Staging of High-Gradient Wakefield Accelerators

C. A. Lindstrøm
Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Abstract

Accelerating particles to high energies with a high-gradient wakefield accelerator may require use of multiple stages. Coupling beams from one stage to another can be difficult due to high divergence and non-negligible energy spreads. We review the challenges, technical requirements and currently proposed methods for solving the staging problem.

Keywords

Staging; wakefield accelerators; plasma wakefields; chromaticity.

1 Introduction

Accelerating particles to high energy in a compact space is the main motivation for high-gradient wakefield accelerator research. What counts as *high energy*, however, depends on the context—free-electron lasers [1] will require energies around 1–10 GeV, whereas linear colliders [2, 3] will require 100 GeV or more. The energy gain in a single wakefield accelerator stage [4–7] is ultimately limited by the energy stored in its driver, which is typically around 1–100 Joules. While a single stage may be enough to drive a free-electron laser, it will not be sufficient 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—also 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 [8], there are still many problems to be solved on a conceptual level. In this Chapter, we will explore why staging is so challenging (Section 2), the detailed technical requirements (Section 3), and finally some currently proposed methods to get there (Section 4).

2 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 high-gradient 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 problem: to accelerate particles with high gradients, strong focusing is required to maintain stable acceleration. This strong focusing results in high-divergence beams in the space between the stages—making it difficult to preserve the beam quality of finite-energy-spread bunches. 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?

2.1 In- and out-coupling of the driver

A defining feature of wakefield accelerators is the co-propagating driver—the source of energy for the accelerating particles. When the driver has been depleted of its energy, it must be swapped out for a fresh one. This extraction and re-injection 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 stages.

Methods for swapping out the driver will depend on the type of driver being used—lasers or charged particle beams. Two devices are available for making the driver collinear with the accelerating particle beam: mirrors and magnetic dipoles. Clearly, mirrors can only be used for laser-drivers, whereas dipoles can be used in both laser- 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, resulting in no burning but taking up more space. Alternatively, one can use a so-called *plasma mirror* [9]—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 [8] (see Fig. 1). However, preserving the emittance of a beam passing through a plasma mirror can be difficult, due to effects such as beam filamentation [10, 11] (for small beams) and Coulomb scattering [12, 13] (for large beams).
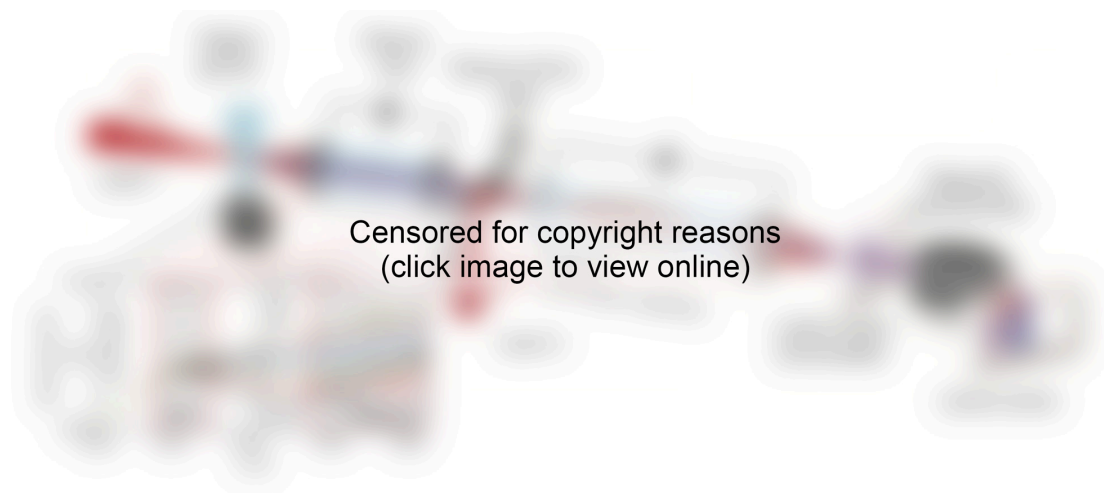
Fig. 1: The BELLA staging experiment performed at LBNL. Electrons from a laser wakefield accelerator stage were focused into a second stage using a plasma lens. Simultaneously, a second laser driver was coupled in using a plasma mirror. Electrons with a central energy of 120 MeV were boosted in the second 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: S. Steinke *et al.*, Nature **530**, 190 (2016) [8].

Magnetic dipoles bend the path of charged particles, and can therefore be used to merge/separate a particle beam and a laser beam, or merge/separate two different particle beams. In the latter case, 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. Effectively, this also applies to laser drivers, which have no charge and therefore behave as if they have a higher (infinite) energy. Unfortunately, the merging/separation process introduces large quality-degrading dispersion to both accelerating beams and beam drivers, which must be carefully cancelled—taking up valuable space along the beamline.

Whichever method is chosen for in- and out-coupling of the driver, it will introduce a separation between the stages—likely between 0.1 and 10 m per stage. This will not only reduce the effective accelerating gradient of a multi-stage accelerator, but also leads to trouble when combined with the strong focusing within each stage.

2.2 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 $1/k_p$:

$$\frac{E_z}{m_e c^2} \simeq k_p, \tag{1}$$

where $k_p = \sqrt{ne^2/m_e\epsilon_0 c^2}$ is the plasma wavenumber, n is the plasma density, ϵ_0 and c are the vacuum permittivity and light speed, 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 was 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 high-gradient wakefield accelerator will have the same effect.
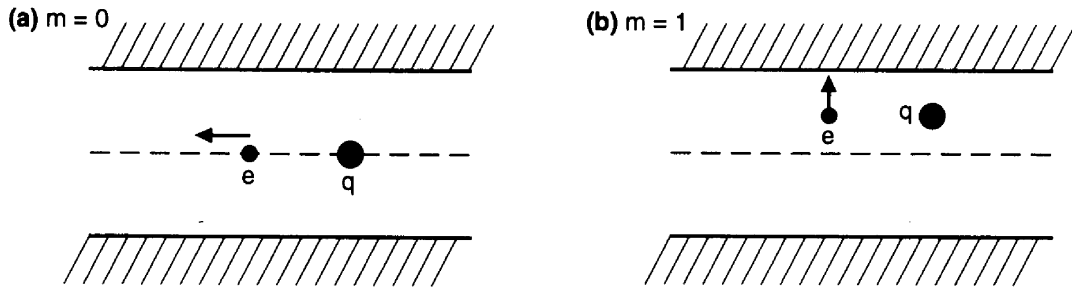

Fig. 2: Two particle bunches in a generic accelerator structure, exhibiting both (a) a longitudinal wakefield of azimuthal mode $m = 0$ and (b) a transverse wakefield of azimuthal mode $m = 1$. Arrows indicate the direction of the force. Source: A. W. Chao, *Physics of Collective Beam Instabilities in High-Energy Accelerators* (Wiley, New York, NY, 1993) [16].

