Path to AWAKE: Evolution of the concept

Caldwell, A (Max Planck Institute for Physics, Munchen, DE) *et al*

29 November 2015

The EuCARD-2 Enhanced European Coordination for Accelerator Research & Development project is co-funded by the partners and the European Commission under Capacities 7th Framework Programme, Grant Agreement 312453.

This work is part of EuCARD-2 Work Package 13: Novel Acceleration Techniques (ANAC2).

Abstract

This report describes the conceptual steps in reaching the design of the AWAKE experiment currently under construction at CERN. We start with an introduction to plasma wakefield acceleration and the motivation for using proton drivers. We then describe the construction status as presented in this conference is given in [1].

Keywords: Plasma wakefield acceleration, Proton driver, Self-modulation instability
1. Introduction

Particle accelerators are the fundamental research tools of the high energy physics community for studying the basic laws that govern our Universe. Experiments conducted at the LHC will give us new insights into the physical world around us. Complementing this, future lepton–lepton and lepton–hadron colliders should reach the TeV scale. Circular electron accelerators are not feasible at these energies; hence future TeV electron accelerator designs are based on linear colliders. However, as the beam energy increases, the scale and cost of conventional accelerators become very large. For a linear accelerator, the size and cost depend on the maximum accelerating gradient in radiofrequency (RF) cavities. At present, metallic cavities achieve maximum accelerating gradients around 100 MV/m. To reach the TeV scale in a linear accelerator, the length of the machine is therefore tens of kilometers.

It is natural to think about how to make future machines more compact, and plasma acceleration is a possible solution. A plasma is a medium consisting of ions and free electrons; therefore, it can sustain very large electric fields (> GV/m) [2, 3]. In the last few decades, more than three orders of magnitude higher acceleration gradients than in RF cavities have been demonstrated with plasmas in the laboratory [4, 5]. Beam-driven plasma wakefield acceleration experiments performed at SLAC [5] successfully doubled the energies of some of the electrons in the initial 42 GeV beam in less than 1 m of plasma.

Generally speaking, a plasma acts as an energy transformer; it transfers the energy from the driver (laser or particle bunch) to the witness bunch that is accelerated. Current proton synchrotrons are capable of producing high energy protons, reaching up to multi TeVs (the LHC), so that a new accelerator frontier would be opened if we could engineer plasma wakefield acceleration [6]. In the PWFA, the mostly transverse space charge field of the relativistic particle bunch displaces the plasma electrons. In the case of a negatively charged particle bunch, the plasma electrons are expelled from the bunch volume. They are then attracted back towards the axis by the net positive charge left behind the bunch head, overshoot and sustain the plasma oscillations. The excited wakefields usually have accelerating (decelerating) longitudinal components and transverse focusing (defocusing) components with comparable amplitudes. In the linear wakefield regime, these fields vary periodically behind the drive bunch and have a $\pi/2$ phase difference.

The angular frequency $\omega_p$ of the plasma wave is fixed by the local plasma density $n_0$: $\omega_p = \sqrt{4\pi n_0 e^2/m_e}$, where $m_e$ is the electron mass and $e > 0$ is the elementary charge. On the time scale of a few wave oscillations, the much heavier plasma ions can be considered immobile. The plasma wave or wake is tied to the drive bunch, and its phase velocity $v_{\phi}$ is close to that of the drive bunch $v_b$ and to the speed of light $c$. Its wavelength is therefore $\lambda_p = 2\pi/k_p = 2\pi c/\omega_p$. The maximum amplitude of the longitudinal electric field in the wave is on the order of the wave-breaking field $E_b = m_e\omega_p^2/e$ [8]. The plasma wave is most effectively driven by a bunch with a length on the order of the wave period: $\sigma_{\phi} \approx \lambda_p/\sqrt{2\pi}$ [9].

Experimental results demonstrating the driving of plasma wakefields by a relativistic electron bunch and the acceleration of a witness bunch were first published in 1988 [10]. In recent years, PWFA research has been led by the program at SLAC [11]. Their experiments with single, 42 GeV electron bunches with $2\times10^{10}$ particles in $\sigma_{\phi} = 20\mu m$ have demonstrated the energy gain by trailing electrons of 42 GeV in 85 cm of plasma [5]. This corresponds to an accelerating gradient in excess of 50 GeV/m sustained over a meter-scale distance. Current experiments at SLAC-FACET aim at demonstrating large energy gain (on the order of the incoming particles energy, $\approx 20$ GeV) with a narrow final energy spread by a separate witness bunch [12, 13].

3. Proton drivers

A future linear electron accelerator, such as the ILC [14], should produce bunches with several $10^{10}$ particles each with $\sim 250$ GeV of energy. These bunches carry about 1 kJ of energy each, and therefore FACET-like drive bunches carrying about 60 J would require staging of many plasma sections to reach the desired energy. An alternative to this staging approach is to use a drive bunch carrying many kilojoules of energy. Such bunches are routinely produced by the CERN Super Proton Synchrotron (SPS, 450 GeV, $3 \times 10^{11}$ protons, $\sim 20$ kJ) or Large Hadron Collider (LHC, 6.5 TeV, $1.2 \times 10^{11}$ protons, $\sim 125$ kJ).

