hat{\phi} = \gamma\phi$, ϕ is the azimuthal angle of the observer electron, $\zeta = (z - z_i)\gamma^3/R$, R is the bending radius of the dipole, and \hat{y} , the normalized distance from the source to the dipole entrance, can be found from the solution of the following equation:

$$\zeta = \frac{\hat{\phi} + \hat{y}}{2} + \frac{\hat{\phi}^3 \hat{\phi} + 4\hat{y}}{24 \hat{\phi} + \hat{y}}. \quad (21)$$

In case B (steady-state) where both the source and the observer are inside the dipole, the integrated CSR wake function is given as:

$$I_{csr} = -\frac{\gamma r_c m c^2}{R} \left[\frac{4(\hat{u}^2 + 8)\hat{u}}{(\hat{u}^2 + 4)(\hat{u}^2 + 12)} \right] \quad (22)$$

where \hat{u} , the normalized angle between the observer and the source, can be found from the solution of the following equation:

$$\zeta = \frac{\hat{u}^3}{24} + \frac{\hat{u}}{2}. \quad (23)$$

In case C (transient at exit) where the source is in front of the dipole entrance while the observer is outside the dipole exit, the integrated CSR wake function is given as:

$$I_{csr} = \frac{\gamma r_c m c^2}{R} \left[\frac{1}{\zeta} - \frac{2(\hat{\phi}_m + \hat{x} + \hat{y} + \hat{\phi}_m^3/2 + \hat{\phi}_m^2 \hat{x})}{(\hat{\phi}_m + \hat{x} + \hat{y})^2 + (\hat{\phi}_m \hat{x} + \hat{\phi}_m^2/2)^2} \right] \quad (24)$$

where $\hat{\phi}_m = \gamma\phi_m$, ϕ_m is the total bending angle of the dipole, $\hat{x} = \gamma x/R$, x the distance from the dipole exit to the observer, and the normalized distance from the source to the dipole entrance \hat{y} can be found from the solution of the following equation:

$$\zeta = \frac{\hat{\phi}_m + \hat{x} + \hat{y}}{2} + \frac{\hat{\phi}_m^2 \hat{\phi}_m^2 + 4\hat{\phi}_m(\hat{x} + \hat{y}) + 12\hat{x}\hat{y}}{24 \hat{\phi}_m + \hat{x} + \hat{y}}. \quad (25)$$

In case D (transient at exit) where the source is inside the dipole while the observer is outside the dipole, the integrated CSR wake function is given as:

$$I_{csr} = \frac{\gamma r_c m c^2}{R} \left[\frac{1}{\zeta} - \frac{2(\hat{\psi} + \hat{x} + \hat{\psi}^3/2 + \hat{\psi}^2 \hat{x})}{(\hat{\psi} + \hat{x})^2 + (\hat{\psi} \hat{x} + \hat{\psi}^2/2)^2} \right] \quad (26)$$

where $\hat{\psi}$, the normalized arc length from the source to the dipole exit, can be found from the solution of the following equation:

$$\zeta = \frac{\hat{\psi} + \hat{x}}{2} + \frac{\hat{\psi}^2 \hat{\psi} + 4\hat{x} \hat{\psi}}{24 \hat{\psi} + \hat{x}}. \quad (27)$$

The transient cases C and D at the exit of the bending dipole magnet are not included in a drift element.

As a benchmark of the above model, we applied this model to the longitudinal beam dynamics simulation of a LCLS-II linac design with 100 pC charge and compared the longitudinal phase space and current profile from this one-dimensional model with those from the fully 3D multiparticle simulation using real number of electrons. 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 [9,10]. Figure 1 shows a schematic of LCLS-II superconducting rf linac. It consists of a laser heater (LH) to suppress microbunching instability, a section of superconducting linac L1 to accelerate the beam to about 270 MeV, a section of a third harmonic cavities [also called harmonic linearizer (HL)] to linearize electron beam longitudinal phase space and decelerate the beam down to about 230 MeV, a bunch compressor BC1, a second section of superconducting linac L2 to accelerate the beam to about 1.6 GeV, another bunch compressor BC2, and a third section of superconducting linac L3 to accelerate the beam to about final 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 superconducting 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.

Figure 2 shows the initial longitudinal phase space and current profile at the entrance of the LCLS-II linac. This initial beam distribution has an energy of about 100 MeV coming out of a high repetition rate photoinjector. A total 1024 macroparticles (corresponding to 1024 grid points/slices) that uniformly distributed along the bunch length coordinate are used in the 1D simulation. The energy deviation of each macroparticle can be obtained from the left longitudinal phase space plot of the Fig. 2. The charge weight of each macroparticle can be obtained from the right current plot, i.e., $w_i = I_i \delta z / c$, where I_i is the current at longitudinal bunch position i and δz is the longitudinal grid size.

