Staging of plasma-wakefield accelerators

Carl A. Lindstrøm[*]

Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, 22607 Hamburg, Germany

Reaching high energies with a plasma-wakefield accelerator may require use of multiple stages. This coupling of beams from one stage to another can be nontrivial due to a combination of high divergence and non-negligible energy spreads. In addition, a number of other requirements must be met, such as space-efficient in- and out-coupling of the driver, matching of beta functions, dispersion cancellation, and ensuring synchronization and alignment to within extremely tight tolerances. Satisfying all these requirements takes up space between the stages, which can significantly reduce the effective gradient of the accelerator. We review the challenges, the technical requirements, and some currently proposed methods for solving (or circumventing) the staging problem.

I. INTRODUCTION

Accelerating particles to high energy in a compact space is the main motivation for plasma-wakefield accelerator research. What counts as *high energy* depends on the context—free-electron lasers [1] will require energies around 1–10 GeV, whereas linear colliders [2,3] will typically require 100 GeV or more. The energy gain in a single wakefield-based accelerator stage is ultimately limited by the energy stored in the driver, which is typically of the order of 1–100 J. While a single stage may be sufficient to drive a free-electron laser, it will not be enough to drive a linear collider. In this case, we need to combine the energy of several individual drivers, by distributing these drivers across many separately driven stages chained together—a concept known as *staging*.

Staging is still largely an unsolved problem with no universal solution, although good progress has been made in recent years. With only one dedicated experimental result to date (by Steinke *et al.* [4]), there are still a number of problems to be solved on a conceptual level. In this review, we will explore why staging is so challenging (Sec. II), the detailed technical requirements (Sec. III), and finally some currently proposed methods to get there (Sec. IV).

II. THE STAGING PROBLEM

On the face of it, staging sounds like a rather simple way to reach higher energies. After all, we have been doing it for a century with conventional accelerating cavities: Just put one after the other. However, in plasma-wakefield accelerators, this turns out to be anything but simple—subtle but fundamental effects complicate matters immensely. Before delving into the details of how to do staging, we need to gain an understanding of why it is so difficult in the first place.

Here is the main problem: To accelerate particles with a high gradient, strong focusing is required to maintain stable acceleration. This strong focusing, present in a plasma wake, results in high-divergence beams in the space between the stages—making it difficult to transport and preserve the beam quality of bunches with a finite energy spread. Solving the problem requires even more space for beam optics, resulting in an overall longer accelerator. Ultimately, while the local accelerating gradient may be high, it is the geometrically averaged (or *effective*) accelerating gradient—staging included—that needs to be high for plasma-wakefield accelerators to compete with conventional accelerators.

This section will attempt to answer three questions: Why do the stages need to be separated in space? Why is strong focusing necessary within stages? And why are high-divergence beams so difficult to handle?

A. In- and out-coupling of the driver

A defining feature of a wakefield accelerator is the copropagating driver—the source of energy for the accelerating particles. This driver can be either a relativistic charged particle beam [5,6] (i.e., electrons, positrons, protons or ions) or an intense laser pulse [7], expelling plasma electrons radially via direct Coulomb repulsion or via the ponderomotive force, respectively. The choice of driver does not significantly alter the shape of the wakefield but will affect the propagation of the driver, with regards to
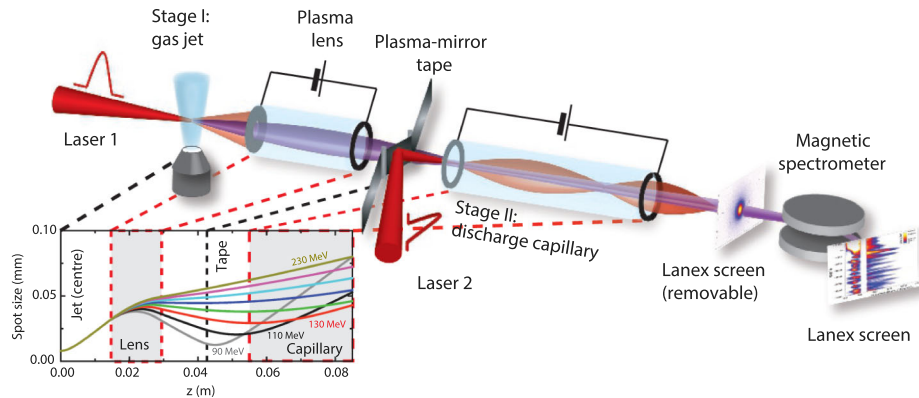
[*]carl.a.lindstroem@desy.de

FIG. 1.   The Berkeley Lab Laser Accelerator (BELLA) staging experiment performed at Lawrence Berkeley National Laboratory. Electrons from a laser wakefield accelerator stage were focused into a second stage using a plasma lens. Simultaneously, a second laser driver was coupled in using a plasma mirror. Electrons with a central energy of 120 MeV were boosted in the second stage by up to 100 MeV. A large energy spread (60% full width at half maximum) resulted in a large chromaticity and, consequently, a low charge-coupling efficiency between the two stages (about 3.5%). Source: Ref. [4].

diffraction and dephasing [8], as well as the overall energy efficiency of the accelerator [9], which is typically significantly higher for beam drivers compared to laser drivers.

If the wakefield-driving process can be sustained for long enough to deplete the driver of its energy, it must be swapped out for a fresh driver. This extraction and reinjection process will inevitably disrupt the wakefield for some time before the next wakefield can be set up. In practice, this corresponds to a gap between the accelerator stages.

The method for swapping out the driver depends on whether a laser or a charged particle beam is used. Two devices are available for making the driver collinear with the accelerating particle beam: mirrors and dipole magnets. Mirrors can be used only for laser drivers, whereas dipole magnets can be used in both laser-driven and beam-driven accelerators.

Mirrors are ideal in that they do not require much longitudinal space along the beam axis—the laser pulse is coupled in and out transversely. A holed mirror can be used to let the beam through undisturbed. However, if placed very close to the stage, the laser intensity is typically too high: A normal optical mirror would burn immediately. One solution is to place the mirror further away from the stage (distances of the order of 0.1–1 m), resulting in no burning but taking up more space. Alternatively, one can use a so-called *plasma mirror* [10]—a thin foil which vaporizes on contact, turning into a solid-density plasma that reflects the laser. Such plasma mirrors were successfully used in the first and currently only staging experiment, performed at the Lawrence Berkeley National Lab [4] (see Fig. 1), placed only approximately 1 cm away from the plasma accelerator. However, preserving the emittance of a beam passing through a plasma mirror can be difficult, due to effects such as beam filamentation (for small beams) [11,12] and multiple Coulomb scattering (for large beams) [13].

Magnetic dipoles bend the path of charged particles and can, therefore, be used to merge or separate a particle beam and a laser beam or merge or separate two different particle beams. The separation induced inside a dipole of magnetic field B and length L is given by

$$\Delta x = \frac{BecL^2}{E} \tag{1}$$

for ultrarelativistic energies E. As an example, for a beam energy of 10 GeV and a magnetic field of 1 T, a 60-cm-long dipole is required to offset the beam by 1 cm. In the case of particle beams, we would ideally use a fast kicker magnet [14]—injecting and extracting the driver bunch while leaving the accelerated bunch undisturbed. However, no kicker exists that can separate bunches at the necessary femtosecond-to-picosecond timescales (the fastest kickers have a nanosecond rise time [15]). Therefore, currently the only way to separate two beams is by energy—i.e., by using beam drivers with lower or higher energy than the accelerated bunch, which makes the required distance even longer. Magnetic dipoles can also be applied to in- and out-coupling of laser drivers, which have no charge and, therefore, behave as if they have a higher (infinite) energy. Unfortunately, the merging and separation process introduces large quality-degrading dispersion to both accelerating beams and beam drivers, which must be carefully canceled—taking up additional space along the beam line for correctional optics.

Whichever method is chosen for in- and out-coupling of the driver, it will introduce a separation between the stages—at least 0.1–1 m per stage, with the exception of plasma mirrors for laser drivers, which may reduce this to centimeter scale. This separation will not only reduce the effective accelerating gradient of a multistage accelerator,
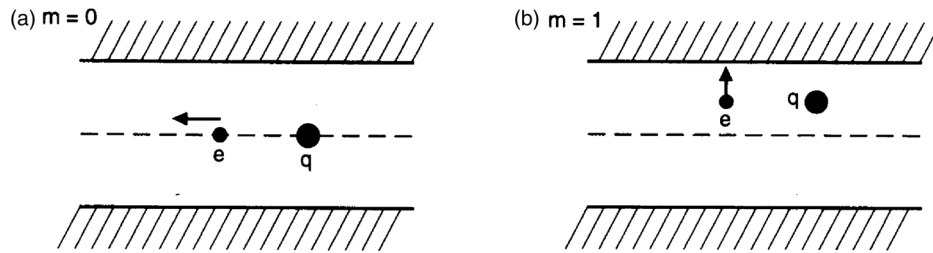
FIG. 2.   Two particle bunches in a generic accelerator structure, exhibiting both (a) a longitudinal wakefield of azimuthal mode $m = 0$ and (b) a transverse wakefield of azimuthal mode $m = 1$. Arrows indicate the direction of the force. Source: Ref. [16].

but also leads to trouble when combined with the strong focusing within each stage.

B. The need for strong focusing

Increasing the accelerating gradient generally requires decreasing the transverse dimensions of the accelerator cavity. This is particularly true for plasma-wakefield accelerators, where the characteristic accelerating field E_z is directly linked to the inverse of the characteristic length scale of the plasma wake—the plasma skin depth $1/k_p$:

$$\frac{eE_z}{m_e c^2} \simeq k_p, \qquad (2)$$

where $k_p = \sqrt{n e^2 / m_e \epsilon_0 c^2}$ is the plasma wave number, n is the plasma density, ϵ_0 and c are the vacuum permittivity and light speed, respectively, and m_e and e are the electron mass and charge, respectively.

In general, while the longitudinal accelerating field makes up the fundamental mode of this wakefield, other modes will also be present [16]—in particular, the transverse wakefield (see Fig. 2). This deflecting force scales with the transverse offset, so, theoretically, if the beam were perfectly aligned on axis, it would be unaffected. However, since the transverse wakefield always pulls the beam away from the axis, a positive feedback loop of larger offsets and larger wakefields is induced for even an infinitesimal initial offset. This instability was discovered in the SLAC linac back in 1960 [17], where transverse wakefields from the beam pipe [18] caused parts of the beam to be lost. An asymmetry in a plasma-wakefield accelerator will have the same effect.

The fundamental challenge is that, when the characteristic cavity dimension (a) decreases, the longitudinal wakefield increases (with a scaling between $1/a$ and $1/a^2$), but the transverse wakefield grows even faster (with a scaling between $1/a^3$ and $1/a^4$). This is dictated by the Panofsky-Wenzel theorem [19] and is always true in the short range (i.e., directly behind the driver), where it is known as the *short-range wake theorem* [20]:

$$\frac{W_x(z)}{\Delta x} \approx -\frac{2}{a^2} \int_0^z W_z(z')dz', \qquad (3)$$

where W_x and W_z are the transverse and longitudinal wakefields (force per charge), respectively, Δx is the beam offset, and z is the comoving longitudinal coordinate.

The good news is that the transverse wakefield instability can be suppressed by introducing a comparatively strong focusing force. If at a certain transverse offset a particle observes a stronger focusing force toward the axis than a transverse wakefield toward the wall, the beam will remain stable. In conventional accelerators like the SLAC linac and Compact Linear Collider (CLIC), this focusing is done with quadrupoles *interleaved* between the accelerator cavities. However, this is not sufficient for a plasma-wakefield accelerator, since the transverse wakefield is too strong [21]—focusing must occur *inside* the accelerating cavity. Fortunately, plasma-wakefield accelerators provide intrinsically strong internal focusing; an exposed ion column (in the case of a nonlinear wake) focuses the electron beam with a focusing strength (proportional to force per offset) of

$$K = \frac{k_p^2}{2\gamma}, \qquad (4)$$

where γ is the relativistic Lorentz factor. In wakefield accelerators without on-axis focusing, such as dielectric-wakefield [22] or hollow-plasma-channel accelerators [23], the proposal is to use strong external focusing (quadrupoles) [24]. Regardless of the focusing mechanism, *in situ* focusing is generally required for stable operation of high-gradient wakefield accelerators.

