

Compensating slice emittance growth in high brightness photoinjectors using sacrificial charge

W. H. Li^{1,*}, A. C. Bartnik,¹ A. Fukasawa², M. Kaemingk¹, G. Lawler,²
N. Majernik^{2,†}, J. B. Rosenzweig² and J. M. Maxson^{1,‡}

¹*Cornell Laboratory for Accelerator-Based Sciences and Education, Ithaca, New York, USA*

²*Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California, USA*



(Received 9 April 2024; accepted 16 July 2024; published 5 August 2024)

Achieving maximum electron beam brightness in photoinjectors requires detailed control of the 3D bunch shape and precise tuning of the beam focusing. Even in state-of-the-art designs, slice emittance growth due to nonlinear space charge forces and partial nonlaminarity often remains non-negligible. In this work, we introduce a new means to linearize the transverse slice phase space: a sacrificial portion of the bunch's own charge distribution, formed into a wavebroken shock front by highly nonlinear space charge forces within the gun, whose downstream purpose is to dynamically linearize the desired bunch core. We show that linearization of an appropriately prepared bunch can be achieved via strongly nonlaminar focusing of the sacrificial shock front, while the inner core focuses laminarly. This leads to a natural spatial separation of the two distributions: a dense core surrounded by a diffuse halo of sacrificial charge that can be collimated. Multiobjective genetic algorithm optimizations of the ultracompact x-ray free electron laser injector employ this concept, and we interpret it with an analytic model that agrees well with the simulations. In simulation, we demonstrate a final bunch charge of 100 pC, peak current ~ 30 A, and a sacrificial charge of 150 pC (250 pC total emitted from cathode) with normalized emittance growth of < 20 nm rad due to space charge. This implies a maximum achievable brightness approximately an order of magnitude greater than existing free electron laser injector designs.

DOI: [10.1103/PhysRevAccelBeams.27.084401](https://doi.org/10.1103/PhysRevAccelBeams.27.084401)

I. INTRODUCTION

High brightness photoinjectors are the electron sources of choice for a wide variety of accelerator applications, ranging from x-ray free electron lasers to MeV-scale electron diffraction and microscopy. Higher brightness performance has the potential to unlock probes with greater spatiotemporal resolution, to enable new modes of operation, or dramatically shrink the cost in size of accelerator components. Critical to the brightness performance in these photoinjectors is the process of emittance compensation [1,2], in which the phase space angles of each longitudinal slice are asymptotically equalized through a judicious

choice of focusing and acceleration. Typical implementations of this technique can result in a recovery of brightness by an order of magnitude or more [2,3].

Provided perfect compensation of the linear space charge forces, emittance growth is then dominated by nonlinear forces that distort the slice phase space. With the exception of uniformly filled ellipsoids and infinitely long uniform cylindrical distributions, all other spatial bunch distributions generate nonlinear forces and are susceptible to slice emittance growth. Ideal distributions with linear space charge forces with zero thermal momentum spread will undergo self-similar evolution in each plane; the space charge forces remain exactly linear in propagation through ideal lenses. In practice, however, the intrinsic momentum distribution from photocathodes is often a 3D Gaussian of finite width in the forward direction [4,5]. This momentum spread causes the evolution of distributions with initially linear space charge to no longer be self-similar, and thus nonlinear space charge emittance growth is in this sense strictly unavoidable. Its magnitude is then dependent on both the slice phase advance in the photoinjector and the size of the intrinsic momentum spread.

For typical transverse distributions with density maxima on axis, the slice emittance is not in general a monotonically increasing function of time. Several studies [3,6,7] have

*Contact author: wli4@bnl.gov

Present address: Brookhaven National Laboratory, Upton, New York, USA.

†Present address: SLAC National Accelerator Laboratory, Menlo Park, California, USA.

‡Contact author: maxson@cornell.edu

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.

demonstrated that in general the slice emittance oscillates in both uniform focusing channels and split photoinjectors. This oscillation can be understood intuitively: particles at the slice edge are in general initially less defocused than the particles in the core. And so, relative to the initial distribution, a density enhancement is eventually generated at the beam's edge. This enhancement induces stronger space charge defocusing, which eventually creates a distribution similar to the initial conditions with an underdense and underdefocused region at the edge. The process then repeats.

For infinitely long beams with weak nonlinearity, this oscillation remains in phase with the slice transverse beam size. Thus, an important rule emerges: weakly nonlinear beams have slice emittance minima at beam size minima, which is exactly the condition for linear space charge emittance compensation. Thus, the process of linear emittance compensation naturally compensates for nonlinearities in the beam as well. However, in practice, there are several complications to achieving the perfect linear emittance compensation, such as overfocusing of edge particles and slice size mismatch [7]. Furthermore, the slice emittance oscillation is not purely periodic: sufficiently nonlinear forces can cause dephasing between the slice emittance and the projected emittance and the initial distribution does not perfectly reform. In practice, this dephasing can occur in half of a plasma period. Nonlinearities strong enough to exhibit this dephasing phenomenon can arise, for example, from 3D effects in bunches with comparable transverse and longitudinal dimensions in the rest frame or sufficiently large transverse density gradients. In photoinjectors, these nonlinearities can additionally lead to *wavebreaking*, wherein the transverse slice phase space becomes double valued. This wavebreaking has been well studied [3,7,8] and forms a density shock at the edge of the beam [9]. While the brightness of such a wavebroken beam can be substantially improved by eliminating the wavebroken portion with an aperture [9], the nonlinear forces experienced by the core during the wavebreaking process may remain uncompensated.

The upper limit of beam brightness is set by the brightness of the electron emission [10], and as this value increases with enhanced electric field at the cathode or by improvements in photoemission momentum spread, new strategies may be required for more precise emittance compensation in future photoinjectors. In this paper, we introduce a new technique to undo slice emittance degradation, where we dynamically utilize a wavebroken shock by focusing it through the core of the bunch before subsequently discarding it with an aperture. Typically, the trajectories of the particles in a beam do not cross, a regime known as laminar motion, even when the beam is focused, as the space charge repulsion of inner particles prevent outer particles from reaching the central region of the beam. Instead, we operate in a regime where some particles have sufficient inward momentum to overcome

the repulsive force of interior particles and cross through the beam central axis, in a process known as a nonlaminar focus. This scenario can be thought of as a manifestation of wavebreaking in phase space; the broken piece of the phase space wave is nonlaminarly focused, passing through the core on its way to the central axis. Meanwhile, the core focuses laminarly, and the net effect is linearization of the bunch core. The nonlaminarity of the focus of the broken component arises from two factors. By definition, the broken component possesses a lower outward radial momentum than the rest of the beam after removal of linear correlations and thus will focus more strongly. Enhancing this effect, the spherical aberration of the solenoid lenses used in this beamline provides the outer particles with a stronger focusing kick than the inner particles [11]. Together, these effects cause the outer particles to possess enough inward radial momentum to cross the axis.

We demonstrate this linearization process in simulations of the next-generation ultracompact x-ray free electron laser (UCXFEL) being developed at the University of California, Los Angeles [12]. This beamline utilizes recently developed rf accelerating cavities with gradients of up to 140 MeV/m [13], including 125 MeV/m accelerating gradient C-band cryogenic linacs [14] and 240 MV/m peak fields in the photoinjector [15], allowing the length of the beamline to be reduced by more than a factor of 10, making these compact XFELs much more available at the university scale. In the gun, the high fields allow for the generation of beams with brightnesses potentially 50 times better than the original Linac Coherent Light Source (LCLS) design, since the maximum four-dimensional brightness scales with the first or the three-halves power of accelerating field for the pancake and cigar regimes respectively [10,16]. It is important to note that this linearization process was first discovered in simulation by the genetic optimization of the UCXFEL injector, and it was understood analytically after the fact.