The fundamental challenge is that when the characteristic cavity dimension (a) decreases, the longitudinal wakefield increases (typically between $1/a$ and $1/a^2$), but the transverse wakefield grows even faster (typically 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]

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

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 co-moving 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 CLIC, this focusing is done with quadrupoles *interleaved* between the accelerator cavities. However, this is not sufficient for a high-gradient wakefield accelerator since the transverse wakefield is too strong—focusing must occur *inside* the accelerating cavity. Fortunately, many wakefield accelerators provide intrinsically strong internal focusing. This is one of the main

advantages of plasma wakefield accelerators, where an exposed ion column (in the case of a nonlinear wake) focuses the electron beam with a focusing strength (force per offset) of

$$K = \frac{k_p^2}{2\gamma},$$

(3)

where γ is the relativistic Lorentz factor. Other wakefield accelerators, such as dielectric wakefield accelerators, plan to use strong external focusing (quadrupoles) [21]. 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* [22]—or similarly a *hose instability* in a plasma accelerator [23, 24]. 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 [25]. The implication is that stable acceleration benefits from a non-zero energy spread [26].

2.3 Chromaticity

Strong focusing results in small beam sizes, but more importantly, highly diverging beams. In terms of Courant-Snyder or Twiss parameters [27], this means small beta functions. To avoid emittance growth in a stage, the beta function must be *matched* to

$$\beta_m = \frac{1}{\sqrt{K}},$$

(4)

where K is the focusing strength—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 (Eq. (3))

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

(5)

which is typically on the mm-to-cm scale for beam energies E on the GeV-level. Capturing and refocusing these rapidly diverging beams can be challenging [28, 29]. High-gradient wakefield acceleration often results in non-negligible energy spread because of the rapidly changing (high-frequency) accelerating field structure—in plasmas 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 cannot all be focused in the same way—an effect known as *chromaticity* .

The chromaticity of the beam focusing is typically defined in terms of the *chromatic amplitude* [30, 31]

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

(6)

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 [32]

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

(7)

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.
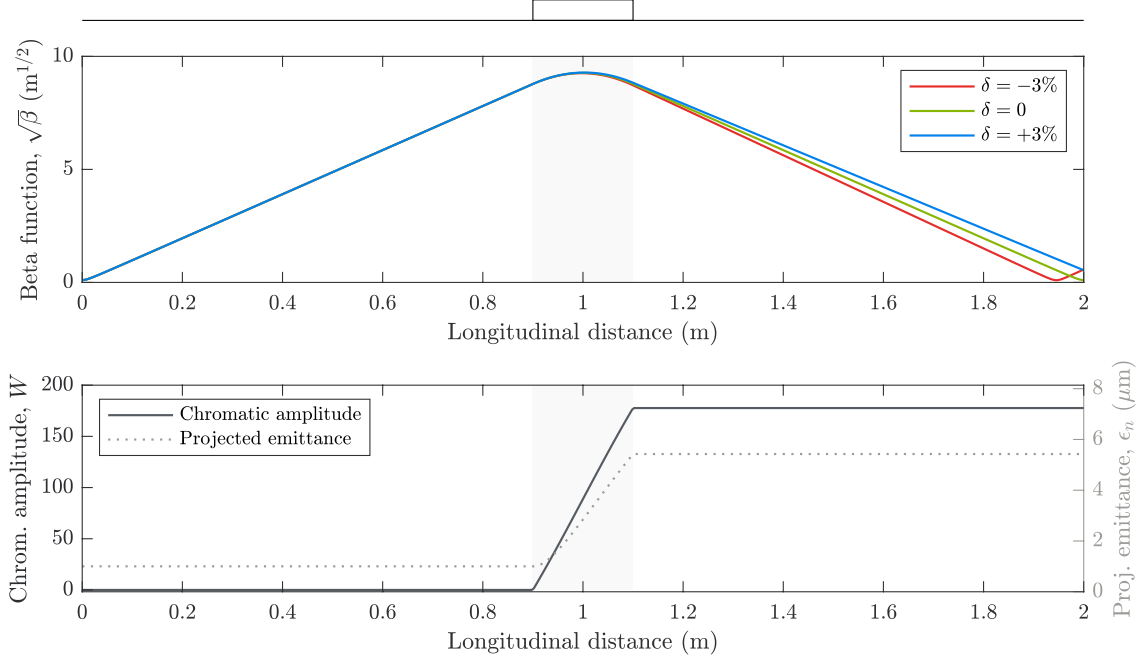
4

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 5.

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, 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 $\Delta W = \beta K l$, 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 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

$$\frac{\Delta \epsilon^2}{\epsilon_0^2} \approx \frac{4L^2}{\beta_m^2} \sigma_\delta^2. \tag{8}$$

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 high-gradient wakefield accelerators if left uncorrected.

3 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.

3.1 Emittance preservation

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

$$\mathcal{L} = H_D \frac{N^2 f \gamma}{4\pi \sqrt{\beta_x \epsilon_{nx}} \sqrt{\beta_y \epsilon_{ny}}}, \tag{9}$$

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, the lasing power is determined by the *6D brightness* [33]

$$B_{6D} = \frac{N}{\epsilon_{nx} \epsilon_{ny} \epsilon_{nz}}, \tag{10}$$

where ϵ_{nz} is the normalized emittance of the longitudinal phase space. Equations 9 and 10 indicate that low emittance is crucial—both applications require normalized transverse emittances of 1 mm mrad or lower. Technically, conservation of charge (N) is equally important, and in practice we will require close to 100% charge-coupling efficiency between stages.

Producing these low-emittance bunches is routinely done using photocathodes [34] or even plasma injection techniques [35–37]. However, preserving the emittance through a large number of stages sets very stringent limits on the emittance growth per stage. It can be useful to imagine an *emittance budget*, where each stage only gets to contribute a certain emittance growth to the final emittance. For wakefield accelerators with many stages, this will typically be on the 0.1 mm mrad-level or less per stage.