The concept of proton-driven plasma wakefield acceleration made its appearance in 2009 after proof-of-principle simulation papers [15, 16]. In these simulations, an incoming 10 GeV electron bunch gained 650 GeV in 400 m of plasma driven by a 100 $\mu$m-long, 1 TeV proton bunch. It was also realized that a high energy transfer efficiency between the driver and the witness was possible for proton energies above 1 TeV [16].

The extremely short driver length required for efficient excitation of the plasma wave presents a serious obstacle to a realization of the concept. The CERN proton bunches available today are approximately 10 centimeters long (the root-mean-square length, $\sigma_{\phi}$) and are ineffective at driving large wakefield amplitudes. From conservation of the longitudinal phase volume we can derive that a factor of $10^4$ longitudinal compression of the 6.5 TeV LHC bunch (energy spread 0.01%) would result in a $100\mu$m bunch with a 10% energy spread, that is 650 GeV.
for this bunch energy. A compressor capable of delivering this huge energy spread to the proton bunch would be prohibitively expensive and maybe as complicated as the ILC itself. Simulations [16] also show that state-of-the-art proton bunches have no safety margin in the transverse emittance, so the longitudinal phase volume cannot be much reduced by blowing up the transverse phase volume. Even for lower energy proton bunches, the longitudinal compression to sub-millimeter scales requires a long RF system to provide the necessary energy chirp along the bunch [17,[20].

An alternative to extreme bunch compression is multi-bunch wave excitation. In this scheme the plasma wave is resonantly driven by a train of short microbunches spaced one wakefield period apart. It is exactly this scheme that was first proposed as the plasma wakefield accelerator in [17]. The multi-bunch excitation was demonstrated experimentally by several groups [21–25] and studied in several theoretical and simulation papers [26–33]. To estimate the required compression rate we assume the wave is driven by N microbunches in the plasma of the same density $n_0 = 10^{15} \text{ cm}^{-3}$ as was used in [16]. The train length is then $N \lambda_p \approx N \text{ mm}$. Each microbunch must fit roughly $1/4$ of the wakefield period to be focused and decelerated simultaneously. The longitudinal space occupied by the driver thus has to be reduced from $\sim 2\sigma_{r, \theta}$ to $N \lambda_p/4$. Correspondingly, the energy spread must increase $8\sigma_{r, \theta}/(N \lambda_p)$ times. For the 6.5 TeV LHC bunch and $N = 10$, the final energy spread is about 0.5% or $\sim 30 \text{ GeV}$, which is still large but will not make the machine prohibitively expensive. If the bunch is not compressed, then $N \sim 2\sigma_{r, \theta}/\lambda_p \gtrsim 100$.

4. The self-modulation instability

The conventional method of beam bunching involves energy chirping along the beam and subsequent longitudinal redistribution of the beam charge in a region with nonzero momentum compaction factor [17, 20]. This method conserves the beam charge and makes optimal use of accelerated protons.

The plasma offers another method that relies on a beam–plasma instability. The instability is caused by mutual amplification of the rippling of the beam radius and the resulting plasma wave, which selectively focuses or defocuses different slices of the beam. Under proper conditions, the instability splits the beam into microbunches spaced by exactly one plasma wavelength (Fig. 1). Beam particles initially located between the microbunches are defocused by the plasma wave and form a wide halo around the bunch train (Fig. 2). Although plasma-based bunching is energy inefficient (as a major fraction of the proton beam energy is lost in the halo), it is relatively cheap and easy, so it is ideally suited for first proof-of-principle experiments on proton driven wakefield acceleration.

The instability of interest is the self-modulation instability (SMI), which belongs to the large family of beam–plasma instabilities (see review [35]). The SMI can be viewed as the axisymmetric mode of the transverse two-stream (TTS) instability [36,37]. The latter is characterized by a low beam density $n_0 \ll n_0$, radial beam non-uniformity, and high relativistic factor of the beam. The SMI is a convective instability that grows both along the bunch and along the plasma.

It was noticed in simulations [38] that the SMI initiated by a small seed perturbation transforms a long particle beam into a bunch train. The seed perturbation is needed to give preference to a single unstable mode. Otherwise a competitive growth of several modes would inevitably destroy the beam even in the fully axisymmetric setup [30,39]. If an externally seeded mode dominates, it suppresses growth of other modes and produces a train of well-separated microbunches. Three-dimensional simulations [40] confirmed that non-axisymmetric modes of the TTS instability (hosing modes [36,37]) are also suppressed if the seed perturbation is strong enough. The formed bunches propagate stably during very long distances, provided that the nonlinear regime is avoided [41]. This result has opened the

---

1 A completely new proton accelerator capable of producing and accelerating short bunches of protons would be even better.
path to experimental verification of proton driven plasma wakefield acceleration.