Beyond the cathode electric field, the maximum brightness is also limited by a photocathode quantity known as the mean transverse energy (MTE), which encapsulates the transverse velocity spread of the electrons as they are emitted from the photocathode. The intrinsic 2D emittance of the photocathode can be put in terms of the MTE as

$$\epsilon_{n,x} = \sigma_x \sqrt{\frac{\text{MTE}}{mc^2}}, \quad (1)$$

where σ_x is the spot size of the photoemitting laser, and mc^2 is the rest energy of the electron.

Typically, FELs, such as LCLS, have used photocathodes with MTEs on the order of 500 meV [17,18]. Recent developments in photocathode growth and production have enabled the production of photocathodes with

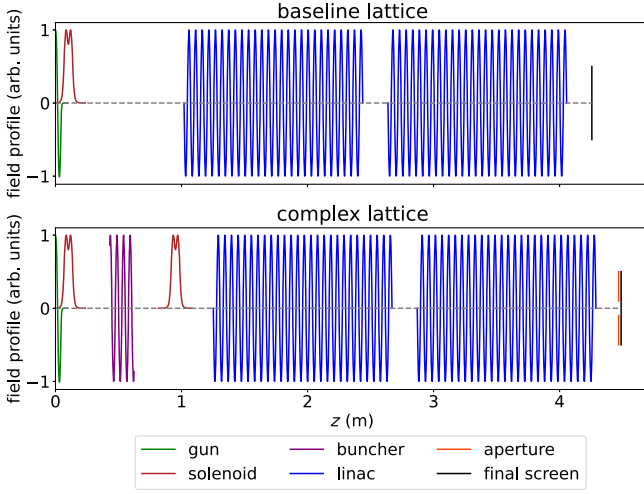


FIG. 1. Schematic of the model XFEL photoinjector with the field profiles of each element. (a) The baseline lattice. (b) The complex lattice, which can be augmented with an aperture.

MTEs as low as 5 meV [19], corresponding to an initial brightness increase of a factor of 100.

Our study begins with a discussion of the performance enhancements achievable when combining the high-gradient accelerating cavities of the UCFEL beamline and a low MTE photocathode. The associated high initial charge density renders this an ideal environment for studying the dynamics of cold, highly space charge dominated beams. Previously [20–22], a lattice with two solenoids and a buncher has been shown to perform well in space charge dominated beamlines. We will examine the performance of such a lattice, comparing its performance with and without a final clipping aperture, and explain the physics behind the qualitatively different dynamics that arise in the presence of the aperture. A simplified lattice with a single solenoid will also be included for reference, and the two lattices are shown in Fig. 1. We will refer to the lattice with the single solenoid as “baseline” and the lattice with two solenoids and a buncher as “complex.”

II. EMITTANCE PRESERVATION

We use a multiobjective genetic algorithm (MOGA) [23,24], coupled to a particle-in-cell code, General Particle Tracer (GPT) [25], to find optimal settings for these three representative high gradient XFEL beamlines and determine the best achievable emittance. In turn, we can use the optimal emittances to determine the degree to which the initial emittance has been recovered.

The final beam energy in all three beamlines is 150 MeV. In the baseline lattice and the complex lattice without the aperture, we will transport 100 pC of charge from the beginning to the end, while in the complex lattice with the aperture, we will start with 250 pC and end with the same 100 pC, clipping down with the final aperture. The

cathode MTE is set to 5 meV for all simulations, as this is the lowest that has been experimentally achieved [19], as well as being near the fundamental limit imposed by disorder induced heating [26].

We set the MOGA to optimize both transverse emittance and bunch length, constraining solutions to reach the desired 150 MeV beam energy at the end of the beamline while losing no particles in transport. The optimizer was allowed to vary the solenoid currents, gun phase, as well as the buncher phase and amplitude. All quantities were constrained to be within physically attainable limits. The initial transverse distribution of the beam was taken to be of a truncated 2D Gaussian form, and the initial temporal distribution was chosen as a truncated 1D Gaussian. The optimizer was allowed to independently vary the widths and truncation locations of these two distributions. Element positions, linac amplitude, and linac phase were fixed. In general, optimizing multiple objectives will generate curves known as a Pareto fronts which shows the trade-off between pairs of objectives. By comparing the Pareto fronts for the baseline lattice and the complex lattice without the aperture, we can see how much of an improvement the extra solenoid and the buncher provides. Subsequently, the effect of the aperture can be identified by looking at the front for the complex lattice with the aperture.

The fronts are shown in Fig. 2. We see immediately that the complex layout improves the emittances by roughly a factor of 2, with the aperture providing yet another factor of 2. The degree to which the cathode contributes to the final emittance can be quantified by defining an *effective MTE* as the MTE required, given the laser spot size, to

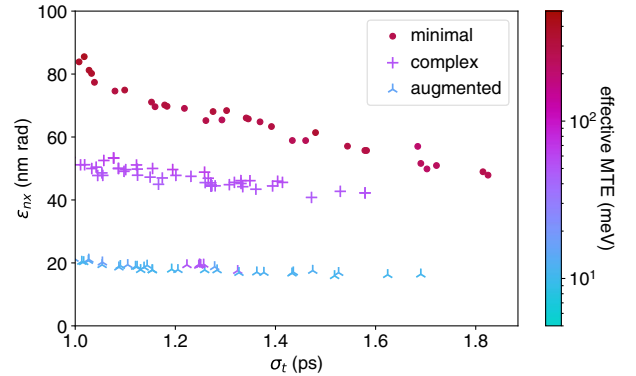


FIG. 2. Pareto fronts for the three lattice configurations: baseline, complex without the aperture, and complex with the aperture (augmented), colored by their effective MTE (defined in the text). The addition of the second solenoid and buncher improves emittance by roughly a factor of 1.4 on average, and the addition of the aperture improves it by more than another factor of 2. The effective MTE of the baseline lattice is approximately 300 meV, showing a substantial emittance growth during transport. The complex lattice without the aperture achieves an effective MTE on the order of 70 meV on average, while adding the aperture achieves an effective MTE of about 10 meV on average.

generate the observed final emittance assuming perfect emittance preservation [27]. Thus, the effective MTE is a heuristic to determine the degree to which emittance is preserved during transport and to determine the MTE scale at which the intrinsic emittance becomes dominant. An ideal beamline would have an effective MTE equal to the real MTE of the cathode. Dilution in transport is captured by an increase in the effective MTE.

The addition of the aperture introduces a small complication to the effective MTE calculation. Directly calculating the effective MTE using the emittance of the core 100 pC, but the laser spot size used to emit the full 250 pC would result in an artificial reduction of the effective MTE. Instead, we use the laser spot size which would be required to emit 100 pC with the same charge density as the full 250 pC beam, which can be approximated by scaling the naive effective MTE up by a factor of 2.5, the ratio of the emitted charge to the final charge. The effective MTEs of all three fronts is represented by the marker colors in Fig. 2. The baseline lattice has effective MTEs on the order of 300 meV showing that the MTE is not a significant contributor to the emittance and that the emittance has been substantially degraded during transport. Such a lattice would see little benefit from upgrading from a 100 meV cathode to a 5 meV cathode. The complex lattice without the aperture does better, with an effective MTE of 70 meV. This lattice would be cathode emittance limited with a 100 meV cathode, but it would not be able to utilize a single-digit MTE cathode to its maximum potential. However, the complex lattice with the aperture has an effective MTE of around 10 meV, indicating excellent emittance preservation. We see that to take full advantage of this beamline, a cathode with single-digit meV MTE is necessary.