3.1.1 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 [38]

$$\frac{\epsilon_{sat}}{\epsilon} = \frac{1}{2}\left((1+\alpha^2)\frac{\beta_m}{\beta} + \frac{\beta}{\beta_m} \right), \tag{11}$$
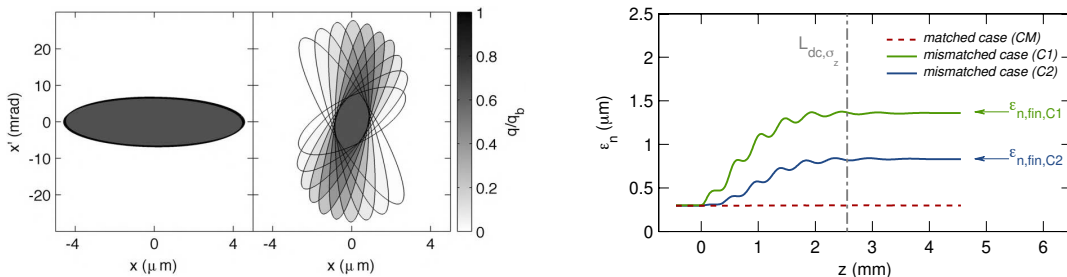

Fig. 4: Mismatching of a beam with finite energy spread. The initial phase space (left panel) is mismatched to the focusing channel, which leads to a smearing in phase space due to the different rates of rotation. This effect is seen to saturate at some point (right panel), when the beam has been fully smeared in phase space, leading to emittance growth. Source: T. Mehrling *et al.*, Phys. Rev. ST Accel. Beams **15**, 111303 (2012) [38] (CC BY 3.0).

where α and β are the Twiss parameters at the plasma entrance (assuming a flat top density profile). As seen in Eq. (11), 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 cancelled.

3.1.2 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

$$\Delta\epsilon_D \approx \frac{1}{2}\left(\frac{D_x^2}{\beta_m} + \beta_m D_{x'}^2\right)\sigma_\delta^2, \tag{12}$$

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 [22] or hosing instability [23, 24], which can lead to more severe emittance growth.

3.2 Isochronicity

Inside a stage, one rarely has to worry about changes to the bunch length—effectively, it is 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 [39, 40].

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). 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:

$$|R_{56}| \ll \frac{\sigma_z}{\sigma_\delta}. \tag{13}$$

In special cases, it can be beneficial to not cancel R_{56} completely—especially if combined with chirped accelerator stages. If one stage produces a chirped bunch (e.g., the front particles have lower energy), a carefully tuned R_{56} can overcompress the 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) [41].

A different, but related problem can occur even if the longitudinal dispersion is cancelled 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
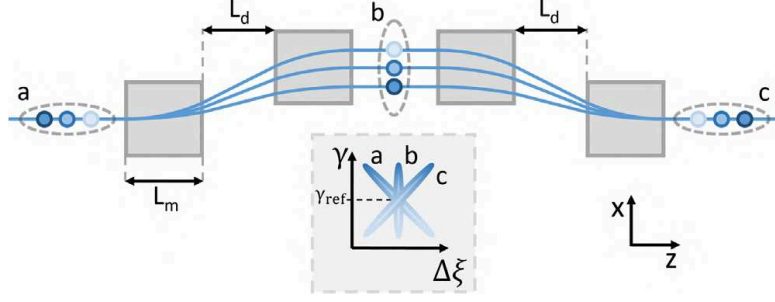
Fig. 5: Longitudinal phase space reversal in a chicane with a fine-tuned R_{56}, which can be used to significantly reduce the energy spread by compensating the chirp from one stage in the following stage. Source: A. Ferran Pousa *et al.*, Phys. Rev. Lett. **123**, 054801 (2019) [41] (reproduced with permission).

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/last optic before and after a stage is therefore restricted by

$$\Delta L \approx \frac{1}{2}\sigma_{x'}^2 L \ll \sigma_z, \tag{14}$$

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 only a problem 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. (14) be captured in much less than 37 cm.

3.3 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.

3.3.1 Synchronization

High-gradient 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

$$\frac{\Delta E_z}{E_z} \approx \omega \Delta t. \tag{15}$$

Random timing jitter therefore results in a corresponding energy jitter—an effective multi-shot 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 [42, 43]. In addition, long-term timing drifts need to be measured and corrected for with a feedback system operating at the same timescales.
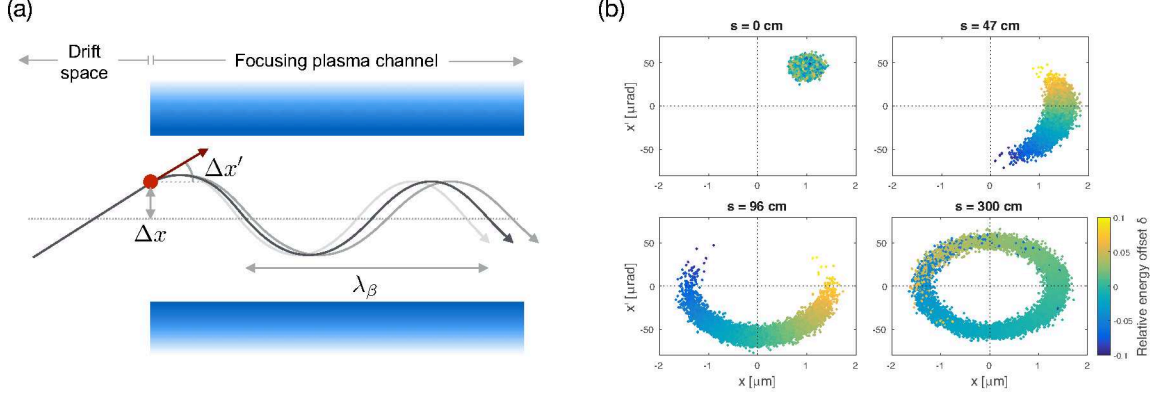
Fig. 6: (a) Illustration of projected emittance growth due to misalignments in a strong-focusing channel. (b) The accelerating bunch starts with an offset and an angle, which due to a finite energy spread smears out to a larger-area ring in phase space. Source: C. A. Lindstrøm *et al.*, *Proceedings of IPAC2016* (JACoW, Geneva, 2016), p. 2561 [44] (CC BY 3.0).

3.3.2 *Transverse misalignments*

Misalignment tolerances will also prove particularly challenging for high-gradient 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 Section 3.1.1) and dispersion (see Section 3.1.2), 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 [44]

$$\Delta\epsilon_j \approx \frac{1}{2}\left(\frac{\Delta x^2}{\beta_m} + \beta_m \Delta x'^2\right). \tag{16}$$

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 sub-μm-scale. If multiple stages are used, the emittance growth per stage is further constrained. For plasma accelerators with relevant gradients (GV/m-scale), the alignment tolerance is estimated to be around 10–50 nm [45–47].

