RESEARCH ARTICLE | MAY 15 2023

Optimization of transformer ratio and beam loading in a plasma wakefield accelerator with a structure-exploiting algorithm

Q. Su ✉; J. Larson; T. N. Dalichaouch; ... et. al

Physics of Plasmas 30, 053108 (2023)
https://doi.org/10.1063/5.0142940

Physics of Plasmas ARTICLE pubs.aip.org/aip/php

Optimization of transformer ratio and beam loading in a plasma wakefield accelerator with a structure-exploiting algorithm

Cite as: Phys. Plasmas **30**, 053108 (2023); doi: 10.1063/5.0142940
Submitted: 18 January 2023 · Accepted: 1 April 2023 ·
Published Online: 15 May 2023

Q. Su,[1,a] J. Larson,[2,b] T. N. Dalichaouch,[3] F. Li,[1] W. An,[4,5] L. Hildebrand,[3] Y. Zhao,[3] V. Decyk,[3] P. Alves,[3] S. M. Wild,[2,6] and W. B. Mori[1,3]

AFFILIATIONS

[1]Department of Electrical and Computer Engineering, University of California, Los Angeles, Los Angeles, California 90095, USA
[2]Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, Illinois 60439, USA
[3]Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California 90095, USA
[4]Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206, China
[5]Department of Astronomy, Beijing Normal University, Beijing 100875, China
[6]Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

[a]**Author to whom correspondence should be addressed:** xpsqq@g.ucla.edu
[b]**Electronic mail:** jmlarson@anl.gov

ABSTRACT

Plasma-based acceleration has emerged as a promising candidate as an accelerator technology for a future linear collider or a next-generation light source. We consider the plasma wakefield accelerator (PWFA) concept where a plasma wave wake is excited by a particle beam and a trailing beam surfs on the wake. For a linear collider, the energy transfer efficiency from the drive beam to the wake and from the wake to the trailing beam must be large, while the emittance and energy spread of the trailing bunch must be preserved. One way to simultaneously achieve this when accelerating electrons is to use longitudinally shaped bunches and nonlinear wakes. In the linear regime, there is an analytical formalism to obtain the optimal shapes. In the nonlinear regime, however, the optimal shape of the driver to maximize the energy transfer efficiency cannot be precisely obtained because currently no theory describes the wake structure and excitation process for all degrees of nonlinearity. In addition, the ion channel radius is not well defined at the front of the wake where the plasma electrons are not fully blown out by the drive beam. We present results using a novel optimization method to effectively determine a current profile for the drive and trailing beam in PWFA that provides low energy spread, low emittance, and high acceleration efficiency. We parameterize the longitudinal beam current profile as a piecewise-linear function and define optimization objectives. For the trailing beam, the algorithm converges quickly to a nearly inverse trapezoidal trailing beam current profile similar to that predicted by the ultrarelativistic limit of the nonlinear wakefield theory. For the drive beam, the beam profile found by the optimization in the nonlinear regime that maximizes the transformer ratio also resembles that predicted by linear theory. The current profiles found from the optimization method provide higher transformer ratios compared with the linear ramp predicted by the relativistic limit of the nonlinear theory.

I. INTRODUCTION

Plasma wakefield acceleration (PWFA),[1] in which a particle beam drives a plasma wave wakefield on which a trailing particle beam surfs, has emerged as a promising candidate for future development of more compact and cost-effective advanced lightsources and linear colliders. In PWFA, plasma electrons are blown out by the space charge force of the relativistic particle beam and attracted back to axis by plasma ions, creating a close to spherical ion bubble surrounded by a sheath of electrons. Experiments have shown great potential for PWFA to simultaneously achieve high accelerating gradients, high-quality beams, and high energy transfer efficiency.[2–6] In a two-bunch PWFA, the drive beam transfers its energy to the wake, and a second electron or positron bunch (the trailing beam) can be placed at an appropriate distance behind the drive beam to gain energy from the

wake. The trailing beam can be externally injected or self-injected. For linear collider applications, both the energy transfer efficiency from the drive beam (η_w) to the wake and from the wake to the trailing beam (η_b) must be high. The preservation of the trailing beam quality, for example, the energy spread and emittance, during the acceleration is also important.

In linear theory, a figure of merit that helps to characterize the energy transfer efficiency is the transformer ratio R. It is defined as the ratio between the maximum absolute values of the accelerating field left behind the driver $|E_z^+|$ (or experienced by the trailing bunch if loaded) and the decelerating field inside the driver $|E_z^-|$,

$$R = \frac{\max|E_z^+|}{\max|E_z^-|}. \tag{1}$$

The meaning of the transformer ratio can be seen as follows. Some particles in the drive beam will lose all their energy in a pump depletion distance, $L_{pd} = \frac{\gamma_b mc^2}{e\max|E_z^-|}$, where γ_b is the Lorentz factor of the drive beam. Thus, the maximum energy that a particle in the trailing beam can obtain is $e\max|E_z^+|L_{pd} = \gamma_b mc^2 R$.

In order to obtain high acceleration efficiency, it is important to maximize the amount of drive beam energy that is transferred to the wake. This occurs when all the particles in the drive beam decelerate at the same rate; in other words, the drive beam feels a uniform decelerating field. Otherwise, some particles in the drive beam will deplete their energy while other particles have substantial remaining energy. A higher transformer ratio can lead to a higher energy gain of the trailing beam for a given drive beam energy over the acceleration distance. However, there is an inherent trade-off between the maximum loaded charge of the trailing beam and the transformer ratio. Assuming that both the loaded E_z inside the drive beam and the trailing beam are constant. If 100% of the drive beam energy were converted into the trailing beam, the trailing beam charge would be smaller than the drive beam charge by a factor proportional to the transformer ratio R.

It has been shown[7] that the transformer ratio cannot exceed 2 in one-dimensional (1D) linear theory for drive beams with symmetric current profiles (with respect to the beam propagation direction). A higher transformer ratio can be obtained by using an asymmetric current distribution.[8] It is not immediately obvious how the transformer ratio is connected to efficiency. However, linear theory leads to the conclusion that for a bunch with fixed length and fixed charge, the maximum transformer ratio for the unloaded wake occurs when the decelerating field is constant,[9] which is also the condition for transferring all of the drive bean energy into the wake. Thus, in the linear regime, the transformer ratio is also a metric for efficiency.

The optimal beam profile for energy transfer efficiency and transformer ratio in the linear regime is, therefore, defined as one that leads to a constant decelerating field. This current profile in 1D was shown to be a delta function precursor followed by a linear ramp. If the decelerating field is parameterized to go from 0 to a constant as $(1 - e^{-\alpha\zeta})E_0$ when $\alpha \to \infty$, then the current profile that achieves this is[8]

$$\rho(\zeta) \sim -\frac{E_0}{4\pi\alpha}\left[\left(\alpha^2 + k_p^2\right)e^{-\alpha\zeta} + k_p^2(\alpha\zeta - 1)\right], \quad \alpha \to \infty, \tag{2}$$

which reduces to a delta function precursor and a linear ramp in the asymptotic limit where $\alpha \to \infty$. Here, ρ is defined as the one-dimensional charge density varying in the longitudinal direction $\zeta = v_b t - z$. In this context, k_p is defined as ω_p/v_b, where ω_p is the plasma oscillation frequency and v_b is the beam velocity. In the limit where $\alpha \to \infty$, the transformer ratio is $R = \sqrt{1 + k_p^2 L^2}$ and the ratio of the charge in the precursor to that in the linear ramp region is $2/k_p^2 L^2$. For such a current profile, there is 100% energy transfer efficiency from beam to wake (to truly achieve 100%, the energy per particle in the precursor should be half that in the wedge-shaped region).

In 3D, a similar analysis can be used because the Green's function for the wake is the response from a delta function charge distribution. In this regime, the decelerating field varies across the beam. Thus, without shaping the driver in the transverse direction, the entire beam cannot slow down together.

Henceforth, we will use normalized units unless we explicitly assign physical units. Charge is normalized to electron charge e; length to k_p^{-1} defined as c/ω_p, where ω_p is the plasma oscillation frequency and c is the speed of light; charge density to en_p; current density to $en_p c$; electric field to $mc\omega_p/e$; and potentials to mc^2/e. We also use the comoving coordinates of a relativistic beam by making a mathematical transformation from the (x, y, z, t) to the $(x, y, \xi = ct - z, s = t)$ variables.

In the 3D blowout (nonlinear) regime, there is currently no precise theoretical formalism to obtain the current profile that flattens the decelerating field and to find out whether this also optimizes R. Phenomenological models for nonlinear wakefields have assumed physical descriptions where an ion channel is surrounded by one or more electron sheaths.[10,11] However, these models tend to break down at the front of the wake where plasma electrons are not completely blown out and the ion channel has not yet fully formed. Thus, the optimal longitudinal shape for the drive beam in the 3D nonlinear regime has not been well studied.

Much of our understanding of the structure and fields of plasma wakefields in the blowout regime comes from nonlinear theory developed by Lu *et al.*[10] In Ref. 10, it was shown that the shape of the ion channel could be completely described by the plasma wake potential $\psi = \phi - A_z$ and the current profiles of the drive and trailing bunches. To use the theory as a predictive tool, Lu *et al.* introduced a single-sheath model by modeling the plasma source term profile $S = -(\rho - J_z)$ to obtain an expression for the wake potential $\psi = \phi - A_z$,

$$\nabla_\perp^2 \psi = -(\rho - J_z). \tag{3}$$

This wake potential ψ defines the focusing field $\vec{E}_\perp + \hat{z} \times \vec{B}_\perp = -\nabla_\perp \psi$ and the accelerating field $E_z = \frac{\partial \psi}{\partial \xi}$ experienced by relativistic beam particles. It has been proven that in the 3D nonlinear regime, the decelerating (and accelerating) fields do not vary across the beam in the transverse direction, so each slice of the drive beam will slow down together.[10,12]

By applying this theory in the very nonlinear (ultrarelativisitic) limit, it was shown that an adiabatically increasing linear current profile would still provide a nearly constant decelerating field[13] even in the nonlinear regime. Based on the nonlinear wakefield theory of Lu *et al.*,[10] the equation for the innermost electron trajectory in the ultrarelativistic limit where the blowout radius $r_b \gg 1$ is given by

$$r_b \frac{d^2 r_b}{d\xi^2} + 2\left[\frac{dr_b}{d\xi}\right]^2 + 1 = \frac{4\Lambda(\xi)}{r_b^2}. \tag{4}$$

Here, $\Lambda(\xi) = \int_0^\infty r n_b(r) dr$ represents a normalized charge per unit length for the beam. By assuming $r_b \gg 1$, the on-axis wake potential $\psi(r, \xi) = \phi - A_z$ can be approximated as $\psi(0, \xi) = \frac{r_b^2(\xi)}{4}$.[10]

For an adiabatic response, we can assume $\frac{d^2 r_b}{d\xi^2} \ll 1, \frac{dr_b}{d\xi} \ll 1$. In this limit, one can predict[13] that a constant decelerating field exists within the drive beam when

$$\Lambda(\xi) = \frac{\xi}{L}\Lambda_0, \quad \psi(0, \xi) \approx \frac{\xi}{L}\Lambda_0, \quad \text{and} \quad E_z(\xi) = \frac{\partial \psi}{\partial \xi} \approx \frac{\Lambda_0}{L}, \quad (5)$$

where L is the beam length and Λ_0 is the maximum Λ. Within the described approximations, this would be the current profile that provides the highest efficiency in the nonlinear regime.

