

31. Accelerator Physics of Colliders

Revised July 2023 by V. Shiltsev (FNAL) and F. Zimmermann (CERN).

This article provides background for the High-Energy Collider Parameter Tables that follow and some additional information; see in-depth review and a comprehensive list of references in [1]; citations below are limited to widely used textbooks and open access seminal papers and reviews.

31.1 Energy and Luminosity

Collisions of two beams of particles accelerated to high energies $E_{1,2}$ provide access to center-of-mass energies (c.m.e.) $E_{\text{cme}} \approx 2\sqrt{E_1 E_2}$, assuming a typically small or zero crossing angle. Most of the 31 colliders that have ever reached the operational stage (seven are operational now) used equal masses and energies of colliding particles, with c.m.e. equal to twice the beam energy $E_{\text{cme}} = 2E_b$. Other machines collide beams of unequal energies, such as electron-proton or electron-ion colliders, or asymmetric B -factories, that produce new short-lived particles, whose decays are more easily detected and analyzed with a Lorentz boost.

In an accelerator, charged particles gain energy from an electric field, which usually varies in time at a high frequency ranging from 100s of kHz to 10s of GHz. With proper phasing to the RF field over distance l , the energy gain of a particle with charge Ze is proportional to the average accelerating gradient G , i.e. $\Delta E_b = ZeGl$. In principle, the highest beam accelerating gradients achieved to date in operational machines or beam test facilities ($G \approx 100$ MV/m in 12 GHz normal-conducting RF cavities and 31.5 MV/m in 1.3 GHz superconducting ones) allow accessing high energies over reasonably long linear accelerators (linacs), but cost considerations often call for minimization of RF acceleration via repeated use of the same RF system which, in that case, would boost the energy in small portions $\Delta E_b = ZeV_{\text{RF}}$ per turn every time a particle passes through the total cavity voltage V_{RF} . Such an arrangement can be realized either in the form of storage-ring circular colliders or also through novel schemes based on, e.g., recirculating linear accelerators (RLAs) with or without energy recovery. Circular colliders are by far the most common; here, the momentum and energy of ultra-relativistic particles are determined by the bending radius inside the dipole magnets, ρ , and by the average magnetic field B of these magnets:

$$p = ZeB\rho \quad \text{or} \quad E_b [\text{GeV}] = 0.3Z(B\rho) [\text{Tm}] . \quad (31.1)$$

Such *synchrotron condition* assures approximately constant radius of the beam orbit during acceleration. Transverse focusing by quadrupole magnets is needed to keep particles inside the rather limited space provided by the accelerator beam pipe passing through the magnet apertures. The maximum field of normal-conducting (NC) magnets is about 2 T, due to the saturation of ferromagnetic materials, and, while this is sufficient for lower energy colliders, such as most e^+e^- storage rings, it is not adequate for frontier-energy hadron (or muon) beams, because of the implied need for excessively long accelerator tunnels and prohibitively high total magnet power consumption. The development of superconducting (SC) magnets that employ high electric current carrying Nb-Ti wires cooled by liquid helium below 5 K, opened up the way towards higher fields and to hadron colliders at record energies [2]. For example, the 14 TeV c.m.e. LHC at CERN, uses double-bore SC magnets with a maximum field of 8.3 T at a temperature of 1.9 K, in a tunnel of $C = 26.7$ km circumference (dipole-magnet bending radius $\rho = 2800$ m). The double-bore design allows acceleration of the same particle type in opposite directions and also the operation with different particle species (e.g., protons and heavy ions) in the two apertures, while a single bore magnet implies the use of particles and antiparticles for the collider application. As the production of anti-particles is energy consuming and therefore limited, this concept opens the door to high-performance hadron

colliders.

