FUTURE ACCELERATOR CHALLENGES IN SUPPORT OF HIGH-ENERGY PHYSICS*

M. S. Zisman†, LBNL, Berkeley, CA 94720 U.S.A.

Abstract
Historically, progress in high-energy physics has largely been determined by development of more capable particle accelerators. This trend continues today with the imminent commissioning of the Large Hadron Collider at CERN, and the worldwide development effort toward the International Linear Collider. Looking ahead, there are two scientific areas ripe for further exploration—the energy frontier and the precision frontier. To explore the energy frontier, two approaches toward multi-TeV beams are being studied, an electron-positron linear collider based on a novel two-beam powering system (CLIC), and a Muon Collider. Work on the precision frontier involves accelerator designs with very high intensity, including a Super-B Factory and a muon-based Neutrino Factory. Without question, one of the most promising approaches is the development of muon-beam accelerators. Such machines have very high scientific potential, and would substantially advance the state-of-the-art in accelerator design. The challenges of the new generation of accelerators, and how these can be accommodated in the accelerator design, are described. To reap their scientific benefits, all of these frontier accelerators will require sophisticated instrumentation to characterize the beam and control it with unprecedented precision.

INTRODUCTION
Scientific progress in high-energy physics has traditionally depended on advances in accelerator design. This trend has taken the community from the early electrostatic accelerators to cyclotrons, to synchrocyclotrons, to synchrotrons, and finally to both circular and linear colliders. Achieving the full performance potential from each generation of accelerators requires corresponding advances in accelerator technology, including magnets, vacuum systems, RF systems, and instrumentation.

The critical scientific role played by accelerators is due to the fact that they permit the study of high-energy physics phenomena under (more or less) controlled conditions. However, the cost of today’s proposed accelerator projects is high, and this has consequences for the community. Indeed, there is a danger at present of “pricing ourselves out of the market” if we are not conscious of costs in each aspect of the design. The practical way of mitigating project costs is to share the burden and, with this in mind, international cooperation and collaboration are key to our ability to successfully launch new projects.

ACCELERATOR DELIVERABLES
Particle accelerators are designed to deliver two parameters to the HEP user—energy and luminosity. Of these, energy is by far the easier parameter to deliver, and the easier one to accommodate for the experimenters. Increased luminosity invariably presents a major challenge, not only to the accelerator builders but to the detector builders. Luminosity, $L$, is a measure of the collision rate per unit area and has dimensions of cm$^{-2}$ s$^{-1}$. For a given event probability (“cross section”), $\sigma$, the event rate at a collider is given by $R = L\sigma$. For a collider having equal beam sizes at the collision point, luminosity is given by

$$L = \frac{N_1 N_2 f_{\text{c}}}{4\pi \sigma_x \sigma_y}$$

where $N_1$ and $N_2$ are the number of particles per bunch in beams 1 and 2, respectively, $f_{\text{c}}$ is the collision frequency, and $\sigma_x^*$ and $\sigma_y^*$ are the horizontal and vertical rms beam sizes at the collision point, respectively. It is obvious from Eq. (1) that high luminosity demands intense beams and small beam sizes at the collision point.

PARTICLE PHYSICS QUESTIONS
In simple terms, there are two main thrusts of accelerator-based high-energy physics. The first of these, corresponding to experiments at the energy frontier, is to understand the origins of mass, that is, the mechanism that gives existing particles their widely different masses. Recent experiments [1] have shown the mass of the top quark to be comparable to that of a gold nucleus, whereas a neutrino mass (thought for a long time to be exactly zero) is likely to be only a fraction of an eV.

The second main thrust is understanding why we live in a matter-dominated universe. This is a quite fundamental question, as it basically addresses our very existence. It is believed that the Big Bang initially created equal amounts of matter and antimatter, yet these did not all annihilate. The survival of matter is related to differences in the reaction rates of particles and antiparticles, referred to as charge-conjugation–parity (CP) violation. CP violation has long been known in the quark sector [2], and two “B Factories” [3, 4] were built to study this phenomenon. Unfortunately, it turns out that CP violation in the quark sector is not enough to explain the observed baryon asymmetry in the universe. The prevalent view is that the additional CP violation needed occurs in the lepton sector. While this has never been observed, neutrinos are considered the hunting ground for finding it.