The parasitic instabilities could originate from shot noise, which is very low for long beams [42], so the seed wakefield does not have to be very strong either. A short electron bunch [42], a powerful laser pulse [43], a sharp cut in the bunch current profile [40, 44], or a relativistic ionization front co-propagating within the drive bunch can seed the SMI quite well. Analytical and numerical calculations, however, have shown that bunches with long rise times (longer than or about the plasma wavelength) do not produce stable bunch trains [30, 38, 39]. A quantitative theory which would determine the minimum acceptable seed strength is still missing. Available theoretical studies are mainly focused on the linear stage of the instabilities in the case of narrow beams with a constant emittance [41, 45, 46]. However, this problem is not of a vital importance now, since a sufficient seeding method was chosen for the first experimental realization, which is a co-propagating ionization front created by a short laser pulse (Fig. 2). In this method, the forward part of the proton bunch freely propagates in the neutral gas and does not contribute to wakefield formation. The plasma interacts with the rear part only (defined as the part of the proton bunch coming after the laser pulse) and this is identical in practice to a sharply cut bunch. This method has an additional advantage of solving the problem of plasma creation as well.

As a long-term prospect, acceleration of electrons in the wake of a self-modulating 7 TeV LHC beam was also studied [47]. A test electron bunch was accelerated to 6 TeV, thus proving the capability of the self-modulation scheme to reach a multi-TeV energy scale with state-of-the-art proton beams. The high energy gain is only possible in a longitudinally non-uniform plasma with a small density step in the region of instability growth [48]. The density step modifies the beam evolution in such a way that the beam shape stops changing at the moment of full microbunching [49]. Otherwise the beam self-organization will not stop at microbunching and will proceed to destroy the microbunches soon after the maximum field is reached. The reason lies in the slow motion of the defocusing field regions with respect to the bunch. The field evolution for the stepped plasma profile is shown in Fig. 3 in comparison with the uniform plasma case for the LHC beam. With the density step, the wakefield is preserved for a long distance at a large fraction of the maximum amplitude. It is particularly remarkable that long acceleration distances are possible without additional focusing of the proton beam by external quadrupoles; these were an essential part of the initial concept [15, 16]. The addition of the plasma density step is thus considered a likely upgrade of the AWAKE experiment.

5. Early outline of the experiment

Two beams of different energies were analyzed as possible candidates for the first experiment on proton driven plasma wakefield acceleration: a 24 GeV beam in the Proton Synchrotron (PS) and 450 GeV beam in the SPS. At low energies (24 GeV), the excited fields turn out to be much lower because of the quick emittance-driven blowup of the beam radius [47, 50]. Therefore the SPS proton beam was chosen. The ten meter long plasma envisaged for the first experiment is too short to produce a reliably measurable energy change of the proton beam [51, 52] (Fig. 4). Therefore, injection of externally produced electrons becomes a must for probing the excited wakefields. With the addition of the electron beam, the broad outlines of the experiment were settled, and the project was proposed for realization at CERN in the Letter of Intent [53], which was submitted to the SPS Committee in May 2011. The experiment was recommended for further review, including preparation of a Design Report.

The first version of the experimental layout is shown in Fig. 5. The proton beam delivered from the SPS ring propagates through the ∼10 m long plasma cell, excites the wakefield, and becomes modulated by this wakefield. The short laser pulse propagates collinearly with the proton beam and serves the dual function of creating the plasma and seeding the SMI. The electron bunch collinear with the proton beam is accelerated by the wakefield and characterized with a magnetic spectrometer. The proposed location for the experiment was the TT4/TT5 hall (in the so called West Area) into which the 450 GeV beam is transported through the TT61 tunnel. Studies underlying this early stage of the project are documented in papers [27, 54] and conference proceedings [51, 52, 55–58]. The main beam and plasma parameters for the earliest vision of the experiment are given in the first data column of Table 1.
Table 1: Evolution of baseline parameters for the AWAKE experiment.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma species</td>
<td>Li, Cs, or Ar</td>
<td>Rb</td>
<td>Rb</td>
</tr>
<tr>
<td>Plasma density, $n_{e0}$</td>
<td>$7 \times 10^{14}$ cm$^{-3}$</td>
<td>$7 \times 10^{14}$ cm$^{-3}$</td>
<td>$7 \times 10^{14}$ cm$^{-3}$</td>
</tr>
<tr>
<td>Plasma source</td>
<td>not decided</td>
<td>gas cell &amp; laser</td>
<td>gas cell &amp; laser</td>
</tr>
<tr>
<td>Proton bunch population, $N_b$</td>
<td>$1.15 \times 10^{11}$</td>
<td>$3 \times 10^{11}$</td>
<td>$3 \times 10^{11}$</td>
</tr>
<tr>
<td>Proton bunch length, $\sigma_{zb}$</td>
<td>12 cm</td>
<td>12 cm</td>
<td>12 cm</td>
</tr>
<tr>
<td>Proton bunch radius, $\sigma_{rb}$</td>
<td>200 µm</td>
<td>200 µm</td>
<td>200 µm</td>
</tr>
<tr>
<td>Proton energy, $W_b$</td>
<td>450 GeV</td>
<td>400 GeV</td>
<td>400 GeV</td>
</tr>
<tr>
<td>Proton bunch normalized emittance, $\epsilon_{bn}$</td>
<td>$3.5 \mu$m</td>
<td>$3.5 \mu$m</td>
<td>$3.5 \mu$m</td>
</tr>
<tr>
<td>Electron injection method</td>
<td>not decided</td>
<td>side</td>
<td>oblique</td>
</tr>
<tr>
<td>Electron bunch radius $\sigma_{re}$</td>
<td>–</td>
<td>200 µm</td>
<td>250 µm</td>
</tr>
<tr>
<td>Electron energy $W_e$</td>
<td>–</td>
<td>16 MeV</td>
<td>16 MeV</td>
</tr>
</tbody>
</table>