Figure 3 shows the longitudinal phase space and current profile after the magnetic bunch compression chicane BC1

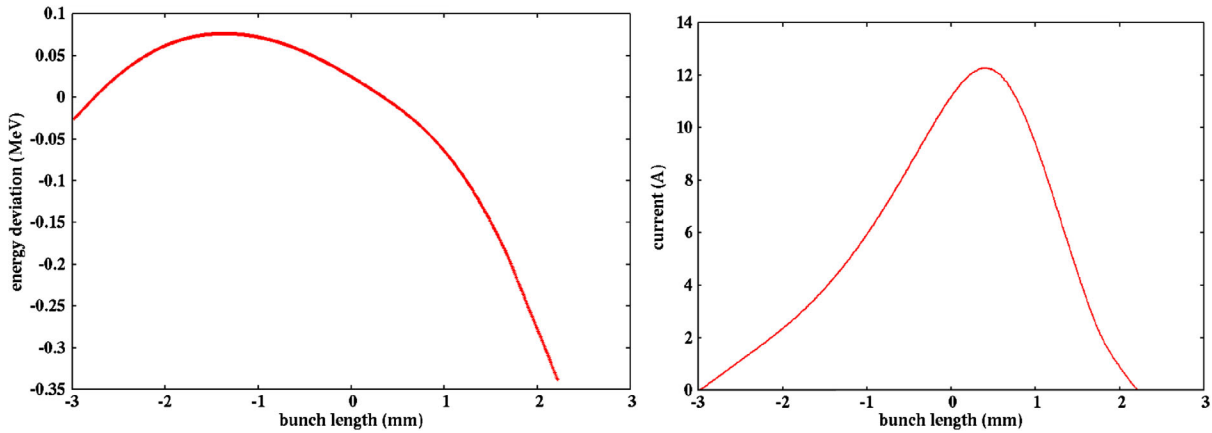


FIG. 2. Initial longitudinal phase space distribution (left) and current profile (right) at the entrance of the linac.

from the tracking using 1024 macroparticles and the above one-dimensional (1D) longitudinal beam dynamics model and the lumped elements and from the three-dimensional (3D) element-by-element tracking using the IMPACT code with the real number of electrons. In the 1D tracking, the longitudinal space-charge effect and the structure wakefields from the rf cavities are included in the simulation. The CSR effect is applied only through the last bending dipole magnet. This is because that the electron beam attains the shortest bunch length (i.e., compressed bunch length) and highest peak current entering the last bending magnet of the bunch compressor chicane. The CSR effect will be strongest inside this magnet [36]. If the peak current before the bunch compressor is already very high, the CSR effect should be applied to the other bending magnets too. It is seen that in this example, the 1D model and the 3D agree with each other very well after BC1.

Figure 4 shows the longitudinal phase space and current profile after magnetic bunch compression chicane BC2 from the above 1D longitudinal beam dynamics model with

lumped beam line elements and from the 3D element-by-element multiparticle tracking using the IMPACT code. It is seen that the longitudinal phase space and the current profile from both models agree with each other quite well. The large current spike around the head of the beam from the 1D model is due to the fact that the 1D model started with a distribution with zero uncorrelated energy spread. Such uncorrelated energy spread in the 3D multiparticle simulation smears the longitudinal phase space somewhat and reduces the current spike near the head of the beam. The longitudinal space-charge effect and the rf cavity structure wakefields are included in the 1D tracking. The CSR effect was applied through the last bending magnet of the BC2.

After the bunch compressor BC2, the electron beam moves through another accelerating section (linac 3), a long transport beam line, and a hard x-ray beam transport line, before entering the undulator section. There are a number of dogleg sections in the transport beam line. Here, we included the longitudinal space-charge effect, the structure

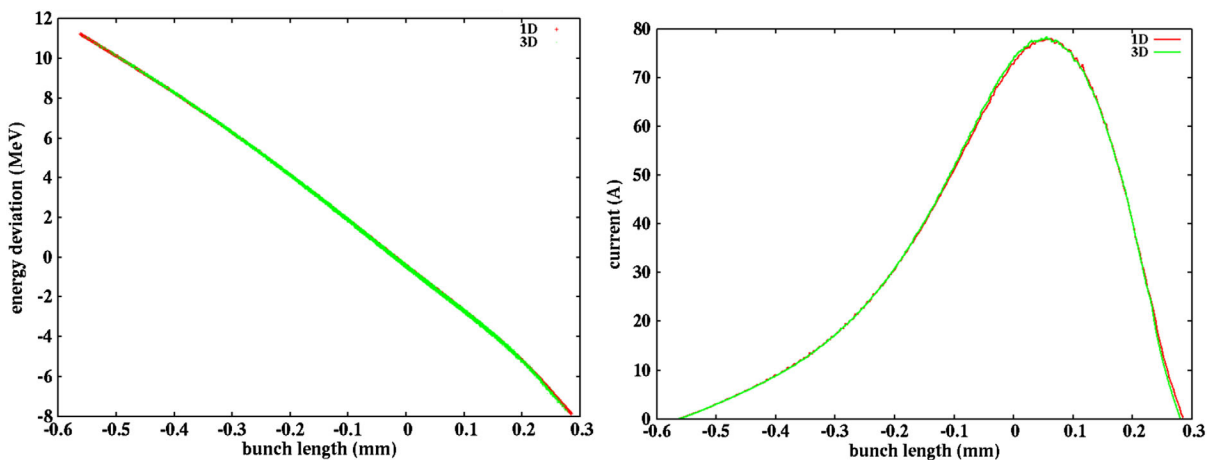


FIG. 3. Longitudinal phase space distribution (left) and current profile (right) after BC1 of the linac from the 1D model (red) and from the 3D model (green).