---

*This work was supported by the Office of Science, U. S. Department of Energy, under Contract No. DE-AC02-05CH11231.
†mszisman@lbl.gov
TODAY’S MACHINES

High-energy physics experiments typically make use of colliders, where counter-propagating beams collide at one or more interaction points (“IPs”). Until recently, colliders were single-ring devices that collided beams of particles and antiparticles, such as e− and e+. However, to get higher intensities and more bunches, modern colliders use two independent rings, and so no longer require the colliding beams to have equal momenta or opposite sign. Colliders are categorized based on the type of particles they utilize. Hadron colliders, exemplified by the Tevatron (p–p̄), LHC (p–p), and RHIC (heavy ions, for nuclear physics), use protons or heavier ions. Lepton colliders, exemplified by CESR-c, PEP-II, and KEKB, all use electrons and positrons.

Hadron Collider Limitations

The main limitation associated with hadron colliders is that they are colliding composite particles. This means that only about 10% of the collision energy is actually available for the hard collisions that lead to new particles. To probe the 1 TeV mass scale, a hadron collider requires a collision energy of order 10 TeV. Such high beam energy has the disadvantage of requiring very strong magnets to store and focus the beams in a reasonable-sized ring. When using antiprotons, there is another disadvantage—these particles are difficult to produce, and it takes many hours to replace them if the beam is unexpectedly lost. The demand for high luminosity led to the LHC choosing a configuration with two separate rings, each containing many proton bunches, that are brought together at a few select locations. This configuration bypasses the antiproton production limitation, but not the other issues.

Lepton Collider Limitations

For these machines, synchrotron radiation emission is the biggest issue. In a circular machine, the emitted power is given by

\[ P_{SR} = \frac{8.85 \times 10^4 I E^4}{\rho} \]  (2)

where \( E \) is the beam energy in GeV, \( I \) is the beam current in A, and \( \rho \) is the bending radius in the ring dipoles, in m. An electron-positron collider in the LHC tunnel with 1 mA of circulating current would radiate roughly 2 GW of synchrotron radiation power into the vacuum chamber walls, and would require this amount of power to be restored by its RF system. This is a daunting requirement, to say the least.

For this reason, the high-energy e+e− colliders on the drawing boards now (ILC [5] and CLIC [6]) are conceived as linear colliders. At a length of 31 km for ILC and 48 km for CLIC, these designs have the drawback of being too large to fit on any existing laboratory site. Furthermore, the single-pass acceleration scheme is relatively inefficient, because there is no reuse of expensive RF hardware. Despite such drawbacks, a linear collider is the only practical approach to an e+e− collider in the TeV energy range.

Luminosity Performance

Electron-positron colliders have made great strides in delivering luminosity in recent years. In particular, as shown in Fig. 1, both KEKB and PEP-II quickly reached luminosities beyond 1 × 10^{34} cm⁻² s⁻¹. It is likely that the performance of future machines will be judged according to this new standard.

FUTURE MACHINES

At present, there are several machines on the drawing board to address the high-priority questions of high-energy physics. For the precision frontier, there are three candidates: ILC, a muon-beam-based Neutrino Factory, and an e+e− Super-B Factory. Candidate energy frontier machines include CLIC and a Muon Collider. Not all of the machines are at the same stage of development; ILC and CLIC are probably furthest along in terms of R&D progress.

In what follows, we will briefly discuss these machine designs and the R&D programs that support them. For reasons of personal taste and familiarity, the emphasis here will be on the muon-based accelerators. Although these are not the most advanced designs, they are undoubtedly the most novel.

Muon Accelerator Advantages

In contrast with other technologies, muon beam accelerators have the ability to address both of the physics questions discussed earlier.