As noted above, in the 1D linear regime, one can show that the transformer ratio is maximized when the efficiency is also maximized. However, it is not clear how the transfer efficiency and transformer ratio are related to the nonlinear regime. The accelerating field in a nonlinear wake has a deep spike leading to very large peak accelerating fields for electrons that may not be useful. Furthermore, some plasma electrons are blown out with sufficient energy that they are lost, so not all the driver energy goes to the wake.

One can, however, use existing nonlinear theory to obtain a scaling of the transformer ratio with a bunch length for a driver with fixed charge. We define a peak accelerating field that ignores the deep spike. For $r_b \gg 1$, the slope of accelerating field $\partial E_z/\partial \xi \approx -1/2$. Assuming a spherical ion cavity, the maximum useful accelerating field at the rear of the wake can, therefore, be estimated using $|E_z^+| \approx \frac{1}{2} r_{\max} \approx \sqrt{\Lambda_0}$, where $r_{max} \approx 2\sqrt{\Lambda_0}$ by assuming the blowout radius r_b reaches its maximum $r_{\max}$ immediately after the drive beam ends.[13] From these assumptions, it follows that the maximum transformer ratio is $R = |E_z^+|/|E_z^-| \approx \sqrt{\Lambda_0}/(\frac{\Lambda_0}{L}) = \frac{L}{\sqrt{\Lambda_0}}$. In linear theory, R depends only on L. In the nonlinear regime, however, R also depends on the peak charge per unit length Λ_0, which can be written for fixed charge Q as $2\pi\Lambda_0 = 2Q/L$. Thus, in the nonlinear regime, the transformer ratio scales as $R \approx \frac{\sqrt{2\pi} L^{3/2}}{(2Q)^{1/2}}$.

Just as in the linear regime, one would expect that a precursor could provide the highest efficiencies. A precursor can rapidly increase the wakefield from which the body of the driver can build upon. However, no theoretical formalism exists for obtaining the response of a precursor and current ramp when the plasma responds nonlinearly. For example, plasma electrons are not fully blown out at the head of the bunch, so it is invalid to extend Eq. (4) to include a precursor. Furthermore, studying nonlinear wakes for which the ultrarelativistic limit is not appropriate is more complicated. We note that experimental evidence indicates that high transformer ratios in the nonlinear regime can be obtained for a triangular shape preceded by a precursor and succeeded by a bump at the tail of the beam.[14]

In order to accurately assess the overall efficiency, it is also important to examine how a trailing beam with charge of interest absorbs the wake energy, which is quantified by η_b. This is the subject of beam loading. Katsouleas *et al.*[15] showed that linear theory predicts that a trapezoidal current profile can minimize the energy spread (flatten the wake) and emittance growth. However, there is a trade-off between efficiency and the acceleration gradient felt by the loaded bunch. Beam loading in nonlinear wakes was analyzed by Tzoufras *et al.*[16] using the nonlinear wakefield theory with a single-sheath model,[10] where it was also found that a nearly trapezoidal current that decreases from front to back is the optimal shape to flatten the wake. In this regime, emittance preservation for beams with finite energy spread can also be achieved through the use of matched beams, and the efficiency can be very high when compared with linear theory. Recently, an improved description for nonlinear wakefields was developed using a multi-sheath model.[11] The multi-sheath model naturally extends the single-sheath model[10] by including a second plasma sheath that captures regions where the source term, $S = -(\rho - J_z)$, for slices at the rear of the bubble is negative outside the ion channel. The source term in Eq. (3) is modeled as three regions: an ion channel with radius $r_b(\xi)$ and $S = -1$, an innermost plasma sheath with width Δ_1 and $S \equiv n_1 > 0$, and an outermost plasma sheath of width Δ_2 for which the source term $S \equiv n_2 < 0$. Integrating Eq. (3), they obtained an expression for the wake potential,[11]

$$\psi(r, \xi) = \psi_0(\xi) - \frac{r^2}{4} = \frac{r_b^2(\xi)}{4}(1 + \beta') - \frac{r^2}{4}, \quad (6)$$

where $\psi_0(\xi) = \psi(0, \xi)$ is the on-axis wake potential. The function β' depends on the parameters of the source term profile and is given by

$$\beta' = 2(1 + n_1)\ln(1 + \alpha_1) - 1 + 2n_2(1 + \alpha_1 + \alpha_2)^2 \ln\left(1 + \frac{\alpha_2}{1 + \alpha_1}\right),$$

where $\alpha_1 \equiv \frac{\Delta_1}{r_b}$ and $\alpha_2 \equiv \frac{\Delta_2}{r_b}$. The parameters n_1, n_2, Δ_1, and Δ_2 are related by the conservation of charge,[10]

$$\int_0^\infty (\rho - J_z) r dr = 0. \quad (7)$$

The single-sheath expression of ψ can be obtained by setting $n_2 = 0$. Including an additional sheath enables the modeling of negative wake potentials at the rear of the wake, which is important for beam loading and self-injection. It was shown that this multi-sheath model provided a more accurate description of the wakefield and could be applied to obtain a beam current with higher acceleration efficiency and lower energy spread compared with the single-sheath model.[11]

Advances in computational power and improved algorithmic development in PWFA simulation tools have opened the possibility to directly determine the optimized current profiles of both the drive beam and the trailing beam from simulation. We define the optimization objective to be those that provide the highest efficiency for a given loaded transformer ratio with the lowest energy spread. Previously, evolutionary algorithms have been applied in accelerator experiments.[17,18] Recently, a slice-by-slice loading algorithm has been used to optimize beam loading in PWFA.[19] Optimization approaches, such as Bayesian optimization[20–22] and neural networks,[23] have also aroused great interest in the community of laser–plasma accelerators.

In this paper, we develop a specialized numerical optimization routine that seeks optimal drive beam and trailing beam current profiles in the nonlinear blowout regime. The method is inspired by the nonlinear least squares solver POUNDERs[24] with modifications to find a current profile that minimizes the deviation of the accelerating field or decelerating field, E_z, about its mean. We couple data from the quasi-static particle-in-cell (PIC) code QuickPIC[25,26] to this modified POUNDERs. We parameterize the beam current profile as a piecewise-linear function for which the beam charge is constant.

The transverse profile is assumed to be Gaussian with a fixed spot size. We also define an optimization objective to quantify variations in the electric field within the beam (either drive or trailing) about an average. The optimization algorithm can work efficiently to minimize the objective function. The procedure also permits including constraints, such as the beam length and/or the beam charge. We first use the optimization method to find the optimal profile for the trailing beam subject to constraining the length. This constraint is also imposed by other optimization methods. As mentioned above, the equations for both the single-sheath[16] and multi-sheath models[11] for nonlinear wakefields[10] can be integrated for a trailing beam inside a nonlinear wake. There is excellent agreement between the predictions from the multi-sheath model for the current profiles that flatten the wakefield and the optimized results from particle-in-cell (OSIRIS and QuickPIC) simulations. The current profiles are nearly inverse trapezoidal in shape. There are slight differences between the algorithm-searched approach current profiles and the theoretical predictions for the beam-loading problem, which will be described shortly. The fact that there is such good agreement gives confidence in both the algorithm-searched process and the theory.

We next use this method to find the optimal drive beam profile that flattens the decelerating field. In this case, we constrain the total charge in the beam. As noted above, it is not straightforward to use existing nonlinear theory because a fully blown-out wake does not exist at the head of the beam. For this problem, it is also not possible to use optimization procedures that rely on a slice-by-slice procedure because the value of the objective decelerating field must be known before the optimization for this kind of approach. Interestingly, the algorithm-searched results find current profiles that are nearly identical to those predicted by 1D linear theory. The optimal current profile has a precursor followed by a triangular shape. We also find that even in the fully nonlinear regime, the profiles that provide the most flattened decelerating field also lead to the largest transformer ratios of the current profiles considered for a given fixed total charge and bunch length.

For the cases examined in this paper, we ignore the effects of ion motion on the focusing and accelerating fields as they are expected to be small.[27] Furthermore, the optimization method described here can straightforwardly be applied to cases where ion motion needs to be included.

II. THE OPTIMIZATION ALGORITHM

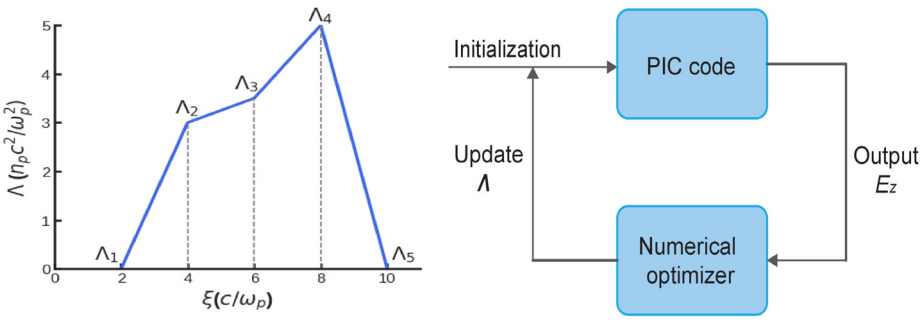
To optimize the shape of the drive beam, we discretize the beam longitudinal profile by initializing a piecewise-linear charge density described by the discretized beam charge per unit length $\boldsymbol{\Lambda}$. In QuickPIC, we initialize a longitudinally piecewise-linear drive beam with a Gaussian transverse profile $n_b \sim e^{-r^2/(2\sigma_r^2)}\Lambda(\xi)$, where $\Lambda(\xi)$ is a piecewise-linear function $\Lambda(\xi_i \le \xi \le \xi_{i+1}) = \Lambda_i + \frac{\Lambda_{i+1}-\Lambda_i}{\xi_{i+1}-\xi_i}(\xi - \xi_i)$ defined by p points of the discretized normalized charge per unit length Λ_i as shown in Fig. 1(a). The collection of the Λ_i is defined as the vector $\boldsymbol{\Lambda}$. The positions of the beginning and the end of the beam are fixed, and the currents at these two points are set to zero $\Lambda_1 = \Lambda_p = 0$.

In the 3D nonlinear regime, the decelerating (and accelerating) fields do not vary across the beam in the transverse direction. Therefore, we design the objective function as the variance of the on-axis accelerating or decelerating field E_z within the bunch length range. Additional constraints are imposed to fix the total charge Q of the drive beam by $\int \Lambda(\xi)d\xi = Q$, or to set upper and lower bounds for the current during the optimization. Under these conditions, the optimization problem can be formulated as the variance of the simulated on-axis $\boldsymbol{E}_z$ field within the bunch length range, where the grid point k of $\boldsymbol{E}_z$ is denoted as $[\boldsymbol{E}_z]_k$,

$$\begin{aligned} \underset{\Lambda_1,\ldots,\Lambda_p}{\text{minimize}} \quad & f(\boldsymbol{\Lambda}) = \sum_{k=1}^{q}([E_z]_k(\boldsymbol{\Lambda}) - \bar{\boldsymbol{E}}_z(\boldsymbol{\Lambda}))^2 \\ \text{subject to} \quad & \sum_{i=1}^{p-1} 0.5(\Lambda_{i+1} + \Lambda_i)\Delta\xi_i = Q, \\ & \boldsymbol{\Lambda}_l \le \boldsymbol{\Lambda} \le \boldsymbol{\Lambda}_u, \end{aligned} \tag{8}$$

where Q is the total charge and $\Delta\xi_i$ is the piecewise-linear bin length. The variables that need to be optimized are the beam currents Λ_i in each bin. We set the lower bound of $\boldsymbol{\Lambda}$ as $\boldsymbol{\Lambda}_l = 0$ since the loaded electron beam density should always be greater than 0, and we set the upper bound $\boldsymbol{\Lambda}_u$ to restrict the search domain to improve the efficiency of the algorithm. The scalar q counts the total grid points of the discretized vector $\boldsymbol{E}_z$, and p denotes the number of current bins we use.