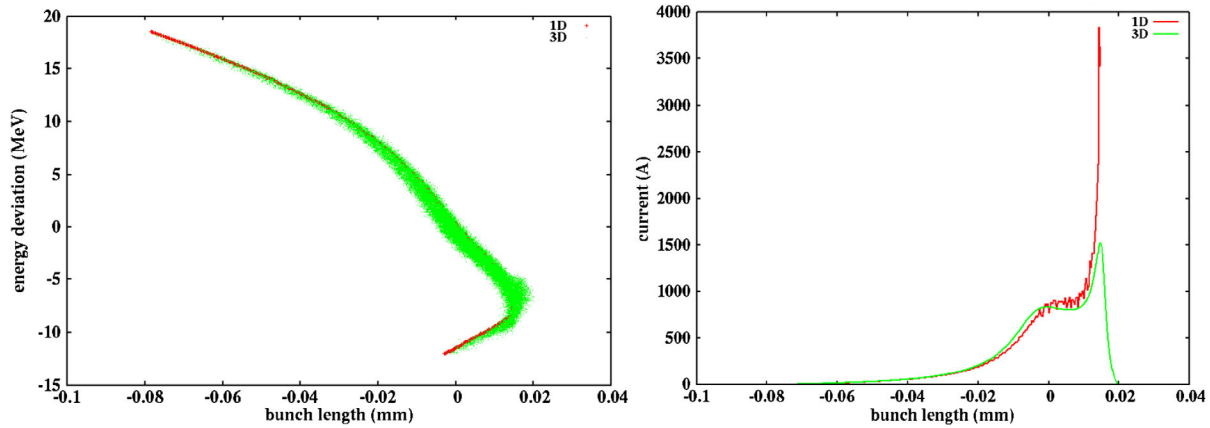


FIG. 4. Longitudinal phase space distribution (left) and current profile (right) after BC2 of the linac from the 1D model (red) and from the 3D model (green).

wakefield in linac3, the CSR effect inside a long bending magnet with a large bending angle, and the resistive wall wakefields. Figure 5 shows the final longitudinal phase space and current profile at the end of accelerator beam delivery system, i.e., at the entrance of undulator section from the above 1D longitudinal beam dynamics model with lumped elements and from the 3D element-by-element tracking using the IMPACT code. The final longitudinal phase space and the current profile from the 1D model agree with those from the 3D model quite well. The longitudinal phase space from the 1D model shows similar shape to that from the 3D model including the fold-over particle distribution near the head of the electron beam. The current spike around the head of the bunch from the 1D model is higher than the spike from the 3D simulation due to the absence of the initial uncorrelated energy spread in the particle distribution.

The benchmark between the above 1D longitudinal beam dynamics model using lumped elements and the 3D element-by-element multiparticle simulation shows good agreement between those two models. This gives us confidence in applying this fast 1D model to final

longitudinal phase space optimization to attain high peak current and flat longitudinal phase space distribution.

III. MULTIOBJECTIVE DIFFERENTIAL EVOLUTION OPTIMIZATION ALGORITHM

In many accelerator applications, one needs to optimize more than one objective function. The problem of multiobjective optimization can be stated in the general mathematical form as:

$$\min \begin{cases} f_1(\vec{x}) \\ \dots \\ f_n(\vec{x}) \end{cases} \quad \text{subject to } g_i(\vec{x}) \leq 0, h_i(\vec{x}) = 0 \quad (28)$$

Here, f_1, \dots, f_n are n objective functions to be optimized, \vec{x} is a vector of control parameters, and g_i and h_i are constraints to the optimization. The goal of multiobjective optimization is to find the Pareto front in the objective function solution space. The Pareto optimal front is a collection of nondominated solutions in the whole feasible

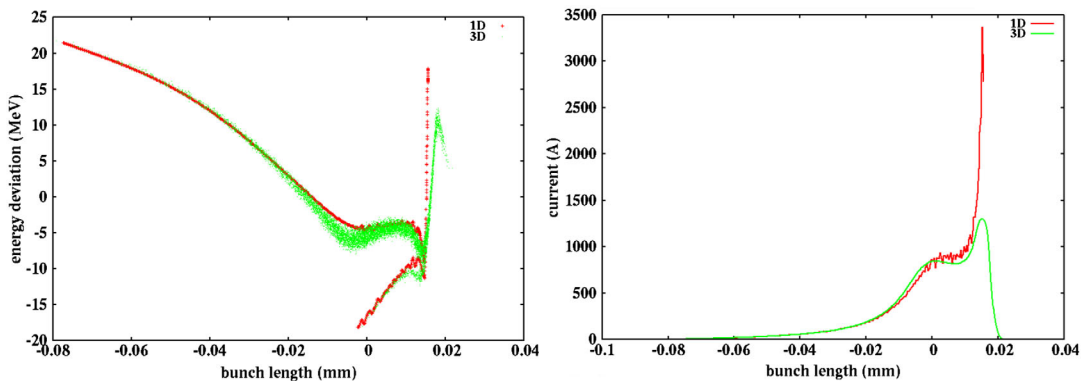


FIG. 5. Longitudinal phase space distribution (left) and current profile (right) at the entrance to undulator from the 1D model (red) and from the 3D model (green).

solution space. Any other solution in the feasible solution space will be dominated by those solutions on the Pareto optimal front. In the multiobjective 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). Here, a component of A corresponds to one objective function value, i.e., $A_i = f_i(\vec{x})$. The solution A is nondominated if it is not dominated by any solution within the group.