Finally, while strong focusing is required, it can still lead to a so-called *beam-breakup instability* [25]—or, similarly, a *hose instability* in a plasma accelerator [26,27]. This is caused by a resonance between different longitudinal slices of the bunch when oscillating in the focusing field. To avoid this effect, the slices must be decohered to oscillate at different frequencies. This is normally done by giving the bunch a head-to-tail energy chirp (a longitudinally correlated energy spread)—a method known as *BNS damping* after Balakin, Novokhatsky, and Smirnov [28]. Using a simple two-particle model of BNS damping and a

simple estimate of the transverse wakefield in a plasma accelerator [29]

$$\frac{d}{dz}\left(\frac{W_x}{\Delta x}\right) \approx \frac{k_p^4}{4\pi\epsilon_0}, \tag{5}$$

we can estimate the relative chirp required to cancel the transverse wakefield to be approximately

$$\frac{d}{dz}\left(\frac{\Delta\gamma}{\gamma}\right) \approx -\frac{eQk_p^2}{2\pi\epsilon_0 m_e c^2}, \tag{6}$$

where Q is the estimated charge of the front particle (representing the head of the bunch). As an example, for a plasma accelerator of density 10^{16} cm^{-3} and bunch of charge 100 pC ($\sim$50 pC in the head) and length 50 μm, an energy chirp of roughly 3% is required from head to tail. The implication is that stable acceleration benefits from a nonzero energy spread [30].

An alternative method proposed to suppress the beam-breakup instability is to make use of ion motion inside the wake to induce nonlinear focusing [31] to avoid the resonance condition, albeit at the cost of direct emittance growth [32,33].

C. Chromaticity

Strong focusing results in small beam sizes but, more importantly, highly diverging beams. In terms of Courant-Snyder or Twiss parameters [34], this means small beta functions. To avoid emittance growth in a stage, the beta function must be *matched* to

$$\beta_m = \frac{1}{\sqrt{K}} = \frac{\sqrt{2\gamma}}{k_p} = \sqrt{\frac{2\epsilon_0 E}{ne^2}}, \tag{7}$$

where K is the focusing strength and E is the beam energy—in this case, the natural divergence of the beam is exactly countered by the focusing field such that the beta function (and, hence, the beam size) stays constant. For a plasma accelerator, the matched beta function is typically on the millimeter-to-centimeter scale for GeV-level beams. In addition, plasma-wakefield acceleration often results in non-negligible energy spread because of the rapidly changing (high-frequency) accelerating field structure—the energy spread is often 1% or more. While this may be good for BNS damping, these beams are difficult to capture and refocus without degrading the beam quality, because the different energy slices are not all focused in the same way [35–37]—an effect known as *chromaticity*.

The chromaticity of the beam focusing is typically defined in terms of the *chromatic amplitude* [38,39]

$$W = \sqrt{\left(\frac{\partial\alpha}{\partial\delta} - \frac{\alpha}{\beta}\frac{\partial\beta}{\partial\delta}\right)^2 + \left(\frac{1}{\beta}\frac{\partial\beta}{\partial\delta}\right)^2}, \tag{8}$$

which measures (to first order) the combined mismatch of the Twiss parameters α and β, for a relative energy offset $\delta = \Delta E/E$. This chromatic amplitude can be related to the projected (energy-averaged) emittance growth via [40]

$$\frac{\Delta\epsilon^2}{\epsilon_0^2} = W^2\sigma_\delta^2 + \mathcal{O}(\sigma_\delta^4), \tag{9}$$

expressed to lowest order in σ_δ (the relative rms energy spread). It is important to note that this projected emittance growth is not a "true" emittance growth, since the emittance of each energy slice is conserved—it can, therefore, in principle, be reversed. However, when entering the next stage and observing strong focusing, such a phase space reversal is practically impossible.

So how large do we expect the emittance growth to be? Consider a simple example of staging: A beam of energy spread σ_δ exits a stage with a beta function β_m and then gets refocused by a (thin) magnetic optics lens after a drift of length L (see Fig. 3 for illustration). In simple cases like this, the chromaticity added in the lens is approximately $\Delta W = \beta Kl$, where Kl is the integrated focusing strength and β is the beta function in the lens. To capture and refocus the beam each in a distance L, the focal length of the lens must be $L/2$—resulting in an integrated focusing strength
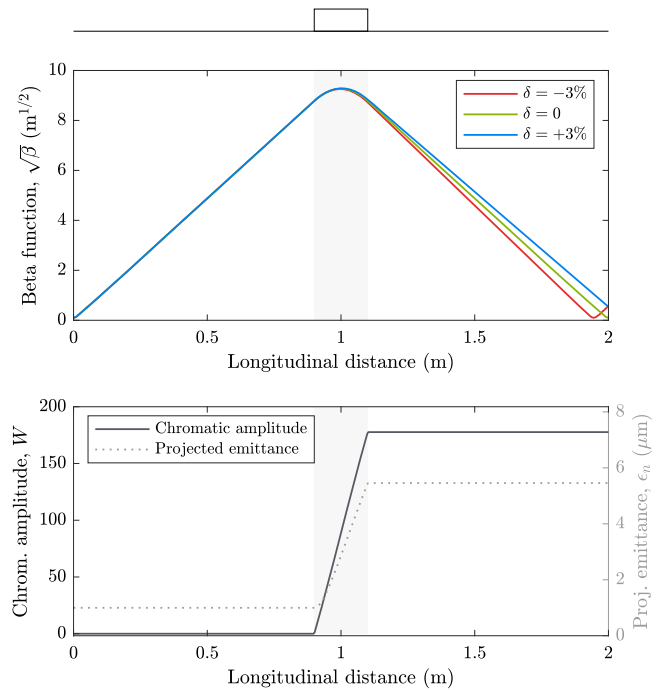

FIG. 3. Example of emittance growth due to chromaticity. A 10 GeV beam with 3% rms energy spread diverges from a plasma accelerator of density 10^{16} cm^{-3} ($\beta_m \approx 10$ mm). A simple beam optics lens captures and refocuses the beam into the next stage, which introduces significant chromaticity. As a result, the projected (energy-averaged) emittance increases by more than a factor of 5.

of $Kl = 2/L$. Outside the stage, the beam will diverge to a beta function $\beta \approx L^2/\beta_m$, assuming a small matched beta function (i.e., $\beta_m \ll L$). Putting it all together, we find that the chromaticity is $W \approx 2L/\beta_m$, and, therefore, the projected emittance growth will be approximately

$$\frac{\Delta\epsilon^2}{\epsilon_0^2} \approx \frac{4L^2}{\beta_m^2}\sigma_\delta^2. \tag{10}$$

This sets strict limits for the acceptable energy spread. Take, for instance, a plasma accelerator stage at energy $E = 10$ GeV with plasma density $n = 10^{16}$ cm^{-3} (giving $\beta_m = 10$ mm), using a capture length of $L = 1$ m and limited to an emittance growth of 1%—the maximum energy spread is only 0.07% rms.

In conclusion, chromaticity places severe constraints on the staging of plasma-wakefield accelerators if left uncorrected. Moreover, while it may be possible to correct this chromaticity, it will likely take up significant space between stages. This, in combination with in- and out-coupling drivers, can significantly reduce the effective gradient of a multistage plasma accelerator. Ultimately, the challenge will be to outperform conventional rf accelerators and remain in the GV/m range.

III. TECHNICAL REQUIREMENTS

While chromaticity is perhaps the biggest challenge, there are also many other considerations to keep in mind when designing a coherent staging scheme. In this section, we will review some of the most important requirements in detail.

A. Emittance preservation

Delivering low-emittance beams is of prime importance for most high-energy accelerator applications. In a linear collider, the *luminosity*—proportional to the collision rate—is given by [3]

$$\mathcal{L} = H_D \frac{N^2 f\gamma}{4\pi\sqrt{\beta_x\epsilon_{nx}}\sqrt{\beta_y\epsilon_{ny}}}, \tag{11}$$
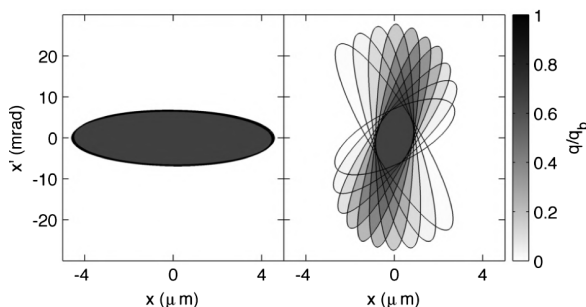
where N is the number of bunch particles, f is the collision frequency, H_D is a numerical factor, β_x and β_y are the interaction point beta functions, and ϵ_{ny} and ϵ_{ny} are the normalized transverse emittances. Similarly, in a free-electron laser (FEL), the lasing power is determined by the *6D brightness* [41]

$$B_{6D} = \frac{N}{\epsilon_{nx}\epsilon_{ny}\epsilon_{nz}}, \tag{12}$$

where ϵ_{nz} is the normalized emittance of the longitudinal phase space (proportional to the longitudinal-phase-space area). Equations (11) and (12) indicate that low emittance is crucial—FELs require normalized transverse emittances of the order of 0.1–1 mm mrad, whereas linear colliders may require emittances as low as 0.01 mm mrad (in one plane). Technically, conservation of charge (N) is equally important, and in practice we will require close to 100% charge-coupling efficiency between stages.

Producing emittances of the order of 1 mm mrad is routinely done using photocathodes [42] and even plasma-injection techniques [43–45], whereas reaching emittances of the order of 0.1 mm mrad or lower (with non-negligible charge) currently requires the use of a damping ring [46]. Preserving this emittance through a large number of stages will set very stringent limits on the emittance growth per stage. It can be useful to imagine an *emittance budget*, where each stage gets to contribute only a certain emittance growth to the final emittance. For plasma-wakefield accelerators with many stages, this growth will therefore be limited to 0.001–0.1 mm mrad or less per stage.

1. Matching

Mismatching bunches with a finite energy spread leads to emittance growth, because the phase space ellipses of different energy slices rotate at different rates—eventually smearing out to a larger area. The end result is a matched beam with a larger emittance (see Fig. 4 for illustration). This process stops (or *saturates*) after a distance $L_{sat} \approx \beta_m/\sigma_\delta$, after which the saturated emittance will be [47]
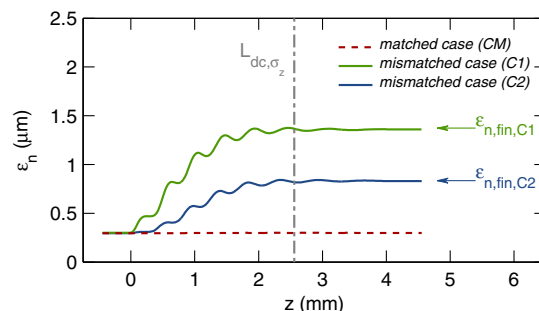

FIG. 4. Mismatching of a beam with finite energy spread. The initial phase space (left panel) is mismatched to the focusing channel, which leads to a smearing in phase space due to the different rates of rotation. This effect is seen to saturate at some point (right panel), when the beam has been fully smeared in phase space, leading to emittance growth. Source: Ref. [47] (CC BY 3.0).

$$\frac{\epsilon_{\text{sat}}}{\epsilon} = \frac{1}{2}\left((1+\alpha^2)\frac{\beta_m}{\beta} + \frac{\beta}{\beta_m}\right), \qquad (13)$$

where α and β are the Twiss parameters at the plasma entrance (assuming a flattop density profile). As seen in Eq. (13), the only way to avoid emittance growth from mismatching is to ensure that $\beta = \beta_m$ and $\alpha = 0$.

Technically, this is also why chromaticity is problematic—if the central energy slice is matched, but other energy slices are not, this will result in an emittance growth. Hence, to avoid mismatching, chromaticity between stages must be canceled.

2. Dispersion cancellation

When dipoles are used for in- and out-coupling of drivers, a correlation between energy and position—*dispersion*—is intentionally introduced to separate beams of different energy. This also disperses the accelerating beam if it has a nonzero energy spread, introducing a projected emittance growth. We can estimate this emittance growth in a stage from an uncorrected (first-order) dispersion to be

$$\Delta\epsilon_D \approx \frac{1}{2}\left(\frac{D_x^2}{\beta_m} + \beta_m D_{x'}^2\right)\sigma_\delta^2, \qquad (14)$$

where D_x is the dispersion, $D_{x'}$ is the dispersion prime (i.e., energy-angle correlation), and β_m is the matched beta function in the stage. Using this relation, we can obtain approximate limits for (first-order) dispersion: $D_x \ll \sqrt{2\epsilon\beta_m}/\sigma_\delta$ and $D_{x'} \ll \sqrt{2\epsilon/\beta_m}/\sigma_\delta$, where ϵ is the geometric emittance of the beam. As an example, a 1 GeV beam with 1% energy spread and 1 mm mrad normalized emittance staged between plasma accelerators of density 10^{16} cm^{-3} will require dispersion and dispersion-prime cancellation to much better than 0.18 mm and 55 mrad, respectively—this can be quite challenging. We should also note that, given the large dispersion often introduced in strong dipoles, we may also need to consider higher-order dispersion.

Moreover, dispersion can cause additional problems beyond just an increased projected emittance. If the longitudinal phase space of the bunch has a correlation—as it often does—a dispersion implies that the bunch has a tilt and/or a curvature. Such an asymmetry can seed a beam-breakup [25] or hosing instability [26,27], which can lead to more severe emittance growth.