6. Plasma uniformity challenge

For the baseline plasma density, the plasma wavelength is rather short, $\lambda_p \approx 1.26$ mm, so the number of micro-bunches is large, $N \sim \sigma_{zb}/\lambda_p \sim 100$. Fields of this number of bunches can add coherently only if the eigenfrequency of plasma oscillations is kept constant along the plasma, otherwise the beam bunches would arrive at the wrong phase of the plasma oscillation. Computer simulations of perturbed density plasmas [59] show that the instability is less sensitive to plasma density non-uniformities than the linear theory [60] or simple estimates suggest. The accelerated electrons, however, are sensitive, and the reason is illustrated in Fig. 6. As the electrons enter a region of detuned plasma density, the plasma wavelength changes, as does the phasing of the plasma wave with respect to accelerated electrons. If the density increases with respect to the design value $n_{e0}$, the plasma wavelength shortens, and the defocusing phase of the wave catches up to the electrons and scatters them transversely [Fig. 6(a)]. If the density reduces, the wavelength gets longer, and the electrons fall into the decelerating phase of the wave [Fig. 6(c)]. These effects are less serious for the protons because of their large longitudinal momentum. These simple arguments lead to a simple engineering formula for maximum acceptable density perturbation,

$$\delta n_e/n_{e0} = 0.25/N,$$

which is also confirmed by simulations [59]. The required density uniformity is thus in the order of 0.25%. This unprecedentedly small number limits the choice of plasma
source to a single option – instant ionization of a highly uniform rubidium vapor by a co-propagating laser pulse [61, 62]. The choice of rubidium is determined by the low ionization potential and heavy atomic mass which was shown necessary to avoid deleterious effects associated with background plasma ion motion [63, 64], which could suppress transverse and longitudinal wakefields leading to early saturation of the self-modulation instability and stop the acceleration of witness electron bunches [63, 64]. In the plasma source, the rubidium vapor is kept in thermodynamic equilibrium with a constant-temperature closed volume.

7. Phase velocity issues

Since the drive beam shape changes in the plasma, the phase velocity of the excited wakefield is not equal to the proton beam velocity. The difference is especially large for the first 4 meters. As the SMI grows, the effective wakefield phase velocity is slower than that of the drive bunch [49, 65, 66] as seen in Fig. 7. The slow wave is problematic for accelerated particles for the same reason as for the plasma non-uniformity: the defocusing phase of the wave can scatter particles while wave crests travel back along the beam. Defocusing of protons does not have such a detrimental effect, as this is how the instability develops. To avoid the phase velocity problem, it was proposed to inject electrons into the plasma wave at the stage of fully developed self-modulation [65]. Tapering the plasma density was also discussed in this context [60, 66].

The wakefield phase velocity approaches the speed of light at \( z \approx 4 \text{ m} \) (Fig. 7). This would be the optimal place to inject electrons if we intended to use only the speed-of-light stage of the wakefield. A vacuum gap in the plasma which could allow electron delivery directly to the axis at this location is difficult to realize without producing a nonuniform region of the plasma. Because of this, the only option seen in the early design of the experiment was side injection at some small angle with respect to the drive beam axis. A fraction of the electrons reaches the beam axis, dephases, accumulates at the peak accelerating wakefield and forms short bunches in several consecutive accelerating buckets. The electrons are then accelerated to high energies with a narrow energy spread (in the order of several percent). The side injection scheme relaxes the timing tolerances for injection and has a particle trapping efficiency up to 50%. The optimum injection energy found from simulation is 16 MeV. This is exactly the energy for which the electron velocity equals the wakefield phase velocity at the end of the self-modulation stage. The minimum injection angle \( \alpha_{in} \) depends on the relativistic factor of electrons \( \gamma_e \): \( \alpha_{in} \approx 0.5 \gamma_e^{-2} \) [67]. Simulations indicate that the angle providing the highest accelerated charge is an order of magnitude higher (~5 mrad).

The design of the electron beam injection to the plasma evolved as attendant effects came to be better understood. To preserve the density uniformity, it is advantageous to let electrons in by the same entrance valve, but with an additional hole [Fig. 8(a)]. Also, this scheme allows freedom in adjusting the place and angle of injection. However, the low energy electron beam may be disrupted by the electric fields induced by the proton bunch which are especially strong near the entrance valve. Screening of the electron beam thus becomes necessary [Fig. 8(b)]. Another deleterious effect is electron scattering on the rubidium gas, which roughly doubles the electron beam radius at the focus point. The solution free from all of the above problems was to transport the electron beam through a narrow vacuum tube separated from the gas volume by a thin foil [Fig. 8(c)]. This solution was considered as the baseline variant until the discovery of better injection methods.
energy depletion length for this beam is about 10.