Recently, we developed a new parallel multiobjective differential evolution algorithm with varying population size of each generation and with external storage to save nondominated solutions [26,27]. The use of variable population from generation to generation is based on the observation that during the early stage of evolution, the number of nondominated solutions is small. There is no need to keep many dominated solutions in the parent population. As the search evolves, more and more nondominated solutions are obtained. Those nondominated solutions are stored in an external storage so that they can be used for selecting the new parent population. The advantage of using a variable population size with external storage is to reduce the number of objective function evaluations and to improve the speed of convergence. The new algorithm is summarized in the following steps: (i) Step 0: Define the minimum parent size, NP_{\min} and the maximum size, NP_{\max} of the parent population. Define the maximum size of the external storage, NP_{ext} . (ii) Step 1: An initial NP_{ini} population of parameter vectors are chosen quasirandomly to cover the entire solution space. (iii) Step 2: Generate the offspring population using a unified differential evolution algorithm. (iv) Step 3: Check the new population against the constraints. (v) Step 4: Combine the new population with the existing parent population from the external storage. Nondominated solutions (N_{dom}) are found from this group of solutions and $\min(N_{\text{dom}}, \text{Next})$ of solutions are put back to the external storage. Pruning is used if $N_{\text{dom}} > \text{Next}$. NP parent solutions are selected from this group of solutions for next generation production. If $NP_{\min} \leq N_{\text{dom}} \leq NP_{\max}$, $NP = N_{\text{dom}}$. Otherwise, $NP = NP_{\min}$ if $N_{\text{dom}} < NP_{\min}$ and $NP = NP_{\max}$ if $N_{\text{dom}} > NP_{\max}$. The elitism is emphasized through keeping the nondominated solutions while the diversity is maintained by penalizing the overcrowded solutions through pruning. (vi) Step 5: If the stopping condition is met, stop. Otherwise, return to Step 2.

The differential evolution algorithm is a simple but powerful method for global parameter optimization [23–25]. Compared with the other evolutionary algorithms such as the genetic algorithm, the differential evolution algorithm makes use of the differences of parent solutions to attain gradient information. This helps improve the convergence speed of the algorithm in comparison to the genetic algorithm. Meanwhile, compared with the particle swarm method, the differential evolution algorithm has a

cross-over stage to enhance the diversity of solutions. This helps the differential evolution algorithm to avoid converging to a premature solution.

The differential evolution algorithm generates new offsprings using two operations: mutation and crossover. During the mutation stage, for each population member (target vector) \vec{x}_i , $i = 1, 2, 3, \dots, NP$ at generation G , a new mutant vector \vec{v}_i is generated by following a mutation strategy. A number of mutation strategies have been proposed for the conventional standard differential evolution algorithm. The presence of multiple mutation strategies complicates the use of the differential evolution algorithm. Recently, we proposed a single mutation expression that can unify most conventional mutation strategies used by the differential evolution algorithm [37]. This single unified mutation strategy can be written as:

$$\vec{v}_i = \vec{x}_i + F_1(\vec{x}_b - \vec{x}_i) + F_2(\vec{x}_{r_1} - \vec{x}_i) + F_3(\vec{x}_{r_2} - \vec{x}_{r_3}). \quad (29)$$

Here, the second term on the right-hand side of Eq. (29) denotes the contribution from the best solution found in the current generation, the third term denotes the rotationally invariant contribution from the random solution [38], and the fourth term is the same terms as that used in the original differential evolution algorithm to account for the contribution from the difference of parent solutions. Those last two terms divert the mutated solution away from the best solution and help to improve the algorithm's exploration of the decision parameter space. The three parameters F_1 , F_2 , and F_3 are the weights from each contribution. This unified expression represents a combination of exploitation (using the best found solution) and exploration (using randomly chosen solutions) when generating the new mutant solution.

Using the Eq. (29), the multiple mutation strategies of the standard differential evolution algorithm can be included in a single expression. For example, the standard differential evolution algorithm such as DE/rand/1 can be attained by setting $F_1 = 0$, and $F_2 = 1$. This new expression provides an opportunity to explore more broadly the space of mutation operators. Using a different set of parameters F_1 , F_2 , F_3 , a new mutation strategy can be achieved. Moreover, by adjusting these parameters during the evolution, the multiple mutation strategies and their combinations can be used during different stages of optimization. Thus, the unified mutation expression has the virtue of mathematical simplicity and also provides users with flexibility for broader exploration of different mutation strategies.

A crossover operation between the new generated mutant vector \vec{v}_i and the target vector \vec{x}_i is used to further increase the diversity of the new candidate solution. This operation combines the two vectors into a new trial vector \vec{U}_i , $i = 1, 2, 3, \dots, NP$, where the components of the trial vector are obtained from the components of \vec{v}_i or \vec{x}_i

according to a crossover probability Cr . In the binomial crossover scheme for a D dimensional control parameter space, the new trial vector $\vec{U}_i, i = 1, 2, \dots, NP$ is generated using the following rule:

$$\vec{U}_i = (u_{i1}, u_{i2}, \dots, u_{iD}) \quad (30)$$

$$u_{ij} = \begin{cases} v_{ij}, & \text{if } \text{rand}_j \leq Cr \text{ or } j = \text{mbr}_i \\ x_{ij}, & \text{otherwise} \end{cases} \quad (31)$$

where rand_j is a randomly chosen real number in the interval $[0, 1]$, and the index mbr_i is a randomly chosen integer in the range $[1, D]$. This ensures that the new trial vector contains at least one component from the new mutant vector.