3. Coulomb scattering

Plasma accelerators require on-axis ionized gases or vapors. This inevitably causes some emittance growth by Coulomb scattering—random collisions between beam particles and atoms or ions. The rate of emittance growth is given by [48–50]

$$\frac{d\epsilon_n}{ds} \approx \frac{2\pi r_e^2 \beta_x}{\gamma}\left[n_i Z_i^2 \ln\Lambda + 1.78 n_0 Z(Z+1)\ln\left(\frac{287}{\sqrt{Z}}\right)\right], \qquad (15)$$

where n_i and n_0 are the ion and neutral-atom densities, respectively, Z_i is the average ionization state of the ions, Z is the atomic number, β_x is the Twiss beta function, γ is the relativistic Lorentz factor, $\ln\Lambda$ is the Coulomb logarithm, and r_e is the classical electron radius.

In a typical plasma-accelerator stage, the emittance growth is negligible—this is because, while the density may be high, the beta function is very small due to strong focusing. However, this is not the case between stages: Here, the beta function increases rapidly, placing constraints on the density of ions and neutral atoms, especially for high-Z gas species. To avoid emittance growth between stages, differential pumping can be used to reduce the gas density outside the stage, but care must be taken to avoid transverse wakefields from small-aperture beam pipes. In addition, rapid beam capture will help in reducing the maximum beta function as well as the overall distance where scattering can occur. Lastly, use of vacuum windows and plasma mirrors can also cause significant scattering, which may make them unviable for use in low-emittance plasma accelerators (depending on the material thickness and density, as well as the beam energy).

B. Isochronicity

During acceleration in the plasma wake, one rarely has to worry about changes to the bunch length—it is effectively conserved. However, outside of the stage, there are multiple ways in which the bunch can be lengthened or compressed. This can be detrimental to beam loading and energy-spread conservation [51,52].

If dipoles are used to separate a driver and an accelerating bunch, perhaps in the form of a chicane, particles of different energy may travel different distances before arriving at the next stage—this is the idea behind a bunch compressor. In technical terms, we talk about the R_{56} matrix element of the accelerator lattice, also known as the *longitudinal dispersion* (R_{16} and R_{36} are the horizontal and vertical dispersions, respectively). For the bunch length to be conserved during staging, we will require that $R_{56} = 0$, in which case the lattice is called *isochronous*. We can place a limit on this condition for a bunch of a given energy spread σ_δ and bunch length σ_z:

$$|R_{56}| \ll \frac{\sigma_z}{\sigma_\delta}. \qquad (16)$$

In special cases, it can be beneficial to not cancel R_{56} completely—especially if combined with chirped accelerator stages and symmetric current profiles. If one stage produces a chirped bunch (e.g., the front particles have lower energy), a carefully tuned R_{56} can overcompress the
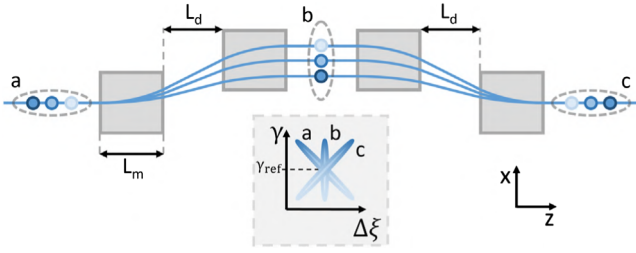
FIG. 5. Longitudinal phase space reversal in a chicane with a fine-tuned R_{56}, which can be used to significantly reduce the energy spread by compensating the chirp from one stage in the following stage. Source: Ref. [53].

bunch in such a way that the bunch length is conserved, but the longitudinal phase space is flipped. In this case, the next stage will exactly dechirp the bunch to give it a significantly lower energy spread (see Fig. 5) [53].

A different but related problem can occur even if the longitudinal dispersion is canceled and the energy spread is negligible. If the bunch is very short (μm-scale) and the divergence is large—as is often the case in laser plasma accelerators—particles with a large angle will travel further before being focused back to a small beam size (related to the R_{25} and R_{45} matrix elements). To avoid bunch lengthening, the distance to the first or last optic before and after a stage is therefore restricted by

$$\Delta L \approx \frac{1}{2}\sigma_{x'}^2 L \ll \sigma_z, \qquad (17)$$

where ΔL is the path length difference for particles at an angle of $\sigma_{x'}$ (i.e., the rms divergence) and L is the distance to the optic. Normally, this is a problem only for lower energy bunches (sub-GeV), because the matched beta function is smaller ($\sim\sqrt{\gamma}$) and the geometric emittance is higher ($\sim 1/\gamma$)—both leading to higher divergence. For example, a 200 MeV bunch of length 1 μm and normalized emittance 1 mm mrad exiting a laser plasma accelerator at density 10^{17} cm^{-3} will have a divergence of 2.3 mrad and, therefore, must, according to Eq. (17), be captured in much less than 37 cm.

C. Tolerances

All of the above considerations have assumed that the staged accelerator is perfectly stable. This is, of course, not the case in practice—everything has a certain level of random jitter. Two particularly important tolerances for jitter are those related to the synchronization and the transverse misalignment between the driver and the accelerating beam.

1. Synchronization

Plasma-wakefield accelerators have high-frequency electromagnetic fields. For stable acceleration, the driver

and the accelerating beam must be synchronized to within a small fraction of the wakefield period. Consider an accelerator with a wakefield that changes from 0 to E_z in a time $1/\omega$—an error in the relative arrival time Δt will result in a relative error of the accelerating gradient of

$$\frac{\Delta E_z}{E_z} \approx \omega\Delta t. \qquad (18)$$

Random timing jitter, therefore, results in a corresponding energy jitter—an effective multishot energy spread. As an example, to maintain a 1% energy stability in a plasma accelerator stage operating at density 10^{17} cm^{-3} (characteristic timescale $1/\omega_p = 177$ fs), one would need to synchronize the driver and the accelerating beam to better than 2 fs. This is a very challenging goal—current state-of-the-art techniques can provide synchronization jitter down to about 10 fs rms [54,55]. In addition, long-term timing drifts need to be measured and corrected for with a feedback system operating at the same timescales.

2. Transverse misalignments

Misalignment tolerances will also prove particularly challenging for plasma-wakefield accelerators. The accelerating beam, the driver, and the accelerating structure all need to be well aligned throughout the full length of the accelerator. Fortunately, this problem is partially mitigated in a plasma accelerator, since the driver defines the location of the accelerating structure—only the driver–accelerating beam offset matters. Assuming instabilities such as hosing and beam breakup can be mitigated, what level of emittance growth do we expect?

Similar to the cases of mismatching (see Sec. III A 1) and dispersion (see Sec. III A 2), beams with finite energy spread will see the centroids of different energy slices rotate in phase space at different rates—leading to a smearing in phase space. Consider a driver–accelerating beam pair with a relative position offset Δx and an angle offset $\Delta x'$, propagating in a stage with a matched beta function β_m (see Fig. 6 for illustration). The projected emittance growth caused by such an offset will gradually increase along the accelerator and then saturate at [56]

$$\Delta\epsilon_j \approx \frac{1}{2}\left(\frac{\Delta x^2}{\beta_m} + \beta_m\Delta x'^2\right). \qquad (19)$$

In simple terms, the driver and the accelerating beam must overlap well in phase space: $\Delta x \ll \sqrt{2\beta_m\epsilon}$ and $\Delta x' \ll \sqrt{2\epsilon/\beta_m}$, where ϵ is the geometric emittance of the accelerating beam. This is particularly challenging for small β_m (i.e., high accelerating gradients) and lower emittances, as the beam is focused to the submicrometer scale. If multiple stages are used, the emittance growth per stage is further constrained—typically by a factor of
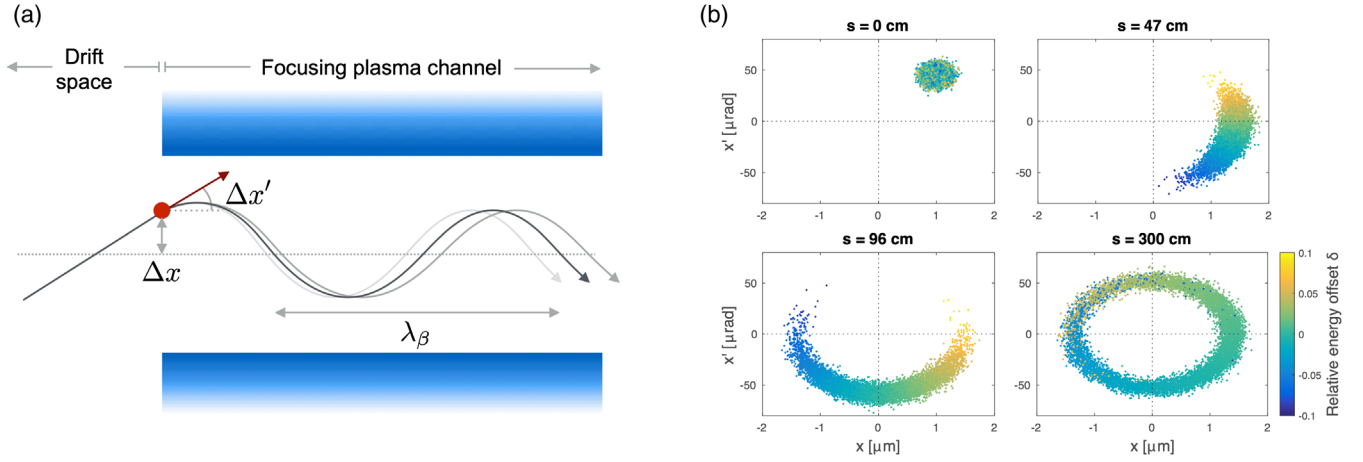
(a)

(b)

FIG. 6.   (a) Illustration of projected emittance growth due to misalignments in a strong-focusing channel. (b) The accelerating bunch starts with an offset and an angle, which due to a finite energy spread smears out to a larger-area ring in phase space. Source: Ref. [56] (CC BY 3.0).

$1/\sqrt{N_s}$, where N_s is the number of independent stages. For plasma accelerators with relevant gradients (GV/m scale), the alignment tolerance is estimated to be around 10–50 nm and 1–5 μrad [57,58]—a fraction of the beam size and divergence in the plasma, respectively. By implication, this is also how well, for instance, a laser in-coupling mirror needs to be aligned.

Furthermore, while the single-bunch emittance is the relevant quantity for a free-electron laser, where the brightness is inversely proportional to the emittance, this is not the case in a particle collider. Instead, since bunches not only need to be very small at the interaction point, but also need to hit the colliding bunches head on, the *multibunch emittance* (i.e., averaged over many shots) is more important. This means that any error in the orbit introduced along the accelerator—induced by a relative positional or angular offset between the accelerating beam and the driver—will need to be suppressed. For linear-collider parameters, this can impose even more stringent alignment tolerances, as low as the nanometer and nanoradian scale [59]. This poses a serious challenge that may require the introduction of dedicated cancellation techniques (e.g., in the optics between stages) to be overcome.

D. Synchrotron radiation

As discussed in Sec. II A, one of the main methods to separate the driver and the accelerating beam is to disperse them using a magnetic dipole—possibly the only way for a beam-driven wakefield accelerator. When accelerating to higher energies—the objective of staging—we will naturally hit the same problem as we tried to avoid by using a linear accelerator (as opposed to a circular accelerator): *synchrotron radiation*. For long bunches, particles emit *incoherent* synchrotron radiation (ISR), with an average power emitted per bunch [60,61]

$$P_{\mathrm{ISR}} = \frac{e^4}{6\pi\epsilon_0 m^2 c} N\gamma^2 B^2, \tag{20}$$

where m and e are the particle mass and charge, respectively, c and ϵ_0 are the vacuum light speed and permittivity, respectively, γ is the relativistic Lorentz factor, B is the magnetic field, and N is the number of particles.