To isolate the effects of the aperture from the rest of the lattice, we will now focus on the two beamlines utilizing the complex lattice. We select the lowest emittance example in the complex lattice without aperture front and compare it with the example in the complex lattice with aperture front that has the closest final bunch length, approximately 1.5 ps. Examining the beam size and emittance evolution of these representative examples reveals a stark qualitative contrast in their respective dynamics, shown in Fig. 3. The complex lattice without the aperture performs precision traditional emittance compensation. Due to residual uncompensated nonlinear space charge effects, the slice emittance never reattains the emittance achieved immediately after the first solenoid. On the other hand, the complex lattice with the aperture allows the slice emittance to rise by nearly an order of magnitude immediately at the first solenoid. The slice emittance then decreases until it is frozen by the linacs. Furthermore, while the beamsizes in the complex lattice without aperture reaches its minimum close to the entrance of the first linac, the beamsizes in the complex lattice with the aperture reaches a minimum in the

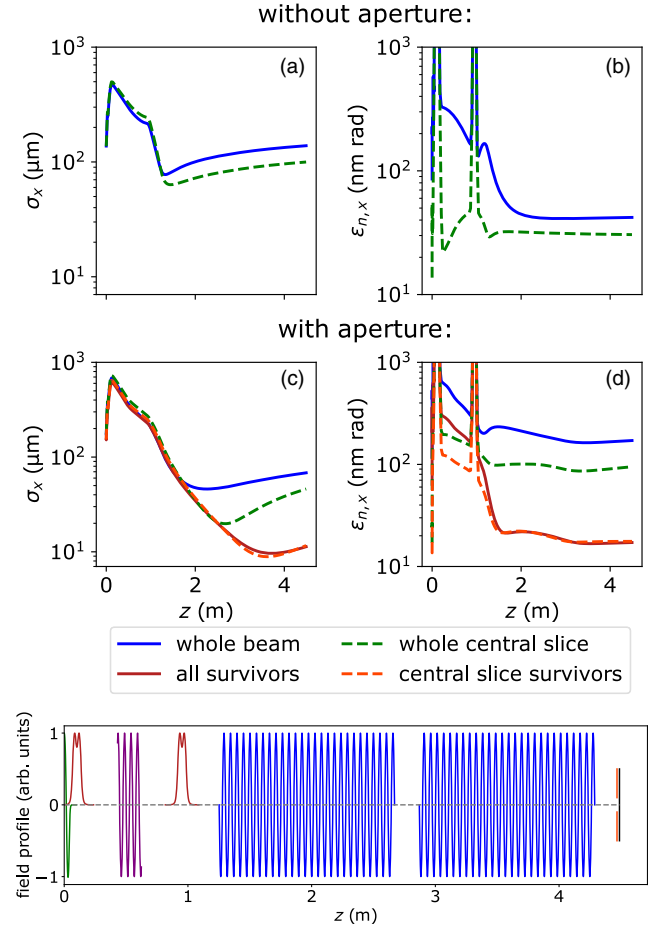


FIG. 3. Transverse beam size (a), (c) and emittance (b), (d) evolution of representative individual simulations from the complex lattice without (a), (b) and with (c), (d) the aperture. The complex lattice without the aperture achieves precise traditional emittance compensation, characterized by the beamsizes and slice emittance minima coinciding in longitudinal position. In contrast, with the aperture, the slice emittance increases drastically immediately upon exiting the gun before rapidly falling as the beam enters the linacs. The beam size evolution of the two lattices is also noticeably different. The lattice schematic from Fig. 1(b) is reproduced here for the reader's convenience.

middle of the first linac, with the beamsizes of the survivors reaching a minimum in the middle of the second linac. Thus, not only is the beam evolution in the two beamlines different, but it also appears that in the beamline with the aperture, there are effectively two separate particle populations.

The two populations can be visually identified in Fig. 4, which shows the beam profile and x phase space immediately before the final aperture. A dense core in both real space and phase space has developed, surrounded by a diffuse halo of electrons. The aperture discards this diffuse halo, reducing the emittance from 150 to 20 nm rad while only losing 60% of the bunch charge.

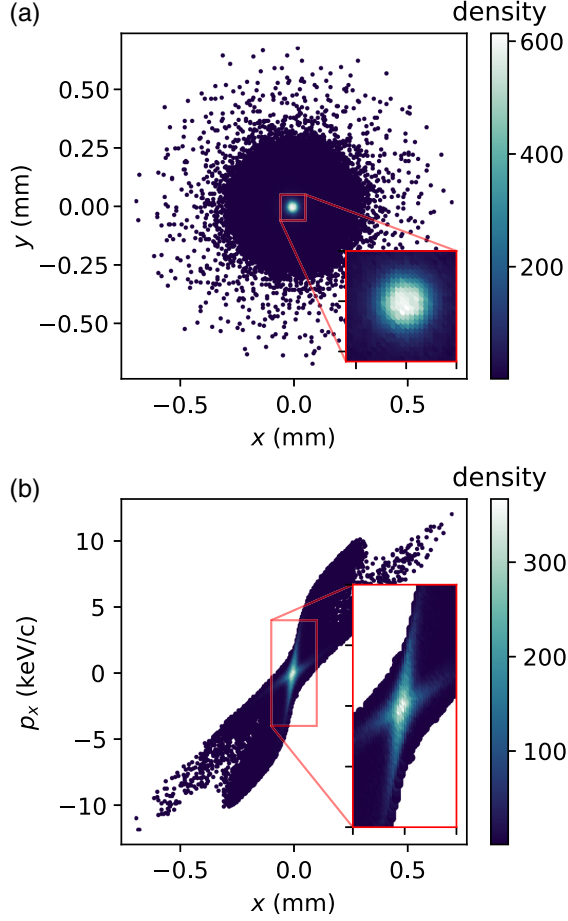


FIG. 4. Beam profile and x phase space immediately before the aperture. (a) Beam profile with linacs. (b) x phase space with linacs. The beam core has become very dense and is surrounded by a diffuse halo. Insets: zoomed in picture of the core.

To understand the process by which this dense core develops, we look at selected snapshots of the beam's phase space, plotted in $r - p_r$ coordinates, for both beamlines, shown in Fig. 5. In all plots, red particles represent particles that are clipped by the aperture, which we call *dead particles* and blue particles represent particles that pass through the aperture, which we call *survivors*. Since the aperture is at the end of the beamline, the coloration reflects knowledge of the eventual trajectories of these particles. As the linear $r - p_r$ correlation is usually large and washes out the details, we plot the phase spaces with this correlation subtracted, with the original correlated phase space being shown in gray for reference. Since the dynamics of the survivors in the central slice and the survivors in the rest of the beam are nearly identical, as can be seen in Fig. 3(c), we look only at the phase space evolution of the central slice.

Figure 5 shows the phase space immediately after the gun, where there is already a significant difference between the two lattices. In general, the curvature of the space charge force varies with the beam's aspect ratio [28]. Without the aperture, the photoemission process is tuned to balance the varying curvature and generate a phase space without large nonlinearity (and a small wave-broken tail). With the aperture, the space charge force is permitted to generate a large negative concavity in phase space, which eventually results in the shock front forming at the edge of the beam.

The beam then travels through the buncher and the two solenoids, with the phase space at the exit of the second solenoid shown in Fig. 5(b). Of particular note here is the much larger radial velocity spread of the outermost particles in the lattice with the aperture, as well as the larger focusing force applied to the beam as a whole. This results in the wave crashing shown in Fig. 5(c), where the