The basic concept is to use a quasi-static PIC code to simulate $\boldsymbol{E}_z$ as an unknown function of the normalized beam charge per unit length $\boldsymbol{\Lambda}$ and then use the numerical optimizer to calculate the objective function $f(\boldsymbol{\Lambda})$ and update $\boldsymbol{\Lambda}$. A schematic of this loop is shown in Fig. 1(b). We use QuickPIC to predict the decelerating field or accelerating field $\boldsymbol{E}_z$ in the target beam. For each run, a piecewise-linear beam is initialized by using parameters produced by the optimization algorithm. Fixing other parameters, QuickPIC gives E_z on the grid,



(a) (b)

FIG. 1. (a) Initialization of a beam with a piecewise-linear profile. The optimization variables are set as beam charge (current) per unit length, $\boldsymbol{\Lambda}$. (b) Optimization workflow. An objective function is evaluated based on the output variables from the PIC simulation. This together with past evaluations is used to determine a new set of values for the optimization variables.

and the optimization algorithm produces new current profiles. We formulate the linear constraints as $\sum a_i\Lambda_i = Q$, $a_i = 0.5(\Delta\xi_i + \Delta\xi_{i-1})$ except the first and last components of Λ_i, which are set to 0. Λ_i is initialized as $Qa_i/\|\boldsymbol{a}\|_2^2$ to satisfy the linear constraint.

Solving Eq. (8) by applying a general-purpose optimization method will usually require passing values of f (as a function of $\boldsymbol{\Lambda}$). While doing so may be easier, it might be inefficient in terms of the number of times that the simulation must be queried. In addition, the objective function f may not be uniquely determined by $\boldsymbol{\Lambda}$, and thus, there may exist situations where for different current profiles $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$, the values of $f(\boldsymbol{\Lambda}_1)$ and $f(\boldsymbol{\Lambda}_2)$ are equal but the axial fields $E_z(\boldsymbol{\Lambda}_1)$ and $E_z(\boldsymbol{\Lambda}_2)$ differ considerably. In this situation, an optimization method that requires the user to return only the value of f will likely need additional evaluations to learn about how f varies with $\boldsymbol{\Lambda}$. In contrast, the optimization method that can access the information of $\boldsymbol{E}_z(\boldsymbol{\Lambda})$ and regards the objective function f as a function of E_z potentially needs fewer simulation evaluations to minimize f. This is highly desired when each evaluation of f requires a computationally expensive call of QuickPIC to produce $\boldsymbol{E}_z(\boldsymbol{\Lambda})$. In fact, our optimization method required only five to ten times the problem dimension p evaluations of QuickPIC to find parameters that produce beams with essentially flat E_z profiles. This is markedly fewer evaluations than required by other commonly employed optimization methods; evolutionary methods, such as those used in Refs. 17 and 18 often require 100 or more objective evaluations in their first generation.

If the gradient of each grid point of $\boldsymbol{E}_z$ with respect to the current $(\nabla_{\boldsymbol{\Lambda}}[E_z]_k)_i = \partial[E_z]_k/\partial\Lambda_i$ were available, it would seem reasonable to search for the descent of f in the reverse direction of gradient of f,

$$-\nabla_{\boldsymbol{\Lambda}} f(\boldsymbol{\Lambda}) = -2\sum_k \left[\left([E_z]_k(\boldsymbol{\Lambda}) - \frac{1}{q}\sum_j [E_z]_j(\boldsymbol{\Lambda})\right) \times \left(\nabla_{\boldsymbol{\Lambda}}[E_z]_k(\boldsymbol{\Lambda}) - \frac{1}{q}\sum_j \nabla_{\boldsymbol{\Lambda}}[E_z]_j(\boldsymbol{\Lambda})\right)\right], \tag{9}$$

or incorporate second-order knowledge using the Hessian matrix $\mathbf{H}$ of f,

$$\begin{aligned}\mathbf{H}(f(\boldsymbol{\Lambda})) = 2\sum_i \Bigg[&\left([E_z]_i(\boldsymbol{\Lambda}) - \frac{1}{q}\sum_j [E_z]_j(\boldsymbol{\Lambda})\right) \times \left(\nabla^2[E_z]_i(\boldsymbol{\Lambda}) - \frac{1}{q}\sum_j \nabla^2[E_z]_j(\boldsymbol{\Lambda})\right) \\ &+ \left(\nabla[E_z]_i(\boldsymbol{\Lambda}) - \frac{1}{q}\sum_j \nabla[E_z]_j(\boldsymbol{\Lambda})\right) \times \left(\nabla[E_z]_i(\boldsymbol{\Lambda}) - \frac{1}{q}\sum_j \nabla[E_z]_j(\boldsymbol{\Lambda})\right)^T\Bigg].\end{aligned} \tag{10}$$

Here, $(\mathbf{H}(f(\boldsymbol{\Lambda})))_{i,j} = \frac{\partial^2 f}{\partial\Lambda_i\partial\Lambda_j}$. Since gradients of E_z to current profile $\boldsymbol{\Lambda}$ are not available, the modified POUNDERs implementation instead builds local quadratic models of each of the q mappings $[E_z]_k$ around candidate points $\boldsymbol{\Lambda}_0$. We build these models by interpolating evaluations of $[E_z]_k$ in a neighborhood of $\boldsymbol{\Lambda}_0$. Because each evaluation of QuickPIC for a given set of parameters $\boldsymbol{\Lambda}$ returns all values of $[E_z]_k$, the information required to build these q models is easily obtained. These approximate gradients and Hessians are used to define a second-order model for the objective; this model is minimized in a neighborhood around the best-known set of parameters $\boldsymbol{\Lambda}$ to generate a candidate point to be evaluated. If this candidate point is an improvement, it becomes the new best-known set of parameters. Otherwise, the previous best-known set of points is kept, and the neighbor size is decreased. In either case, the models of each $[E_z]_i$ can be updated by using any new evaluations. For details, see Ref. 24.

III. OPTIMIZATION OF THE TRAILING BEAM AND COMPARISON WITH THEORY

As a benchmark test, we use our algorithm to obtain the optimal current profile of a trailing beam that flattens the accelerating field of a nonlinear plasma wake. The algorithm results are compared to the optimal profile determined from the nonlinear theory using multi-sheath model (Ref. 11). The simulations use normalized units so that each simulation is general and corresponds to a family of different plasma densities. When discussing results for absolute units in this paper, the plasma densities are assumed to be $n_p = 1.0 \times 10^{17}$ cm^{-3}, for which $k_p^{-1} = 16.83\ \mu$m.

A. Review of the multi-sheath model

For results obtained using the multi-sheath model, we follow the methodology outlined in Ref. 11. We constrain the variables of the multi-sheath model by integrating Eq. (7) to solve for n_1 in terms n_2, α_2, and α_1,

$$n_1 = \frac{1 - n_2\left(\alpha_2^2 + 2\alpha_2\alpha_1 + 2\alpha_2\right)}{(1+\alpha_1)^2 - 1}, \tag{11}$$

where $\alpha_1 \equiv \frac{\Delta_1}{r_b}$ and $\alpha_2 \equiv \frac{\Delta_2}{r_b}$. The other phenomenological sheath parameters $\Delta_1(r_b)$, $\Delta_2(r_b)$, and $n_2(r_b)$ are assumed to be functions of the channel radius $r_b(\xi)$ and can be determined empirically from PIC simulations. Here, we employ the same profiles used by Dalichaouch *et al.*[11] for the first plasma sheath with $\Delta_1 = \Delta_{10} + \varepsilon r_b$ and the second plasma sheath with $n_2 = n_{20}e^{-sr_b^2/r_m^2}$ and $\Delta_2 = \Delta_{20}$, where r_m is the maximum blowout radius and n_{20} is determined by calculating the limit of ψ, which approaches ψ_{min} when $r_b \to 0$,

$$n_{20} = \frac{2\psi_{\min}}{(\Delta_{10}+\Delta_{20})^2 \ln\left(1 + \frac{\Delta_{20}}{\Delta_{10}}\right)}. \tag{12}$$

In Ref. 11, it was shown that the differential equation of innermost particle trajectory can be written as

$$A'(r_b)\frac{d^2 r_b}{d\xi^2} + B'(r_b)r_b\left(\frac{dr_b}{d\xi}\right)^2 + C'(r_b)r_b = \frac{\Lambda(\xi)}{r_b}, \tag{13}$$

where the coefficients $A'(r_b)$, $B'(r_b)$, and $C'(r_b)$ are

$$\begin{aligned} A'(r_b) &= 1 + \left[\frac{1}{4} + \frac{\beta'}{2} + \frac{1}{8}r_b\frac{d\beta'}{dr_b}\right]r_b^2, \\ B'(r_b) &= \frac{1}{2} + \frac{3}{4}\beta' + \frac{3}{4}r_b\frac{d\beta'}{dr_b} + \frac{1}{8}r_b^2\frac{d^2\beta'}{dr_b^2}, \\ C'(r_b) &= \frac{1}{4}\left[1 + \frac{1}{\left(1 + \frac{\beta' r_b^2}{4}\right)^2}\right]. \end{aligned} \tag{14}$$

By setting $n_2 = 0$, Eq. (13) reduces to the equation of r_b trajectory described by the single-sheath model (Ref. 10), where $\beta' = \beta = \frac{(1+\alpha_1)^2 \ln(1+\alpha_1)^2}{(1+\alpha_1)^2-1} - 1$. After calculating the trajectory of r_b, the longitudinal electric field can be obtained by taking derivatives of ψ at the central axis,

$$E_z(\xi) = \frac{d}{d\xi}\psi_0(\xi) = D'(r_b) r_b \frac{dr_b}{d\xi}, \tag{15}$$

where $D'(r_b) \equiv \frac{1}{2} + \frac{\beta'}{2} + \frac{1}{4} r_b \frac{d\beta'}{dr_b}$. The derivative of E_z in the ξ direction can then be calculated by

$$\frac{dE_z}{d\xi} = D'(r_b) r_b \frac{d^2 r_b}{d\xi^2} + F'(r_b)\left(\frac{dr_b}{d\xi}\right)^2, \tag{16}$$

where $F'(r_b) \equiv D'(r_b) + \frac{3}{4} r_b \frac{d\beta'}{dr_b} + \frac{1}{4} r_b^2 \frac{d^2\beta'}{dr_b^2}$. In Ref. 11, it was shown that Eqs. (13)–(16) can also be used to design a beam current profile that produces a desired wakefield $f(\xi) = E_z(\xi_t \le \xi \le \xi_f)$. For a given function $f(\xi)$, Eq. (15) can be integrated to obtain the loaded bubble trajectory $\tilde{r}_b(\xi)$. Equations (13)–(16) can then be used to reverse engineer the current profile $\Lambda(\xi)$ of the beam load in terms of $\tilde{r}_b(\xi)$ and $f(\xi)$,

$$\Lambda(\xi) = C'\tilde{r}_b^2 + \left(\frac{B'}{D'^2} - \frac{A'F'}{D'^3\tilde{r}_b^2}\right) f(\xi)^2 + \left(\frac{A'}{D'}\right)\frac{df(\xi)}{d\xi}. \tag{17}$$