However, plasma-wakefield accelerators often operate with short bunches, in which case the bunch will emit *coherent* synchrotron radiation (CSR) [62,63]. In the case of full coherence, the electric fields of all particles add linearly such that the radiated power scales quadratically with particle number

$$P_{\mathrm{CSR}} = NP_{\mathrm{ISR}} \sim N^2. \tag{21}$$

The radiation is fully coherent only if the bunch length is less than $\sigma_{\mathrm{SR}} = \rho/\gamma^3 = mc/Be\gamma^2$, where ρ is the bending radius of the magnetic field. Conversely, the radiation will be completely incoherent if the bunch length is longer than $\sigma_{\mathrm{SR}}N^{3/4}$. In the intermediate, partially coherent regime, the radiated power depends critically on the current profile of the bunch, such that simulations are normally required for an accurate prediction of the effect. However, for a longitudinally Gaussian bunch, we can estimate the emitted power to be

$$P_{\mathrm{CSR}} = \frac{\kappa e^2 c}{\epsilon_0} \frac{N^2}{\rho^{2/3}\sigma_z^{4/3}}, \tag{22}$$

where $\kappa \approx 0.0279$ is a numerical factor. Consider the example of a 10 GeV bunch of length 10 μm rms and charge 1 nC being bent by a magnetic field of 1 T—the radiation would be partially coherent, and the bunch would radiate about 0.3% of its energy per meter of dipole.
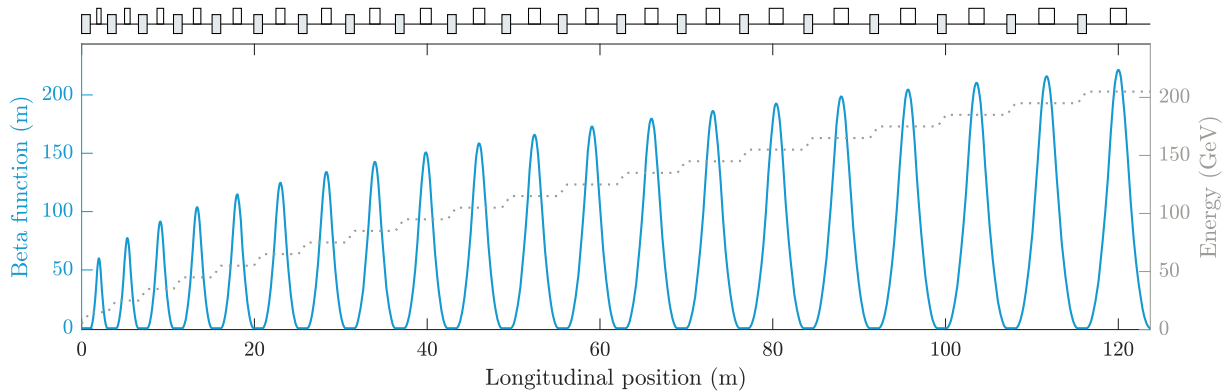
FIG. 7. Illustration of how the staging length increases with energy: The accelerator stages (gray boxes) remain the same length, whereas the length of the staging optic (white boxes) increases as $\sqrt{\gamma}$. As a consequence, the effective gradient decreases with energy. In this example, the initial beam energy is 5 GeV, and each stage adds 10 GeV over 1 m (i.e., 10 GV/m). The beam is matched to a plasma density 10^{16} cm^{-3} by an optic with a magnetic field gradient 240 T/m, which takes up 25% of the space between stages. Accelerating by 200 GeV across 20 stages requires approximately 120 m in total (1.7 GV/m)—a sixfold reduction in the average accelerating gradient compared to the in-plasma gradient.

While this is not dramatic, energy spread and chirp may be induced in a multidipole chicane.

Equally important to the energy loss are the *transverse CSR effects*—strong intrabunch transverse wakefields that can lead to significant emittance growth [64]. Highly sensitive to the full 6D phase space of the bunch, these effects normally need to be studied using numerical simulations. Fortunately, with careful chicane design, transverse CSR effects can be partially mitigated [65]. Lastly, CSR can also lead to a microbunching instability, resulting in emittance growth as well as noise in the current profile of the bunch [66,67].

E. Effective gradient

With so many constraints to satisfy, how much space will be needed between stages? What is the effective gradient of the accelerator? If, say, 10 m of space is required between 10-cm-long stages operating at 5 GV/m, the effective gradient would be only 50 MV/m—not much better than a conventional machine. Although the ultimate metric will probably be the *cost per collision* (linear collider) or the *cost per photon* (free-electron laser) over the lifetime of the machine, the effective gradient is an approximate metric for the construction cost—it estimates the overall footprint of the machine. Staging will, therefore, be a crucial part of the optimization process. The current goal of the advanced accelerator community is to demonstrate an effective gradient of at least 1 GV/m [68].

A potentially worrying aspect of staged plasma-wakefield accelerators is how the staging length scales with energy [69]. To capture and refocus the highly divergent beams, increased focusing strength will be required as the beam energy increases stage by stage. If we already use the strongest focusing optics available, the only solution is to make the optics *longer*. How much

longer? The staging length turns out to scale with $\sqrt{\gamma}$—the square root of the energy—i.e., slowly increasing with energy. This means that the effective gradient will go down as the energy goes up.

To understand this $\sqrt{\gamma}$ scaling, let us consider a simple beam optic between two stages (see Fig. 7). Its focusing strength reduces with energy as $K \sim 1/\gamma$. Using a constant-length optic between stages would result in a focal length (and, therefore, staging length) scaling of $f \approx 1/Kl \sim \gamma$. However, if the length of the optic instead scales as $l \sim \sqrt{\gamma}$, so does the focal length: $f \sim \sqrt{\gamma}$. Since both the focal length and the optic length scale similarly, the optic will always take up a constant fraction of the staging length. Furthermore, the matched beta function also scales as $\beta_m \sim \sqrt{\gamma}$, and, therefore, the entire optics solution (the evolution of the beta function) will scale as $\sqrt{\gamma}$—i.e., if you have found an optics solution for one energy, you have found it for them all.

F. Beam diagnostics

Based on many of the above requirements, it is clear that a plasma accelerator requires high-precision tuning, which, in turn, requires highly sensitive diagnostics. This may increase the distance between stages, further reducing the effective gradient. A number of diagnostics will be required—here, we will consider only diagnostics needed between stages (for in-plasma diagnostics, see Ref. [70]). (1) First and foremost, beam-position monitors (BPMs) are required for measurement of the beam orbit. For submicrometer precision, these will likely need to be cavity-based BPMs [71]. (2) The energy spectrum of the beam exiting the plasma is crucial to the operation of downstream stages and so likely needs to be measured accurately after each stage. This typically requires the beam to be dispersed with

a dipole magnet, ideally in combination with point-to-point beam imaging for improved energy resolution. (3) The current profile of the accelerated bunch determines its beam loading and, therefore, its energy spread and acceleration efficiency. The current profile can be measured with femtosecond-scale resolution using coherent transition radiation spectroscopy [72] or a transverse deflecting structure (TDS) [73]. Using a TDS (in combination with a dipole) also allows for measurement of the longitudinal phase space, although this setup can take up significant space. (4) Emittance measurements require imaging of the beam's transverse profile. This can be combined with an imaging spectrometer for spectrally resolved emittance measurements. Screen resolution will be an important consideration: Low-emittance beams are very small inside the plasma accelerator (nanometer-to-micrometer scale), which necessitates large beam magnification during imaging—suitable magnetic optics may take up considerable space. An alternative may be to measure emittance via the spatial distribution of x rays radiated by the beam inside the plasma accelerator (betatron radiation) [74].

Ideally, all these measurements should be performed continuously and nondestructively in order to perform online feedback and correction. For some diagnostics, this is not possible, requiring instead an intermittent sampling only. For other techniques, the measurement is too space inefficient to be performed after every stage. It will, therefore, be vital to strike the right balance between measurement resolution and space efficiency, as well as developing new and better methods to diagnose the accelerated beam.

G. Driver distribution and dumping

The scheme used for distributing and dumping the driver may appear peripheral, but it can greatly influence the staging solution required. Clearly, the schemes will vary drastically with the choice of driver.

For laser drivers, in- and out-coupling can, in principle, be very space efficient if it can be done transversely using mirrors (as discussed in Sec. II A). The driver distribution is, therefore, relatively straightforward and can be branched off from a central laser production complex (good for synchronization) or a number of smaller complexes (less overall laser transport) [75]. Dumping the laser energy after out-coupling may require special attention, given the potentially large amount of laser power. To recover the energy of either the partly depleted laser or the remnant wakefield (by frequency upshifting a trailing laser pulse), photovoltaic laser dumps have been proposed [76,77].

For beam drivers, the distribution scheme is more critical: Using a higher drive-beam energy reduces the number of stages but also increases the size of the production and distribution complex, as well as the in- and out-coupling distance. A conventional beam-production facility, like that
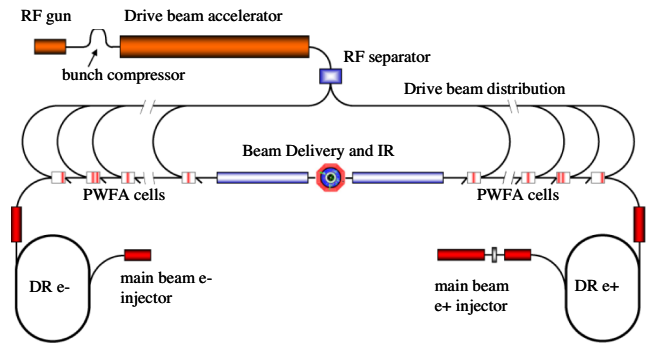

FIG. 8. Driver distribution scheme using U turns in a beam-driven plasma accelerator linear-collider concept. Source: Ref. [78] (CC BY 3.0).

developed for CLIC [3], can have a high wall-plug-to-beam-driver efficiency (5%–10%) and is, therefore, considered as the starting point for any future staged beam-driven plasma-accelerator facility. In this case, the challenge is to synchronize the accelerating bunch sequentially with each bunch in the driver train (with a nanosecond-scale spacing)—clever kicker and delay design is required. A number of conceptual designs have been proposed, including the use of U turns [78] (see Fig. 8) and an undulating delay chicane [79], as well as branching arcs [59,80]—all striking a different balance of tunnel length, synchrotron radiation, and delay time. While some concepts exist for handling the depleted drive beam [81,82], little thought has gone into how to safely transport and dump a high-power, 100%-energy-spread beam without causing a disruptive amount of irradiation.

IV. PROPOSED METHODS

Faced with all the above challenges and requirements, it is clear that we need to innovate in order to succeed. Several new ideas are attempting to either tackle the problems head on or to circumvent them by avoiding staging altogether. This section will go through a few of the most notable proposals so far.

A. Plasma-density ramps

An obvious solution to the strong-focusing conundrum is to simply reduce the strength of the focusing and thereby increase the matched beta function—in a plasma accelerator, that corresponds to reducing the plasma density [see Eq. (7)]. This mitigates the divergence and chromaticity problem but comes at the cost of reducing the accelerating gradient. However, the plasma density does not need to be the same everywhere—we can use a higher density throughout the middle of the stage for high-gradient acceleration and a lower density at the entrance and exit for reduced beam divergence. This longitudinal density tailoring is often called a *plasma-density ramp* [83–85]. Figure 9 illustrates this concept.
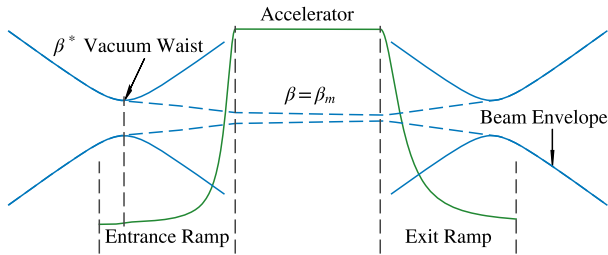
FIG. 9. Schematic of a plasma accelerator with density ramps. The beam is focused by external focusing to the vacuum waist (β^*) close to the start of the entrance ramp. If perfectly matched to the ramp, the beam stays matched throughout the accelerator ($\beta = \beta_m$ in the flattop). Finally, the beam is transported through the exit ramp, which reduces the divergence before exiting into the outside vacuum. Source: Ref. [85] (CC BY 4.0).

Calculating the evolution of the beta function through a tailored plasma-density profile is relatively straightforward via the *betatron equation* [86]

$$\frac{1}{2}\beta''(s)\beta(s) - \frac{1}{4}\beta'(s)^2 + K(s)\beta(s)^2 = 1, \quad (23)$$

where the focusing force is given by

$$K(s) = \frac{e^2}{2\epsilon_0}\frac{n(s)}{E(s)}. \quad (24)$$

Both the exposed charge density $n(s)$ and the particle energy $E(s)$ are changing with the longitudinal position s. While it is possible to solve Eq. (23) analytically in certain cases [87], in general, it needs to be integrated numerically.

It is important to note that, although most plasma-density ramps can successfully reduce the divergence, they do not necessarily solve the chromaticity problem—while one energy slice might be matched and emittance preserved, this does not mean that every energy slice will. To ensure that all energies remain matched throughout the accelerator stage, the ramps must be *adiabatic* [88,89]. This means that the plasma density is changing sufficiently slowly,

$$\left|\frac{n'(s)}{n(s)}\right| \ll \frac{1}{\beta_m(s)}, \quad (25)$$

such that $\alpha \approx 0$ throughout the entire ramp—true for all energy slices, making the adiabatic ramp quasiachromatic. An example of such a ramp would be $n(s) = n_0(1 + s/l_r)^{-2}$, where n_0 is the flattop density and $l_r \gg 2\beta_{m0}$ is the characteristic ramp decay length (β_{m0} is the matched beta function in the flattop). To reach a beta function β^* at the entrance or exit of such an adiabatic ramp, the overall ramp length must be $L_r \gg 2\beta^*$. As an example, to avoid significant emittance growth from chromaticity in a meter-scale gap between stages [see Eq. (10)], the ramp entrance or exit beta function must

be around 100 mm or more for a beam of 1% rms energy spread, resulting in a meter-scale plasma-density ramp. Therefore, while adiabatic ramps are desirable due to their insensitivity to energy spread or slight mismatching, this comes at the price of significantly longer ramp sections. Long ramps will reduce both the energy efficiency and the effective gradient of the accelerating structure and can introduce potentially non-negligible and nonuniform decelerating fields, which must be compensated for in the main accelerating section.