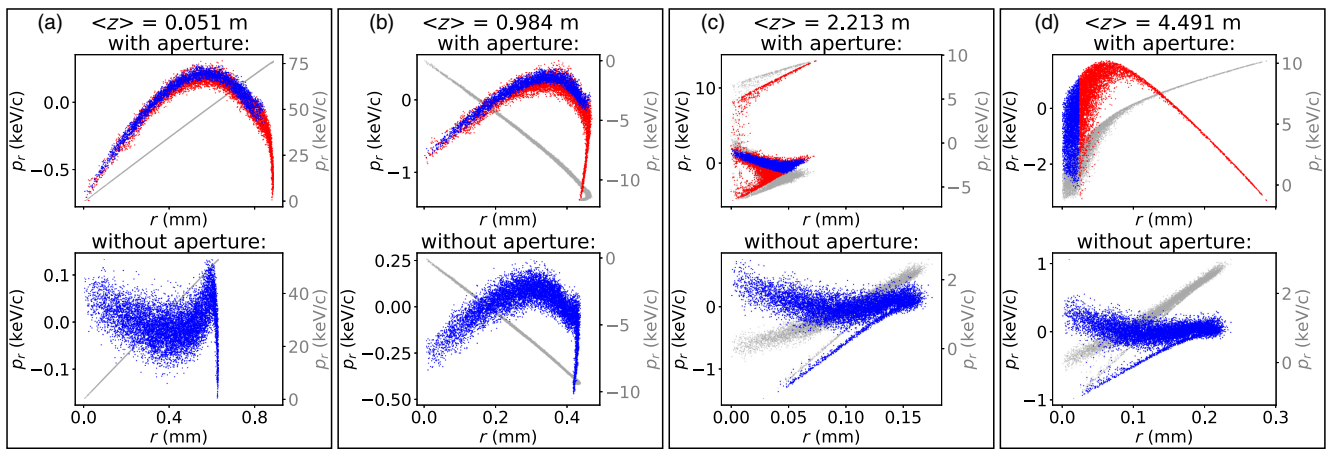


FIG. 5. Evolution of the $r - p_r$ phase space of the central longitudinal slice of the beam in the complex lattice. Top row of each subfigure: with the aperture. Bottom row of each subfigure: without the aperture. The dead particle (red) and survivor (blue) phase spaces are displayed with the linear $r - p_r$ correlation removed. The original, correlated phase space is shown in light gray. Note that the phase spaces of the two lattices are plotted with the same horizontal axis scale but different vertical scales. (a) Immediately after the exit of the gun. (b) Immediately after the second solenoid. (c) During transport through the linacs. (d) Immediately before the final screen.

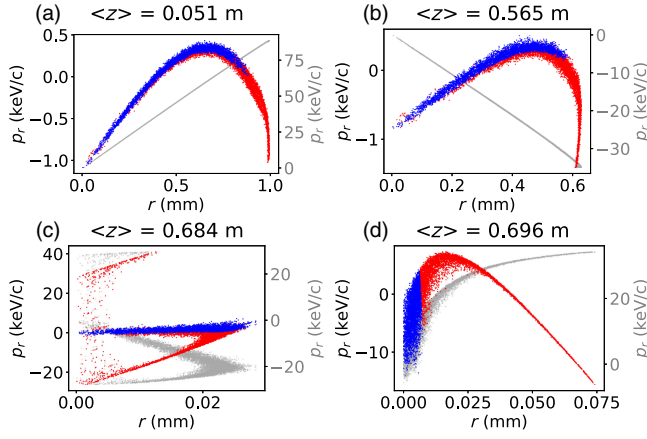


FIG. 6. Evolution of the $r - p_r$ phase space in a beamline with the aperture reoptimized without the linacs. The coloring has the same meaning as in Fig. 5. The similarities in the phase space evolution to the beamline with the linacs are evident. (a) Immediately after the exit of the gun. (b) Immediately after the second solenoid. (c) Transport to the final screen. (d) Final screen.

fastest falling particles in the shock front have been nonlaminarly focused through the core of the beam. Contrast this to the lattice without the aperture below, where the nonlinear tail has merely been folded under the rest of the beam and no nonlaminarity has occurred.

Finally, Fig. 5(d) shows the phase space immediately before the final screen, where the aperture is placed for the lattice that utilizes it. We see again that the lattice with the aperture produces a beam with a very dense spatial core surrounded by a diffuse halo. Meanwhile, the lattice without the aperture has generated a roughly flat phase space, but residual curvature is still visible, as well as the folded nonlinear tail.

Our goal for the remainder of this paper will be to understand the forces at play during the wave crashing, which linearize the phase space of the core and generate the high core density. To aid in understanding the process by which this dense core develops, we simplify the simulation by removing the two linacs and the buncher and moving the other elements closer to the gun, reoptimizing the lattice and initial electron beam parameters. Figure 6 shows the phase space evolution at approximately equivalent points on the beamline as shown in Fig. 5. The basic dynamics are identical, with a large curvature developing inside the gun, a shock front forming at the edge of the beam which is subsequently nonlaminarly focused through the core, and a dense core appearing at the end of the beamline. The following model is based on this linac-less simulation.

III. WAVE-CRASHING LINEARIZATION

This section focuses on understanding the physics that causes these highly nonlinear dynamics to improve, rather than dilute, the emittance of the core. We begin with a simplified analytic model of the space charge forces

imparted by the shock front on the beam as it collapses inward. In the beam's rest frame, it is on the order of 20 times longer than it is wide after exiting the second solenoid, so we treat it as an infinitely long cylinder of charge, surrounded by an infinitely long, infinitely thin, cylindrical shell of charge. Due to the infinite length of the beam, there is no net longitudinal space charge force on any particle. At $t = 0$, the shell has radius a_0 , defined to be the radius of the outermost particle in the beam. We will assume the particles that form the shell have velocities uniformly distributed from v_s to v_f , all directed radially inward. We will perform all calculations in the rest frame of the beam.

The beam is being focused, which we will model as the beam receiving a linear velocity kick $v = -Kr$ at $t = 0$, directed radially inward. From Fig. 5(b), we see that all particles in the shell are focusing faster than the particles in the beam, so we have $v_s > Kr$ for all r .

We see from the phase space pictures that the linearization is already completed by the time, the fastest particles in the shell reach a focus, and after this point, which we call t_f , we will assume that the force from the shell is negligible due to the large decrease in the shell's charge density. We want to find the impulse provided by the shell to a particle with charge q at a start position $\vec{r}_0$, i.e., the force from the shell integrated from $t = 0$ to $t = t_f$. As a simplification, we will assume that the velocities of all particles are constant, i.e., the space charge force only provides small perturbations to their trajectory. Thus, the shell will undergo a nonlaminar focus as it collapses past the center and starts expanding outward.

The impulse can be written

$$\Delta \vec{p} = \int_0^{a_0/v_f} q \vec{E}(t) dt, \quad (2)$$

where $\vec{E}(t)$ is the field at $\vec{r}(t)$.

To calculate $\vec{E}$, consider the subset of shell particles with average velocity $\vec{v}$ and velocity spread $d\vec{v}$. Since we assume the velocities are uniformly distributed and all points radially inward, this subset is itself a cylindrical shell with a linear charge density given by

$$\lambda = \frac{\lambda_r dv}{v_f - v_s}, \quad (3)$$

where we will begin writing only the magnitudes of the velocities and λ_r is the total linear charge density of the collapsing shell.

The field from an infinitely long cylinder of surface charge density σ and radius a at a point $\vec{r}$ outside the shell is

$$\vec{E} = \frac{a\sigma}{\epsilon_0 r} \hat{r}. \quad (4)$$

We note that $a\sigma$ represents the linear charge density of the shell and is identical to the previously defined λ , which

is constant with time, so

$$\vec{E}(t) = \frac{a\sigma}{\epsilon_0 r(t)} \hat{r} = \frac{a\sigma}{\epsilon_0 r_0(1 - Kt)} \hat{r}. \quad (5)$$

The field inside the cylinder is identically 0, so a subset with a given velocity will only contribute to the field once it has collapsed past $\vec{r}$.

We will define $t_p(v)$ to be the time that a particle with velocity v has a radial position equal to that of the test charge. Thus, it must satisfy the following equation:

$$a_0 - vt_p(v) = r_0(1 - Kt_p(v)). \quad (6)$$

We can solve this to obtain

$$t_p(v) = \frac{a_0 - r_0}{v - Kr_0}. \quad (7)$$

For $t < t_p(v)$, the field is identically 0, because the subset has a radius greater than the radial position of the test charge. Since the upper limit of the integration is when the fastest shell particles reach the center of the beam and all shell particles have a higher velocity than all beam particles, the subset of shell particles with velocity v has a radius smaller than that the radial position of the test charge for times between $t_p(v)$ and a_0/v_f , and its field contribution is nonzero during this time. Thus, the integral in Eq. (2) becomes, as a function of the subset velocity v ,

$$\Delta\vec{p}(v) = \int_{t_p(v)}^{t_f} \frac{q\lambda}{\epsilon_0 r_0(1 - Kt)} \hat{r} dt. \quad (8)$$

We will define the critical velocity v_c to be the velocity for which $t_p(v_c) = t_f$, i.e., a shell particle with velocity v_c attains the same radial position as the test charge at the same time that the fastest shell particle reaches the center of the beam. All particles with velocity lower than v_c will not contribute to the field experienced by the test charge. Note that v_c is a function of the test charge starting position r_0 . We have

$$v_c(r_0) = v_f - r_0 \left(\frac{v_f}{a_0} - K \right). \quad (9)$$

Figure 7 shows a cartoon of particle trajectories to clarify the physical meaning of the quantities we have defined. All shell particles begin with radial position a_0 , while the test particle begins with radial position r_0 . t_f is the x-intercept of the trajectory of the fastest shell particle, and $t_p(v)$ is the time when the trajectory of the test particle and the shell particle with velocity v cross. Note that $t_p(v_c)$ is defined to be identical to t_f .