3.4 Synchrotron radiation

As discussed in Section 2.1, 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 radiaton (ISR), with an average power emitted per bunch [48, 49]

$$P_{\text{ISR}} = \frac{e^4}{6\pi\epsilon_0 m^2 c} N\gamma^2 B^2, \tag{17}$$

9

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

However, high-gradient wakefield accelerators often operate with short bunches, in which case the bunch will emit *coherent* synchrotron radiation (CSR) [50,51]. In the case of full coherence, the electric fields of all particles add linearly such that the radiated power scales quadratically with particle number

$$P_{\mathrm{CSR}} = NP_{\mathrm{ISR}} \sim N^2. \tag{18}$$

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

$$P_{\mathrm{CSR}} = \frac{\kappa e^2 c}{\epsilon_0} \frac{N^2}{\rho^{2/3}\sigma_z^{4/3}}, \tag{19}$$

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

Equally important to the energy loss are the *transverse* CSR effects—strong intra-bunch transverse wakefields that can lead to significant emittance growth. 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 [52].

3.5 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), the effective gradient is a good approximate metric—it estimates the overall footprint of a 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 [53].

An interesting, and potentially worrying aspect of staged high-gradient wakefield accelerators is how the staging length scales with energy [54]. 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 capture and refocusing 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.
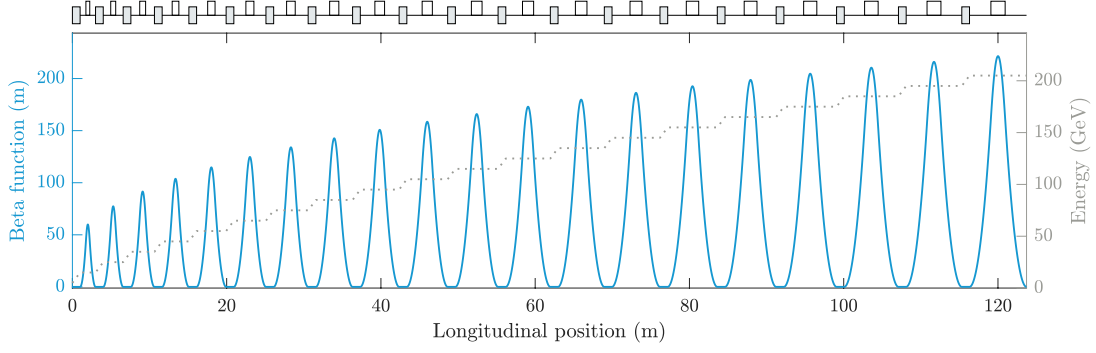
Fig. 7: Illustration of how the staging length increases with energy: the accelerator stages (gray boxes) remain the same length, whereas the length of the staging optic (white boxes) increases as $\sqrt{\gamma}$. As a consequence, the effective gradient decreases with energy. In this example, the initial beam energy is 5 GeV, and each stage adds 10 GeV over 1 m. 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.

4 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.

4.1 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, this corresponds to reducing the plasma density (see Eq. (5)). This mitigates the divergence/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 most 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* [55–57]. Figure 8 illustrates this concept.
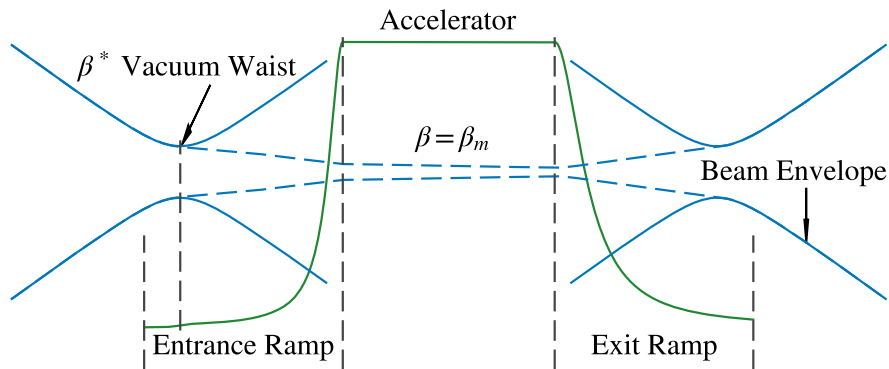

Fig. 8: Schematic of a plasma accelerator with density ramps. The beam is focused by external focusing to the vacuum waist (β^*) close to the start of the entrance ramp. If perfectly matched to the ramp, the beam stays matched throughout the accelerator ($\beta = \beta_m$ in the flat-top). Finally, the beam is transported through the exit ramp, which reduces the divergence before exiting into the outside vacuum. Source: R. Ariniello *et al.*, Phys. Rev. Accel. Beams **22**, 041304 (2019) [57] (CC BY 4.0).

Calculating the evolution of the beta function through a tailored plasma density profile is relatively straightforward via the *betatron equation* [58]

$$\frac{1}{2}\beta''(s)\beta(s) - \frac{1}{4}\beta'(s)^2 + K(s)\beta(s)^2 = 1, \tag{20}$$

where the focusing force is given by

$$K(s) = \frac{e^2}{2\epsilon_0}\frac{n(s)}{E(s)}. \tag{21}$$

Note that 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. (20) analytically in certain cases [59], in general it needs to be integrated numerically.

It is important to note that although plasma density ramps 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, the ramps must be *adiabatic* [60,61]. This effectively means that the plasma density is changing sufficiently slowly,

$$\left|\frac{n'(s)}{n(s)}\right| \ll \frac{1}{\beta_m(s)}, \tag{22}$$

such that $\alpha \approx 0$ throughout the entire ramp. An example of such a ramp would be $n(s) = n_0(1 + s/l_r)^{-2}$, where n_0 is the flat-top density and $l_r \gg 2\beta_{m0}$ is the characteristic ramp decay length (β_{m0} is the matched beta function in the flat-top). To reach a beta function β^* at the entrance/exit of such an adiabatic ramp, the overall ramp length must be $L_r \gg 2\beta^*$.

Adiabatic ramps are desirable due to their insensitivity to energy spread or slight mismatching, but this comes at the price of significantly longer ramp sections. Long ramps will both reduce the energy efficiency and the effective gradient of the accelerating structure, and can introduce potentially non-negligible and non-uniform 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.

4.2 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 (actively) driven or not.