In short, plasma-density ramps constitute a crucial tool for reducing high divergence and chromatic effects but will likely not be able to entirely solve the problem.

B. Plasma lenses

Another useful tool for capturing and refocusing beams between stages is the *plasma lens*—a charged-particle optics device that provides strong focusing in both planes simultaneously (as opposed to the quadrupole). Plasma lenses fall into two categories: *passive* and *active* plasma lenses, referring to whether the focusing force is externally (i.e., actively) driven or not.

1. Passive plasma lenses

Passive plasma lenses utilize the same mechanism as plasma-density ramps—plasma-wakefield focusing. The wakefield can be driven either by the beam itself or by a separate driver (laser or particle beam). Typically, such lenses can provide very strong focusing fields—in the MT/m range—and so can be made very compact. While the concept dates all the way back to 1922 [90–93], passive plasma lenses in their modern form were proposed in 1989 [94] and have been successfully demonstrated for both beam drivers [95,96] and laser drivers [97]. Figure 10 illustrates how a passive plasma lens might be applied in practice.

Two regimes are often identified: the *underdense* and the *overdense* regime, referring to whether the plasma density in the lens is lower or higher than the beam density, respectively. If the lens is underdense, a nonlinear plasma wakefield (a blowout) forms with a fully exposed ion column—this provides a linear focusing force and is therefore, in principle, emittance preserving. Although the focusing force for electrons is exerted by an electric field, we can calculate the equivalent magnetic field gradient (as $E_r \equiv cB_\phi$ for ultrarelativistic particles)

$$g_{\text{PPL}} = \frac{en}{2c\epsilon_0}, \quad (26)$$

where n is the plasma density of the lens. This focusing is uniform for all particles inside the wake.

Compare that to a lens in the overdense regime, where a linear wakefield forms—in this case, the *local* focusing force will also be given by Eq. (26), but the exposed charge
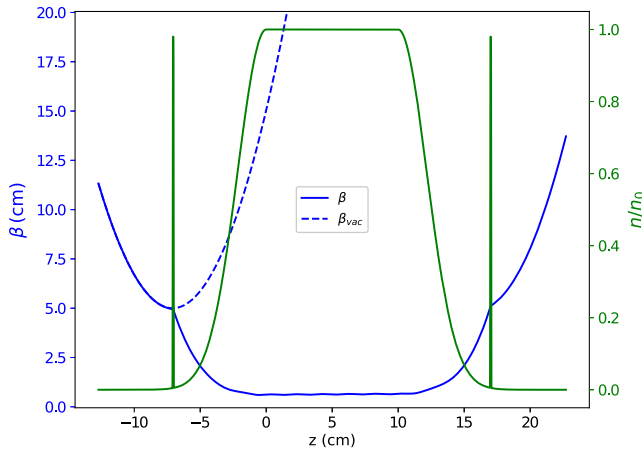
FIG. 10.   Passive plasma lensing scheme proposed for FACET-II [98], where the lenses are integrated into the plasma-density ramps. A gas jet produces a locally higher gas density that is ionized by a laser. The lens is driven by the same beam driver as is used in the main accelerating section. Source: Ref. [99] (CC BY 4.0).

density can vary throughout the beam, both transversely and longitudinally. This nonuniform focusing force can result in emittance growth. The same effect can also occur in underdense plasma lenses for self-focused beams, where the wake builds up longitudinally along the bunch, resulting in a projected emittance growth.

Note that, since the lenses are usually very short, energy changes from the longitudinal wakefield are typically ignored.

2. Active plasma lenses

An alternative way to focus beams is to use the plasma as a conductor and use large currents to produce strong magnetic fields. This actively driven plasma lens can provide a uniform focusing field for the entire bunch without the need for a driver, which makes them compact and simple to operate. On the other hand, the focusing strength is typically limited to the kT/m range [100]—orders of magnitude weaker than passive plasma lenses but still very strong compared to conventional quadrupoles. This is the type of plasma lens that was used for the BELLA staging experiment (see Fig. 1) [4].

The history of active plasma lenses started in 1950 [101,102], during which they were studied for various purposes such as ion focusing [103] and antimatter capture [104], before their recent revival for use in plasma accelerators [105]. Modern active plasma lenses consist of a thin (millimeter-scale) gas-filled capillary with electrodes on either side [106] (see Fig. 11 for a schematic overview). A high-voltage discharge ionizes the gas before a large current passes through the plasma. By Ampere's law, the azimuthal magnetic field at each radius inside the lens is proportional to the total current enclosed at that radius.
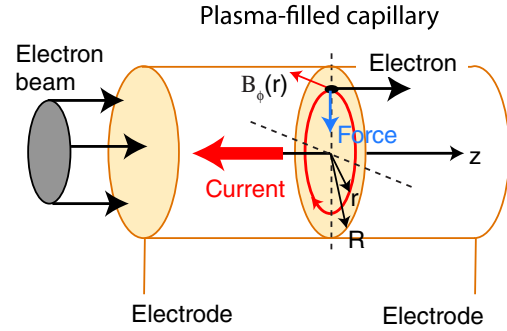

FIG. 11.   Schematic of an active plasma lens: An electron beam enters a gas-filled capillary, which is discharged via two high-voltage electrodes. The current flowing between the electrodes induces an azimuthal magnetic field B_ϕ that increases with radius r—a radial focusing force that focuses the electron beam in both planes. Source: Ref. [105].

If, ideally, the current density inside is uniform, the resulting magnetic field is linear and has a field gradient

$$g_{\mathrm{APL}} = \frac{\mu_0 I}{2\pi R^2}, \qquad (27)$$

where I is the overall current, R is the capillary radius, and μ_0 is the permittivity of free space. As an example, plugging in numbers for a typical lens of radius 500 μm and current 500 A, we obtain a magnetic field gradient of 400 T/m—a relatively strong lens. Equation (27) also indicates an advantage of active plasma lenses: By inverting the direction of the current, the lens can also focus positively charged particles—this is not easily done with a passive plasma lens.

Unfortunately, active plasma lenses are not always emittance preserving. There are three principal ways in which these lenses can degrade the beam quality. (1) *Nonuniform current density* [107] leads to nonlinear focusing fields, which causes emittance growth. This can be caused by a temperature gradient between the core and the wall (where the heat escapes) [108,109]. Interestingly, while this aberration is present in light gases like hydrogen and helium [110], it is possible to fully suppress it in a heavier gas like argon, where the heat transfer to the wall is significantly slower [111]. Another effect which can cause nonuniform focusing fields is the *z-pinch* effect [112], where the magnetic field of the lens is strong enough to self-focus its own current. (2) *Coulomb scattering* (see Sec. III A 3) can cause emittance growth due to the atomic or ionic density on axis [48–50]. Since the emittance growth rate [see Eq. (15)] scales as $\frac{d\epsilon_n}{ds} \sim \beta_x Z^2$, where β_x is the beta function in the lens and Z is the atomic number, we observe that the beta function should be kept to a minimum (also good for reduced chromaticity) and that scattering is more severe for heavier gases. The effect is usually negligible in hydrogen or helium but can be problematic for an argon-based active plasma lens. (3) *Passive*

plasma lensing will also occur in an active plasma lens if the beam density is sufficiently high, which it often is if the lens is placed close to the exit of a wakefield accelerator. Currently, this appears to be the main limiting factor for the application of active plasma lenses to beams relevant to FELs or linear colliders [113]. Clearly, any application of active plasma lensing needs to take all the above effects into account.

C. Achromatic beam transport

Chromaticity in accelerators is nothing new—several strategies for mitigating it exist. In particular, collider final focusing shares many of the same challenges—how to deal with highly diverging or converging beams with non-negligible energy spread—and, therefore, much thought has already gone into solving these issues. The bottom line is that, even though beam transport with a single focusing optic is chromatic, we can often construct lattices of multiple elements that are effectively *achromatic*.

1. Apochromatic correction

Using linear optics elements only, it is fundamentally impossible to provide fully achromatic beam transport for all energies [114]. However, it *is* possible to cancel chromaticity for a limited energy spread at certain locations in the lattice. This is inspired by how focusing of multiple colors is done in camera lenses: Introduce more degrees of freedom (i.e., more lenses) and tune the system such that several colors are all in focus. In beam optics, this corresponds to different energies traversing the lattice with a different beta-function evolution but eventually converging to the same focus (see Fig. 12). This idea was first introduced for the final focusing of CLIC in 1987 and is known as *apochromatic correction* [115].

A beam transport lattice can be apochromatically corrected to arbitrary order [40], such that not only the first-order chromatic amplitude is zero but also higher orders—at the cost of introducing more degrees of freedom
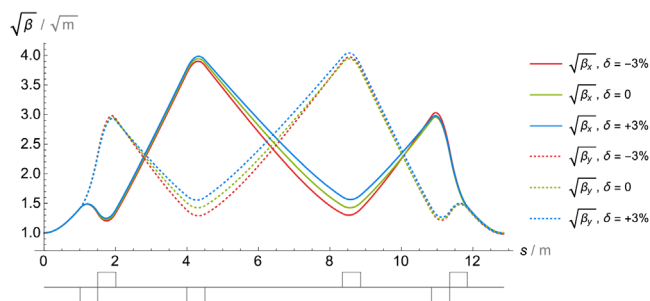

FIG. 12. Illustration of apochromatic correction in a lattice of only linear focusing optics (six quadrupole magnets). The lattice is constructed such that beta functions evolve differently for the various energy slices but eventually converge before entering the next stage. This method also works for plasma lenses and plasma density ramps. Source: Ref. [40] (CC BY 3.0).

(optics elements). However, this achromatic behavior applies only to a limited range of energies around the nominal energy. This transportable energy range is roughly $\sigma_\delta \approx 1/W$, where W is the chromatic amplitude of a more basic lattice where no apochromatic correction is applied. Beyond this characteristic energy spread, full emittance preservation is not possible. Therefore, apochromatic correction should be used in combination with other methods that reduce the intrinsic chromaticity—such as plasma-density ramps and plasma lenses.

2. Sextupoles in dispersive sections

In order to truly increase the energy acceptance of a beam line, nonlinear optics must be introduced. The conventional solution is to use sextupole magnets in regions of large dispersion. This is the solution employed in collider final focusing, where the chromaticity is so large ($W \gtrsim 10^4$) that apochromatic focusing is incapable of correcting for the required energy spread ($\sigma_\delta \approx 1\%$). Many plasma-wakefield accelerators have similar parameters and may, therefore, require nonlinear optics lattices.

The underlying concept of chromaticity correction with sextupoles is relatively straightforward. The local focusing gradient in a sextupole is proportional to the transverse offset. Therefore, if the beam is dispersed such that different energies enter the sextupole at different offsets, the chromaticity can be canceled. If the beam is horizontally dispersed ($x \to x + D_x\delta$), the focusing forces are given by

$$F_x \sim xD_x\delta + \frac{1}{2}(x^2 - y^2) + \frac{1}{2}D_x^2\delta^2, \qquad (28)$$

$$F_y \sim yD_x\delta + yx \qquad (29)$$

for the energy slice δ, where the first terms ($x\delta$ and $y\delta$) can be used for chromaticity correction and the rest are nonlinear geometric (x^2, y^2, and yx) and chromatic terms (δ^2). These nonlinear forces cause emittance growth—an effect that must be mitigated by introducing another sextupole elsewhere in the lattice to exert exactly the opposite nonlinear forces. See Fig. 13 for an illustration of this concept.