...are those of the plasma wave:

ω

simulation codes. Parameters of the experiment fall far beyond

in the real geometry turned out to be a dif-

simulations. However, self-modulation of a long proton beam

plasma wakefield acceleration has been guided by computer

8. Supporting simulations

From the very beginning, the development of proton driven

plasma wakefield acceleration has been guided by computer

simulations. However, self-modulation of a long proton beam

in the real geometry turned out to be a difficult task for sim-

ulation codes. Parameters of the experiment fall far beyond

the area for which most codes were originally developed and

tuned. The smallest scales that must be resolved in simulations

are those of the plasma wave: ω

−1 for time and c/ω

p for length.

If compared with the plasma wavelength, beams and interaction
distances are very long. In AWAKE, proton bunches of length

up to 3000 c/ω

p must propagate 50000 c/ω

p in the plasma. The energy depletion length for this beam is about 10⁶ c/ω

p. In [27], electrons propagate 1.2 × 10⁸ c/ω

p to gain 6 TeV. For comparison, the electron beam used in the SLAC experiments [21] was shorter than 10c/ω

p and propagated up to 85000 c/ω

p.

Because of the complexity of the problem, several well

benchmark codes were used in AWAKE related studies: ki-

netic LCODE [65, 70], fluid LCODE [23], OSIRIS [34], Quick-

PIC [71], and VLPL [72, 73]. Different plasma models im-

plemented in these codes made it possible to choose the op-

timum simulation tool for each task. The interplay of the

SMI and non-axisymmetric (hosing) perturbations was stud-

ied with three-dimensional particle-in-cell codes VLPL and

OSIRIS. Axisymmetric beam perturbations in the axisymmet-

ric plasma wave were mainly simulated with two-dimensional

LCODE and OSIRIS and cross-checked with QuickPIC. Pa-

rameter scans were made with LCODE, as it is quasi-static and

fast. The simulations of beam–plasma interactions presented in

this paper were produced with kinetic LCODE unless stated

otherwise.

To get confidence in simulation results, two special tests were

formulated that bear on two key physical effects of interest.

Test 1 is the long term evolution of a small amplitude plasma

wave. In this test, the proton beam density is

\[ n_b = 0.5 n_{b0} e^{-r^2/(2σ_b^2)} \left[ 1 + \cos \left( \sqrt{2} \frac{ξ}{σ_p}; |ξ| < σ_p \right) \right] \]

(2)

and zero otherwise. Here ξ = z − ct, σ

p = c/ω

p and n

b0 = 0.1 n

b0; the proton beam is assumed to be unchangeable, and plasma ions are immobile. We follow the excited wakefield up to the distance 3000 c/ω

p behind the driver and give attention to the average wave period and conservation of the wave-field amplitude. The wave-field amplitude must stay constant at approximately 0.0725E

b0. The average wave period must be close to 1.0005 λ

p; it must exceed λ

p because of nonlinear ef-

fects [74]. The driver(2) excites a wave of approximately the

same amplitude as the hard cut edge of the proton beam does. The distance 3000 c/ω

p is 5σ

p, for the baseline plasma density. The test shows how well the initial seed perturbation is repro-

duced by the code. Figure 9(a–c) illustrates that the kinetic

LCODE passes the test when set to high resolution and fails at lower resolutions. Figure 9(d,e) shows agreement between fluid

LCODE and QuickPIC in reproducing the spatial profile of the

wave at a large distance behind the driver.

Test 2 concerns the growth of the seeded self-modulation in-

stability. The beam parameters are those from the first data col-

umn of Table 1. At the entrance to the plasma (at z = 0) the beam density is

\[ n_b = 0.5 n_{b0} e^{-r^2/(2σ_b^2)} \left[ 1 + \cos \left( \sqrt{2} \frac{ξ}{σ_p}; |ξ| < σ_p \right) \right] \]

(2)

and zero otherwise. Here

\[ n_{b0} = \frac{N_b}{(2π)^{1/2}σ_p σ_b} \]

(3)

the beam energy spread is δW

b = 135 MeV, the transverse emit-
tance is e

b = 8 μm mrad, and plasma ions are immobile. We look at the maximum wavefield amplitude excited at various z

(Fig. 10), irrespective of the position of the maximum in rela-
tion to the beam head. There is no analytical prediction for the

wavefield amplitude, so the result is characterized by agreement

with high resolution runs and between the codes. Although the
dependencies produced with different codes do not coincide ex-
actly, the agreement in Fig. 10 is considered to be good, as the

process is very sensitive to simulation accuracy because of the

exponential growth of perturbations. Compliance with this test

is necessary for reliable simulations of the seeded instability.

9. High charge driver

The proton bunch population N
b = 1.15 × 10¹¹ discussed in early studies is typical for multi-bunch operation of the SPS
but can be increased. In the single-bunch operation regime, bunches with up to $N_b = 3 \times 10^{11}$ protons can be stably produced at the same bunch length, and this value became the baseline choice in 2013. The denser driver not only produces a stronger wakefield (Fig. 11), but also brings the beam–plasma interaction into a qualitatively new regime. Several effects have appeared at these higher densities in simulations: limitation on the wakefield amplitude caused by nonlinear wavelength elongation [74, 75], motion of rubidium ions [63, 64], breaking of the plasma wave, and positive plasma charging after the wave breaks [76]. This new regime of beam-plasma interactions turned out to be more difficult for computer simulations, as the breaking wakefield is always accompanied by numerical noise in available codes. Because of this, the wakefield in most theoretical papers is characterized by a scaled wakefield potential

$$\Phi(r,\xi, z) = \omega_p \int_{-\infty}^{\xi/c} E_z(r, z, \tau) d\tau,$$  

which behaves more smoothly than the electric field (Fig. 11).