To load a constant wakefield, one can simply set $f(\xi) = -E_t$ and $df(\xi)/d\xi = 0$, where E_t is the desired constant loaded wakefield started at ξ_t. We note that the multi-sheath model fundamentally differs from analytic theory (Ref. 16), which relies on solving Eq. (4) in the ultrarelativistic limit (i.e., $r_b \gg 1$) where all sheath terms are neglected. In the ultrarelativistic limit, it was shown that the wakefield can be flattened by using an trailing beam with a trapezoidal current profile $\Lambda(\xi) = \sqrt{E_t^4 + \frac{R_b^4}{2^4}} - E_t(\xi - \xi_t)$, where R_b is the maximum blowout radius. To calculate the maximum loaded beam length, we can integrate ψ from ξ_t to ξ_f to get $\psi_{\min} - \psi_t = \int_{\xi_t}^{\xi_f} E_z d\xi$, where at ξ_f, we assume $r_b = 0$, and ψ becomes ψ_{min}. Assuming $r_b \gg 1$ at ξ_t in the ultrarelativistic limit, we can ignore β' and obtain $\psi(0,\xi) = r_b^2/4$ from Eq. (6). This model for the wake potential exhibits the same asymptotic behavior as the single-sheath model,[16] where $\psi_{\min} \to 0$ as $r_b \to 0$, and the maximum loaded length becomes $\Delta\xi = \frac{r_t^2}{4E_t}$. On the other hand, the multi-sheath model is valid for all r_t and predicts a longer loaded length $\Delta\xi = \frac{r_t^2}{4E_t} + \frac{\beta'(r_t) r_t^2/4 - \psi_{\min}}{E_t}$.

B. Comparison of the optimized trailing beam current with the profile obtained from the multi-sheath model

To check that the optimization procedure agrees with the multi-sheath model, we consider the case studied both by Tzoufras *et al.*[16] and by Dalichaouch *et al.*,[11] in which a nonlinear plasma wakefield is excited by a drive beam with a bi-Gaussian density distribution $n_b(r,\xi) = \{N_b/[(2\pi)^{3/2}\sigma_r^2\sigma_z]\} e^{-r^2/(2\sigma_r^2)} e^{-\xi^2/(2\sigma_z^2)}$, where N_b is the number of particles in the beam and $n_{b0} \equiv N_b/[(2\pi)^{3/2}\sigma_r^2\sigma_z]$. The normalized beam parameters are $\sigma_r = 0.5$, $\sigma_z = 1.414$, $N_b = 139$ (note that N_b is normalized to n_p/k_p^3). The normalized charge per unit length is $\Lambda = \int dr r n_b = n_{b0}\sigma_r^2 e^{-\xi^2/(2\sigma_z^2)}$. The peak charge per unit length Λ of the driver is $\Lambda_0 = 6.24$, and drive beam energy is $\gamma_b = 20\,000$ with 0 energy spread. As seen in Fig. 2(a), the drive beam excites a plasma wakefield with maximum blowout radius $r_m \simeq 2\Lambda_0^{1/2} \simeq 5.0$. We aim to load the trailing beam starting at $\xi_t = 8.27$, where the bubble radius is $r_t \simeq 3.91$, and the wakefield is $E_t \simeq 1.75$. To understand how much charge can be loaded for parameters of interest, we assume a plasma density of $n_p = 1.0 \times 10^{17}\,\mathrm{cm}^{-3}$. For the normalized parameters given above, $k_p^{-1} = 16.83\,\mu$m, $\sigma_z = 23.8\,\mu$m, $\sigma_r = 8.4\,\mu$m, the charge of the drive beam is 10.6 nC, the peak current of the drive beam is 53.3 kA, and the loaded wakefield is $E_t = 53.2\,\mathrm{GV/m}$. From Tzoufras *et al.*[16] $\Delta\xi_{tr} = \frac{r_t^2}{4E_t} \simeq 2.18$, while the multi-sheath model[11] predicts a longer loading length of $\Delta\xi_{tr} \simeq 3.09$ by setting $\psi_{\min} = -0.9$ and setting the integration parameters in the multi-sheath model as $\Delta_1 = 0.825 + 0.05 r_b$, $\Delta_2 = 3$, and $s = 3$.

We obtain the optimized trailing beam profile both from the multi-sheath model and from the optimization algorithm by initializing the trailing beam with a longitudinally piecewise-linear profile from

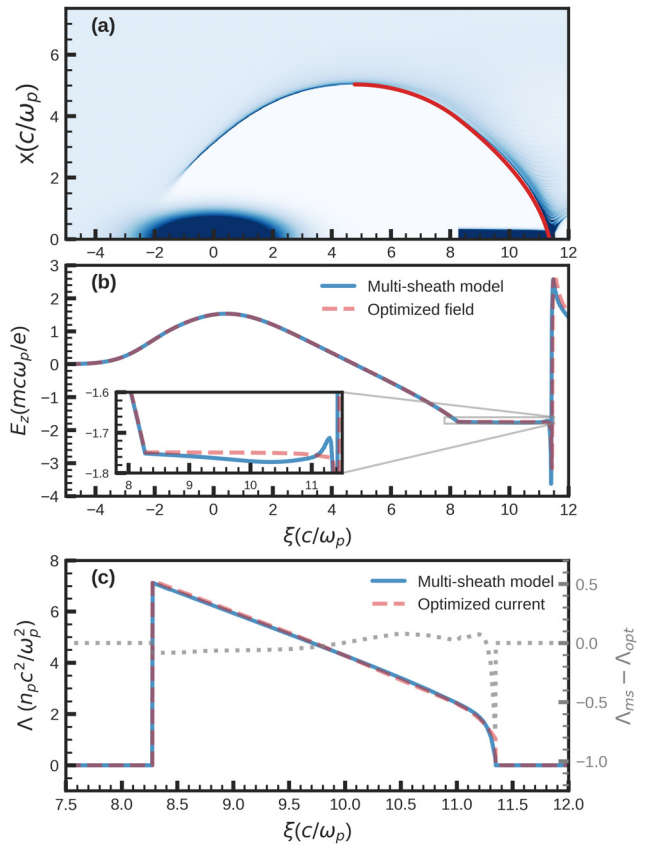


FIG. 2. (a) Two-dimensional cut of the three-dimensional data for the charge density of plasma electrons excited by a bi-Gaussian drive beam and a trailing beam. The charge density of the drive beam and that of the trailing beam are also shown. (b) Comparison of the PIC simulation results for the on-axis wakefield using the optimized profile and the theoretically predicted profiles shown in (c). (c) Comparison of beam profile calculated by the multi-sheath model (blue) and by the optimization procedure using POUNDERs (dashed red). The difference of these profiles is shown as the gray-dotted line.

$\xi_t = 8.274$ to $\xi = 11.3498$. For the multi-sheath model, the loaded bubble trajectory is numerically integrated from Eq. (13) for the desired electric field profile $f(\xi > \xi_t) = E_t$ and $df/d\xi = 0$. The current profile that produces this wakefield is then calculated directly from Eq. (17). For the optimization algorithm, we set 22 bins[28] to resolve the profile. We ran QuickPIC for a single 3D time step to get the acceleration field within the trailing beam. At each iteration within the numerical method, we calculated the optimization objective as Eq. (8) and used the method discussed in Sec. II to optimize the trailing beam profile. To reduce the computational expense, we first set a resolution of $0.11 \times 0.11 \times 0.03$ in the QuickPIC simulation to obtain a low-fidelity optimization result. We then increased the resolution of the QuickPIC simulation up to as fine as $0.03 \times 0.03 \times 0.009$ and used the previously obtained low-fidelity solution as the initialization to further refine the optimal profile.

The optimized accelerating field (wakefield) and optimized current profile (Λ_{opt}) are shown in Figs. 2(b) and 2(c), respectively, as dashed-red lines. For comparison, the accelerating field and current profiles predicted from the multi-sheath model are shown as solid blue lines. One can also see in Figs. 2(b) and 2(c) that the on-axis wakefield is flattened at the location of the trailing beam for the profile predicted by the multi-sheath model, and the optimized current profile is slightly different from the profile predicted by the multi-sheath model (Λ_{ms}). To make these subtle differences clearer, we also plot $\Lambda_{opt} - \Lambda_{ms}$ in Fig. 2(c). In the blowout regime, the accelerating field does not vary across the transverse cross section of the beams,[10,12] so we can assume the whole beam feels the same accelerating field when the on-axis E_z field is flattened. In the inset of Fig. 2(b), we show a zoomed-in plot of the flattened accelerating field showing that the current profiles obtained from the optimization method provide flatter accelerating fields than those obtained from the theoretical framework and that the theoretical framework works well. On this scale, we can see that the relative variation of the E_z field within the trailing beam, defined by $\sigma_{E_z}/\overline{E_z}$, where σ_{E_z} is defined as the standard deviation of E_z and $\overline{E_z}$ is defined as the average of the of E_z, is less than 1% in both cases. Analysis of the data shows that the relative variation of E_z is 0.1% for the optimized trailing beam profile and 0.6% for a simulation with the current profile based on the theoretical prediction. Therefore, the optimization algorithm improves the variation of the accelerating field with a small modification to the trailing beam current profile. The excellent agreement between the predicted and algorithm-obtained optimal current profiles confirm both the theory and the optimization method. We also note that the relative differences between the current profiles obtained through optimization and theory may differ slightly depending on strength of the driver (Λ_0) and the location of the head of the trailing beam.