During the mutation stage and the crossover stage at generation, each individual solution $\vec{x}_i, i = 1, 2, 3, \dots, NP$ has a set of control parameters $F_{1,i}, F_{2,i}, F_{3,i}$ and Cr_i associated with it. Before generating a new mutant solution using the unified differential evolution expression (29), a set of control parameters $F_{1,i}, F_{2,i}, F_{3,i}$, and Cr_i are calculated as:

$$F_{j,i} = F_{j\min} + r_{ji}(F_{j\max} - F_{j\min}) \quad (32)$$

$$Cr_i = Cr_{\min} + r_i(Cr_{\max} - Cr_{\min}) \quad (33)$$

where $r_{ji}, r_i, j = 1, 2, 3$ are uniform random values in the interval $[0, 1]$, $F_{j\min}$ and $F_{j\max}$ for $j = 1, 2, 3$ are the minimum and the maximum allowed values of those control parameters, Cr_{\min} and Cr_{\max} are the minimum and the maximum cross-over probability. The values of $F_{j\min}$ and $F_{j\max}$ are set to 0 and 1 respectively in this study. We also set $Cr_{\min} = 0.5$ and $Cr_{\max} = 1$. The selection of these values is based on the consideration that the various conventional differential evolution mutation strategies can be covered by the settings of those parameters, and in the literature, F_3 is rarely greater than one.

IV. LONGITUDINAL BEAM DYNAMICS OPTIMIZATION OF A LCLS-II DESIGN

We applied the above fast one-dimensional longitudinal beam dynamics model together with the multiobjective optimization algorithm to an existing LCLS-II design optimization. The two objective functions in this optimization are the final negative fraction of charge and rms energy spread inside a longitudinal phase space window. Here, the window's bunch length is set between -5 microns and 5 microns, while the energy spread is set between -8 MeV and 8 MeV. The smaller the negative fraction of charge, the higher the peak current will be. The smaller the rms energy spread inside the window, the flatter the longitudinal phase space will be. In general, the higher peak current might result in the larger energy spread.

By simultaneously optimizing those two conflict objectives, we hope to attain the best achievable solutions, i.e., the Pareto optimal front. The choice of these two objectives for optimization is because they play an important role in the performance of x-ray FEL radiation. The higher peak electron current and the smaller energy spread will result in shorter FEL radiation saturation distance, higher radiation power, and narrower radiation bandwidth [39].

We used 10 control parameters of the linear accelerator in this optimization. These 10 control parameters in the LCLS-II linac are the linac section one rf cavity accelerating gradient amplitude and rf phase, 3rd harmonic cavity amplitude and phase, bending angle in bunch compressor one, linac section two rf cavity amplitude and phase, bending angle in bunch compressor two, and linac three rf cavity amplitude and phase. The maximum rf cavity accelerating gradient is constrained around 16 MV/m and the final energy around 4 GeV. The initial population size is 256. The minimum and the maximum population size are set as 128 and 1024 respectively. The maximum number of nondominated solutions in the external storage is set as 2000. The Pareto front converges after about 460 generations with about 76 thousand objective function evaluations. It took one and half hour computing time on 64 Intel Xeon Phi Processor 7250 at National Energy Research Supercomputing Center. A major section of the Pareto front (with fraction of charge less than 0.6) has converged with less than one hour of computing time after 300 generations and 40 thousand function evaluations. Figure 6 shows the Pareto front evolution of the final rms energy spread and the negative fraction of charge inside the defined window from the above multiobjective longitudinal beam dynamics optimization. It is seen that after 200 generations, the Pareto front is already close to the final optimal front solution. The optimal front suggests that the more charge inside the

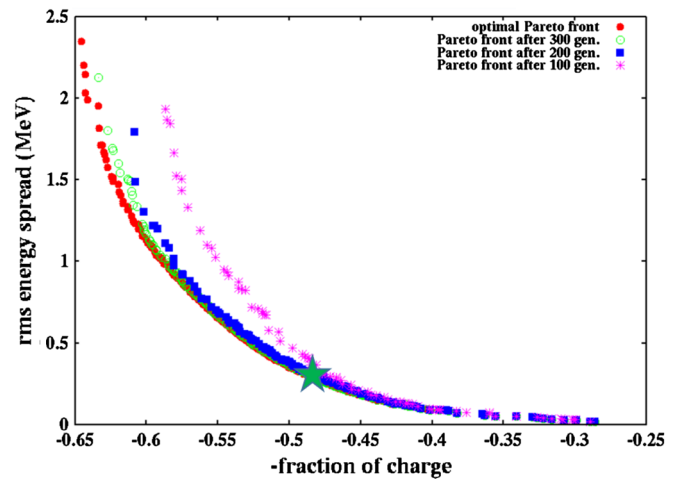


FIG. 6. The Pareto front after 100 generations (magenta), after 200 generations (blue), after 300 generations (green) and the final optimal Pareto front (red) of longitudinal phase space optimization.