A particularly important concept for mitigating chromaticity is that of *local chromaticity correction* [116,117]. While it is possible to correct chromaticity *globally* by distributing sufficiently many sextupoles across the lattice to be able to cancel chromaticity at the end [118], it will always be better to correct the chromaticity *locally* by placing a dedicated sextupole next to each chromaticity-inducing focusing optic. This greatly increases the energy acceptance of the lattice, because chromaticity never really develops in the first place. All modern collider designs [2,3] employ local chromaticity correction in their final-focusing systems.
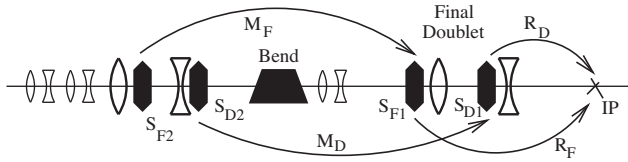
FIG. 13. Local chromaticity correction for a linear collider using nonlinear optics and dispersive elements. The beam is dispersed with a dipole (black trapezoid) onto the final doublet of quadrupoles (white). Each quadrupole has a corresponding sextupole (black hexagons) to locally correct the chromaticity. Prior to the dipole is a similar doublet that compensates the geometric terms introduced by each of the two final sextupoles. Source: Ref. [116].
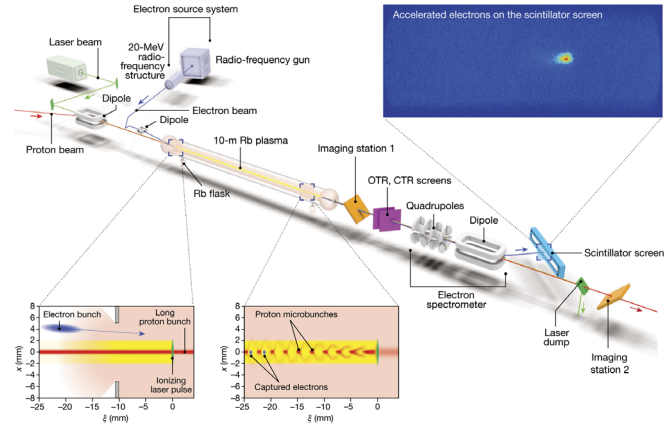

FIG. 14. The AWAKE experimental setup, using a proton bunch in a laser-ionized plasma to accelerate an electron beam to 2 GeV (upper right). The long proton bunch self-modulated (lower left) into a train of short bunches, resonantly driving a plasma wakefield, into which electrons were externally injected and then accelerated. Source: Ref. [124] (CC BY 4.0).

The drawbacks of chromaticity correction using non-linear optics are (i) the introduction of large dispersion and (ii) long and complex lattices. Installing two collider-style final-focusing systems back to back between the stages of a wakefield accelerator would take up a large amount of space—defeating the purpose of the high-gradient acceleration. However, if local chromaticity correction can be applied in a simpler and more optimized scheme (e.g., utilizing plasma lenses and mirror symmetry), this might hold the key to compact and chromaticity-free staging.

D. Single-stage acceleration

Sometimes, the only way to win is not to play. We should therefore briefly consider some alternative methods for accelerating to high energies without using multiple stages—i.e., *single-stage acceleration*. Before delving in, it is worth noting that all these techniques will also require some form of emittance-preserving out-coupling—this means that many of the above considerations are still relevant, but the requirements for compactness can be relaxed.

1. Proton-driven plasma accelerators

The main hurdle to achieving single-stage wakefield acceleration to high energy is the overall energy content of the driver—it needs to be very high. While lasers and electron beams rarely go beyond 100 J, large synchrotrons can provide proton beams with 1–100 kJ of energy per bunch—more than enough to accelerate 1 nC to 1 TeV. Unfortunately, these proton bunches are not short enough to drive high-frequency, high-gradient wakefield accelerators. The solution is to transform the long proton beam into a train of many short bunches and let the wakefield build up resonantly along the train. Interestingly, this can be done in a plasma using a process known as *self-modulation* [119,120].

Self-modulation is a process whereby a beam in a plasma self-interacts with its own focusing and defocusing wake-fields. Where the beam is focused, it gets denser; where the beam is defocused, it gets ejected—amplifying the on-axis

charge modulation and increasing the amplitude of the wakefield. This instability gradually builds up until the proton beam is fully self-modulated. The AWAKE experiment [121] at CERN, utilizing 400 GeV proton bunches from the Super Proton Synchrotron, has successfully demonstrated this self-modulation [122,123] and used it to accelerate electron bunches up to 2 GeV [124] in a 10-m-long plasma stage (see Fig. 14).

While self-modulated, proton-driven wakefield acceleration constitutes one path toward realizing high-gradient, high-energy acceleration, the concept has some serious limitations. First, the extremely limited repetition rate of such machines (of the order of 0.01–0.1 Hz) implies that high integrated luminosity or brightness is not possible. Second, a very large proton synchrotron is required—it may make sense only to build such accelerators in the vicinity of existing infrastructure. Third, the external injection into the wakefield is nontrivial—it will be particularly challenging to preserve the emittance of the electron bunch. Moreover, if a separate proton self-modulator stage is required before the injection and acceleration of electrons, many of the above-mentioned staging challenges will apply.

2. Traveling-wave electron acceleration

A proposal for reaching high energies in a single-stage laser-driven wakefield accelerator is using a *traveling* (or *flying*) laser focus. This may allow laser-wakefield accelerators to overcome the dephasing limit, i.e., where the electron beam outruns the wake because the laser group velocity in plasma is subluminal.

The scheme has been proposed implemented in several different ways. One idea is to couple in two laser drivers transversely—one from each side—and use their super-position to drive the wakefield [125]. Furthermore, if their
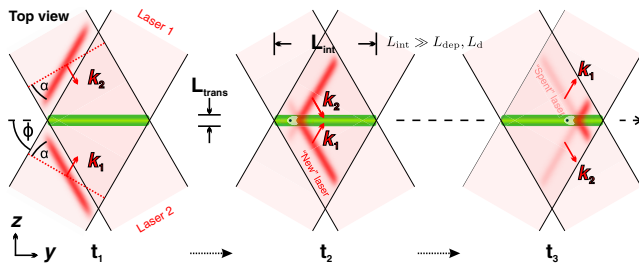
FIG. 15. Illustration of the traveling-wave electron acceleration scheme: Two lasers are coupled in transversely into an on-axis plasma channel. The superposition of the two lasers drives a wakefield, and, since the laser pulse fronts are tilted, the traveling wave can be made to move at speed c. Source: Ref. [125] (CC BY 4.0).
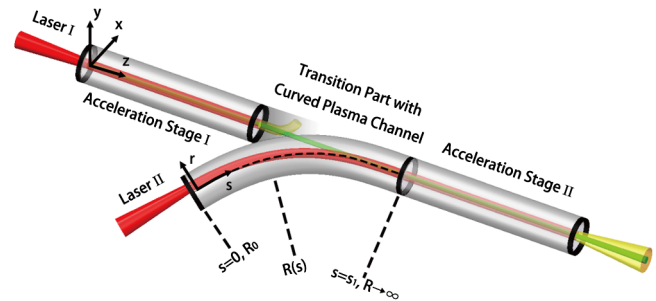

FIG. 16. Illustration of the curved plasma channel scheme: After an initial laser plasma accelerator stage, a fresh laser pulse (red) is in-coupled using a curved plasma channel. This channel guides the laser around the curve but lets the electron beam (green) pass straight through until it is trapped by the new wakefield. The depleted laser (yellow) from the first stage is also out-coupled in this transition region. Source: Ref. [128].

pulse fronts are tilted just right, the laser focus can be made to travel at the (vacuum) speed of light c to match the speed of the relativistic electron beam—overcoming the dephasing problem faced in regular laser-driven plasma accelerators. Figure 15 illustrates this so-called traveling-wave electron acceleration scheme. This is interesting, because it exploits the ability of lasers to be transversely in-coupled and may allow seamless blending of multiple lasers with no need for staging. On the other hand, it will require an extreme level of control over the laser parameters that may be difficult to achieve in practice.

Another scheme is the use of a chromatic flying focus, where different frequencies are focused at different longitudinal locations using a diffractive lens [126]. The relative timing of the peak intensity of each frequency can then be adjusted by the chirp of the laser pulse, effectively decoupling the speed at which the focal spot travels. A similar scheme proposes to instead use a stepped echelon optic in combination with an axiparabola in order to delay and focus different radial portions of the laser pulse separately [127].

While these methods overcome the dephasing limit, potentially allowing for TeV energy gains in a few meters only, the energy-transfer efficiency can be relatively low. Each part of the laser beam only reaches high intensity over a short distance (approximately the Rayleigh length)—not sufficiently long to significantly deplete its energy. Reference [127] estimates an energy-transfer efficiency on the few-percent level, which may prove too low to compete with a staged accelerator.

3. Curved plasma channels

Finally, another laser-based method has been proposed for swapping out the driver without staging: coupling in fresh laser drivers directly using *curved plasma channels* [128]. Similar to how laser pulses are guided in straight plasma channels with a transverse (parabolic) density profile [129,130], a laser pulse can also follow a curved channel. In contrast, a high-energy electron bunch will pass

straight through such a plasma profile and can, therefore, be handed off from one driver to the next with minimal disruption. Figure 16 illustrates this scheme.

While the beam never really exits the plasma during acceleration, transitioning from one stage to the next will necessarily introduce transverse oscillations to the beam centroid, which would cause emittance growth and potentially a beam-breakup instability. The scheme therefore appears too simplistic for applications such as a linear collider. However, if mitigation strategies could be found to suppress these effects, a matured version of the scheme could potentially offer a path toward high-energy gains with laser-driven plasma wakefields.

V. CONCLUSIONS

The application of plasma wakefields to truly high-energy accelerators is currently held back by the practical problems of staging. In the conventional view, the driver must be in- and out-coupled in the space between stages. Combined with the strong focusing and finite energy spread of such accelerators, the capture and refocusing of beams become highly chromatic—resulting in large emittance growth and potentially beam loss. For the beam quality to remain high throughout a multistage accelerator, many requirements must be met: matching of beta functions, dispersion cancellation, isochronicity, and avoiding too much scattering and synchrotron radiation, as well as tight synchronization and misalignment tolerances. At the same time, sufficient diagnostics must be included, and the whole setup needs to be compact to ensure a high effective acceleration gradient.

Many good ideas have been proposed for how to mitigate or avoid the challenges of staging. The concept of plasma-density ramps is key to reducing the high divergence—and, if made adiabatic, they can help reduce the chromaticity. Passive and active plasma lenses can take the next step: capturing and refocusing the beams more compactly than is

possible with conventional quadrupoles. Nevertheless, specialized beam transport optics will likely still be required to handle the residual chromaticity—this can either be done with the use of apochromatic staging or, better yet, nonlinear optics with local chromaticity correction. Finally, alternative methods have been proposed that sidestep staging altogether by employing single-stage acceleration. Examples include self-modulated proton-driven plasma wakefields, laser-based traveling-wave electron acceleration, and laser-driven acceleration in curved plasma channels—none of which are currently able to avoid staging without significantly sacrificing repetition rate, energy efficiency, or beam quality, respectively.

In summary, staging is a difficult and important problem that needs to be tackled before plasma-wakefield accelerators can be useful for large-scale applications such as a particle collider. Clearly, the field is young, and the search for solutions is still in its infancy—in other words, the best is surely yet to come.

[1] J. M. J. Madey, Stimulated emission of bremsstrahlung in a periodic magnetic field, J. Appl. Phys. **42,** 1906 (1971).

[2] The International Linear Collider technical design report, edited by T. Behnke *et al.* 2013.

[3] A multi-TeV linear collider based on CLIC technology: CLIC conceptual design report, edited by M. Aicheler *et al.*, CERN, Geneva, 2013.

[4] S. Steinke, J. van Tilborg, C. Benedetti, C. G. R. Geddes, C. B. Schroeder, J. Daniels *et al.*, Multistage coupling of independent laser-plasma accelerators, Nature (London) **530,** 190 (2016).

[5] 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 (1985).

[6] R. D. Ruth, A. W. Chao, P. L. Morton, and P. B. Wilson, A plasma wake field accelerator, Part. Accel. **17,** 171 (1985).

[7] T. Tajima and J. M. Dawson, Laser Electron Accelerator, Phys. Rev. Lett. **43,** 267 (1979).

[8] E. Esarey, C. B. Schroeder, and W. P. Leemans, Physics of laser-driven plasma-based electron accelerators, Rev. Mod. Phys. **81,** 1229 (2009).

[9] M. Litos, E. Adli, W. An, C. I. Clarke, C. E. Clayton, S. Corde *et al.*, High-efficiency acceleration of an electron beam in a plasma wakefield accelerator, Nature (London) **515,** 92 (2014).

[10] C. Thaury, F. Quéré, J.-P. Geindre, A. Levy, T. Ceccotti, P. Monot *et al.*, Plasma mirrors for ultrahigh-intensity optics, Nat. Phys. **3,** 424 (2007).

[11] B. Allen, V. Yakimenko, M. Babzien, M. Fedurin, K. Kusche, and P. Muggli, Experimental Study of Current Filamentation Instability, Phys. Rev. Lett. **109,** 185007 (2012).

[12] G. Raj *et al.*, Probing ultrafast magnetic-field generation by current filamentation instability in femtosecond relativistic laser-matter interactions, Phys. Rev. Research **2,** 023123 (2020).

[13] H. Wiedemann, *Particle Accelerator Physics* (Springer, Berlin, 2007), p. 321.

[14] M. J. Barnes, L. Ducimetiére, T. Fowler, V. Senaj, and L. Sermeus, Injection and extraction magnets: Kicker magnets, in *Proceedings of CERN Accelerator School 2009: Specialised Course on Magnets* (CERN, Geneva, 2010), p. 141.

[15] F. Arntz, M. Gaudreau, M. Kempkes, D. Technologies, A. Krasnykh, and A. Kardo-Sysoev, A kicker driver for the international linear collider, in *Proceedings of the 22nd Particle Accelerator Conference, PAC-2007, Albuquerque, NM* (IEEE, Piscataway, NJ, 2007), p. 2972.