IV. OPTIMIZATION OF THE DRIVE BEAM CURRENT PROFILE BY FIXING THE TOTAL CHARGE

A. Optimized driver current profile for different length

Now that we have verified that the optimization procedure works, we next turn our attention to the drive beam. To optimize the drive beam current profile, we initialize a piecewise-linear monoenergetic drive beam with $\gamma_b = 55\,773$, $\sigma_r = 0.1$. We assume a fixed total particle number $2\pi n_p (k_p^{-1})^3 \int_0^L \Lambda(\xi) d\xi$, which corresponds to $N_b = 65.55 \times n_p (k_p^{-1})^3 = 3.125 \times 10^{10}$ or 5 nC for a plasma density of $n_p = 1.0 \times 10^{17}\,\mathrm{cm}^{-3}$. Constraining the charge to be fixed is a linear constraint to the current profile of the driver in Eq. (8). Without this constraint, the optimized current tends to converge to all zero values, which is indeed a trivial (but uninteresting) solution for flattening the decelerating field. In this example, we set the transverse resolution to be 0.12×0.12 in the QuickPIC simulations. To fully resolve the current profile and reduce the number of optimization variables, we use nonuniform bins where the bin size equals the simulation resolution, $\Delta\xi$, at the beginning and the end of the beam, and is set to a larger bin size for the middle of the beam. We also fix the density for the first and last bin to be 0 and examine several cases with different drive beam lengths $L \in \{4, 8, 11, 14\}$. We choose an axial simulation window size of 5.5, 8.5, 11.5, and 14.5, respectively, for the drive beam lengths of $L \in \{4, 8, 11, 14\}$. The number of grids chosen for each simulation is fixed as $512 \times 512 \times 512$. The resulting optimized current profiles are shown in Fig. 3(a). When using several bins with a size equal to $\Delta\xi$ at the beginning and end of the beam, we get nearly a "perfectly" flattened decelerating field, as seen in Figs. 3(b) and 3(c). Because we are using direct assignment of the piecewise-linear density on the grid when doing simulation (rather than depositing the current from particles), it is possible to have an instantaneous rise of the density in a single grid. Thus, the wakefield can directly rise to the desired constant value at the head of the beam in a length given by the chosen simulation resolution.

We found that in all cases, the optimized current profile includes a narrow precursor at the head of the drive bunch. The precursor is always found to have a length of one grid cell no matter the resolution in ξ direction, so one can assume that the optimized precursor is a delta function. This was also predicted by 1D linear theory. Although we are operating in a nonlinear regime, an impulse response still is the most efficient method to create the most rapid rise in E_z. Furthermore, as seen in the inset of Fig. 3(a), the amount of charge in the precursor is largest for the shorter beam sizes. This is due to the fact that the decelerating field is smaller for longer beams so the jump in the E_z field required to be generated by the precursor is less. Under the current resolution, we calculated the normalized total charge Q of the optimized precursor when changing the length of the drive beam, then divided the charge Q of each optimized beam by the average decelerating field E_z, where E_z/Q is close to a constant $68.2 \pm 4.5\,\mathrm{GV/(m\,nC)}$.

One also can see in Fig. 3(a) that for the shortest beam length, $L = 4$ (gray line), the optimal current profile discovered by the algorithm has a gradually increasing slope at the front of the beam (i.e., a nonlinear transition) immediately after the precursor (from $\xi = 0.5$ to $\xi = 1.5$), which deviates from the prediction of the ultrarelativistic limit of the theory of Lu *et al.*[10] as described earlier. One can also see in Fig. 3(a) that as the length of the drive bunch is increased while keeping the charge fixed, the profile of the beam after the precursor becomes almost a perfect linear ramp. As seen in Fig. 3(b), both the decelerating and accelerating fields are reduced as the length L is increased. To better visualize the decelerating field, in Fig. 3(c), we plot only the positive region for E_z. As can be seen, the decelerating field is nearly flat within the location of the drive beam, and its value decreases as L is increased.

B. Transformer ratios of optimized driver current profiles for different beam lengths

In nonlinear wakes, there is a large negative electron density and a wakefield spike at the rear of the first bubble. Since this spike occurs

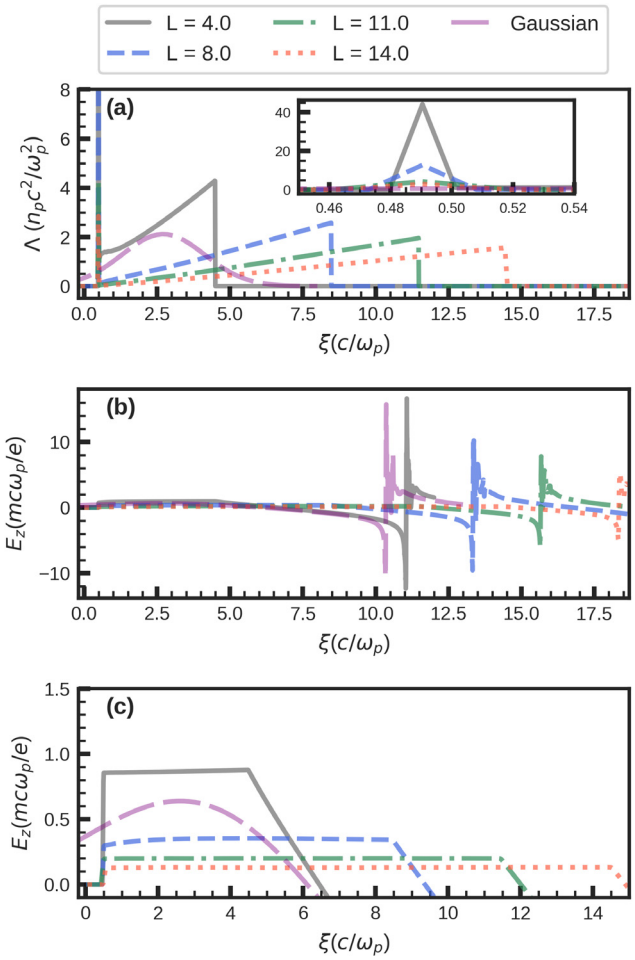


FIG. 3. (a) Optimized current profiles obtained when fixing the total charge of the drive beam as 5 nC, and varying the beam length from 4.0 (gray), 8.0 (dashed blue), 11.0 (dash-dotted green), and 14.0 (dotted red). The bin size $\Delta\xi$ of the piecewise-linear profile is set to the simulation grid size at the beginning and end of the beam and is larger in between. The current profile for a Gaussian with charge of 5 nC and pulse length of $\sigma = 2$ is also presented (long dashed-purple). (b) On-axis electric field $E_z(\xi)$ using the optimized current profiles and for the Gaussian profile in (a). (c) On-axis electric field $E_z(\xi)$ at the head of the drive beam for each case in (b).

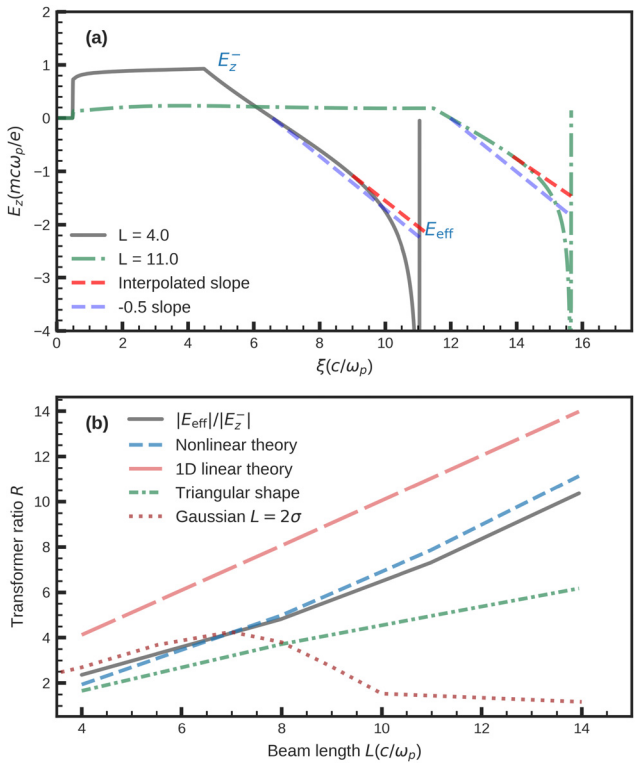


FIG. 4. (a) Optimized on-axis electric field $E_z(\xi)$ for $L = 4.0$ (gray) and $L = 11.0$ (dash-dotted green). The two methods used to define $E_{\rm eff}$ and $|E_z^-|$ are also shown. The dashed blue line starts at $E_z = 0$ at the maximum blowout radius with a slope of $\partial E_z/\partial\xi = -1/2$, and the dashed red line has a slope extrapolated from the simulated E_z field where it has a nearly linear slope. $E_{\rm eff}$ is defined as the wakefield where the dashed red line and simulated E_z field meet at the back of the bubble. (b) Simulated transformer ratio $R = |E_{\rm eff}|/|E_z^-|$ as a function of the drive beam lengths $L \in \{4, 8, 11, 14\}$ for the optimized beam profile (gray), triangular shape longitudinal profile (dash-dotted green), and Gaussian shape (dotted red) with a pulse length $\sigma = L/2$. The theoretical transformer ratio $R = \frac{L}{\sqrt{\Lambda_0}}$ (dashed blue) is calculated by using the value of the current Λ_0 at the rear of the beam in the simulation; the transformer ratio for linear theory (long-dashed red) is calculated by $R = \sqrt{(1 + L^2)}$.

in a very small region, the peak accelerating field is generally not a good figure of merit when characterizing the transformer ratio. We, therefore, define the effective wakefield $E_{\rm eff}$ by extrapolating the part of the wakefield E_z to the rear of the wake [Fig. 4(a)], assuming that it has a slope approximately given by $\partial E_z/\partial\xi = -1/2$. We then define the transformer ratio as the ratio of the effective wakefield $|E_{\rm eff}|$ to the maximum decelerating field felt by the drive beam, $R = |E_{\rm eff}|/|E_z^-|$. In Fig. 4(b), we plot the transformer ratio from simulations with the optimized beam profile (gray), a triangular-shaped profile (dash-dotted green), and Gaussian shape (dotted red) for the four different bunch lengths $L \in \{4, 8, 11, 14\}$. For the Gaussian profiles, we use $L = 2\sigma_z$. The transformer ratio R of the optimized current profile obtained from simulation is also compared with the linear theory using $R = \sqrt{(1 + L^2)}$ (long-dashed line), and with the nonlinear theory (dashed blue line) using $|E_z^+|/|E_z^-| \approx \sqrt{\Lambda_0}/(\frac{\Lambda_0}{L}) = \frac{L}{\sqrt{\Lambda_0}}$, where Λ_0 is the normalized charge per unit length at the end of the optimized beam. We see in Fig. 4(a) that for longer beam lengths, both the peak decelerating field and effective accelerating field become smaller for a decreased peak current. As can be seen in Fig. 4(b), the transformer ratio increases with bunch length. Thus, the accelerating field decreases at a lower rate than does the decelerating field.