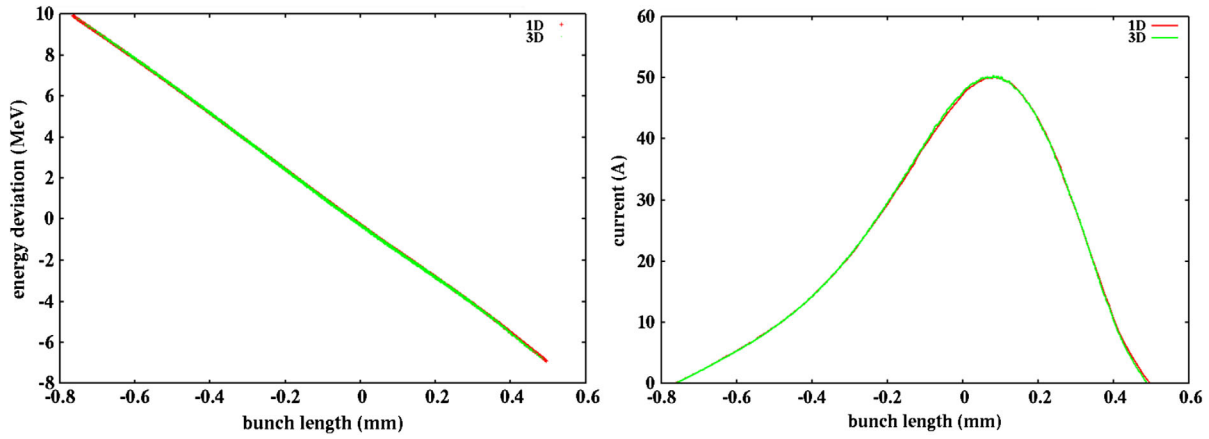


FIG. 7. Longitudinal phase space distribution (left) and current profile (right) after BC1 of the LCLS-II linac from the 1D model (red) and from the 3D model (green) using an optimal solution settings.

core of the beam, the larger correlated energy spread the longitudinal phase space would be.

From those optimal solutions, we selected one optimal solution, the green star in Fig. 6. The selection of this

solution is somewhat based on trial and error. We would like to have final peak current inside the core of the electron beam beyond 1 kA while keeping the relative rms energy spread below 0.01%. For this solution, the optimized linac

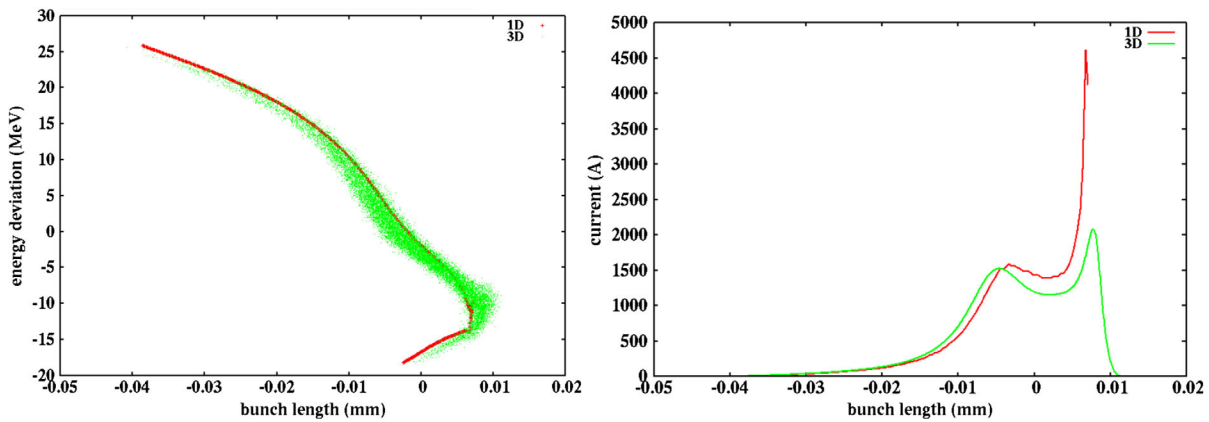


FIG. 8. Longitudinal phase space distribution (left) and current profile (right) after BC2 of the linac from the 1D model (red) and from the 3D model (green) using an optimal solution settings.

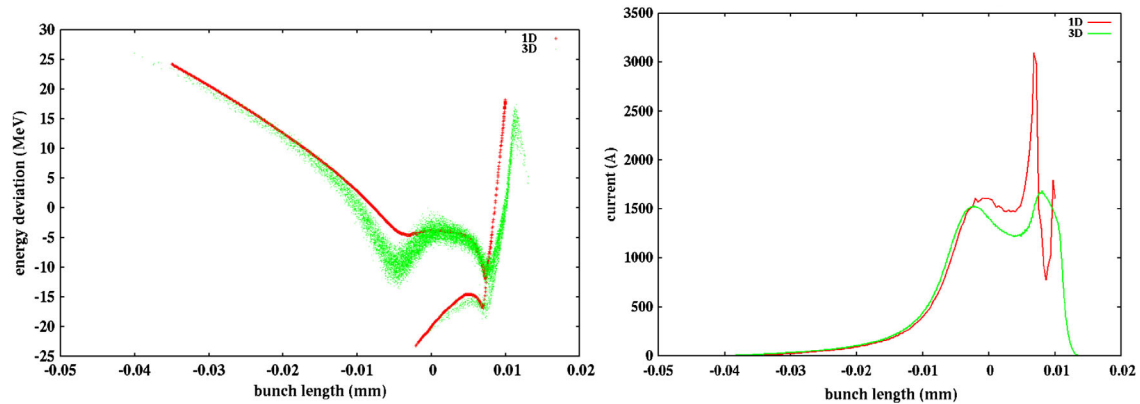


FIG. 9. Longitudinal phase space distribution (left) and current profile (right) at the entrance to undulator from the 1D model (red) and from the 3D model (green) using an optimal solution settings.