A further helpful visualization is provided in Fig. 8, which presents four snapshots of the collapsing shell process in two dimensions. The test charge, the fastest

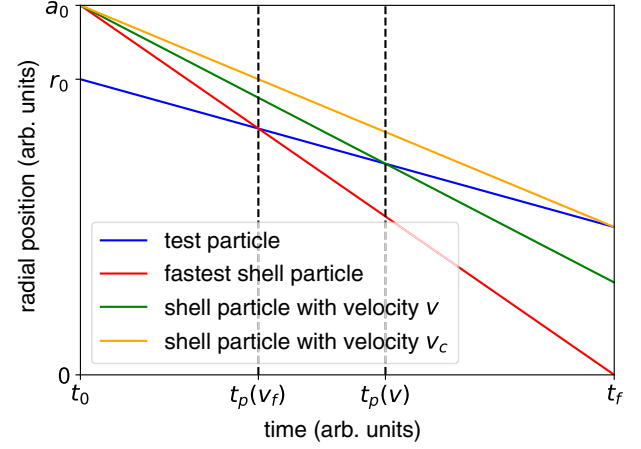


FIG. 7. Cartoon of sample particle trajectories and the associated quantities defined in our model.

particle in the shell, and a particle with the critical velocity are all highlighted to clarify their motion throughout the process.

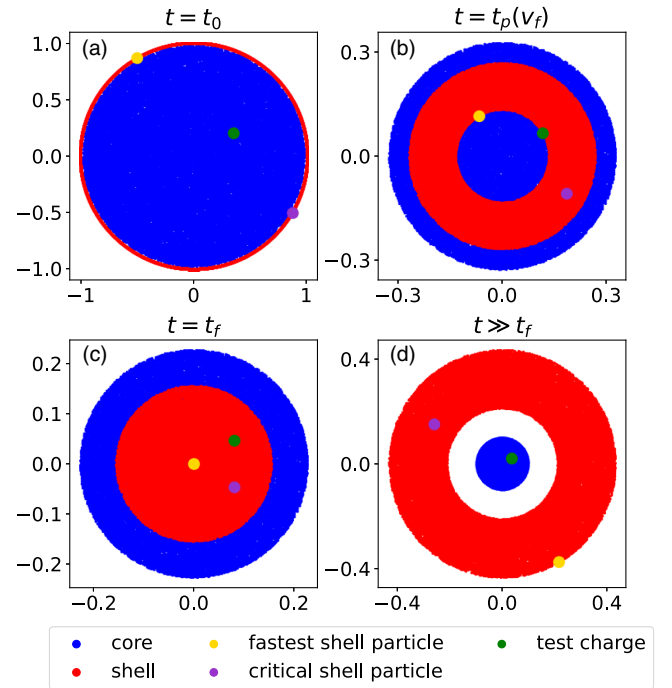


FIG. 8. Two-dimensional sketch of the collapsing shell, with the test charge, the fastest particle in the shell, and a shell particle with the critical velocity for this test charge highlighted. The subplots represent four snapshots of the process: (a) the initial state, (b) when the fastest particle in the shell passes the test charge, (c) when the fastest particle reaches the center and, by definition, the shell particle with the critical velocity has just reached the radial location of the test charge, (d) and some time significantly after the end of the collapsing shell process, showing the diffuse halo that the shell forms. The axes represent x and y position with arbitrary units, and numeric labels are included solely for providing a scale between the subplots.

We now have two cases: if $v_c < v_s$, then all shell particles will contribute to the impulse on the test charge; if $v_c > v_s$, then some particles have radial position greater than that of the test charge at all times in the integration and do not contribute to the impulse. Therefore, to calculate the total impulse provided to the test charge by the entire shell, we integrate Eq. (8) setting the maximum of v_c and v_s as the lower limit and v_f as the upper limit:

$$\Delta \vec{p}(r_0) = \frac{q\lambda_r}{r_0\epsilon_0(v_f - v_s)} \int_{\max(v_c(r_0), v_s)}^{v_f} \int_{t_p(v)}^{t_f} \frac{1}{1 - Kt} dt dv. \quad (10)$$

These integrals can be carried out analytically, to yield the following piecewise function (note that the subscript on r has been dropped):

$$\Delta \vec{p}(r) = \begin{cases} \frac{q\lambda_r}{Kre_0(v_f - v_s)} L_s \hat{r}, & r < \frac{a_0(v_f - v_s)}{v_f - Ka_0} \\ \frac{q\lambda_r}{Kre_0(v_f - v_s)} L_b \hat{r}, & r \geq \frac{a_0(v_f - v_s)}{v_f - Ka_0} \end{cases}, \quad (11)$$

where we define

$$L_s = Kr \log\left(\frac{a_0(v_f - Kr)}{v_f(a_0 - r)}\right) + Ka_0 \log\left(1 - \frac{r}{a_0}\right) + v_f \log\left(\frac{v_f}{v_f - Kr}\right) \quad (12)$$

and

$$L_b = Kr \log\left(\frac{v_f - Kr}{v_s - Kr}\right) + Ka_0 \log\left(\frac{v_s - Ka_0}{v_f - Ka_0}\right) + v_s \log\left(\frac{(v_f - Ka_0)(v_s - Kr)}{v_f(v_s - Ka_0)}\right) + v_f \log\left(\frac{v_f}{v_f - Kr}\right). \quad (13)$$

Note that L_s corresponds to the case where $v_c(r) > v_s$, while L_b corresponds to the case where $v_c(r) \leq v_s$.

We can now calculate the impulse provided to the beam by the collapsing shell in our simulation. We will consider the timestep shown in Fig. 5(b) to be the start of the collapse. We define the shell to be the portion of the beam with double-valued phase space. We can then extract the remaining parameters of the shell, which are summarized in Table I. Notice that the velocity spread increases by about a factor of 10 between the timesteps shown in (b) and (c). This is due to the slower particles experiencing the space charge field from a larger enclosed charge and being slowed more than the faster particles. For now, we will continue to neglect the space charge force applied to the shell and instead examine the behavior of the beam for a range of velocity spreads. It is worth noting that the large

TABLE I. Table of input parameters for the analytic model. The three values for v_f represent a low velocity spread, average velocity spread, and high velocity spread, respectively. The charge contained in the shell is roughly 5% of the total beam charge.

Parameter	Value
a_0	0.6 mm
K	$3.2 \times 10^{10} \text{ s}^{-1}$
λ_r	380 pC/m
v_s	$1.99 \times 10^7 \text{ m/s}$
v_f	$(2.08, 2.24, 2.45) \times 10^7 \text{ m/s}$

increase in velocity spread indicates that the space charge force is significant and should not be neglected, but, remarkably, the correct behavior arises regardless.