To understand the reason that the optimized transformer ratio deviates from the theoretical estimate, we show in Figs. 5 and 6 more detailed analysis for the two cases that have lengths $L = 4$ and $L = 11$. The equation for evolution of the blowout radius r_b, Eq. (13), is integrated both forward to a larger ξ and backward to the front in ξ for both bunch lengths starting from the initial conditions at $\xi = 2.0$. The theoretical trajectories for r_b are shown in Figs. 5(a) and 6(a) where

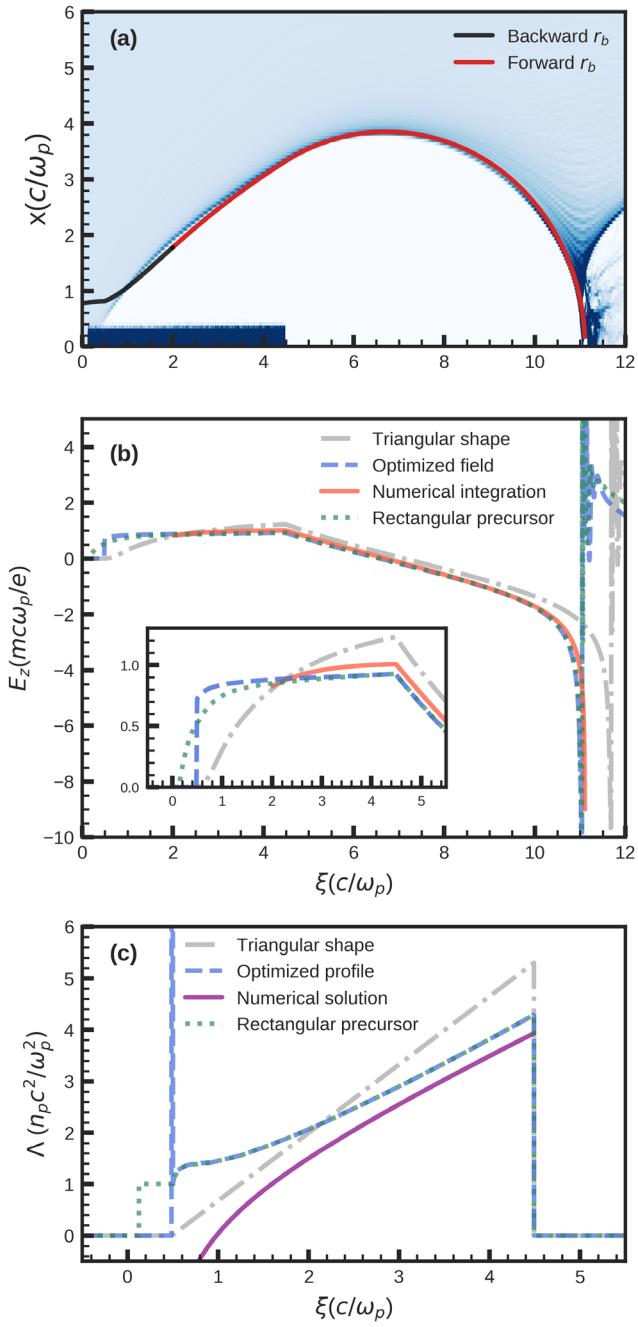


FIG. 5. (a) Charge density of plasma electrons excited by the optimized drive beam with length $L = 4.0$. The red line represents the bubble radius integrated forward from $\xi = 2.0$ by nonlinear theory including both single-sheath (until r_{bmax} is reached) and multi-sheath (after r_{bmax} is reached) models, and the black line is obtained by integrating r_b backward from $\xi = 2.0$ with the single-sheath model. (b) On-axis accelerating field $E_z(\xi)$ for a triangular profile (dash-dotted gray), optimized profile (dashed blue), and rectangular precursor extension of optimized profile (dotted green). The red line represents the predicted E_z field for $\xi > 2$ from Eq. (15) given the predicted r_b shown as the red line in (a). (c) Profiles that produce $E_z(\xi)$ shown in (b), and a profile (purple line) calculated from Eq. (17) with the bubble radius r_t and a constant decelerating field $f(\xi) = E_t$ at the ξ_t where the beam ends.

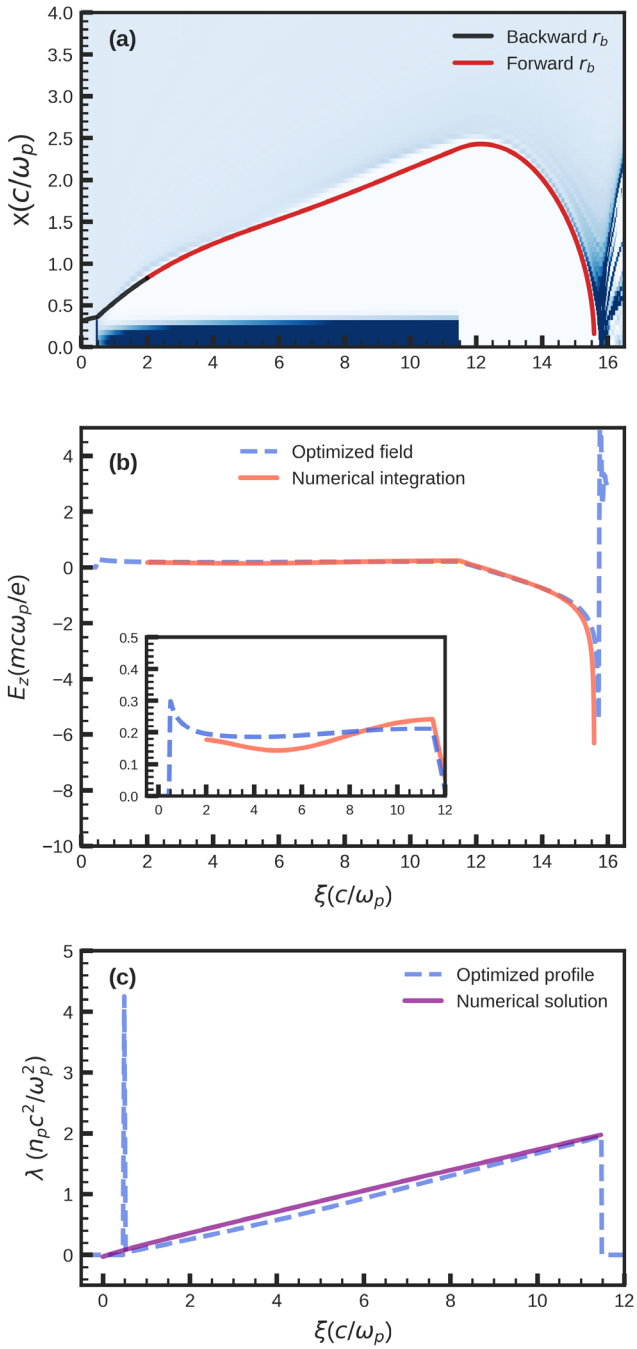


FIG. 6. (a) Charge density of plasma electrons excited by the optimized drive beam with length $L = 11.0$. The red line represents the bubble radius integrated forward (increasing ξ) from $\xi = 2.0$ by the nonlinear theory including both single-sheath and multi-sheath models, and the black line is obtained by integrating r_b backward from $\xi = 2.0$ with the single-sheath model. (b) On-axis accelerating field $E_z(\xi)$ for the optimized trailing beam profile (dash-dotted blue). The red line represents the predicted E_z field for $\xi > 2$ from Eq. (15) given the predicted r_b shown as the red line in (a). (c) Profiles that produce the $E_z(\xi)$ shown in (b), and a profile (purple line) calculated from Eq. (17) with the bubble radius r_t and a constant decelerating field $f(\xi) = E_t$ at the ξ_t where the beam ends.

the red line corresponds to integrating forward to a larger ξ and the black line corresponds to integrating backward in ξ. We use the single-sheath model for values of ξ before r_b reaches its maximum value, r_{max}. However, the wake potential for the first half bubble is positive definite so only one sheath is required. We then use the multi-sheath model for ξ after r_b reaches its maximum for an accurate description of the negative wake potential at back of the bubble. For both models, we set $\Delta_1 = 0.825 + 0.05r_b$ for $L = 4.0$, $\Delta_1 = 1.0 + 0.1r_b$ for $L = 11.0$, respectively. We set $\Delta_2 = 3$ and $s = 3$ when using the multi-sheath model. The theoretically obtained r_b agrees well with the QuickPIC simulation, and the predicted E_z field for $\xi > 2$ from Eq. (15) given the predicted r_b also agrees well except for some small oscillations on a smaller scale [red line in Figs. 5(b) and 6(b)].

For $L = 4$, the transformer ratio for the optimized current is higher than that estimated from theory. The difference arises because the estimation of Lu *et al.* for r_{max} is based on assuming an adiabatic response for r_b from which it follows that the maximum E_z will be reached when the current profile has reached its maximum (at the end of the drive beam). As can be seen in Fig. 5(a), however, for the $L = 4$ case, r_b continues to increase after the drive beam so that r_{max} is higher than expected, resulting in a higher effective wakefield E_{eff} than expected. Furthermore, the optimized profile also gives a lower decelerating field E_z^- [Fig. 5(c)]. These two factors contribute to the slightly higher transformer ratio obtained from optimization.

For longer pulse $L = 11.0$ (Fig. 6), the optimized transformer ratio is lower than the theoretical estimates. This occurs because as Λ_0 of the drive beam decreases r_{max} also decreases, and the contribution of β' can no longer be neglected. In this case, the slope of E_z deviates from the nonlinear limit $\partial E_z/\partial\xi = -1/2$ that is used to estimate E_{eff} in Fig. 4.

C. Discussion of why nonlinear theory cannot get the precursor

In reality, we cannot create a precursor with an infinitesimal length, nor can we ramp the drive beam current directly from a peak value to 0 in an infinitesimal distance. Thus, it is useful to consider constraints on the designs where the beam parameters are more realizable. We consider the $L = 4$ case and constrain the precursor into a rectangular shape (referred to as a doorstep[9]) with a length of 0.37 [dotted green line in Fig. 5(c)] and with same charge as in the "delta" function precursor [see the gray line in the inset of Fig. 3(a)]. The new shape has a smoother transition at the beginning of the beam, as shown in Fig. 5(c). The transformer ratio does not change; and, as seen in Fig. 5(b), the decelerating field of the drive beam is nearly constant except for a smooth transition region of length of approximately k_p^{-1} at the beginning of the beam. For comparison, the results for a linear triangular beam current (with the same total charge) are shown [dash-dotted gray line in Fig. 5(c)], which is the optimal shape predicted from ultrarelativistic nonlinear theory.[13] As we have mentioned, we integrate r_b from $\xi = 2.0$ both forward to a larger ξ and backward to the front with Eq. (14) by setting $\beta' = \beta$ and using optimized beam current $\Lambda(\xi)$. In Fig. 5(a), we can see that the integrated r_b trajectory agrees well with the simulation result when moving forward to the rear of the bubble. On the other hand, when moving backward to the front of the wake, the prediction of the single-sheath model deviates from the simulation result. The lack of agreement in this region arises because there is not a well-defined bubble radius at the front of the bubble because of particle crossing, so that the r_b trajectory we get from the nonlinear theory is not the innermost electron sheath near the front of the bubble. This explains why the numerical solution for $\Lambda(\xi)$ with Eq. (17) does not recover the precursor [purple line in Figs. 5(c) and 6(c)]. The same comparison of the optimized profile between theory and simulation is shown in Fig. 6(c) for the $L = 11$ case. The agreement is better because the precursor has less charge.