The exploration of rare nuclear and high energy particle physics phenomena requires not only an appropriately high energy, but also a sufficiently large number of detectable reactions. The number of events of interest N_{exp} is given by the product of the cross section of the reaction under study, σ_{exp} , and the time integral over the instantaneous *luminosity*, \mathcal{L} :

$$N_{\text{exp}} = \sigma_{\text{exp}} \cdot \int \mathcal{L}(t) dt. \quad (31.2)$$

In the Tables, luminosity is stated in the units of $\text{cm}^{-2}\text{s}^{-1}$. The integral on the right is referred to as integrated luminosity \mathcal{L}_{int} , and, reflecting the smallness of typical particle-interaction cross-sections is often reported in units of inverse femto- or attobarn, e.g., $1 \text{ ab}^{-1} = 10^{42} \text{ cm}^{-2}$. Colliders usually employ bunched beams of particles with approximately Gaussian distributions, and for two bunches containing N_1 and N_2 particles colliding head-on with frequency f_{coll} , a basic expression for the luminosity is

$$\mathcal{L} = f_{\text{coll}} \frac{N_1 N_2}{4\pi\sigma_x^*\sigma_y^*} \mathcal{F} \quad (31.3)$$

where σ_x^* and σ_y^* characterize the rms transverse beam sizes in the horizontal and vertical directions at the interaction point, and \mathcal{F} is a factor of order 1, that takes into account inefficient geometric overlapping of the beams due to a crossing angle and finite bunch length, and dynamic effects, such as the mutual focusing of the two beam during the collision (see below). Having n_b bunches per beam increases the frequency of collisions $f_{\text{coll}} = n_b f_0$ where f_0 is either the revolution frequency of a circular collider or the repetition rate of a linear one. To achieve a high luminosity, one, therefore, has to maximize the population and number of bunches, either producing these narrowly or focusing them tightly, and colliding them at high frequencies at dedicated locations, where products of their reactions can be registered by particle detectors.

Subsequent sections in this report briefly expand on the beam dynamics behind collider design, comment on the realization of collider performance in a selection of today's facilities, and end with some remarks on future possibilities.

31.2 Beam Dynamics

Given the enormous and highly concentrated power carried by modern high energy particle beams, the main concern of beam dynamics in colliders is stability of motion of i) individual particles in accelerators, ii) single high-intensity beams of many particles moving together, and iii) colliding beams [3–5].

31.2.1 Single Particle Dynamics

While a reference particle at the nominal energy proceeds along the design trajectory (reference orbit) mostly determined by transverse magnetic dipole fields, other particles in the bunch are kept close by through the focusing effect of quadrupole fields. Assume that the reference particle carries a right-handed Cartesian coordinate system, with the co-moving z -coordinate pointed in the direction of motion along the reference trajectory, $z = s - vt$ (with v the reference particle velocity, and t time). The independent variable is the distance s of the reference particle along this trajectory, rather than time t , and for simplicity this reference path is taken to be planar. The transverse coordinates are x (horizontal) and y (vertical), where $\{x, z\}$ defines the plane of the reference trajectory.

Several time scales are involved, and this is reflected in the approximations used in formulating the equations of motion. All of today's high-energy colliders are alternating gradient synchrotrons or, respectively, storage rings and the shortest time scale is set by so-called *betatron oscillations*. The linearized equations of motion of a particle displaced from the reference trajectory are:

Highest-energy circular colliders face a serious impediment in the form of synchrotron radiation (SR) that causes an energy loss per turn of

$$\Delta E_{\text{SR}} = \frac{1}{3\varepsilon_0} \frac{Z^2 e^2 \beta^3 \gamma^4}{\rho}, \quad (31.14)$$

here, ε_0 is the permittivity of vacuum. For electrons and positrons it is equal to $\Delta E_{\text{SR}} = 88.5 \text{ [keV/turn]} E_b^4 [\text{GeV}] / \rho [\text{m}]$ and requires correspondingly high total RF voltage per turn to replenish the loss. Above a few hundred GeV, the SR energy loss becomes comparable to beam energy $\Delta E_{\text{SR}} \sim E_b$, which makes circular e^+e^- colliders impractical for c.m.e. above ~ 500 GeV.

Dynamics of the particle spin and sophisticated methods to maintain beam polarization along the acceleration chain, from the polarized sources to collisions, dedicated spin matching procedures to enable self polarization in e^+/e^- storage rings and the *resonant depolarization* method of ultra-precise c.m.e. calibration are described in [9].