4.2.1 Passive plasma lenses

Passive plasma lenses utilize the same mechanism as plasma density ramps—plasma wakefield focusing. The wakefield can either be driven 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 [62–65], passive plasma lenses in their modern form were proposed in 1989 [66] and have been successfully demonstrated for both beam drivers [67,68] and laser drivers [69]. Figure 9 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
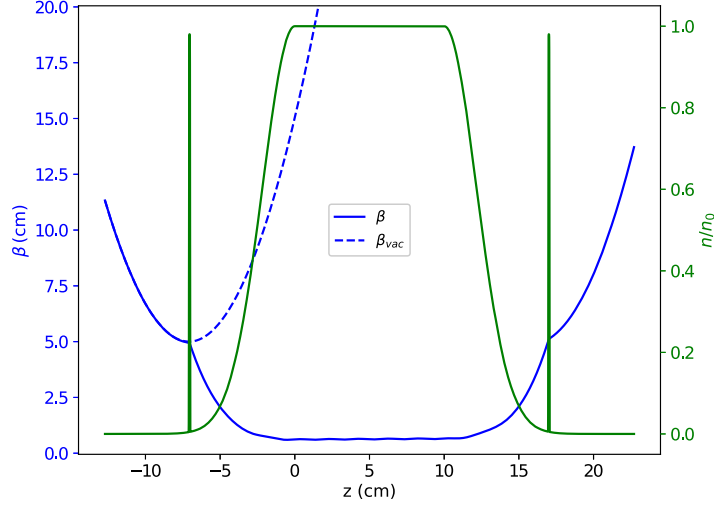
Fig. 9: Passive plasma lensing scheme proposed for FACET-II [70], where the lenses are integrated into the plasma density ramps. A gas jet produces a locally higher gas density that is ionized by a laser. The lens is driven by the same beam driver as is used in the main accelerating section. Source: C. E. Doss *et al.*, Phys. Rev. Accel. Beams **22**, 111001 (2019) [71] (CC BY 4.0).

force for electrons is exerted by an electric field, we can calculate the equivalent magnetic field gradient (as $E_r \equiv cB_\phi$ for ultra-relativistic particles)

$$g_{\text{PPL}} = \frac{en}{2c\epsilon_0},$$ (23)

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. (23), but the exposed charge 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.

4.2.2 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 make them compact and simple to operate. On the other hand, the focusing strength is typically limited to the kT/m-range—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) [8].

The history of active plasma lenses started in 1950 [72, 73], during which they were studied for various purposes such as ion focusing [74] and antimatter capture [75] before their recent revival for use in plasma accelerators [76]. Modern active plasma lenses consist of a thin (mm-scale) gas-filled capillary with electrodes on either side [77] (see Fig. 10 for a schematic overview). A high-voltage discharge ionizes the gas before a large current passes through the plasma. By Ampère's law, the azimuthal magnetic field at each radius inside the lens is proportional to the total current enclosed at that radius. If, ideally,
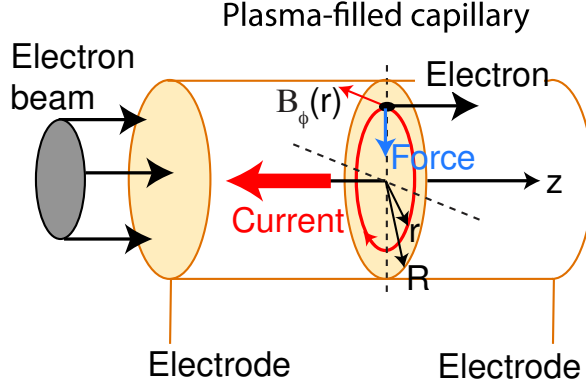
Fig. 10: Schematic of an active plasma lens: an electron beam enters a gas-filled capillary, which is discharged via two high-voltage electrodes. The current flowing between the electrodes induces an azimuthal magnetic field B_ϕ that increases with radius r—a radial focusing force that focuses the electron beam in both planes. Source: J. van Tilborg *et al.*, Phys. Rev. Lett. **115**, 184802 (2015) [76] (reproduced with permission).

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

$$g_{\text{APL}} = \frac{\mu_0 I}{2\pi R^2},$$ (24)

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 (24) 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* [78] 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) [79, 80]. Interestingly, while this aberration is present in light gases like hydrogen and helium [81], it is possible to fully suppress it in a heavier gas like argon where the heat transfer to the wall is significantly slower [82]. Another effect which can cause nonuniform focusing fields is the *z-pinch* effect [83], where the magnetic field of the lens is strong enough to self-focus its own current.

2. *Coulomb scattering* [12, 13] can cause emittance growth due to the atomic/ionic density on axis. Since the emittance growth rate scales as

$$\frac{d\epsilon_n}{ds} \sim \frac{n\beta Z^2}{\gamma},$$ (25)

   where Z is the atomic number and β is the beta function in the lens, scattering is much more severe for heavier gases—the effect is usually negligible in hydrogen and helium, but can be a problem 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 [84].

Clearly, any application of active plasma lensing needs to take all the above effects into account.

4.3 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/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*.

4.3.1 Apochromatic correction

Using linear optics elements, it is fundamentally impossible to provide fully achromatic beam transport for all energies [85]. 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 different beta function evolution, but eventually converging to the same focus (see Fig. 11). This idea was first introduced for the final focusing of CLIC in 1987, and is known as *aprochromatic correction* [86].
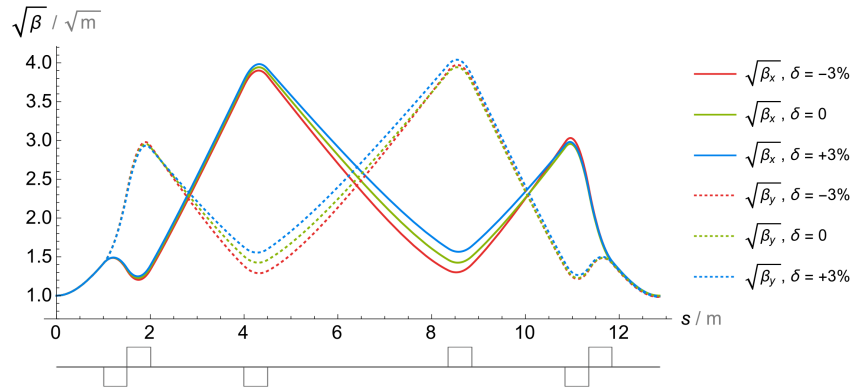

Fig. 11: Illustration of apochromatic correction in a lattice of only linear focusing optics (six quadrupole magnets). The lattice is constructed such that beta functions evolve differently for the various energy slices, but eventually converge before entering the next stage. This method also works for plasma lenses and plasma density ramps. Source: C. A. Lindstrøm and E. Adli, Phys. Rev. Accel. Beams **19**, 071002 (2016) [32] (CC BY 3.0).

A beam transport lattice can be apochromatically corrected to arbitrary order [32], such that not only the first order chromatic amplitude is zero, but also higher orders—at the cost of introducing more degrees of freedom (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.