one RF cavity accelerating gradient amplitude and phase are 11.5 M/m and -14.9 degrees respectively, harmonic linearizer RF cavity amplitude and phase are 9.0 M/m and 153.6 degrees respectively, bunch compression one bending angle is 0.10 radian, linac two RF cavity amplitude and phase are 16.1 M/m and -30.4 degrees, bunch compressor two bending angle is 0.047 radian, linac three amplitude and phase are 16.0 M/m and 0.0 degrees. With those parameter settings of the LCLS-II linac, we ran the fast 1D longitudinal beam dynamics simulation and the fully 3D element-by-element multiparticle simulation using the IMPACT code. The longitudinal phase spaces and current profiles after BC1, after BC2, and at the entrance to undulator are given in Figs. 7–9. It is seen that the predictions of the 1D longitudinal beam dynamics model agree with those of the 3D model quite well. There is about a factor of four compression after the first bunch compression chicane. After the second bunch compressor chicane, the electron beam is further compressed by more than a factor of 20. The optimized solution results in a final core peak current greater than 1.2 kA, which is about 50% improvement from the previous design of around 800 A peak current. Such a higher core peak current can result in higher x-ray FEL radiation power.

V. CONCLUSIONS

In this paper, we proposed a lumped one-dimensional longitudinal beam dynamics model that includes nonlinear effects from drift, rf acceleration, magnetic bunch compression, longitudinal space-charge effect, CSR effect, and structure and resistive wall wakefields, for fast longitudinal phase space optimization in x-ray FEL linear accelerators. Benchmarking using a LCLS-II design, this simplified model shows good agreement with the fully 3D element-by-element multiparticle tracking using the IMPACT code. This model, implemented in a recently developed multiobjective differential evolution optimization program, provides a fast computational tool for the longitudinal phase space optimization and resulted in an improved solution on the existing LCLS-II design.

In the above application example, we started the simulation from the initial current profile and the correlated slice energy distribution (zero uncorrelated energy spread). This could result in large current spike near the end of the beam distribution and potential longitudinal microbunching instability. The uncorrelated energy spread can be included in the 1D model if one uses an initial 2D phase space distribution with nonzero uncorrelated energy spread and more macroparticles to represent variation along the energy dimension. The microbunching instability can be mitigated by smoothing the longitudinal current profile using coarser grid points. This can also be attained through using more macroparticles. Using a large number of macroparticles will be useful to study microbunching

instability (for example to predict the instability gain), but it will slow down the computation significantly.

The 1D longitudinal optimization model helps attain a better solution with higher core peak current. However, this does not guarantee that this solution would be useful for the final x-ray FEL radiation if the transverse emittance is substantially degraded by the CSR effect due to large compression after the BC2. This issue can be addressed in two ways: First, the compression factors in BC1 and BC2 can be constrained in the above longitudinal beam dynamics optimization to keep these compressions within acceptable levels. Second, the accelerator beam line around the bunch compressors can be designed with small horizontal beam size to reduce the CSR induced emittance growth. The latter method was employed in the LCLS-II design. For the above optimal solution, without retuning the transverse lattice settings, the projected rms horizontal emittance growth after BC2 is around 30% while the slice emittance stays about the same.

Given the fast computational speed of the proposed 1D model, in future study, we plan to apply this tool to the optimization of using control parameters from both the linac parameter setting and the initial longitudinal phase space and current profile. Those optimization results will help set up the requirements for the electron beam longitudinal distribution out of a photoinjector in order to achieve the optimal final electron beam quality at the entrance of undulator.

ACKNOWLEDGMENTS

We would like to thank the LCLS-II physics design team for the LCLS-II application study. This work was supported by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 and used computer resources at the National Energy Research Scientific Computing Center.

-
- [1] P. Emma *et al.*, First lasing and operation of an ångstrom-wavelength free-electron laser, *Nat. Photonics* **4**, 641 (2010).
 - [2] E. Allaria *et al.*, Highly coherent and stable pulses from the FERMI seeded free-electron laser in the extreme ultraviolet, *Nat. Photonics* **6**, 699 (2012).
 - [3] H. Weise and W. Decking, in *Proceedings of the 2017 Free Electron Laser Conference* (JACoW, Geneva, 2017), p. 913.
 - [4] H. Tanaka *et al.*, A compact x-ray free-electron laser emitting in the sub-ångstrom region, *Nat. Photonics* **6**, 540 (2012).
 - [5] H.-S. Kang *et al.*, PAL-XFEL technical design report, Pohang Accelerator Laboratory, 2014.
 - [6] C. Milne *et al.*, SwissFEL: The Swiss x-ray free electron laser, *Appl. Sci.* **7**, 720 (2017).
 - [7] I. Zagorodnov and M. Dohlus, Semianalytical modeling of multistage bunch compression with collective effects, *Phys. Rev. ST Accel. Beams* **14**, 014403 (2011).