Both the potential (5) and the electric field $E_z$ oscillate with the plasma frequency. Envelopes of on-axis potential oscillations (at $r = 0$) are shown in Fig. 12 for two beam populations. The horizontal direction in Fig. 12 is the distance along the beam, the vertical direction is the distance along the plasma, and color is the wakefield amplitude. Figure 12 gives an idea of how the wakefield evolves in space and time and also shows the qualitative difference between the two baseline cases. The higher amplitude wave in Fig. 12(b) quickly decays after reaching the maximum along the beam, while the lower amplitude wave [Fig. 12(a)] persists long after the beam passage. In both cases the wakefield decays at the plasma end because of bunch train destruction [48, 49]. Figure 12(c) in comparison with Fig. 12(b) shows the effect of ion motion. As the initially uniform ion background is perturbed by the wave (at $|\xi| > 25\text{cm}$), the wave almost fully disappears, as Ref. [64] predicts. However, this is not the only reason for the high amplitude wave to disappear: the wave amplitude can quickly go down even with immobile ions, because of wave nonlinearity [74] (at $z \sim 3\text{m}$ in Fig. 12(c)).

Figure 9: Results of the test 1 for kinetic LCODE with the square grid of size (a) 0.01$c/\omega_p$, (b) 0.025$c/\omega_p$, and (c) 0.05$c/\omega_p$; maps of the electric field $E_z$ far behind the driver produced by (d) fluid LCODE (Lf) and QuickPIC.

Figure 10: Results of the test 2 for several codes in (a) normal and (b) semi-logarithmic scales.

Figure 11: Calculated maximum amplitudes of the accelerating field $E_{z,max}$ and of the wakefield potential $\Phi_{max}$ excited along the bunch plotted as functions of position along the plasma for proton bunch populations $N_b = 1.15 \times 10^{11}$ (lower curves) and $N_b = 3 \times 10^{11}$ (upper curves). The curves overlap for the low population.
10. From conceptual design to technical design

The AWAKE experiment came closer to reality as a more suitable site for it was found in the CNGS (CERN Neutrinos to Gran Sasso) beam line. The CNGS deep-underground area [77] is designed for running an experiment with high proton beam energy, just like AWAKE, without any significant radiation issue. The facility has a 750 m long proton beam line designed for a fast extracted beam at 400 GeV. Installing the AWAKE experiment upstream of the CNGS target (Fig. 13) was determined to be possible with only minor modifications to the end of the proton beam line; these include changes to the final focusing system and the integration of the laser and electron beam with the proton beam. At energies above 75 GeV, the maximum field generated in the plasma weakly depends on the driver energy [75], and the length of the high field region is roughly proportional to the square root of the driver energy. Therefore, reduction of the proton energy from 450 GeV to 400 GeV is of no significance.

The efforts to develop the AWAKE project are summarized in the Design Report (DR) [78] and its numerous supplements. The background physics is also presented in papers [61, 79, 80] and conference proceedings [69, 81–85]. Refs. [86–89] are earlier status papers describing evolution of the project in general. On the basis of Design Report, the AWAKE experiment was approved in August 2013 and now is under construction. The main parameters of the experiment, as they appear in the Design Report, are given in the second data column of Table 1. The status of the AWAKE experiment has been presented at this conference [1].

11. On-axis injection of electrons

Side injection of electrons is technically challenging [90] and has a serious disadvantage in that the parameter window in which both trapping and acceleration are good is rather narrow. Figure 14 shows how the final energy spectrum of electrons changes if injection parameters deviate from the optimal values. For the Design Report, the following parameter values were used: injection angle for electron beam $\alpha_0 = 9$ mrad, electron beam energy $W_e = 16$ MeV, injection delay with respect to the ionizing laser pulse $\xi_e = 13.6$ cm, electron beam trajectory intersects the axis at $z_0 = 3.9$ m. These parameters were obtained from computer simulations. It is expected that the actual optimum injection parameters will differ from these so that experimental flexibility is required, which is difficult to achieve in the side-injection scheme. Therefore, the possibility of using on-axis injection in search of a good operation regime for first experiments was investigated.

The term “on-axis injection” refers to propagation of both electron and proton beams along the same line starting from the entrance to the plasma. This injection method was initially
considered ineffective because of wakefield phase velocity issues, and early computer simulations confirmed this assessment \cite{54,55,67,81}. However, a parameter window for high trapping rate and efficient acceleration was found \cite{91,92}. The better performance is possible because of a supraluminal wake wave that appears at the stage of developed self-modulation at a certain delay behind the ionizing laser pulse. If the velocity of the injected electrons is close to or greater than the phase velocity of the wave at the driver self-modulation stage, then the electrons are trapped by the wakefield and kept in the potential wells until the driver beam is fully bunched. After that, the electrons are continuously accelerated with the rate that depends on the distance between the electron bunch and the seed laser pulse.