[16] A. W. Chao, *Physics of Collective Beam Instabilities in High-Energy Accelerators* (Wiley, New York, 1993).

[17] M. G. Kelliher and R. Beadle, Pulse-shortening in electron linear accelerators, Nature (London) **187,** 1099 (1960).

[18] K. L. F. Bane and M. Sands, The short-range resistive wall wakefields, AIP Conf. Proc. **367,** 131 (1996).

[19] W. K. H. Panofsky and W. A. Wenzel, Some considerations concerning the transverse deflection of charged particles in radiofrequency fields, Rev. Sci. Instrum. **27,** 967 (1956).

[20] V. Lebedev, A. Burov, and S. Nagaitsev, Efficiency versus instability in plasma accelerators, Phys. Rev. Accel. Beams **20,** 121301 (2017).

[21] C. A. Lindstrøm, E. Adli, J. M. Allen, W. An, C. Beekman, C. I. Clarke *et al.*, Measurement of Transverse Wakefields Induced by a Misaligned Positron Bunch in a Hollow Channel Plasma Accelerator, Phys. Rev. Lett. **120,** 124802 (2018).

[22] W. Gai, P. Schoessow, B. Cole, R. Konecny, J. Norem, J. Rosenzweig, and J. Simpson, Experimental Demonstration of Wakefield Effects in Dielectric Structures, Phys. Rev. Lett. **61,** 2756 (1988).

[23] S. Gessner, E. Adli, J. M. Allen, W. An, C. I. Clarke, C. E. Clayton *et al.*, Demonstration of a positron beam-driven hollow channel plasma wakefield accelerator, Nat. Commun. **7,** 11785 (2016).

[24] C. Li, W. Gai, C. Jing, J. G. Power, C. X. Tang, and A. Zholents, High gradient limits due to single bunch beam breakup in a collinear dielectric wakefield accelerator, Phys. Rev. Accel. Beams **17,** 091302 (2014).

[25] W. K. H. Panofsky and M. Bander, Asymptotic theory of beam break-up in linear accelerators, Rev. Sci. Instrum. **39,** 206 (1968).

[26] D. H. Whittum, W. M. Sharp, S. S. Yu, M. Lampe, and G. Joyce, Electron-Hose Instability in the Ion-Focused Regime, Phys. Rev. Lett. **67,** 991 (1991).

[27] C. Huang, W. Lu, M. Zhou, C. E. Clayton, C. Joshi, W. B. Mori *et al.*, Hosing Instability in the Blow-Out Regime for Plasma-Wakefield Acceleration, Phys. Rev. Lett. **99,** 255001 (2007).

[28] V. E. Balakin, A. V. Novokhatsky, and V. P. Smirnov, VLEPP: Transverse beam dynamics, in *Proceedings of HEACC1983, Fermilab, Batavia, IL* (Fermilab, Batavia, IL, 1983), p. 119.

[29] G. Stupakov, Short-range wakefields generated in the blowout regime of plasma-wakefield acceleration, Phys. Rev. Accel. Beams **21,** 041301 (2018).

[30] T. J. Mehrling, R. A. Fonseca, A. Martinez de la Ossa, and J. Vieira, Mitigation of the Hose Instability in Plasma-Wakefield Accelerators, Phys. Rev. Lett. **118,** 174801 (2017).

[31] T. J. Mehrling, C. Benedetti, C. B. Schroeder, E. Esarey, and W. P. Leemans, Suppression of Beam Hosing in Plasma Accelerators with Ion Motion, Phys. Rev. Lett. **121,** 264802 (2018).

[32] J. B. Rosenzweig, A. M. Cook, A. Scott, M. C. Thompson, and R. B. Yoder, Effects of Ion Motion in Intense Beam-Driven Plasma Wakefield Accelerators, Phys. Rev. Lett. **95,** 195002 (2005).

[33] 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).

[34] E. D. Courant and H. S. Snyder, Theory of the alternating-gradient synchrotron, Ann. Phys. (Paris) **3,** 1 (1958).

[35] P. Antici, A. Bacci, C. Benedetti, E. Chiadroni, M. Ferrario, A. R. Rossi *et al.*, Laser-driven electron beamlines generated by coupling laser-plasma sources with conventional transport systems, J. Appl. Phys. **112,** 044902 (2012).

[36] M. Migliorati, A. Bacci, C. Benedetti, E. Chiadroni, M. Ferrario, A. Mostacci, L. Palumbo, A. R. Rossi, L. Serafini, and P. Antici, Intrinsic normalized emittance growth in laser-driven electron accelerators, Phys. Rev. Accel. Beams **16,** 011302 (2013).

[37] T. André *et al.*, Control of laser plasma accelerated electrons for light sources, Nat. Commun. **9,** 1334 (2018).

[38] H. Zyngier, Strategy for correcting for chromaticity (in French), Report No. LAL-77/35, Laboratoire de l'Accélérateur Linéaire, Orsay, 1977.

[39] B. W. Montague, Linear optics for improved chromaticity correction, LEP Note No. 165, CERN, Geneva, 1979.

[40] C. A. Lindstrøm and E. Adli, Design of general apochromatic drift-quadrupole beam lines, Phys. Rev. Accel. Beams **19,** 071002 (2016).

[41] C. Lejeune and J. Aubert, Emittance and brightness: Definitions and measurements, *Applied Charged Particle Optics* (Academic Press, New York, 1980), p. 159.

[42] D. T. Pierce, R. J. Celotta, G. C. Wang, W. N. Unertl, A. Galejs, C. E. Kuyatt, and S. R. Mielczarek, The GaAs spin polarized electron source, Rev. Sci. Instrum. **51,** 478 (1980).

[43] W. P. Leemans, B. Nagler, A. J. Gonsalves, C. Tóth, K. Nakamura, C. G. R. Geddes, E. Esarey, C. B. Schroeder, and S. M. Hooker, GeV electron beams from a centimetre-scale accelerator, Nat. Phys. **2,** 696 (2006).

[44] A. Buck, J. Wenz, J. Xu, K. Khrennikov, K. Schmid, M. Heigoldt *et al.*, Shock-Front Injector for High-Quality Laser-Plasma Acceleration, Phys. Rev. Lett. **110,** 185006 (2013).

[45] A. Deng, O. S. Karger, T. Heinemann, A. Knetsch, P. Scherkl, G. G. Manahan *et al.*, Generation and acceleration of electron bunches from a plasma photocathode, Nat. Phys. **15,** 1156 (2019).

[46] P. Emma and T. Raubenhemier, Systematic approach to damping ring design, Phys. Rev. Accel. Beams **4,** 021001 (2001).

[47] T. Mehrling, J. Grebenyuk, F. S. Tsung, K. Floettmann, and J. Osterhoff, Transverse emittance growth in staged laser-wakefield acceleration, Phys. Rev. Accel. Beams **15,** 111303 (2012).

[48] B. W. Montague, Emittance growth from multiple scattering in the plasma beat-wave accelerator, in *Proceedings of the CAS-ECFA-INFN Workshop: Generation of High Fields for Particle Acceleration to Very-high Energies* (CERN, Geneva, 1984), p. 208.

[49] N. Kirby, M. Berry, I. Blumenfeld, M. J. Hogan, R. Ischebeck, and R. Siemann, Emittance growth from multiple Coulomb scattering in a plasma wakefield accelerator, in *Proceedings of the 22nd Particle Accelerator Conference, PAC-2007, Albuquerque, NM* (IEEE, New York, 2007), p. 3097.

[50] Y. Zhao, R. Lehe, A. Myers, M. Thévenet, A. Huebl, C. B. Schroeder, and J.-L. Vay, Modeling of emittance growth due to Coulomb collisions in plasma-based accelerators, Phys. Plasmas **27,** 113105 (2020).

[51] T. Katsouleas, Physical mechanisms in the plasma wakefield accelerator, Phys. Rev. A **33,** 2056 (1986).

[52] M. Tzoufras, W. Lu, F. S. Tsung, C. Huang, W. B. Mori, T. Katsouleas, J. Vieira, R. A. Fonseca, and L. O. Silva, Beam Loading in the Nonlinear Regime of Plasma-Based Acceleration, Phys. Rev. Lett. **101,** 145002 (2008).

[53] A. F. Pousa, A. M. de la Ossa, R. Brinkmann, and R. W. Assmann, Compact Multistage Plasma-Based Accelerator Design for Correlated Energy Spread Compensation, Phys. Rev. Lett. **123,** 054801 (2019).

[54] S. Schulz, I. Grguraš, C. Behrens, H. Bromberger, J. T. Costello, M. K. Czwalinna *et al.*, Femtosecond all-optical synchronization of an x-ray free-electron laser, Nat. Commun. **6,** 5938 (2015).

[55] R. J. Shalloo *et al.*, Measurement of femtosecond-scale drift and jitter of the delay between the North and South beams of Gemini, Central Laser Facility Annual Report, 2015, p. 36.

[56] C. A. Lindstrøm, E. Adli, J. Pfingstner, E. Marín, and D. Schulte, Transverse tolerances of a multi-stage plasma wakefield accelerator, in *Proceedings of IPAC2016, Busan, Korea* (JACoW, Geneva, 2016), p. 2561.

[57] R. Assmann and K. Yokoya, Transverse beam dynamics in plasma-based linacs, Nucl. Instrum. Methods Phys. Res., Sect. A **410,** 544 (1998).

[58] S. Cheshkov, T. Tajima, W. Horton, and K. Yokoya, Particle dynamics in multistage wakefield collider, Phys. Rev. Accel. Beams **3,** 071301 (2000).

[59] D. Schulte, Application of advanced accelerator concepts for colliders, Rev. Accel. Sci. Techol. **09,** 209 (2016).

[60] J. Larmor, LXIII: On the theory of the magnetic influence on spectra; and on the radiation from moving ions, Philos. Mag. **44,** 503 (1897).

[61] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).

[62] E. L. Saldin, E. A. Schneidmiller, and M. V. Yurkov, On the coherent radiation of an electron bunch moving in an arc of a circle, Nucl. Instrum. Methods Phys. Res., Sect. A **398**, 373 (1997).

[63] A. Novokhatski, Coherent synchrotron radiation: Theory and simulations, Report No. SLAC-PUB-14893; ICFA Beam Dyn. Newslett. **57**, 127 (2012).

[64] Y. Cai and Y. Ding, Three-dimensional effects of coherent synchrotron radiation by electrons in a bunch compressor, Phys. Rev. Accel. Beams **23**, 014402 (2020).

[65] Y. Jing and V. N. Litvinenko, Design of a bunch compressor with CSR suppression to achieve hundreds of kA peak current, in *Proceedings of IPAC2019, Melbourne, Australia* (JACoW, Geneva, 2019), p. 382.

[66] S. Heifets, G. Stupakov, and S. Krinsky, Coherent synchrotron radiation instability in a bunch compressor, Phys. Rev. Accel. Beams **5**, 064401 (2002).

[67] A. D. Brynes, I. Akkermans, E. Allaria, L. Badano, S. Brussaard, G. De Ninno *et al.*, Characterisation of microbunching instability with 2D Fourier analysis, Sci. Rep. **10**, 5059 (2020).

[68] B. Cros and P. Muggli, Towards a proposal for an Advanced Linear Collider, Report on the advanced and novel accelerators for high energy physics roadmap workshop, CERN, Geneva, 2017.

[69] C. A. Lindstrøm, E. Adli, J. M. Allen, J. P. Delahaye, M. J. Hogan, C. Joshi, P. Muggli, T. O. Raubenheimer, and V. Yakimenko, Staging optics considerations for a plasma wakefield acceleration linear collider, Nucl. Instrum. Methods Phys. Res., Sect. A **829**, 224 (2016).

[70] M. C. Downer, R. Zgadzaj, A. Debus, U. Schramm, and M. C. Kaluza, Diagnostics for plasma-based electron accelerators, Rev. Mod. Phys. **90**, 035002 (2018).

[71] D. Lipka, Cavity BPM designs, related electronics and measured performances, in *Proceedings of DIPAC'09, Basel, Switzerland, 2009* (CERN, Geneva, 2009), p. 280.

[72] R. Lai, U. Happek, and A. J. Sievers, Measurement of the longitudinal asymmetry of a charged particle bunch from the coherent synchrotron or transition radiation spectrum, Phys. Rev. E **50**, R4294 (1994).

[73] P. Emma, J. Frisch, and P. Krejcik, A transverse RF deflecting structure for bunch length and phase space diagnostics, SLAC LCLS Technical Report No. LCLSTN0012, SLAC, California, 2000.

[74] P. S. M. Claveria, E. Adli, L. D. Amorim, W. An, C. E. Clayton, S. Corde *et al.*, Betatron radiation and emittance growth in plasma wakefield accelerators, Phil. Trans. R. Soc. A **377**, 20180173 (2019).

[75] W. Leemans and E. Esarey, Laser-driven plasma-wave electron accelerators, Phys. Today **62**, No. 3, 44 (2009).