4.3.2 Sextupoles in dispersive sections

In order to truly increase the energy acceptance of a beamline, 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 high-gradient wakefield accelerators have similar parameters, and may therefore require nonlinear optics lattices.

The underlying concept of chromaticity correction with sextupoles is relatively straightforward.

15

The local focusing fields in a sextupole are proportional to the transverse offset, so if the beam is dispersed such that different energies enter the sextupole at different offsets, the chromaticity can be cancelled. If the beam is horizontally dispersed ($x \rightarrow x + D_x \delta$), the focusing forces are given by

$$F_x \quad \sim \quad x D_x \delta + \frac{1}{2}\left(x^2 - y^2\right) + \frac{1}{2}D_x^2 \delta^2 \tag{26}$$

$$F_y \quad \sim \quad y D_x \delta + yx \tag{27}$$

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. 12 for an illustration of this concept.

A particularly important concept for mitigating chromaticity is that of *local chromaticity correction* [87, 88]. 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 [89], 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.
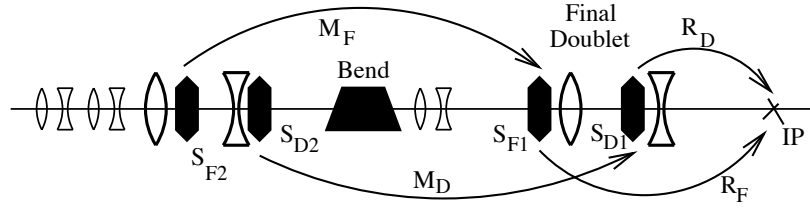

Fig. 12: Local chromaticity correction for a linear collider using nonlinear optics and dispersive elements. The beam is dispersed with a dipole (black trapezoid) onto the final doublet of quadrupoles (white). Each quadrupole has a corresponding sextupole (black hexagons) to locally correct the chromaticity. Prior to the dipole is a similar doublet that compensates the geometric terms introduced by each of the two final sextupoles. Source: P. Raimondi and A. Seryi, Phys. Rev. Lett. **86**, 3779 (2001) [87] (reproduced with permission).

The drawbacks of chromaticity correction using nonlinear optics are: (1) the introduction of large dispersion, and (2) 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.

4.4 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 important to note that all these techniques will also require some form of emittance-preserving out-coupling—which means that many of the above considerations are still relevant, but the requirements for compactness can be relaxed.

4.4.1 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 Joules, 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* [90, 91].

Self-modulation is a process whereby a beam in a plasma self-interacts with its own focusing and defocusing wakefields. 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 [92] at CERN, utilizing 400 GeV proton bunches from the Super Proton Synchrotron has successfully demonstrated this self-modulation [93, 94], and used it to accelerate electron bunches up to 2 GeV [95] in a 10 m long plasma stage (see Fig. 13).
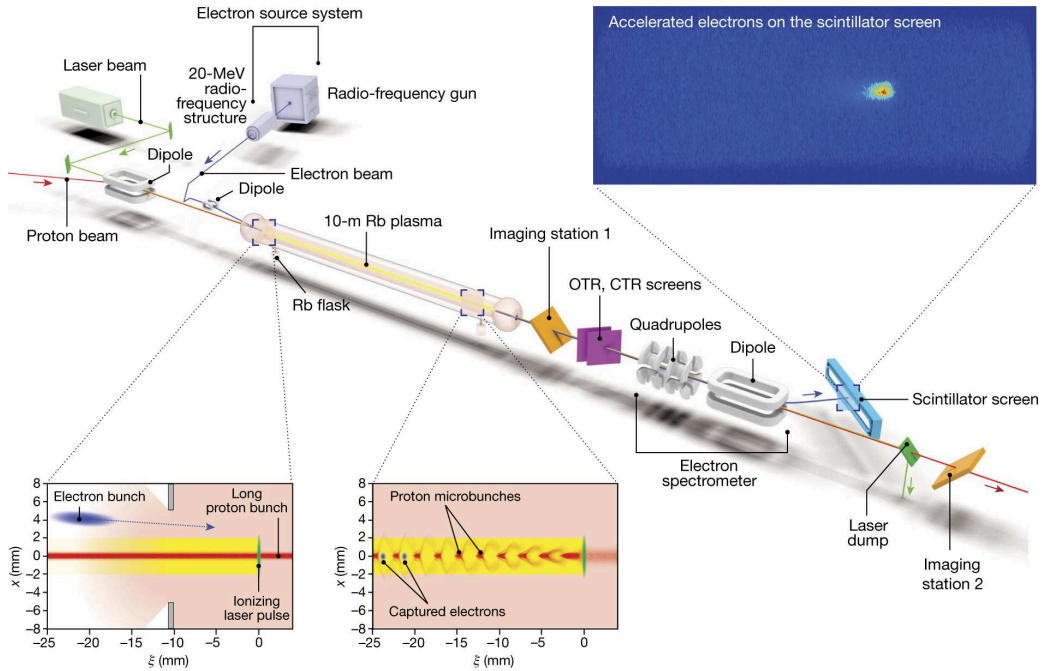

Fig. 13: The AWAKE experimental setup, using a proton bunch in a laser-ionized plasma to accelerate an electron beam to 2 GeV (upper right). The long proton bunch self-modulated (lower left) into a train of short bunches, resonantly driving a plasma wakefield, into which electrons were externally injected and then accelerated. Source: E. Adli *et al.* (AWAKE Collaboration), Nature **561**, 363 (2018) [95] (CC BY 4.0).

While self-modulated proton-driven wakefield acceleration constitutes one of the most promising ways of realizing high-gradient, high-energy acceleration, there are also a few limitations. Firstly, a very large proton synchrotron is required—it may only make sense to build such accelerators in the vicinity of existing infrastructure. The limited repetition rate of such machines is also of concern. Secondly, the external injection into the wakefield is non-trivial—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.
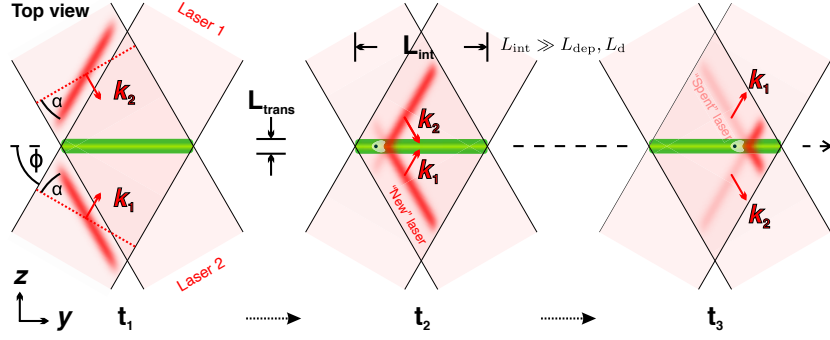
Fig. 14: Illustration of the travelling-wave electron acceleration scheme: two lasers are coupled in transversely into an on-axis plasma channel. The superposition of the two lasers drives a wakefield, and since the laser pulse fronts are tilted, the travelling wave can be made to move at speed c. Source: A. Debus *et al.*, Phys. Rev. X **9**, 031044 (2019) [96] (CC BY 4.0).