The parameter window for good on-axis injection is also narrow, but unlike side injection it depends on a single parameter – injection delay $\xi$ – that is easily controlled by timing. Although the simulated final electron energy spectrum for on-axis injection is not as narrow as for the best side injection variants (Fig. 15), there are enough electrons to characterize the accelerating ability of the wakefield. Simulations \cite{22} also indicate that electrons injected at radii up to 0.4 mm are still trapped by the wave, so the requirements on electron beam focusing are somewhat relaxed for the first stages of the experiment. This fact is of particular value, as the electron beam size in some regimes may be blown up due to interaction with the proton beam in the common beamline upstream of the plasma cell \cite{93}.

In the first stages of the experiment, it has been decided to inject long bunches of electrons so that the exact phasing with the proton bunch modulation is not an issue. The electron bunches will be in the order of 10 ps long and will thus cover several modulation cycles. Once the SMI is better understood and optimal parameters are found, it is planned to inject short electron bunches at the desired phase. The details on the electron source and injection system are given in \cite{94,95}.

The on-axis scenario became the baseline injection option for AWAKE in April 2014, and the facility subsystems were designed for this scheme. Elements of this work are documented in conference proceedings \cite{93,100}.

\section{Density transitions at the plasma cell ends}

The entrance region in which the plasma density gradually increases from zero to the nominal value is potentially dangerous for the axially propagating electron beam. The effect is similar to the plasma lens effect \cite{101}. There is a radial force from the magnetic field of the proton beam which is partially neutralized by the plasma. This force always focuses protons and defocuses the electrons. The electric force can overcome the defocusing, but only in a slowly-varying density plasma. Otherwise, the radial electric force acting on an electron oscillates as the electron moves to regions of different plasma density, and its average becomes negligibly weak. For the parameters of the AWAKE experiment, a transition region of length about 10 cm is sufficient to defocus the electrons \cite{22}.

In its early versions, the AWAKE baseline design had fast valves located at the ends of the vapor column. With the valves closed, the temperature uniformity ensured the vapor density uniformity along the gas cell. Fluid simulations of the rubidium vapor flow showed that the opening time (10 ms) of state-of-the-art fast valves (developed for AWAKE with VAT, Switzerland) cannot ensure a short enough density ramp. With this opening time, the density ramp is about 1 m long (Fig. 16). To shorten the transition area, the solution with a continuously leaking flow through orifices at each end of the vapor cell
was proposed (Fig. 17). The rubidium sources should be placed as close as possible to the orifices to minimize the density ramp length. Thus there is a continuous flow of rubidium from the sources to the plasma cell and afterwards from the plasma cell to the expansion volumes through the orifices. The walls of the expansion volumes should be cold enough to condense all rubidium atoms. The residual pressure in the expansion volumes decreases with the temperature decrease. From the practical point of view, it is desirable to keep the walls below 39°C, the melting temperature of rubidium.

Simulations with an in-house DSMC code and COMSOL Multiphysics (a Finite Element Method based software) confirmed that the density transition area in the case of continuous flow is as short as several centimeters (Fig. 18). The on-axis gas density in the expansion volume near the orifice is $n_0 = n_{\text{avg}} \left(1 - \frac{\delta z / D}{\sqrt{\left(\delta z / D\right)^2 + 0.25}}\right)$, where $D = 10 \text{ mm}$ is the orifice diameter, and $\delta z$ is the distance to the orifice. The plasma density follows the same density profile.

13. Oblique injection

The continuous gas flow through the orifices reduces the length scale of the density transition area, but does not completely dispose of the problem of electron defocusing. The low density gas in the expansion volume is ionized by the laser pulse in the same manner as in the plasma cell. The proton beam excites the seed wakefield in this low-density plasma, and this wakefield is defocusing for electrons (Fig. 19). The amplitude of the defocusing force is approximately constant in a wide (almost four orders of magnitude) interval of plasma densities from $2 \times 10^{11} \text{ cm}^{-3}$ to the nominal value of $7 \times 10^{14} \text{ cm}^{-3}$. According to formula (6), the defocusing region is 15 cm long. This distance is sufficient to deliver the radial momentum of about 0.5 MeV/c to electrons thus preventing their trapping by the plasma wave.

Fortunately, the defocusing region does not extend beyond the radial plasma boundary (Fig. 19). The electrons that pass the upstream expansion volume outside the ionized area propagate almost freely and some of them can even receive a small focusing push of several mrad. The loss of electrons at the density transition region thus can be avoided with the oblique injection (Fig. 20). In this scenario, the electrons approach the axis in the region of constant plasma density and therefore get trapped into the established plasma wave. The required injection angle $\alpha_i$ and radial offset of the electron beam at the orifice are small enough, so the oblique injection does not require any changes in the facility design, as compared to the on-axis injection. The optimum values found in simulations are: electron delay $\xi_e = 11.5 \text{ cm}$, injection angle $\alpha_i = 2.8$ mrad, and focusing point $z_f = 140 \text{ cm}$.