For the neutrino experiments, the stored muon beam produces neutrinos as indicated in Eq. (3).

\[ \mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu \]  (3a)

\[ \mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu \]  (3b)

The decay kinematics is well understood, and there are minimal hadronic uncertainties in either the spectrum or the flux. Moreover, \( \nu_e \rightarrow \nu_\mu \) oscillations result in “wrong sign” muons, which are easily observed in the detector.

Figure 1: Peak luminosity trends. Both KEKB and PEP-II reached high luminosity much more quickly than any previous generation of collider.
For energy frontier experiments at a Muon Collider, the fact that muons are point particles makes the full center-of-mass energy available to produce new particles. Because the muon is much heavier than the electron, it produces negligible synchrotron radiation. Compared with an electron-positron collider, a Muon Collider has a reduced energy spread due to “beamstrahlung” and can be designed as a circular facility that fits on existing laboratory sites (see, e.g., Fig. 2).

**Muon Beam Challenges**

There are two major challenges associated with a muon beam accelerator. Firstly, muons are created as a tertiary beam, via the process

\[ p + \text{target} \rightarrow \pi \rightarrow \mu. \]

This process results in a relatively low production rate, so achieving the desired muon intensity requires a very intense proton beam, nominally 4 MW. Finding a target material that can tolerate such a beam intensity is not easy. Moreover, the decay process produces a muon beam with a very large transverse phase space and large energy spread. Handling this beam is difficult with standard quadrupole focusing devices, so strong, large-aperture solenoids are used to focus in both planes simultaneously. A rapid means to decrease the large transverse emittance is required and, even so, a large acceptance acceleration system and decay ring are needed.

Secondly, muons have a very short lifetime, only 2.2 µs at rest. Obviously, the short lifetime puts a high premium on rapid beam manipulations. Reducing the beam emittance requires the presently untested technique of “ionization cooling.” This technique, described below, makes use of high-gradient RF cavities immersed in a strong solenoidal field—a configuration that has turned out to make it difficult to achieve high gradients.

Thirdly, the decay electrons produced give difficulties. In a Neutrino Factory, they create a substantial heat load in the midplane of the decay ring dipoles, and in a Muon Collider they give rise to substantial background in the detector.

For all these reasons, muon beams are difficult to produce and manipulate, and will make the accelerator designers work hard to obtain their obvious benefits.

**Ionization Cooling**

Because ionization cooling plays a critical role in muon beam facilities, we describe the concept briefly here. Ionization cooling is analogous to the more familiar process of synchrotron radiation damping in an electron storage ring. In each case, there is an energy loss mechanism that reduces \( p_x, p_y, \) and \( p_z \), along with an energy gain mechanism that restores only \( p_z \). Repeated application of the process reduces \( p_x/p_z \) and \( p_y/p_z \), and thus decreases the 4D emittance. In the electron case, the energy loss mechanism is synchrotron radiation emission, whereas for muons it is energy loss \((dE/dx)\) in low-Z material. The energy gain mechanism, in both cases, is RF acceleration in the longitudinal (z) direction. Figure 3 illustrates schematically the ionization cooling concept.

For both electrons and muons, there is also a heating term. For electrons it is quantum excitation, whereas for muons it is multiple scattering in the absorber material. The balance between cooling and heating leads to an equilibrium emittance, given approximately by:

\[
\varepsilon_{x,y,\text{equil}} = \frac{\beta \{0.014 \text{GeV}\}^2}{2 \beta m_\mu X_0 \frac{dE_\mu}{dx}}.
\]

where \( \beta_i \) is the beta function at the absorber location, \( \beta \) is the relativistic velocity, \( m_\mu \) is the muon mass, \( X_0 \) is the radiation length and \( dE_\mu/dx \) the rate of muon ionization energy loss in the absorber material. Looking at Eq. (4), we see that a low equilibrium emittance results from a low \( \beta_i \), large radiation length and large energy loss. This implies that hydrogen, though far from convenient, is the best absorber material.

**ACCELERATOR DESCRIPTIONS**

**ILC**

The ILC [5] is aimed at the 0.5–1 TeV energy range, with an initial energy of 0.5 TeV. As shown in Fig. 4, it comprises two superconducting main linacs, electron and positron sources, and a central complex with two damping rings. The damping rings must provide beams with a 2 pm-rad vertical emittance.