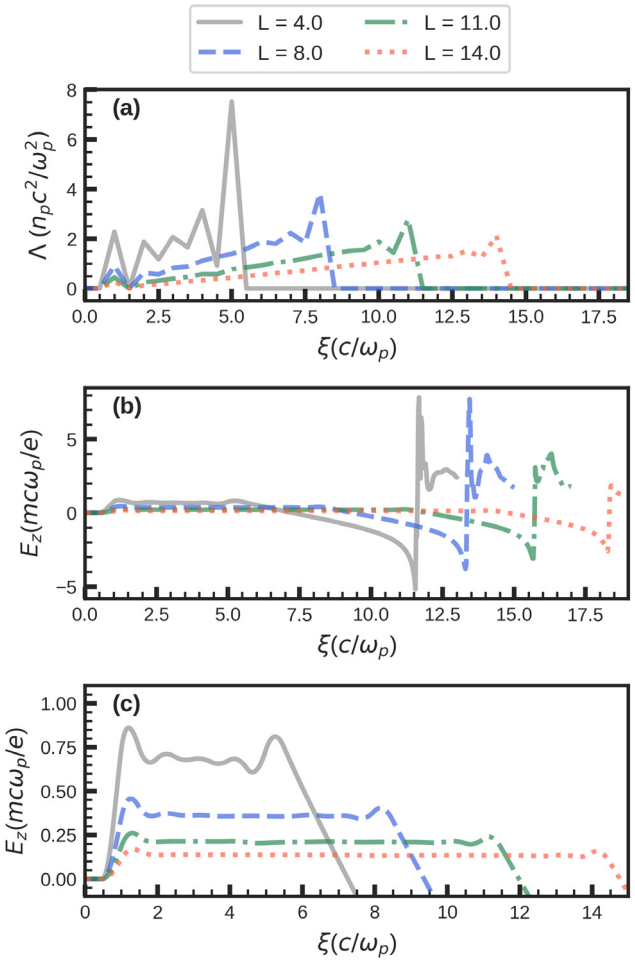


FIG. 7. (a) Optimized current profiles for larger bin sizes when fixing the total charge of drive beam as 5 nC and varying the beam length over 4.0 (gray), 8.0 (dashed blue), 11.0 (dash-dotted green), and 14.0 (dotted red). Here, $k_p\Delta\xi$ is fixed to be 0.5. (b) On-axis electric field $E_z(\xi)$ using the optimized current profile in (a). (c) On-axis electric field $E_z(\xi)$ at the head of the drive beam.

D. Optimized current profiles for larger bin sizes

When generating current profiles in experiments, providing precise control of the current profile is difficult. We, therefore, investigate how the optimized profiles change when the bin sizes are increased. Figure 7 presents results for the varying bunch lengths $L \in \{4, 8, 11, 14\}$ where the bin sizes are now uniformly set to be $\Delta\xi = 0.5$, while other parameters are kept the same as in Fig. 3. In each case, the optimized current profile includes a precursor with a length of $\Delta\xi$

followed by a quasi-linear ramp, which resembles the solution in Fig. 3. For the larger bin sizes, however, the "linear" ramp now has oscillations that get smaller for the large bunch sizes. For each case shown in Fig. 7(c), the wakefield now smoothly ramps up at the head of the drive beam and then overshoots the average decelerating field because of the finite width of the precursor. The small oscillation in the current profile at the rear of each beam is now present to compensate for the fact that the beam current cannot fall sharply to 0. The optimal shapes found for the coarser bin sizes are close to the beam current profiles used in an experiment that achieved a high transformer ratio.[14]

V. A LONG-RANGE SIMULATION USING THE OPTIMIZED DRIVE BEAM AND TRAILING BEAM

To confirm that the optimized drive and trailing beam profiles can be used to sustain efficient acceleration over pump depletion distances, we use QPAD,[29] a highly efficient quasi-static code based on the azimuthal Fourier decomposition method. We initialize the plasma density as $n_p = 1.0 \times 10^{17}$ cm^{-3}, and $k_p^{-1} = 16.83\ \mu$m. We set up azimuthally symmetric drive and trailing beams and keep the $m = 0, 1, 2$ modes. Instabilities can, thus, grow from noise. We do not address instabilities and tolerance to initial offsets and spot size asymmetries in this paper. To prevent head erosion and also to ensure that the plasma electrons are fully blown out, we initialize the precursor with a normalized emittance ε_N of 0.0167 and initialize the main body of the driver with a normalized emittance 1.67. The trailing beam also has a normalized emittance of 1.67. Both the drive and trailing beams are initialized with a Lorentz factor $\gamma = 55\,773$ and matched spot size $\sigma_r = (2\varepsilon_N^2/\gamma)^{1/4}$. We use a drive beam with $L = 4.0$ and an optimized current profile (from Fig. 3) and load a trailing beam at $x_t = 8.68$ and $E_t = 0.91$ based on the optimization algorithm [see Fig. 8(a)]. We note that at this location, the bubble radius is $r_t = 3.31$. The decelerating field on the drive beam is flattened to 0.91, so the loaded transformer ratio is $R \simeq 1.0$. From the analytic theory shown by Tzoufras *et al.*, the maximum loaded beam length is $\Delta\xi_{tr} = \frac{r_t^2}{4E_t} \simeq 3.0$. The multi-sheath model predicts a much longer loaded beam length $\Delta\xi_{tr} = \frac{r_t^2}{4E_t} + \frac{\beta'(r_t)r_t^2/4 - \psi_{\min}}{E_t} \simeq 4.8$ assuming $\psi_{\min} = -1$. We initialize 21 bins of a piecewise-linear profile and load a trailing beam with a length of 3.8. The current profile is then obtained by the described optimization method. The charge in the trailing beam is 4.3 nC, which corresponds to 86% of the drive beam charge. In the simulation, we use a fine resolution $dr = 0.015$ and $d\xi = 0.0058$ to fully resolve the matched spot size and very short duration of the low emittance precursor. The optimized wakefield is shown in Fig. 8(b).

We initialize the energy spread of both the drive beam and the trailing beam to 0 to simplify the analysis. The trailing beam is accelerated for 1 m, at which point some drive beam particles have nearly pump depleted to energies around several MeV. The average particle energy of the drive beam and trailing beam is shown in Fig. 9. The energy of the trailing beam nearly doubles and sustains a stable acceleration with the acceleration efficiency of 84% from the drive beam to the trailing beam. The projected energy spread grows from 0 to less than 0.7%.

VI. CONCLUSION

PWFA has emerged as a promising candidate for the accelerator technology for a future linear collider and/or light source. For the linear collider application, the energy transfer efficiency from the drive beam to the wake and from the wake to the trailing beam must be efficient, and the energy spread of the trailing bunch should be kept low. One way to achieve this is to use longitudinally shaped bunches. In the linear regime, there is an analytical formalism to obtain optimal shapes, for which the transformer ratio is maximized. In the nonlinear blowout regime, however, the theoretical framework is not as well defined for the driver. We, thus, use a novel optimization method to efficiently find optimized drive beam profiles and trailing beam profiles for PWFA. We parameterize the beam currents as a piecewise-linear longitudinal profile with N bins and define optimization objectives. We use the particle-in-cell code QuickPIC to evaluate the objective, and we use a modified version of POUNDERs (a derivative-free optimization method) to determine a new longitudinal profile. The optimization method required very few evaluations of QuickPIC to identify the best set of parameters. The algorithm is shown to converge quickly and finds trailing beam shapes that are similar to those calculated by a recent multi-sheath model for the nonlinear wakefield.

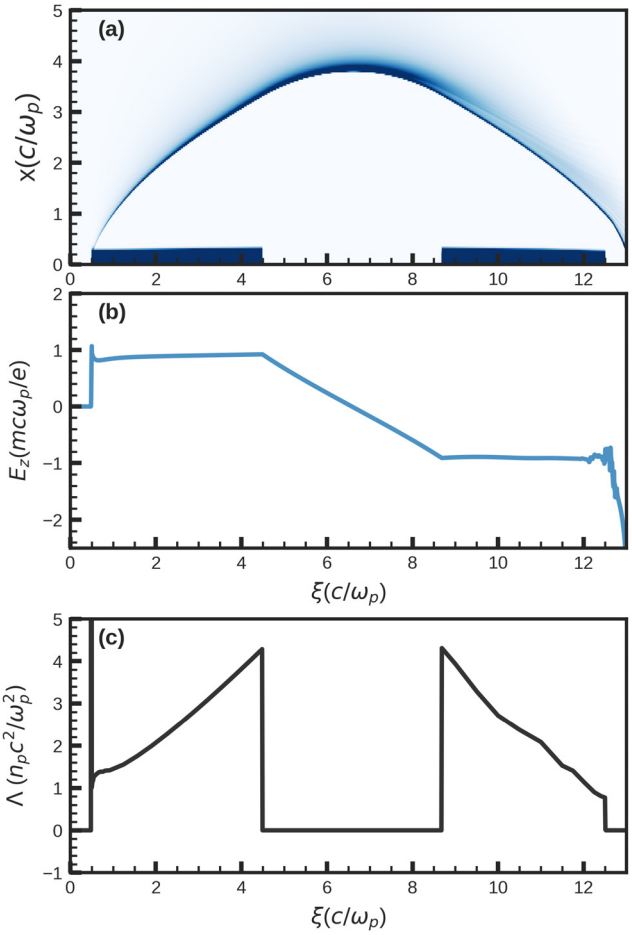


FIG. 8. (a) Charge density of plasma electrons excited by the optimized drive beam and the optimized trailing beam. (b) On-axis accelerating field $E_z(\xi)$ given the optimized current profile. (c) The optimized beam density per unit length $\Lambda(\xi)$ for the drive and trailing beams.

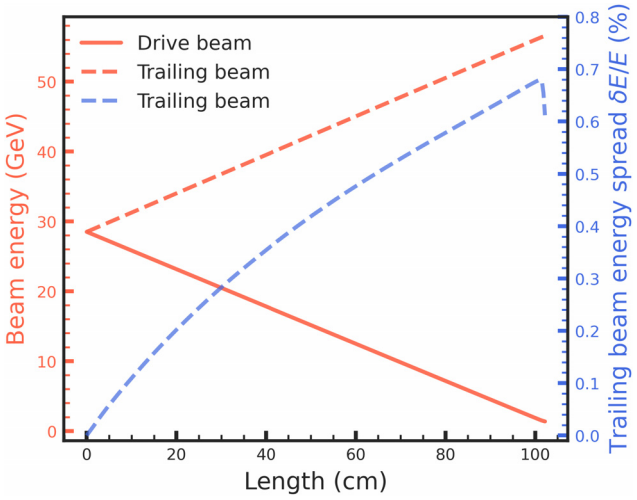


FIG. 9. Drive beam energy depletion (red) and trailing beam energy growth (red dashed). The growth of the energy spread of the trailing beam is shown in blue dashed line.

All results were produced when limiting the optimization method objective evaluations to ten times N, the number of bins. We also found that even in the nonlinear regime, current profiles of the drive beam that optimized the efficiency for a fixed charge also provided the highest transformer ratio. In this study, we keep the ions fixed, which will not be the case for the very narrow matched spot sizes of trailing beams for linear collider parameters.[27] The optimization procedure described here should also work well for finding the optimal beam shapes, including the transverse shape, when ion motion is included. There are numerous other applications for the described method within the field of plasma-based acceleration. These include finding optimal shapes for positron acceleration (other optimization methods have already been applied to aspects for this problem[19]), finding optimal density profiles for matching sections and finding optimal spatial-temporal couplings for drive beams[30,31] (both lasers and particle beams) that provide the highest quality self-inject beam or efficiency. Another area of interest is to determine optimal conditions when there is statistical uncertainty in the problem, such as shot-to-shot changes to the density profile and beam distribution functions.