31.2.2 High Intensity Beams

Ultimate collider luminosity calls for high beam currents $I_b = Zef_0n_bN$. Three related major difficulties include growing RF demands to compensate the synchrotron-radiation power loss $P = I_b\Delta E_{\text{SR}}$ in e^+/e^- beams, the advent of so-called *coherent* (or *collective*) *beam instabilities*, and growing demands for minimization of radiation due to inevitable particle losses. Many types of single- and multi-bunch instabilities are caused by beam interactions with electromagnetic fields induced by the beam itself due to the *impedance* of the vacuum chambers and RF cavities [4], or caused by unstable clouds of secondary particles, like electrons or ions, which are formed around the circulating beams [10]. These instabilities can develop as quickly as within tens to thousands of turns and need to be controlled. Mechanisms that are routinely employed to avoid coherent instabilities include the use of nonlinear magnets to generate sufficient spread of the tunes and therefore, provide *Landau damping*, fast beam-based transverse and longitudinal feedback systems, and electron/ion clearing (either by weak magnetic or electric fields or by modulation of the primary beam current profile rendering secondaries unstable, or by reducing the yield of secondary electrons via either a special coating or extensive *beam scrubbing* of the vacuum chamber walls).

High current beam operation is sensitive to even minuscule fractional intensity losses caused by particles' scattering at a large angle or with a large energy loss, sufficient for either the particle amplitudes $\sqrt{2J_{x,y}\beta_{x,y}(s)}$, or their dispersive position deviations $\Delta x = D_x(s)(\delta p/p)$ to exceed the available transverse aperture, usually set by collimators (otherwise, by the vacuum chamber and magnet apertures). This can be due to residual vacuum molecules near the beam orbit or Compton scattering off thermal photons, due to Coulomb scattering off other particles within the same bunch (*Touschek effect*), or due to collisions with opposite beam particles and fields, such as inelastic interaction of protons, Bhabha scattering $e^+e^- \rightarrow e^+e^-$, or radiative Bhabha scattering $e^+e^- \rightarrow e^+e^-\gamma$ (see corresponding chapters in [11]).

Particles can also get lost on the aperture as a result of much slower mechanisms of diffusion caused either by the above processes with smaller scattering amplitudes, but stochastically repeated many times, such as intensity-dependent multiple Coulomb *intrabeam scattering* [12], by external noises such as ground motion or magnetic field fluctuations, or via chaotic mechanisms like Arnold diffusion, modulational diffusion, or resonance streaming in nonlinear fields, enhanced by minor tune modulations. Diffusion leads to a slow evolution of the beam distribution function and appearance of highly unwanted large-amplitude tails and beam emittance growth. The only way to counteract it is to arrange *beam cooling* (damping of particle oscillations). The cooling requires a reaction force opposite to particle momentum arranged such that, on average, the corresponding dissipative particle energy loss is compensated for by external power [13, 14].

In the case of electron or positron storage rings, such cooling occurs naturally due to synchrotron radiation and provides an automatic route to achieve small equilibrium emittances through a balance between radiation damping and excitation of oscillations by random radiation of individual photons. Fast radiation damping allows *top-up injection* of new particles without removing existing ones, a useful method to maximize the integrated luminosity of circular e^+e^- colliders. Synchrotron radiation damping will also be an important cooling mechanism for future energy-frontier hadron colliders, like the proposed FCC-hh and SppC (see below). Four other methods of beam cooling have been developed and successfully employed to attain low emittances, namely *electron cooling* and *stochastic cooling* of heavy particles (ions and antiprotons), *laser cooling* of ion beams, and the *ionization cooling* of muons.

To avoid damage or excessive irradiation of accelerator components so that these remain accessible for maintenance in the tunnel, sophisticated collimation systems are utilized. These systems usually employ a series of targets or primary collimators which scatter the halo particles, and numerous absorbers (sometimes as many as a hundred, which intercept particles in dedicated locations) [15, Ch.9.7]. In the highest energy modern and future colliders, extreme total beam energies $n_b N E_b$ ranging from MJ to GJ and impacting surface energy densities reaching many GJ/mm² pose one of the biggest challenges for high efficiency and robust particle collimation.