- [8] S. DiMitri and M. Cornacchia, Electron beam brightness in linac drivers for free-electron-lasers, *Phys. Rep.* **539**, 1 (2014).
- [9] P. Emma *et al.*, in *Proceedings of the 2014 Free Electron Laser Conference, Basel, Switzerland* (JACoW, Geneva, 2014), pp. 743–747.
- [10] T. O. Raubenheimer *et al.*, in *Proceedings of the 2015 Free Electron Laser Conference, Daejeon, Korea* (JACoW, Geneva, 2015), pp. 618624.
- [11] K. L. F. Bane and P. Emma, LiTrack: A fast longitudinal phase space tracking code with graphical user interface, Report No. SLAC-PUB-11037, 2005.
- [12] M. Borland, Elegant, Report No. APS LS-287, presented at ICAP 2000, Darmstadt, Germany.
- [13] J. Qiang, R. D. Ryne, S. Habib, and V. Decyk, An object-oriented parallel particle-in-cell code for beam dynamics simulation in linear accelerators, *J. Comput. Phys.* **163**, 434 (2000).
- [14] J. Qiang, R. D. Ryne, M. Venturini, A. A. Zholents, and I. V. Pogorelov, High resolution simulation of beam dynamics in electron linacs for x-ray free electron lasers, *Phys. Rev. Accel. Beams* **12**, 100702 (2009).
- [15] I. V. Bazarov and C. K. Sinclair, Multivariate optimization of a high brightness dc gun photoinjector, *Phys. Rev. ST Accel. Beams* **8**, 034202 (2005).
- [16] C. F. Papadopoulos *et al.*, in *Proceedings of FEL2010, Malmö, Sweden*, p. 479, <http://accelconf.web.cern.ch/AccelConf/FEL2010/papers/wepb37.pdf>.
- [17] R. Bartolini, M. Apollonio, and I. P. S. Martin, Multi-objective genetic algorithm optimization of the beam dynamics in linac drivers for free electron lasers, *Phys. Rev. Accel. Beams* **15**, 030701 (2012).
- [18] A. Hofler, B. Terzić, M. Kramer, A. Zvezdin, V. Morozov, Y. Roblin, F. Lin, and C. Jarvis, Innovative applications of genetic algorithms to problems in accelerator physics, *Phys. Rev. Accel. Beams* **16**, 010101 (2013).
- [19] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. Evol. Comput.* **6**, 182 (2002).
- [20] J. Holland, *Adaptation In Natural and Artificial Systems* (University of Michigan Press, Ann Arbor, 1975).
- [21] <https://ewh.ieee.org/conf/cec/>.
- [22] R. Storn and K. Price, Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces, ICSI, USA, Tech. Report No. TR-95-012, 1995.
- [23] R. Storn and K. Price, Differential evolution—a simple and efficient heuristic for global optimization over continuous Spaces, *J. Global Optim.* **11**, 341 (1997).
- [24] M. M. Ali and A. Torn, *Computers and Operations Research* (Elsevier, New York, 2004), no. 31, p. 1703.
- [25] K. Price *et al.*, *Differential Evolution - A Practical Approach to Global Optimization* (Springer, Berlin, 2005).
- [26] J. Qiang, Y. Chao, C. Mitchell, S. Paret, and R. Ryne, in *Proceedings of the 4th International Particle Accelerator Conference, IPAC-2013, Shanghai, China, 2013* (JACoW, Shanghai, China, 2013), p. 1031.
- [27] J. Qiang, Advances in global optimization of high brightness beams, *Int. J. Mod. Phys. A* (to be published).
- [28] J. Qiang, C. Mitchell, R. Ryne, and M. Venturini, in *Proceedings of ICAP2015, Shanghai, China, 2015*, <http://accelconf.web.cern.ch/AccelConf/ICAP2015/papers/tuaji1.pdf>.
- [29] M. Dohlus, T. Limberg, and P. Emma, Bunch compression for linac-based FELs, *ICFA Beam Dynamics Newsletter* **38**, 15 (2005).
- [30] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (Adam Hilger, New York, 1988).
- [31] J. Qiang, S. Lidia, R. D. Ryne, and C. Limborg-Deprey, Three-dimensional quasistatic model for high brightness beam dynamics simulation, *Phys. Rev. Accel. Beams* **9**, 044204 (2006).
- [32] T. Weiland and I. Zagorodnov, TESLA Report No. 2003-19, 2003.
- [33] E. L. Saldin, E. A. Schneidmiller, and M. V. Yurkovb, On the coherent radiation of an electron bunch moving in an arc of a circle, *Nucl. Instrum. Methods Phys. Res., Sect. A* **398**, 373 (1997).
- [34] J. Qiang, C. Mitchell, and R. Ryne, A fast high-order method to calculate wakefields in an electron beam, *Nucl. Instrum. Methods Phys. Res., Sect. A* **682**, 49 (2012).
- [35] C. Mitchell, J. Qiang, and R. Ryne, A fast method for computing 1-D wakefields due to coherent synchrotron radiation, *Nucl. Instrum. Methods Phys. Res., Sect. A* **715**, 119 (2013).
- [36] C. Mitchell, J. Qiang, and P. Emma, Longitudinal pulse shaping for the suppression of coherent synchrotron radiation-induced emittance growth, *Phys. Rev. Accel. Beams* **16**, 060703 (2013).
- [37] J. Qiang, C. Mitchell, and A. Qiang, in *Proceedings of CEC2016, Vancouver*, p. 4061, 2016, <https://doi.org/10.1109/CEC.2016.7744305>.
- [38] K. V. Price, An introduction to differential evolution, in *New Ideas in Optimization*, edited by D. Corne, M. Dorigo, and V. Glover (McGraw-Hill, London, 1999), pp. 79–108.
- [39] Z. Huang and K. J. Kim, Review of x-ray free-electron laser theory, *Phys. Rev. Accel. Beams* **10**, 034801 (2007).