[76] C. B. Schroeder, E. Esarey, C. Benedetti, and W. P. Leemans, Efficiency considerations for high-energy physics applications of laser-plasma accelerators, AIP Conf. Proc. **1777**, 020001 (2016).

[77] J. Cowley, C. Thornton, C. Arran, R. J. Shalloo, L. Corner, G. Cheung *et al.*, Excitation and Control of Plasma Wakefields by Multiple Laser Pulses, Phys. Rev. Lett. **119**, 044802 (2017).

[78] S. Pei, M. J. Hogan, T. O. Raubenheimer, A. Seryi, H. H. Braun, R. Corsini, and J. P. Delahaye, Conceptual design of the drive beam for a PWFA-LC, in *Proceedings of the 23rd Particle Accelerator Conference, Vancouver, Canada, 2009* (IEEE, Piscataway, NJ, 2009), p. 2682.

[79] E. Adli, J.-P. Delahaye, S. J. Gessner, M. J. Hogan, T. O. Raubenheimer, W. An, C. Joshi, and W. B. Mori, A beam driven plasma-wakefield linear collider: From Higgs factory to multi-TeV, SLAC Report No. SLAC-PUB-15426, 2013.

[80] J. Pfingstner, E. Adli, C. Lindstrøm, E. Marín, and D. Schulte, Considerations for a drive beam scheme for a plasma wakefield linear collider, in *Proceedings of IPAC2016, Busan, Korea* (JACoW, Geneva, 2016), p. 2565.

[81] A. Bonatto, C. B. Schroeder, J.-L. Vay, C. G. R. Geddes, C. Benedetti, E. Esarey, and W. P. Leemans, Passive and active plasma deceleration for the compact disposal of electron beams, Phys. Plasmas **22**, 083106 (2015).

[82] R. Pompili, E. Chiadroni, A. Cianchi, A. Del Dotto, L. Faillace, M. Ferrario, P. Iovine, and M. R. Masullo, Plasma lens-based beam extraction and removal system for plasma wakefield acceleration experiments, Phys. Rev. Accel. Beams **22**, 121302 (2019).

[83] K. A. Marsh *et al.*, Beam matching to a plasma wake field accelerator using a ramped density profile at the plasma boundary, in *Proceedings of the 21st Particle Accelerator Conference, Knoxville, TN, 2005* (IEEE, Piscataway, NJ, 2005), p. 2702.

[84] I. Dornmair, K. Floettmann, and A. R. Maier, Emittance conservation by tailored focusing profiles in a plasma accelerator, Phys. Rev. Accel. Beams **18**, 041302 (2015).

[85] R. Ariniello, C. E. Doss, K. Hunt-Stone, J. R. Cary, and M. D. Litos, Transverse beam dynamics in a plasma density ramp, Phys. Rev. Accel. Beams **22**, 041304 (2019).

[86] M. Martini, An introduction to transverse beam dynamics in accelerators, Report No. CERN/PS 96-11, CERN, Geneva, 1996, p. 16.

[87] X. L. Xu, J. F. Hua, Y. P. Wu, C. J. Zhang, F. Li, Y. Wan *et al.*, Physics of Phase Space Matching for Staging Plasma and Traditional Accelerator Components Using Longitudinally Tailored Plasma Profiles, Phys. Rev. Lett. **116**, 124801 (2016).

[88] R. H. Helm, Adiabatic approximation for dynamics of a particle in the field of a tapered solenoid, Report No. SLAC-Report-4, 1962.

[89] K. Floettmann, Adiabatic matching section for plasma accelerated beams, Phys. Rev. Accel. Beams **17**, 054402 (2014).

[90] J. B. Johnson, A low voltage cathode ray oscillograph, J. Opt. Soc. Am. **6**, 701 (1922).

[91] B. von Borries and E. Ruska, The short space-charge field of an auxiliary discharge as a collecting lens for cathode rays, Z. Phys. **76**, 649 (1932).

[92] W. H. Bennett, Magnetically Self-Focussing Streams, Phys. Rev. **45**, 890 (1934).

[93] D. Gabor, A space-charge lens for the focusing of ion beams, Nature (London) **160**, 89 (1947).

[94] P. Chen, S. Rajagopalan, and J. Rosenzweig, Final focusing and enhanced disruption from an underdense plasma lens in a linear collider, Phys. Rev. D **40,** 923 (1989).

[95] J. B. Rosenzweig, P. Schoessow, B. Cole, C. Ho, W. Gai, R. Konecny, S. Mtingwa, J. Norem, M. Rosing, and J. Simpson, Demonstration of electron beam self-focusing in plasma wake fields, Phys. Fluids B **2,** 1376 (1990).

[96] J. S. T. Ng, P. Chen, H. Baldis, P. Bolton, D. Cline, W. Craddock et al., Observation of Plasma Focusing of a 28.5 GeV Positron Beam, Phys. Rev. Lett. **87,** 244801 (2001).

[97] C. Thaury, E. Guillaume, A. Döpp, R. Lehe, A. Lifschitz, K. T. Phuoc et al., Demonstration of relativistic electron beam focusing by a laser-plasma lens, Nat. Commun. **6,** 6860 (2015).

[98] C. Joshi, E. Adli, W. An, C. E. Clayton, S. Corde, S. Gessner et al., Plasma wakefield acceleration experiments at FACET II, Plasma Phys. Controlled Fusion **60,** 034001 (2018).

[99] C. E. Doss, E. Adli, R. Ariniello, J. Cary, S. Corde, B. Hidding et al., Laser-ionized, beam-driven, underdense, passive thin plasma lens, Phys. Rev. Accel. Beams **22,** 111001 (2019).

[100] K. N. Sjobak, E. Adli, R. Corsini, W. Farabolini, G. Boyle, C. A. Lindstrøm et al., Multi-kT/m focusing gradient in a linear active plasma lens, arXiv:2012.10680.

[101] W. K. H. Panofsky and W. R. Baker, A focusing device for the external 350-MeV proton beam of the 184-inch cyclotron at Berkeley, Rev. Sci. Instrum. **21,** 445 (1950).

[102] E. B. Forsyth, L. M. Lederman, and J. Sunderland, The Brookhaven-Columbia plasma lens, IEEE Trans. Nucl. Sci. **12,** 872 (1965).

[103] E. Boggasch, J. Jacoby, H. Wahl, K.-G. Dietrich, D. H. H. Hoffmann, W. Laux, M. Elfers, C. R. Haas, V. P. Dubenkov, and A. A. Golubev, z-Pinch Plasma Lens Focusing of a Heavy-Ion Beam, Phys. Rev. Lett. **66,** 1705 (1991).

[104] R. Kowalewicz, M. L. di Scampamorte, S. Milner, F. Pedersen, H. Riege, J. Christiansen, K. Frank, M. Stetter, R. Tkotz, and E. Boggasch, Performance of the CERN plasma lens in laboratory and beam tests at the antiproton source, in Proceedings of the 1991 Particle Accelerator Conference, San Francisco, CA (IEEE, Piscataway, NJ, 1991), p. 2631.

[105] J. van Tilborg, S. Steinke, C. G. R. Geddes, N. H. Matlis, B. H. Shaw, A. J. Gonsalves et al., Active Plasma Lensing for Relativistic Laser-Plasma-Accelerated Electron Beams, Phys. Rev. Lett. **115,** 184802 (2015).

[106] C. A. Lindstrøm, K. N. Sjobak, E. Adli, J.-H. Röckemann, L. Schaper, J. Osterhoff et al., Overview of the CLEAR plasma lens experiment, Nucl. Instrum. Methods Phys. Res., Sect. A **909,** 379 (2018).

[107] J. van Tilborg, S. K. Barber, H.-E. Tsai, K. K. Swanson, S. Steinke, C. G. R. Geddes et al., Nonuniform discharge currents in active plasma lenses, Phys. Rev. Accel. Beams **20,** 032803 (2017).

[108] N. A. Bobrova, A. A. Esaulov, J.-I. Sakai, P. V. Sasorov, D. J. Spence, A. Butler, S. M. Hooker, and S. V. Bulanov, Simulations of a hydrogen-filled capillary discharge waveguide, Phys. Rev. E **65,** 016407 (2001).

[109] B. H. P. Broks, K. Garloff, and J. J. A. M. van der Mullen, Nonlocal-thermal-equilibrium model of a pulsed capillary discharge waveguide, Phys. Rev. E **71,** 016401 (2005).

[110] J.-H. Röckemann, L. Schaper, S. K. Barber, N. A. Bobrova, G. Boyle, S. Bulanov et al., Direct measurement of focusing fields in active plasma lenses, Phys. Rev. Accel. Beams **21,** 122801 (2018).

[111] C. A. Lindstrøm, E. Adli, G. Boyle, R. Corsini, A. E. Dyson, W. Farabolini et al., Emittance Preservation in an Aberration-Free Active Plasma Lens, Phys. Rev. Lett. **121,** 194801 (2018).

[112] J. Christiansen, K. Frank, H. Riege, and R. Seeböck, Studies of a plasma lens with pseudo-spark geometry for application in high energy particle accelerators, CERN PS Note 84-10-AA, CERN, Geneva, 1984.

[113] C. A. Lindstrøm and E. Adli, Analytic plasma wakefield limits for active plasma lenses, arXiv:1802.02750.

[114] A. Wolski, *Beam Dynamics in High Energy Particle Accelerators* (Imperial College Press, London, 2014).

[115] B. W. Montague and F. Ruggiero, Apochromatic focusing for linear colliders, CLIC Note No. 37, CERN, Geneva, 1987.

[116] P. Raimondi and A. Seryi, Novel Final Focus Design for Future Linear Colliders, Phys. Rev. Lett. **86,** 3779 (2001).

[117] G. R. White et al. (ATF2 Collaboration), Experimental Validation of a Novel Compact Focusing Scheme for Future Energy-Frontier Linear Lepton Colliders, Phys. Rev. Lett. **112,** 034802 (2014).

[118] J. Irwin, K. Brown, F. Bulos, D. Burke, R. Helm, G. Roy, R. Ruth, N. Yamamoto, and K. Oide, The optics of the final focus test beam, *Proceedings of the 1991 Particle Accelerator Conference, San Francisco, CA* (IEEE, Piscataway, NJ, 1991), p. 2058.

[119] N. Kumar, A. Pukhov, and K. Lotov, Self-Modulation Instability of a Long Proton Bunch in Plasmas, Phys. Rev. Lett. **104,** 255003 (2010).

[120] C. B. Schroeder, C. Benedetti, E. Esarey, F. J. Grüner, and W. P. Leemans, Growth and Phase Velocity of Self-Modulated Beam-Driven Plasma Waves, Phys. Rev. Lett. **107,** 145002 (2011).

[121] E. Gschwendtner, E. Adli, L. Amorim, R. Apsimon, R. Assmann, A.-M. Bachmann et al., AWAKE, The advanced proton driven plasma wakefield acceleration experiment at CERN, Nucl. Instrum. Methods Phys. Res., Sect. A **829,** 76 (2016).

[122] E. Adli et al. (AWAKE Collaboration), Experimental Observation of Proton Bunch Modulation in a Plasma at Varying Plasma Densities, Phys. Rev. Lett. **122,** 054802 (2019).

[123] M. Turner et al. (AWAKE Collaboration), Experimental Observation of Plasma Wakefield Growth Driven by the Seeded Self-Modulation of a Proton Bunch, Phys. Rev. Lett. **122,** 054801 (2019).

[124] E. Adli et al. (AWAKE Collaboration), Acceleration of electrons in the plasma wakefield of a proton bunch, Nature (London) **561,** 363 (2018).

[125] A. Debus, R. Pausch, A. Huebl, K. Steiniger, R. Widera, T. E. Cowan, U. Schramm, and M. Bussmann, Circumventing the Dephasing and Depletion Limits of Laser-Wakefield Acceleration, Phys. Rev. X **9,** 031044 (2019).

[126] D. H. Froula, D. Turnbull1, A. S. Davies, T. J. Kessler, D. Haberberger, J. P. Palastro *et al.*, Spatiotemporal control of laser intensity, Nat. Photonics **12**, 262 (2018).

[127] J. P. Palastro, J. L. Shaw, P. Franke, D. Ramsey, T. T. Simpson, and D. H. Froula, Dephasingless Laser Wakefield Acceleration, Phys. Rev. Lett. **124**, 134802 (2020).

[128] J. Luo, M. Chen, W. Y. Wu, S. M. Weng, Z. M. Sheng, C. B. Schroeder *et al.*, Multistage Coupling of Laser-Wakefield Accelerators with Curved Plasma Channels, Phys. Rev. Lett. **120**, 154801 (2018).

[129] C. G. Durfee and H. M. Milchberg, Light Pipe for High Intensity Laser Pulses, Phys. Rev. Lett. **71**, 2409 (1993).

[130] D. J. Spence and S. M. Hooker, Investigation of a hydrogen plasma waveguide, Phys. Rev. E **63**, 015401(R) (2000).