The results are shown in Fig. 9 for low velocity spread (half of the average), average velocity spread, and high velocity spread (double the average). The “impulse scaling factor” is an overall hand-tuned correction factor multiplying the applied impulse. Since the impulse is linear in the shell charge density, this scaling factor can be thought of as representing differences in the optimal shell charge density for each velocity spread. The shape of the impulse is generally the same in all cases and the phase space becomes straightened out. The only change that is needed for a different velocity spread is an overall scaling factor for the impulse, shown in the titles of each plot. The final emittances are 23, 21, and 22 nm rad for the low, average, and high velocity spreads, respectively, down from 53 nm rad at the exit of the second solenoid. The emittance from the simulation is 19 nm rad, very good agreement considering the many simplifications we have made in the model. The deviance of the scaling factor from unity is largely due to ignoring the effects of space charge on the motion of the shell. As we shall see in the next section, incorporating these effects into the model numerically yields a scaling factor of nearly unity.

Since the impulse scaling factor is hand-tuned for each case individually and varies from 1.2 for a low velocity spread to 1.6 for a high velocity spread, it is reasonable to ask if fine-tuning this factor is necessary for the emittance compensation to be effective. We can see how much of an effect the scaling factor has by plotting all three of these plots together with the scaling factor used for the average velocity spread, namely, 1.3. This plot is shown in Fig. 10, and we see that all of the final phase spaces have, for the most part, been made nearly linear, despite the scaling factor being suboptimal for the low and high velocity spreads. Indeed the emittances for these two cases have only increased by 2 nm rad (10%) from their optimal values.

Based on these results, it is reasonable to assume that the changing velocity spread will not have a large effect on the final shape of the phase space. Changing the velocity

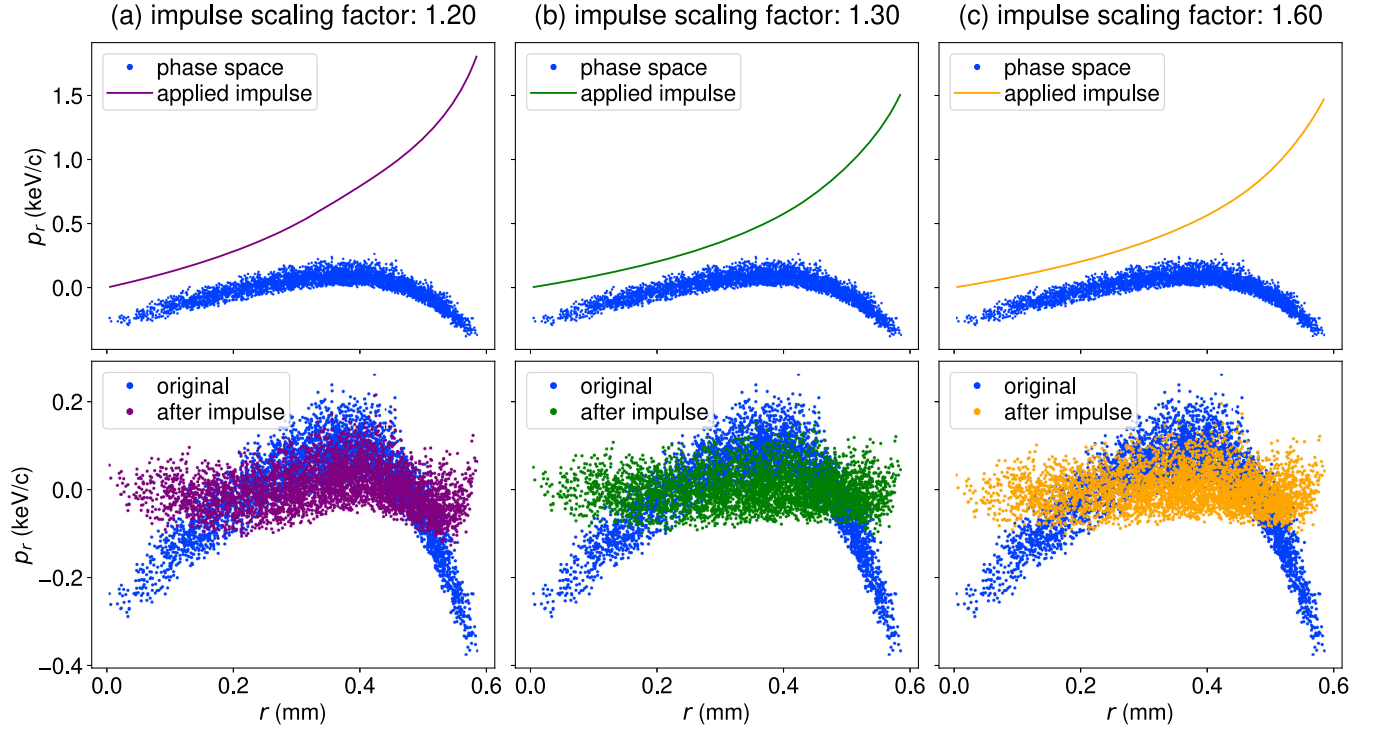


FIG. 9. Analytically calculated impulses and final phase spaces. Phase spaces are displayed with linear $r - p_r$ correlations removed. Top row: impulses calculated from Eq. (11). Bottom row: final phase spaces. (a) Low velocity spread. (b) Average velocity spread. (c) High velocity spread. The “impulse scaling factor” is an overall factor multiplying the applied impulse, which can be thought of as representing different optimal shell charge densities in each case.

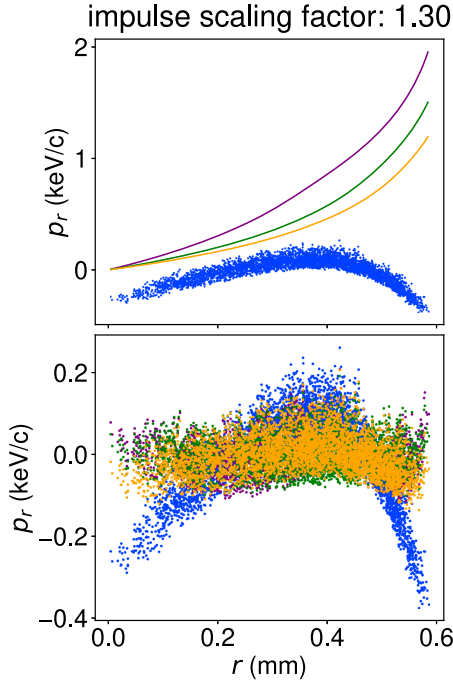


FIG. 10. Impulses and phase spaces for all the velocity spreads, using the same scaling factor for all of them. Each color represents the same data as they do in Fig. 9. The phase spaces for the three cases with different velocity spreads are all roughly straight, despite their optimal scaling factors being different.

spread by a total factor of 4 from the lowest to the highest results in emittances that are within 10% of each other, and the change is caused essentially entirely by a scaling factor on the impulse.

Therefore, the impulse provided by the collapsing shell is precisely the right shape to undo the curvature caused by space charge in the gun. It provides a larger kick to the radial tails than to the radial center as it collapses, and straightens out the phase space, resulting in an emittance decrease of more than a factor of 2. Recall from the discussion in the introduction that the shell is precisely the wavebroken shock front caused by the strongly nonlinear space charge forces in the gun. Said another way, the phase space nonlinearity in the gun develops such that this final impulse from the collapsing shell compensates it.

IV. THE COLLAPSING SHELL MODEL INCLUDING SPACE CHARGE

Now we extend our model to include the slowing of the collapsing shell by the space charge forces of the beam, where we will model the beam as being uniformly distributed. This effect is responsible for the order of magnitude increase in velocity spread during the collapse that was not accounted for in the analytic model, which we rectify in this section.