![Figure 2: Schematic layout of 1.5–4 TeV Muon Collider on the Fermilab site.](image)

![Figure 3: Schematic diagram of an ionization cooling channel. The best absorber material to use is liquid hydrogen. Its proximity to the RF cavities is an engineering challenge, as this represents a potential safety hazard if not implemented properly.](image)
The design effort has been ongoing for a number of years, so the facility design is reasonably well defined. The main technical challenges for the design, which will be covered below, include reaching the specified gradient for the production cryomodules of 31.5 MV/m and obtaining the very low vertical emittance required by the linac.

**Neutrino Factory**

A schematic diagram of a Neutrino Factory [7, 8, 9] is shown in Fig. 5. It comprises many different systems. First, there is a proton driver capable of providing 4 MW of proton beam on a production target. The target—a free Hg jet immersed in a 20-T solenoid—is dispersed by the beam but “remade” by the time the next proton pulse arrives. A tapered solenoid channel captures the pions and the muons into which they decay.

The capture section is followed by a bunching and phase rotation section, where the muons are bunched into 201-MHz RF buckets and phase rotated in longitudinal phase space to transform the initial bunch having a small time duration and a large energy spread into a bunch train with a long time duration and smaller energy spread. The phase-rotated bunch train, with a mean momentum of about 200 MeV/c, is then cooled transversely in a cooling channel.

After cooling, the beam is accelerated to its working energy of a few tens of GeV using a linac, a pair of dog-bone recirculating linear accelerators (RLAs), and finally one or more fixed-field, alternating gradient (FFAG) accelerators. The last element is a decay ring with a long straight section aimed at a detector located 3000–7500 km from the ring.

**Super-B Factory**

The goal of this e⁺e⁻ collider facility is to greatly improve on the data sample at the $\Upsilon(4S)$ resonance accumulated by the original B Factories [3, 4]. To do so, it aims at a peak luminosity of $\sim 1 \times 10^{36}$ cm⁻² s⁻¹. The proposed design [10] involves two rings, with a single interaction region. The interaction region is designed to use a new approach, the “crab waist,” to substantially reduce the effective beam size at the interaction point. The beams collide at a large crossing angle (see Fig. 6), but the overlap region is short in the $z$ direction, permitting a low beta function without requiring a correspondingly short bunch length. A pair of sextupoles in each ring is used to suppress synchro-betatron resonances. The likely site for the facility is Frascati, and their plan is to reuse most of the PEP-II components to reduce costs.

**CLIC**

CLIC [6] is an e⁺e⁻ linear collider designed to reach 3 TeV collision energy. It has a novel two-beam powering concept, as indicated in Fig. 7. A low-energy drive beam is created and then sent through passive structures that extract microwave energy at high frequency to power a second beam, with lower intensity but much higher energy. Using room-temperature accelerating structures for the main linac, the design goal is to operate at a gradient of 100 MV/m, some three times higher than that used in the ILC. The footprint of this machine, 48 km, is somewhat larger than that of ILC, but it is designed to operate at six times higher energy than ILC.
Figure 7: Schematic diagram of CLIC complex, showing drive beam generation system and main linacs.

Muon Collider

A scheme for a Muon Collider that would fit on the Fermilab site is shown in Fig. 8. The concept is based on using the existing Recycler and Main Ring at Fermilab, fed with protons from the Project X linac [11]. The early portion of the facility is essentially the same as that described above for a Neutrino Factory. However, for a Muon Collider, the 4D cooling process described earlier is not sufficient to prepare the desired beam. The beam must be cooled in all six dimensions, that is, longitudinal as well as transverse. The concept of 6D cooling is shown in Fig. 9.