ACKNOWLEDGMENTS

This material is based upon work supported by CAMPA and ComPASS-4, projects of the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research and Office of High Energy Physics, Scientific Discovery through Advanced Computing (SciDAC) program under Contract Nos. DE-AC02-06CH11357 and DE-AC02-05CH11231, FNAL subcontract 644405 and DOE HEP Grant No. DE SC0010064, NSF Award No. 2108970, and the National Natural Science Foundation of China (NSFC) Grant No. 12075030. We gratefully acknowledge the computing resources provided on Bebop, a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Qianqian Su: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Stefan Wild:** Funding acquisition (equal); Resources (supporting); Software (supporting); Supervision (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). **Warren Bicknell Mori:** Funding acquisition (lead); Project administration (lead); Supervision (lead); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Jeffrey M. Larson:** Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Thamine N. Dalichaouch:** Formal analysis (equal); Methodology (equal); Supervision (equal); Visualization (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). **Fei Li:** Conceptualization (supporting); Formal analysis (supporting); Software (equal); Validation (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). **Weiming An:** Conceptualization (equal); Methodology (supporting); Software (supporting); Supervision (supporting). **Lance Hildebrand:** Investigation (supporting); Resources (supporting); Validation (supporting). **Yujian Zhao:** Investigation (supporting); Resources (supporting). **Viktor Decyk:** Software (supporting); Supervision (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). **E. Paulo Alves:** Methodology (supporting); Resources (supporting); Supervision (supporting).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]P. Chen, J. M. Dawson, R. W. Huff, and T. Katsouleas, "Acceleration of electrons by the interaction of a bunched electron beam with a plasma," Phys. Rev. Lett. **54**, 693–696 (1985).

[2]M. J. Hogan, R. Assmann, F.-J. Decker, R. Iverson, P. Raimondi, S. Rokni, R. H. Siemann, D. Walz, D. Whittum, B. Blue, C. E. Clayton, E. Dodd, R. Hemker, C. Joshi, K. A. Marsh, W. B. Mori, S. Wang, T. Katsouleas, S. Lee, P. Muggli, P. Catravas, S. Chattopadhyay, E. Esarey, and W. P. Leemans, "E-157: A 1.4-m-long plasma wake field acceleration experiment using a 30 GeV electron beam from the Stanford Linear Accelerator Center Linac," Phys. Plasmas **7**, 2241–2248 (2000).

[3]M. J. Hogan, C. D. Barnes, C. E. Clayton, F. J. Decker, S. Deng, P. Emma, C. Huang, R. H. Iverson, D. K. Johnson, C. Joshi, T. Katsouleas, P. Krejcik, W. Lu, K. A. Marsh, W. B. Mori, P. Muggli, C. L. O'Connell, E. Oz, R. H. Siemann, and D. Walz, "Multi-GeV energy gain in a plasma-wakefield accelerator," Phys. Rev. Lett. **95**, 054802 (2005).

[4]I. Blumenfeld, C. E. Clayton, F.-J. Decker, M. J. Hogan, C. Huang, R. Ischebeck, R. Iverson, C. Joshi, T. Katsouleas, N. Kirby, W. Lu, K. A. Marsh, W. B. Mori, P. Muggli, E. Oz, R. H. Siemann, D. Walz, and M. Zhou, "Energy doubling of 42 GeV electrons in a metre-scale plasma wakefield accelerator," Nature **445**, 741–744 (2007).

[5]M. Litos, E. Adli, W. An, C. I. Clarke, C. E. Clayton, S. Corde, J. P. Delahaye, R. J. England, A. S. Fisher, J. Frederico, S. Gessner, S. Z. Green, M. J. Hogan, C. Joshi, W. Lu, K. A. Marsh, W. B. Mori, P. Muggli, N. Vafaei-Najafabadi, D. Walz, G. White, Z. Wu, V. Yakimenko, and G. Yocky, "High-efficiency acceleration of an electron beam in a plasma wakefield accelerator," Nature **515**, 92–95 (2014).
[6]S. Corde, E. Adli, J. M. Allen, W. An, C. I. Clarke, C. E. Clayton, J. P. Delahaye, J. Frederico, S. Gessner, S. Z. Green, M. J. Hogan, C. Joshi, N. Lipkowitz, M. Litos, W. Lu, K. A. Marsh, W. B. Mori, M. Schmeltz, N. Vafaei-Najafabadi, D. Walz, V. Yakimenko, and G. Yocky, "Multi-gigaelectronvolt acceleration of positrons in a self-loaded plasma wakefield," Nature **524**, 442–445 (2015).
[7]K. L. Bane, P. B. Wilson, and T. Weiland, "Wake fields and wake field acceleration," AIP Conf. Proc. **127**, 875–928 (1985).
[8]P. Chen, J. J. Su, J. M. Dawson, K. L. F. Bane, and P. B. Wilson, "Energy transfer in the plasma wake-field accelerator," Phys. Rev. Lett. **56**, 1252–1255 (1986).
[9]K. L. Bane, P. Chen, and P. B. Wilson, "Collinear wake field acceleration," Report No. SLAC-PUB-3662 (Stanford Linear Accelerator Center, 1985).
[10]W. Lu, C. Huang, M. Zhou, M. Tzoufras, F. Tsung, W. Mori, and T. Katsouleas, "A nonlinear theory for multidimensional relativistic plasma wave wakefields," Phys. Plasmas **13**, 056709 (2006).
[11]T. N. Dalichaouch, X. Xu, A. Tableman, F. Li, F. S. Tsung, and W. B. Mori, "A multi-sheath model for highly nonlinear plasma wakefields," Phys. Plasmas **28**, 063103 (2021).
[12]J. B. Rosenzweig, B. Breizman, T. Katsouleas, and J. J. Su, "Acceleration and focusing of electrons in two-dimensional nonlinear plasma wake fields," Phys. Rev. A **44**, R6189–R6192 (1991).
[13]W. Lu, W. An, C. Huang, C. Joshi, W. B. Mori, M. Hogan, T. Raubenheimer, A. Seryi, P. Muggli, and T. Katsouleas, "High transformer ratio PWFA for application on XFELs," in Proceedings of PAC09, 2009.
[14]G. Loisch, G. Asova, P. Boonpornprasert, R. Brinkmann, Y. Chen, J. Engel, J. Good, M. Gross, F. Grüner, H. Huck, D. Kalantaryan, M. Krasilnikov, O. Lishilin, A. M. de la Ossa, T. J. Mehrling, D. Melkumyan, A. Oppelt, J. Osterhoff, H. Qian, Y. Renier, F. Stephan, C. Tenholt, V. Wohlfarth, and Q. Zhao, "Observation of high transformer ratio plasma wakefield acceleration," Phys. Rev. Lett. **121**, 064801 (2018).
[15]T. C. Katsouleas, S. Wilks, P. Chen, J. M. Dawson, and J. J. Su, "Beam loading in plasma accelerators," Part. Accel. **22**, 81–99 (1987).
[16]M. Tzoufras, W. Lu, F. Tsung, C. Huang, W. Mori, T. Katsouleas, J. Vieira, R. Fonseca, and L. Silva, "Beam loading by electrons in nonlinear plasma wakes," Phys. Plasmas **16**, 056705 (2009).
[17]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. Spec. Top.-Accel. Beams **16**, 010101 (2013).
[18]B. Mustapha and P. Ostroumov, "Optimization algorithms for accelerator physics problems," in Proceedings of ICAP09, San Francisco, CA, 2009.
[19]S. Diederichs, C. Benedetti, E. Esarey, J. Osterhoff, and C. B. Schroeder, "High-quality positron acceleration in beam-driven plasma accelerators," Phys. Rev. Accel. Beams **23**, 121301 (2020).
[20]J. Duris, D. Kennedy, A. Hanuka, J. Shtalenkova, A. Edelen, P. Baxevanis, A. Egger, T. Cope, M. McIntire, S. Ermon, and D. Ratner, "Bayesian optimization of a free-electron laser," Phys. Rev. Lett. **124**, 124801 (2020).
[21]S. Jalas, M. Kirchen, P. Messner, P. Winkler, L. Hübner, J. Dirkwinkel, M. Schnepp, R. Lehe, and A. R. Maier, "Bayesian optimization of a laser-plasma accelerator," Phys. Rev. Lett. **126**, 104801 (2021).
[22]R. J. Shalloo, S. J. D. Dann, J.-N. Gruse, C. I. D. Underwood, A. F. Antoine, C. Arran, M. Backhouse, C. D. Baird, M. D. Balcazar, N. Bourgeois, J. A. Cardarelli, P. Hatfield, J. Kang, K. Krushelnick, S. P. D. Mangles, C. D. Murphy, N. Lu, J. Osterhoff, K. Põder, P. P. Rajeev, C. P. Ridgers, S. Rozario, M. P. Selwood, A. J. Shahani, D. R. Symes, A. G. R. Thomas, C. Thornton, Z. Najmudin, and M. J. V. Streeter, "Automation and control of laser wakefield accelerators using Bayesian optimization," Nat. Commun. **11**, 6355 (2020).
[23]M. Kirchen, S. Jalas, P. Messner, P. Winkler, T. Eichner, L. Hübner, T. Hülsenbusch, L. Jeppe, T. Parikh, M. Schnepp, and A. R. Maier, "Optimal beam loading in a laser-plasma accelerator," Phys. Rev. Lett. **126**, 174801 (2021).
[24]S. M. Wild, "Solving derivative-free nonlinear least squares problems with POUNDERS," in *Advances and Trends in Optimization with Engineering Applications*, edited by T. Terlaky, M. F. Anjos, and S. Ahmed (SIAM, 2017), pp. 529–540.
[25]C. Huang, V. K. Decyk, C. Ren, M. Zhou, W. Lu, W. B. Mori, J. H. Cooley, T. M. Antonsen, Jr., and T. Katsouleas, "QUICKPIC: A highly efficient particle-in-cell code for modeling wakefield acceleration in plasmas," J. Comput. Phys. **217**, 658–679 (2006).
[26]W. An, V. K. Decyk, W. B. Mori, and T. M. Antonsen, "An improved iteration loop for the three dimensional quasi-static particle-in-cell algorithm: QuickPIC," J. Comput. Phys. **250**, 165–177 (2013).
[27]W. An, W. Lu, C. Huang, X. Xu, M. J. Hogan, C. Joshi, and W. B. Mori, "Ion motion induced emittance growth of matched electron beams in plasma wakefields," Phys. Rev. Lett. **118**, 244801 (2017).
[28]While fewer than ten bins did not fully capture the beam dynamics, increasing/decreasing the number of bins beyond 20 had only a marginal effect on the simulation outputs.
[29]F. Li, W. An, V. K. Decyk, X. Xu, M. J. Hogan, and W. B. Mori, "A quasi-static particle-in-cell algorithm based on an azimuthal Fourier decomposition for highly efficient simulations of plasma-based acceleration: QPAD," Comput. Phys. Commun. **261**, 107784 (2021).
[30]J. R. Pierce, J. P. Palastro, F. Li, B. Malaca, D. Ramsey, J. Vieira, K. Weichman, and W. B. Mori, "Arbitrarily structured laser pulses," Phys. Rev. Res. **5**, 013085 (2023).
[31]F. Li, T. N. Dalichaouch, J. R. Pierce, X. Xu, F. S. Tsung, W. Lu, C. Joshi, and W. B. Mori, "Ultrabright electron bunch injection in a plasma wakefield driven by a superluminal flying focus electron beam," Phys. Rev. Lett. **128**, 174803 (2022).