Let $a(t)$ be the position of a shell particle with initial velocity v as a function of time. The field from the beam at a , assuming that a is inside the beam, is given by Gauss's Law:

$$\vec{E}(a) = \hat{r} \frac{\rho a}{2\epsilon_0}, \quad (14)$$

where ρ is the charge density of the beam. Taking $r_b(t)$ to be the outer radius of the beam as a function of time, which we will determine later, we have

$$\rho(t) = \frac{\lambda_b}{\pi r_b(t)^2}, \quad (15)$$

so

$$\vec{E}(a) = \hat{r} \frac{\lambda_b a}{2\pi\epsilon_0 r_b(t)^2}, \quad (16)$$

where λ_b is the longitudinal charge density of the beam. Thus, we can write the differential equation

$$\ddot{a}(t) = \frac{q}{m} \frac{\lambda_b a(t)}{2\pi\epsilon_0 (r_b(t))^2}, \quad (17)$$

with initial conditions

$$a(0) = a_0, \quad \dot{a}(t) = v. \quad (18)$$

We now need to determine r_b , which is determined by the effects of space charge on the beam itself. We will assume for simplicity that the beam stays as a uniform cylinder as it focuses. We start with the envelope equation in K-V form for an infinitely long, uniform cylinder with radius $r(0) = r_0$:

$$\ddot{r}(t) - \frac{\epsilon^2}{r(t)^3} - \frac{qE(t)}{m} = 0, \quad (19)$$

where ϵ is the normalized emittance in units of length $\times$ momentum.

By Gauss's Law, we can compute $E(t)$ as

$$E(t) = \frac{\rho(t)r(t)}{2\epsilon_0}, \quad (20)$$

where $\rho(t)$ is the charge density. We have

$$\rho(t) = \frac{\lambda}{\pi r^2}, \quad (21)$$

where λ is the linear charge density, which remains constant through the process. We can now write the field

$$E(t) = \frac{\lambda}{2\pi\epsilon_0 r(t)}. \quad (22)$$

Equation (19) becomes

$$\ddot{r}(t) - \frac{\epsilon^2}{r(t)^3} - \frac{q\lambda}{2\pi\epsilon_0 m r(t)} = 0. \quad (23)$$

Since we are modeling the beam as an infinitely long cylinder, charge that is at a radial position greater than that of the test charge has no effect. Thus, the envelope equation here holds true for all charges inside the cylinder as well, since we can treat all the charge with radial position less than r as an independent beam. Accounting for space charge thus requires replacing the RHS of Eq. (6), the $r_0(1 - Kt)$ in the denominator of Eq. (8), and the $r_b(t)$ in Eq. (17) with the solution of Eq. (23).

This becomes intractable analytically, so we solve these equations numerically instead, using the average velocity spread from before. The results are shown in Fig. 11. Even with the complications introduced from space charge forces affecting the dynamics of the shell collapse, the overall shape of the impulse is maintained. With an appropriate scaling factor, the phase space can be linearized again. With a scale factor of 1.1, the final emittance is only slightly worse than the best case emittance from the previous

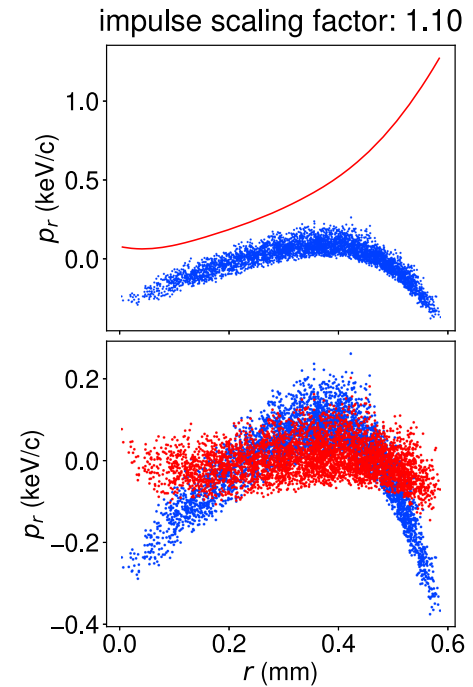


FIG. 11. Impulse and final phase space of a collapsing shell with space charge forces taken into account in the dynamics of the shell. We see that, even here, the overall shape is maintained, and appropriate scaling can straighten out the phase space. The final emittance (23.1 nm rad) is only slightly higher than the best case from before (21.4 nm rad).

examples (23.1 vs 21.4 nm rad). The difference is most likely due to several factors including 3D effects, the initial transverse extent of the shell, and forces applied by the shell after t_f .

We can perform a simple back-of-the-envelope calculation to determine the correction from the beam's finite longitudinal extent. The on-axis electric field from a uniform disk of charge with surface charge density σ is given by

$$E_z = \frac{\sigma}{2\epsilon_0} \left(1 - \frac{z}{(z^2 + a_0^2)^{1/2}} \right). \quad (24)$$

We slice the beam longitudinally into disks, each with charge density $\sigma = \frac{\lambda_b d\xi}{\pi a_0^2}$, where ξ is the longitudinal coordinate in the beam's rest frame, with $\xi = 0$ defined as the longitudinal center of the beam. The total electric field at a point $\xi = z$ for a uniform cylinder of length L is then given by

$$E_z(z) = \frac{\lambda_b}{2\pi a_0^2 \epsilon_0} \left[\int_{-\frac{L}{2}}^z \left(1 - \frac{z - \xi}{((z - \xi)^2 + a_0^2)^{1/2}} \right) d\xi - \int_z^{\frac{L}{2}} \left(1 - \frac{\xi - z}{[(\xi - z)^2 + a_0^2]^{1/2}} \right) d\xi \right]. \quad (25)$$

The contribution to the radial field from the longitudinal field is given by Gauss's Law. Due to superposition, we can calculate the free space effect of the longitudinal field derivative, which can then be added on to the previously calculated radial field as a correction. We have

$$\frac{\partial E_z}{\partial z} = \frac{\lambda_b}{2\pi a_0^2 \epsilon_0} \left[2 + \frac{z - \frac{L}{2}}{((z - \frac{L}{2})^2 + a_0^2)^{1/2}} - \frac{z + \frac{L}{2}}{((z + \frac{L}{2})^2 + a_0^2)^{1/2}} \right]. \quad (26)$$

At $z = 0$, this becomes

$$\frac{\partial E_z}{\partial z} = \frac{\lambda_b}{2\pi a_0^2 \epsilon_0} \left(2 - \frac{L}{((\frac{L}{2})^2 + a_0^2)^{1/2}} \right). \quad (27)$$

Since this expression has no r dependence, near the center of the beam, we can approximate the field as being linear in r . Comparing this correction coefficient to the expression in Eq. (16), we see that the correction is the magnitude of the factor

$$2 - \frac{L}{((\frac{L}{2})^2 + a_0^2)^{1/2}}. \quad (28)$$

With our beam parameters, this represents a percent-level correction to the space charge field.

Therefore, we see that a simple model of an infinitely long cylindrical shell collapsing into an infinitely long

cylindrical beam produces the correct shape to fix the phase space curvature, even after accounting for complications, such as velocity spread and space charge dynamics.

V. SUMMARY

In summary, we have demonstrated that in simulation a novel approach to nonlinear slice emittance compensation, utilizing the wavebreaking characteristic of nonlinear space charge forces in dense beams to undo the emittance degradation caused by those very same nonlinear forces on the core of the beam. The wavebroken shock front at the edge of the beam is nonlaminarly focused, crashing through the core of the beam and linearizing it in the process. We have found that in simulated photoinjectors utilizing cutting-edge low MTE photocathodes, the employment of this wave-crashing linearization can reduce the emittance by more than a factor of 2.

Furthermore, we have constructed a simple analytic model of the wave-crashing process which agrees well with simulation, showing that the curvature of the linearizing force arises from the large velocity spread inherent to the wavebroken shock front. The extreme nonlinearity at the edge of the beam, instead of being an undesired artifact, becomes critical to the linearization of the core.

The charge forming this nonlinearity can subsequently be discarded by an aperture. While the presence of a physical aperture can exhibit wakefield effects that are beyond the scope of this paper and are the subject of future study, we note that in some applications a physical aperture may not be required. In certain applications of high brightness beams, such as FELs or inverse Compton scattering, the sacrificial charge may not contribute meaningfully to the radiation production and might be ignored without the necessity of physically discarding it.