4.4.2 *Travelling-wave electron acceleration*

A proposal for reaching high energies in a single-stage laser-driven wakefield accelerator is using a *travelling* laser focus. The idea is to couple in two laser drivers transversely—one from each side—and use their superposition to drive the wakefield. Further, if their pulse fronts are tilted just right, the laser focus can be made to travel at the speed of light c—to avoid dephasing between the laser and the electron beam. Figure 14 illustrates this so-called travelling-wave electron acceleration scheme [96].

The scheme 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.

4.4.3 *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* [97]. Similar to how laser pulses are guided in straight plasma channels with a transverse (parabolic) density profile [98,99], 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 15 illustrates this scheme.

While the beam never really exits the plasma during acceleration, transitioning from one stage to the next will introduce transverse oscillations to the beam centroid. This causes emittance growth and may lead to a beam-breakup instability. Nevertheless, mitigation strategies may be found—in which case this scheme promises an achievable path towards high energy acceleration with high-gradient wakefields.
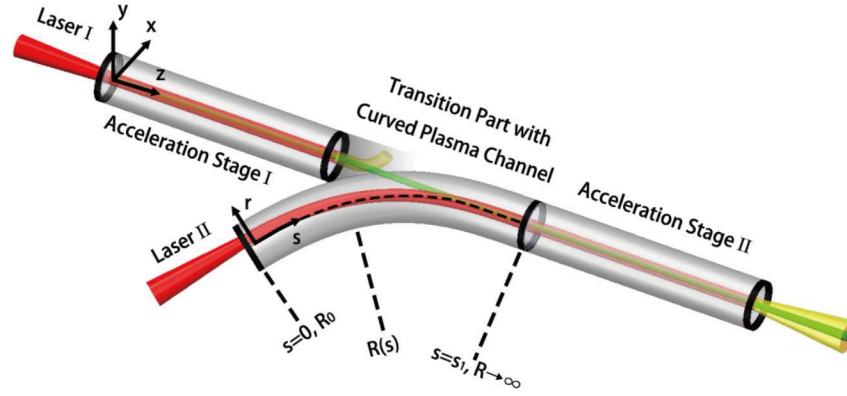
Fig. 15: Illustration of the curved plasma channel scheme: after an initial laser plasma accelerator stage, a fresh laser pulse (red) is in-coupled using a curved plasma channel. This channel guides the laser around the curve, but lets the electron beam (green) pass straight through until it is trapped by the new wakefield. The depleted laser (yellow) from the first stage is also out-coupled in this transition region. Source: J. Luo *et al.*, Phys. Rev. Lett. **120**, 154801 (2018) [97] (reproduced with permission).

5 Conclusions

The application of high-gradient 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 multi-stage accelerator, many requirements must be met: matching of beta functions, dispersion cancellation, isochronicity, avoiding too much synchrotron radiation, as well as tight synchronization and misalignment tolerances. At the same time, 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 travelling-wave electron acceleration, and laser-driven acceleration in curved plasma channels.

In summary, staging of high-gradient accelerators is still a young field, and the best solutions are surely still to come!

Acknowledgements

References

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

[2] T. Behnke (editor) *et al.*, *The International Linear Collider Technical Design Report*, (International Linear Collider, 2013).

[3] M. Aicheler (editor) *et al.*, *A Multi-TeV linear collider based on CLIC technology: CLIC Conceptual Design Report* (CERN, Geneva, 2013).

[4] T. Tajima and J. M. Dawson, "Laser electron accelerator," Phys. Rev. Lett. **43**, 267 (1979).

[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] W. Gai *et al.*, "Experimental demonstration of wake-field effects in dielectric structures," Phys. Rev. Lett. **61**, 2756 (1988).

[8] S. Steinke *et al.*, "Multistage coupling of independent laser-plasma accelerators," Nature (London) **530**, 190 (2016).

[9] C. Thaury *et al.*, "Plasma mirrors for ultrahigh-intensity optics," Nat. Phys. **3**, 424 (2007).

[10] B. Allen *et al.*, "Experimental study of current filamentation instability," Phys. Rev. Lett. **109**, 185007 (2012).

[11] G. Raj *et al.*, "Probing ultrafast magnetic-field generation by current filamentation instability in femtosecond relativistic laser-matter interactions," arXiv:1907.12052.

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

[13] N. Kirby *et al.*, "Emittance growth from multiple Coulomb scattering in a plasma wakefield accelerator," *Proceedings of PAC2007, Albuquerque, NM, USA* (IEEE, Piscataway, NJ, 2007), p. 3097.

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

[15] F. Arntz *et al.*, "A kicker driver for the international linear collider," *Proceedings of PAC2007, Albuquerque, NM, USA* (IEEE, Piscataway, NJ, 2007), p. 2972.

[16] A. W. Chao, *Physics of Collective Beam Instabilities in High-Energy Accelerators* (Wiley, New York, NY, 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 radioâĂŘfrequency 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. 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. ST Accel. Beams **17**, 091302 (2014).

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

[23] 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).

[24] C. Huang *et al.*, "Hosing instability in the blow-out regime for plasma-wakefield acceleration," Phys. Rev. Lett. **99**, 255001 (2007).

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

[26] 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).

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

[28] P. Antici *et al.*, "Laser-driven electron beamlines generated by coupling laser-plasma sources with conventional transport systems," J. Appl. Phys. **112**, 044902 (2012).

[29] M. Migliorati *et al.*, "Intrinsic normalized emittance growth in laser-driven electron accelerators," Phys. Rev. ST Accel. Beams **16**, 011302 (2013).

[30] H. Zyngier, "Strategy for correcting for chromaticity" (in French), *LAL-77/35* (Laboratoire de l'Accélérateur Linéaire, Orsay, 1977).

[31] B. W. Montague, "Linear optics for improved chromaticity correction," *LEP Note 165* (CERN, Geneva, 1979).

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

[33] C. Lejeune and J. Aubert, "Emittance and brightness: definitions and measurements," *Applied Charged Particle Optics* (Academic Press, 1980), p. 159.