R&D ACTIVITIES

All of the potential projects described above have embarked on R&D activities to prove the physics concepts, validate technology choices, and develop realistic, defensible cost estimates. The choices for what must be studied can be complicated, as there are many “audiences” for the results and they do not all have the same concerns. These audiences include:

- project advocates
- the HEP community
- laboratory directors (at least one)
- funding agencies or governments (at least one)

Although it is not possible here to do justice to the large volume of R&D work in progress throughout the world, an attempt will be made in what follows to give a flavor of the R&D that is under way.

ILC

An noted earlier, a primary thrust of this program is to demonstrate that the cryomodule design gradient of 31.5 MV/m can be reached reliably in a production setting. Figure 10 shows results from a number of recent cryomodule tests. The trend is clearly headed in the right direction, but there remains work to be done. Because a large number of cryomodules must be built, the ability to demonstrate reproducibility is mandatory, and this requires a substantial, and thus expensive, production run.

Another technical concern for the ILC is the effects of the electron-cloud instability on the vertical beam emittance in the positron damping ring. Such concerns initially prompted a design with two positron damping rings (each with half the current). For cost reasons, this approach was eventually abandoned in favor of a single ring with appropriate mitigating features. Both clearing electrodes and “grooved” chambers (see Fig. 11) have been proposed and are being tested. Simulations indicate that solutions can be found, and plans for a detailed test program at Cornell (denoted “CESR-TA”) are well along.

Figure 9: Schematic approaches to longitudinal cooling. On the left, a magnet creates dispersion so there is a correlation between horizontal position and momentum. The beam then passes through a wedge-shaped absorber such that low-energy particles have reduced energy loss compared with high-energy ones. On the right, the magnet is filled with a gaseous absorbing medium. Higher energy particles have a longer path length and hence experience more energy loss than do low-energy particles.

Figure 10: Recent results for ILC cryomodule gradients.
Figure 11: Grooved chamber approach to suppress e-cloud effects. Simulations indicate that the suppression is very effective if the sides of the grooves are steep. Unfortunately, this geometry is unattractive from an impedance perspective and trade-off studies must be made.

**CLIC**

The primary effort for CLIC is to demonstrate the feasibility of its two-beam technology and estimate its cost. The third generation of a CLIC test facility, CTF3, is now available at CERN, the injection portion of which is shown in Fig. 12. While centered at CERN, there are presently 19 countries represented in the CLIC R&D effort. There is technical coordination with the ILC program on areas of common interest, such as the damping ring instability referred to above.

Another key area of R&D is the study of high-gradient effects on the CLIC room-temperature RF structures. When using iris devices made from “hard” materials like Mo or W, gradients up to 190 MV/m have been reached. The limit, however, is the breakdown rate, which is specified to be below $10^{-6}$. As shown in Fig. 13, to achieve acceptable breakdown rates, the gradient must be reduced substantially, which makes the distinction between harder materials and copper start to disappear. Moreover, the apparent gain in gradient performance at high frequencies also tends to disappear if one compares the performance at the same breakdown rate. This conclusion has led to the CLIC design frequency being reduced from 30 GHz to 12 GHz, which permits taking advantage of much of the work done at SLAC in support of the NLC design [12].

**Super-B Factory**

There are two primary questions to answer for this project:

- Does the crab waist scheme work as expected?
- Can the IP beta functions be made low enough to get a hundred-fold increase in luminosity with reasonable beam currents?

Tests of the crab waist scheme are getting under way now at DAΦNE to determine this. These tests have required significant modifications of the ring design, as shown in Fig. 14. Thus far, the results are encouraging. The beneficial effects of the crab waist sextupoles can clearly be seen in both reducing the beam size and increasing the luminosity.
Muon Beam Program

There is a broad program under way worldwide to study the issues related to developing an intense muon beam facility. In large measure, the R&D program at this stage applies equally to a Neutrino Factory or a Muon Collider, so there is no need to make a distinction. In Europe, the work is supported by the BENE network and the UKNF program. In Japan, the NuFact-J group, supported by several universities and sometimes U.S.-Japan funds, has been active. In the U.S., the work is supported by the Neutrino Factory and Muon Collider Collaboration (NFMCC), augmented recently by effort from Fermilab’s Muon Collider Task Force (MCTF).