The region of good trapping is quite large in the space of injection parameters, as compared to the electron beam portrait (Fig. 21), so no sharp tuning of the injection angle or focus point is required for the best performance. Figure 21 was obtained for a realistic density profile with the transition area described by formula (6) and for $\xi_e = 11.5 \text{ cm}$. It also shows that electrons that propagate along the axis of the proton beam have no chance to be trapped.
14. Ramped density

The new plasma cell design offers the opportunity of creating plasma density profiles with a constant density gradient along the plasma cell. This gradient naturally appears if the continuous flow of rubidium vapor through the entrance and exit orifices is unbalanced. The gradient could have a value of several percent over 10 m. Unlike shorter scale perturbations [59], the gradient has no direct detrimental effect on the accelerated electrons, and the limitation [1] is not applicable to non-uniformities of this long scale. The reason is that the change of the plasma wavelength is so slow that the proton beam itself has enough time to respond to this change. The resulting change of the wakefield structure is favorable for electron acceleration if the gradient is positive (Fig. 22), as it controls the phase of the plasma wave [104]. This effect is illustrated by Fig. 22 with the case of strong gradients. The amplitude of the electric field is approximately the same in both cases. However, electrons gain energy only in case of positive density gradient. A negative density gradient results in a gradual increase of the plasma wavelength and continuous drift of the wakefield phase towards the tail of the bunch. A positive density gradient makes the phase velocity of the plasma wake equal to or slightly faster than the speed of light (at some specific delay $\xi_e$ behind the laser pulse). This makes it possible for the electrons to stay in phase with the wakefield and gain energy continuously until the end of the plasma section.

The best electron energy spectrum found in simulations so far is shown in Fig. 24 in comparison with earlier baseline results. Higher electron energies (Fig. 22) are also possible at larger $\xi_e$, but at the expense of reduced trapping efficiency. These most recent simulations include the oblique injection, realistic plasma boundaries [6] at both ends, and the linear growth of the plasma density by 1% over 10 m. About 40% of injected electrons are trapped and accelerated. This current baseline is listed in the last column of Table 1.

15. Longer term perspectives

The optimal proton-driven plasma wakefield accelerator would use single short proton bunches to drive the plasma wave. An accelerator capable of producing high-energy bunches of protons with about 100 $\mu$m length does not exist yet and the technology for realizing such an accelerator is not currently in hand. We therefore plan to push the modulation approach as far as possible to understand what is feasible.

As discussed in this report, it is advantageous to separate the modulation of the bunch from the acceleration stage. A later phase of the AWAKE experiment would therefore likely have separate plasma cells – the first allowing for seeding and development of the SMI (likely still the rubidium cell) and the second...
for acceleration of externally injected electrons [105, 106]. The latter may need to be 10–100 m long or more, and can be based on other ionization techniques than laser ionization since it does not require seeding of the SMI. Three other concepts have been therefore explored that would allow scaling plasma sources to hundreds of meters.

The discharge cell is based on a pulsed argon plasma produced in a glass tube by a microsecond pulsed electric current [105]. With this source, stable creation of 6 meter long, almost 100% ionized plasmas of the density up to $10^{15}$ cm$^{-3}$ has already been demonstrated, and the work continues towards measuring and reducing the density nonuniformity to the sub-percent level.

In the helicon source [107, 108], the magnetized plasma is created and heated by a right-hand circularly polarized low-frequency RF wave (the helicon wave, or the whistler) propagating along magnetic field lines in a frequency regime between the lower hybrid and the electron cyclotron frequency. Since external helical antennas are used, the heating power can be spatially distributed, allowing for arbitrary plasma lengths. Helicon sources usually operate at lower plasma densities than that required for AWAKE, so the recent demonstration of plasma densities as high as $7 \times 10^{14}$ cm$^{-3}$ [108] was an important milestone towards suitability of this approach.

The third concept relies on the so-called Resonance Enhanced Multi-Photon Ionization scheme [109]. It is a three-photon process which requires significantly lower laser power than the non-resonant barrier suppression ionization currently used for rubidium ionization.

In addition to staging the acceleration process, the use of more than one plasma cell will also allow us to implement the density step required to freeze in the high gradient acceleration required to reach very high energies.

It is important to study the properties of the electron bunch after acceleration in a plasma cell, since this will eventually limit attainable luminosities with a plasma based accelerator. Simulation studies are ongoing with the objective of establishing optimal injection parameters for applications, as well as establishing the best technology for a future electron injector [106]. For example, how to best preserve electron beam quality during the acceleration process must be studied in detail. The continued developments of theoretical and simulation efforts, together with experimental results, will pave the way for the proposal of a high energy accelerator project based on proton beams and the SMI. First ideas on an energy frontier electron-proton collider based on our scheme has been described in [110, 111].

16. Acknowledgements

The contribution of Novosibirsk team to this work is supported by The Russian Science Foundation, grant No. 14-12-00043. LCODE computer simulations are mostly made at
Siberian Supercomputer Center SB RAS. QuickPIC simulations are performed on the Institutional Computing resources at Los Alamos National Laboratory. This work was supported in parts by: EU FP7 EuCARD-2, Grant Agreement 312453 (WP13, ANAC2); and EPSRC and STFC, United Kingdom. M. Wing acknowledges the support of DESY, Hamburg and the Alexander von Humboldt Stiftung.

References

[1] E. Gschwendtner et al., this issue.
[94] M. Bernardini et al., Proceedings of IPAC2015 (Richmond, USA), p.34.
[104] A.Petrenko et al., this issue.