[34] D. T. Pierce *et al.*, "The GaAs spin polarized electron source," Rev. Sci. Instrum. **51**, 478 (1980).

[35] W. P. Leemans *et al.*, "GeV electron beams from a centimetre-scale accelerator," Nat. Phys. **2**, 696 (2006).

[36] A. Buck *et al.*, "Shock-front injector for high-quality laser-plasma acceleration," Phys. Rev. Lett. **110**, 185006 (2013).

[37] A. Deng *et al.*, "Generation and acceleration of electron bunches from a plasma photocathode," Nat. Phys. **15**, 1156 (2019).

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

[39] T. Katsouleas, "Physical mechanisms in the plasma wake-field accelerator," Phys. Rev. A **33**, 2056 (1986).

[40] M. Tzoufras *et al.*, "Beam loading in the nonlinear regime of plasma-based acceleration," Phys. Rev. Lett. **101**, 145002 (2008).

[41] A. Ferran Pousa, A. Martinez 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).

[42] S. Schulz *et al.*, "Femtosecond all-optical synchronization of an X-ray free-electron laser," Nat. Commun. **6**, 5938 (2015).

[43] 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.

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

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

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

[47] D. Schulte, "Application of advanced accelerator concepts for colliders," Rev. Accel. Sci. Technol. **9**, 209 (2017).

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

[49] J. D. Jackson, *Classical Electrodynamics, 3rd edition* (New York, NY, 1999).

[50] 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. A **398**, 373 (1997).

[51] A. Novokhatski, "Coherent synchrotron radiation: theory and simulations," SLAC-PUB-14893; ICFA Beam Dyn. Newslett. **57**, 127 (2012).

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

[53] 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).

[54] C. A. Lindstrøm *et al.*, "Staging optics considerations for a plasma wakefield acceleration linear collider," Nucl. Instrum. Methods Phys. Res. A **829**, 224 (2016).

[55] K. A. Marsh *et al.*, "Beam matching to a plasma wake field accelerator using a ramped density profile at the plasma boundary," *Proceedings of PAC2005, Knoxville, TN, USA* (JACoW, Geneva, 2005), p. 2702.

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

[57] 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).

[58] M. Martini, "An introduction to transverse beam dynamics in accelerators," *CERN/PS 96-11* (CERN, Geneva, 1996), p. 16.

[59] X. L. Xu *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).

[60] R. H. Helm, "Adiabatic approximation for dynamics of a particle in the field of a tapered solenoid," *SLAC-Report-4* (1962).

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

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

[63] 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).

[64] W. H. Bennett, "Magnetically self-focussing streams," Phys. Rev. **45**, 890 (1934).

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

[66] 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).

[67] J. B. Rosenzweig *et al.*, "Demonstration of electron beam self-focusing in plasma wake fields," Phys. Fluids B **2**, 1376 (1990).

[68] J. S. T. Ng *et al.*, "Observation of plasma focusing of a 28.5 GeV positron beam," Phys. Rev. Lett. **87**, 244801 (2001).

[69] C. Thaury *et al.*, "Demonstration of relativistic electron beam focusing by a laser-plasma lens," Nat. Commun. **6**, 6860 (2015).

[70] C. Joshi *et al.*, "Plasma wakefield acceleration experiments at FACET II," Plasma Phys. Control. Fusion **60**, 034001 (2018).

[71] C. E. Doss, *et al.*, "Laser-ionized, beam-driven, underdense, passive thin plasma lens," Phys. Rev. Accel. Beams **22**, 111001 (2019).

[72] 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).

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

[74] E. Boggasch *et al.*, "z-pinch plasma lens focusing of a heavy-ion beam," Phys. Rev. Lett. **66**, 1705 (1991).

[75] R. Kowalewicz *et al.*, "Performance of the CERN plasma lens in laboratory and beam tests at the antiproton source," *Proceedings of PAC1991, San Francisco, CA, USA* (IEEE, New York, NY, 1991), p. 2631.

[76] J. van Tilborg *et al.*, "Active plasma lensing for relativistic laser-plasma-accelerated electron beams," Phys. Rev. Lett. **115**, 184802 (2015).

[77] C. A. Lindstrøm *et al.*, "Overview of the CLEAR plasma lens experiment," Nucl. Instrum. Methods Phys. Res. A **909**, 379 (2018).

[78] J. van Tilborg *et al.*, "Nonuniform discharge currents in active plasma lenses," Phys. Rev. Accel. Beams **20**, 032803 (2017).

[79] N. A. Bobrova *et al.*, "Simulations of a hydrogen-filled capillary discharge waveguide," Phys. Rev. E **65**, 016407 (2001).

[80] 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).

[81] J.-H. Röckemann *et al.*, "Direct measurement of focusing fields in active plasma lenses," Phys. Rev. Accel. Beams **21**, 122801 (2018).

[82] C. A. Lindstrøm *et al.*, "Emittance preservation in an aberration-free active plasma lens," Phys. Rev. Lett. **121**, 194801 (2018).

[83] 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).

[84] C. A. Lindstrøm and E. Adli, "Analytic plasma wakefield limits for active plasma lenses," arXiv:1802.02750 (2018).

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

[86] B. W. Montague and F. Ruggiero, "Apochromatic focusing for linear colliders," *CLIC Note 37* (CERN, Geneva, 1987).

[87] P. Raimondi and A. Seryi, "Novel final focus design for future linear colliders," Phys. Rev. Lett. **86**, 3779 (2001).

[88] 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).

[89] J. Irwin *et al.*,, "The optics of the final focus test beam," Conf. Proc. C**910506**, 2058 (1991).

[90] N. Kumar, A. Pukhov and K. Lotov, "Self-modulation instability of a long proton bunch in plasmas," Phys. Rev. Lett. **104**, 255003 (2010).

[91] 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).

[92] E. Gschwendtner *et al.*, "AWAKE, The Advanced Proton Driven Plasma Wakefield Acceleration Experiment at CERN," Nucl. Instrum. Methods Phys. Res. A **829**, 76 (2016).

[93] 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).

[94] 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).

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

[96] A. Debus *et al.*, "Circumventing the dephasing and depletion limits of laser-wakefield acceleration," Phys. Rev. X **9**, 031044 (2019).

[97] J. Luo *et al.*, "Multistage coupling of laser-wakefield accelerators with curved plasma channels," Phys. Rev. Lett. **120**, 154801 (2018).

[98] C. G. Durfee and H. M. Milchberg, "Light pipe for high intensity laser pulses," Phys. Rev. Lett. **71**, 2409 (1993).

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