Much of the work is being done by international collaborations involving participants from all these regions. These include:

- MERIT (beam test of a mercury-jet target in a 15 T solenoidal field at CERN) [13]
- MICE (muon ionization cooling demonstration at Rutherford Appleton Laboratory (RAL)) [14]
- EMMA (non-scaling FFAG demonstration at Daresbury Laboratory) [15]
- IDS-NF (international Neutrino Factory design study) [16]

In addition, there is component R&D under way aimed at development of suitable muon beam RF cavities and LH2 absorbers.

The main component R&D at present involves the study of RF breakdown in a magnetic field. Interest in this is motivated by the NFMCC results for an 805-MHz test cavity, shown in Fig. 15. Although the cavity reached 40 MV/m in the absence of a magnetic field, the maximum gradient decreased markedly when an axial solenoidal field was applied. The effect is believed to be related to focusing the dark currents tightly enough to do damage to copper surfaces when the field is applied.

As the actual cooling channel design calls for 201-MHz cavities, a prototype has been built (see Fig. 16) at this frequency. The plan is to first test it in the fringe field of an existing solenoid at Fermilab, and then to test it with a more realistic field from a large diameter coupling coil identical to those to be used in MICE. In tests without magnetic field, the cavity easily reached 19 MV/m, limited only by the power of the RF amplifier tube.

The MERIT experiment [13] was designed to test the behavior of a mercury-jet target under realistic conditions. The experiment was carried out in October, 2007 using 24 and 14 GeV proton beams from the CERN PS. Figure 17 shows a diagram of the system and Fig. 18 shows the actual hardware being installed in the TT2A tunnel at CERN. Although the data are still being analyzed, preliminary indications are that the technology will work well at a beam power in excess of 4 MW and a proton beam repetition rate in excess of 50 Hz.
The international Muon Ionization Cooling Experiment (MICE) [14] is intended to demonstrate, for the first time, the process of ionization cooling of a muon beam. The layout of the experiment is shown in Fig. 19. The experiment uses the single-particle counting techniques of particle physics experiments to measure the incoming and outgoing emittance of each muon. Over the next several years, the experiment will test beams having a range of incoming emittance values and momenta. An important goal of this work is to validate the simulation tools used to estimate cooling channel performance. These tools will ultimately be used to design a facility and estimate its performance.

One of the challenges of MICE will be to operate liquid-hydrogen absorbers in close proximity to high-power RF cavities. With this in mind, safety has been a design criterion from the outset. There have already been two international reviews of the safety aspects of the experiment, and the RAL safety group has been kept in close contact with both the design and the R&D efforts.

There are some 140 scientists and engineers from Europe, Asia, and the U.S. collaborating on the experiment. Commissioning of the muon beam line, along with some of the upstream detectors, has recently gotten under way. The first pions were detected in the beam line in early April 2008. This beam was produced by operating a plunging Ti target in the halo of the ISIS synchrotron beam at a rate of roughly 1 Hz. Losses in ISIS must be kept to a minimum in order to keep radiation levels low enough for hands-on maintenance.

**SUMMARY**

In this paper, we have seen that HEP facilities now in the planning stage offer great potential to address the key outstanding questions in our field. There is little doubt that such new facilities will raise new questions as well. R&D toward the design of these new facilities is progressing on many fronts. While not all of these machines will be built, the work being done provides the HEP community and the funding agencies with the technical input needed to make good decisions about what should be done.

As with all accelerator R&D, success will depend on a synergy between accelerator physics and accelerator technology. In particular, control of instabilities and emittance will require state-of-the-art instrumentation. The skills of the instrumentation designers and builders will undoubtedly be crucial to turning the accelerator physicists’ dreams into the cutting edge scientific tools of the future.

From the U.S. perspective, going forward with Project X at Fermilab is a necessary step to provide for the long-term future. This accelerator provides scientific capability for carrying out improved neutrino experiments and paves the way for future muon-beam facilities. The technology being developed and used will also provide valuable technical information in support of ILC construction, should that be the eventual choice of the community.

**REFERENCES**