As higher brightnesses become necessary to push the performance of the next generation of FELs and colliders, the commensurate charge densities at the cathode demand novel techniques of mitigating space charge-induced emittance degradation. We have demonstrated one such technique, providing a potential path to taking full advantage of the next generation of rf guns and photocathode.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Grant No. DE-SC0020144, and U.S. National Science Foundation Grant No. PHY-1549132, the Center for Bright Beams.

-
- [1] L. Serafini and J.B. Rosenzweig, Envelope analysis of intense relativistic quasilaminar beams in rf photoinjectors: mA theory of emittance compensation, *Phys. Rev. E* **55**, 7565 (1997).

- [2] B. E. Carlsten, New photoelectric injector design for the Los Alamos National Laboratory XUV FEL accelerator, *Nucl. Instrum. Methods Phys. Res., Sect. A* **285**, 313 (1989).
- [3] H. Chen, L. Zheng, P. Huang, C. Song, Y. Du, R. Li, W. Huang, and C. Tang, Analysis of slice transverse emittance evolution in a very-high-frequency gun photoinjector, *Phys. Rev. Accel. Beams* **24**, 124402 (2021).
- [4] T. Vecchione, I. Ben-Zvi, D. H. Dowell, J. Feng, T. Rao, J. Smedley, W. Wan, and H. A. Padmore, A low emittance and high efficiency visible light photocathode for high brightness accelerator-based X-ray light sources, *Appl. Phys. Lett.* **99**, 034103 (2011).
- [5] J. Feng, J. Nasiatka, W. Wan, S. Karkare, J. Smedley, and H. A. Padmore, Thermal limit to the intrinsic emittance from metal photocathodes, *Appl. Phys. Lett.* **107**, 134101 (2015).
- [6] S. G. Anderson and J. B. Rosenzweig, Nonequilibrium transverse motion and emittance growth in ultrarelativistic space-charge dominated beams, *Phys. Rev. ST Accel. Beams* **3**, 094201 (2000).
- [7] K. Floettmann, Emittance compensation in split photoinjectors, *Phys. Rev. Accel. Beams* **20**, 013401 (2017).
- [8] R. K. Li, K. G. Roberts, C. M. Scoby, H. To, and P. Musumeci, Nanometer emittance ultralow charge beams from rf photoinjectors, *Phys. Rev. ST Accel. Beams* **15**, 090702 (2012).
- [9] B. S. Zerbe, X. Xiang, C.-Y. Ruan, S. Lund, and P. Duxbury, Dynamical bunching and density peaks in expanding Coulomb clouds, *Phys. Rev. Accel. Beams* **21**, 064201 (2018).
- [10] I. V. Bazarov, B. M. Dunham, and C. K. Sinclair, Maximum achievable beam brightness from photoinjectors, *Phys. Rev. Lett.* **102**, 104801 (2009).
- [11] P. Loschialpo, W. Namkung, M. Reiser, and J. D. Lawson, Effects of space charge and lens aberrations in the focusing of an electron beam by a solenoid lens, *J. Appl. Phys.* **57**, 10 (1985).
- [12] J. B. Rosenzweig *et al.*, An ultra-compact x-ray free-electron laser, *New J. Phys.* **22**, 093067 (2020).
- [13] S. Tantawi, M. Nasr, Z. Li, C. Limborg, and P. Borchard, Design and demonstration of a distributed-coupling linear accelerator structure, *Phys. Rev. Accel. Beams* **23**, 092001 (2020).
- [14] J. B. Rosenzweig, A. Cahill, B. Carlsten, G. Castorina, M. Croia, C. Emma, A. Fukusawa, B. Spataro, D. Alesini, V. Dolgashev, M. Ferrario, G. Lawler, R. Li, C. Limborg, J. Maxson, P. Musumeci, R. Pompili, S. Tantawi, and O. Williams, Ultra-high brightness electron beams from very-high field cryogenic radiofrequency photocathode sources, *Nucl. Instrum. Methods Phys. Res., Sect. A* **909**, 224 (2018).
- [15] A. D. Cahill, J. B. Rosenzweig, V. A. Dolgashev, S. G. Tantawi, and S. Weathersby, High gradient experiments with X-band cryogenic copper accelerating cavities, *Phys. Rev. Accel. Beams* **21**, 102002 (2018).
- [16] D. Filippetto, P. Musumeci, M. Zolotarev, and G. Stupakov, Maximum current density and beam brightness achievable by laser-driven electron sources, *Phys. Rev. ST Accel. Beams* **17**, 024201 (2014).
- [17] R. Akre, D. Dowell, P. Emma, J. Frisch, S. Gilevich, G. Hays, Ph. Hering, R. Iverson, C. Limborg-Deprey, H. Loos, A. Miahnahri, J. Schmerge, J. Turner, J. Welch, W. White, and J. Wu, Commissioning the Linac Coherent Light Source injector, *Phys. Rev. ST Accel. Beams* **11**, 030703 (2008).
- [18] G. Shu, Y. Chen, S. Lal, H. Qian, S. H. Shaker, and F. Stephan, First design studies of a NC CW rf gun for European XFEL, in *Proceedings of the 10th International Particle Accelerator Conference, IPAC-2019, Melbourne, Australia* (JACoW Geneva, Switzerland, 2019), <https://accelconf.web.cern.ch/ipac2019/papers/tuprb010.pdf>.
- [19] S. Karkare, G. Adhikari, W. A. Schroeder, J. K. Nangoi, T. Arias, J. Maxson, and H. Padmore, Ultracold electrons via near-threshold photoemission from single-crystal Cu (100), *Phys. Rev. Lett.* **125**, 054801 (2020).
- [20] A. Bartnik, C. Gulliford, I. Bazarov, L. Cultera, and B. Dunham, Operational experience with nanocoulomb bunch charges in the Cornell photoinjector, *Phys. Rev. ST Accel. Beams* **18**, 083401 (2015).
- [21] C. Gulliford, A. Bartnik, I. Bazarov, L. Cultera, J. Dobbins, B. Dunham, F. Gonzalez, S. Karkare, H. Lee, H. Li, Y. Li, X. Liu, J. Maxson, C. Nguyen, K. Smolenski, and Z. Zhao, Demonstration of low emittance in the Cornell energy recovery linac injector prototype, *Phys. Rev. ST Accel. Beams* **16**, 073401 (2013).
- [22] C. Gulliford, A. Bartnik, I. Bazarov, B. Dunham, and L. Cultera, Demonstration of cathode emittance dominated high bunch charge beams in a dc gun-based photoinjector, *Appl. Phys. Lett.* **106**, 094101 (2015).
- [23] I. V. Bazarov and C. K. Sinclair, Multivariate optimization of a high brightness dc gun photoinjector, *Phys. Rev. ST Accel. Beams* **8**, 034202 (2005).
- [24] C. Gulliford, A. Bartnik, and I. Bazarov, Multiobjective optimizations of a novel cryocooled dc gun-based ultrafast electron diffraction beam line, *Phys. Rev. Accel. Beams* **19**, 093402 (2016).
- [25] S. B. van der Geer and M. J. de Loos, Applications of the general particle tracer code, in *Proceedings of the Particle Accelerator Conference, Vancouver, BC, Canada, 1997* (IEEE, New York, 1997), Vol. 2, Cat. No. 97CH36167, pp. 2577–2579.
- [26] J. M. Maxson, I. V. Bazarov, W. Wan, H. A. Padmore, and C. E. Coleman-Smith, Fundamental photoemission brightness limit from disorder induced heating, *New J. Phys.* **15**, 103024 (2013).
- [27] C. M. Pierce, M. B. Andorf, E. Lu, C. Gulliford, I. V. Bazarov, J. M. Maxson, M. Gordon, Y.-K. Kim, N. P. Norvell, B. M. Dunham, and T. O. Raubenheimer, Low intrinsic emittance in modern photoinjector brightness, *Phys. Rev. Accel. Beams* **23**, 070101 (2020).
- [28] K.-J. Kim, Rf and space-charge effects in laser-driven rf electron guns, *Nucl. Instrum. Methods Phys. Res., Sect. A* **275**, 201 (1989).