31.2.3 High Luminosity Collisions

Eq. (31.3) for luminosity can be recast in terms of normalized transverse emittances Eq. (31.10) and the beta-functions β^* at the IP as:

$$\mathcal{L} = f_0 \gamma n_b \frac{N^2}{4\pi \sqrt{\varepsilon_{nx} \beta_x^* \varepsilon_{ny} \beta_y^*}} \mathcal{F}. \quad (31.15)$$

Here, equal bunch populations N are assumed in two Gaussian beams with the same emittances. Naturally, to achieve a high luminosity, one has to maximize the total beam populations $n_b N$ within the lowest possible emittances, and collide the beams at high frequency at locations where the focusing beam optics provides the lowest possible values of the amplitude functions β^* , the so-called *low-beta insertion*. The latter requires sophisticated systems of strong focusing elements, sometimes occupying quite a significant fraction of the collider's total length. The lowest $\beta_{x,y}^*$ is determined by the maximum field gradients and apertures in the interaction region (IR) magnets and the effectiveness of compensation of chromatic and nonlinear aberrations.

The typical geometric reduction factor is $\mathcal{F} \approx 1$, and it rarely drops below 0.5 for the majority of colliders, unless this is specifically required by physics processes under study. The reduction due to the *hourglass effect* is caused by the increase in transverse beam sizes as one proceeds away from the IP, where $\beta(s)$ grows parabolically, as in Eq. (31.8). For long round bunches the effect scales as $\mathcal{F} \approx \sqrt{\pi} A \exp(A^2) \text{erfc}(A)$, where $A = \beta^*/\sigma_z$. Nonzero beam crossing angles θ_c in the horizontal plane and long bunches (rms bunch length σ_z) will reduce the luminosity, too, by a factor $\mathcal{F} \approx 1/(1 + \Phi^2)^{1/2}$, where the parameter $\Phi = \sigma_z \tan(\theta_c/2)/\sigma_x^*$ is known as the *Piwinski angle*.

One of the most common limits to producing high luminosity arises from electric and magnetic forces of the opposite bunch at the IPs, characterized by a dimensionless *beam-beam parameter*:

$$\xi_{x,y} = \frac{r_0 N \beta_{x,y}^*}{2\pi \gamma \sigma_{x,y}^* (\sigma_x^* + \sigma_y^*)}, \quad (31.16)$$

where $r_0 = Z^2 e^2 / (4\pi \varepsilon_0 m c^2)$ is the classical radius of the colliding particle (with charge Ze and mass m). From Eqs. (31.3) or (31.15) and (31.16), one can note that the path to higher luminosity via higher beam intensity and smaller beam sizes almost automatically calls for a higher beam-beam

parameter as $\mathcal{L} \propto \xi$. Cited in the Tables, the beam-beam parameter is roughly equal to the betatron tune shift experienced by small-amplitude particles – positive in the case of opposite charge beams, like e^+e^- , and negative for same charge beams as in pp collisions. Beam-beam forces can lead to coherent effects, such as unstable beam oscillations or blow-up of one beam’s size while the other beam remains small or even shrinks (*flip-flop effect*). The tune spread arising from ξ and the nonlinear nature of beam-beam interactions results in strong diffusion along high-order transverse resonances $kQ_x + lQ_y = m$ and, ultimately, in beam size growth and beam losses. Operational experience indicates that the aforementioned effects are tolerable below certain *beam-beam limit* of $\xi_{x,y} \approx 0.003 - 0.012$ in hadron colliders [16], and – due to strong synchrotron radiation damping – an order of magnitude higher one in e^+e^- colliders, with maximum $\xi_{x,y} \approx 0.03 - 0.12$ [17, 18]. The accessible beam-beam parameter range can also be restricted by coherent beam-beam instabilities. These various limits translate into a maximum allowed single bunch intensity N and call for an increase of the number of bunches n_b to achieve higher luminosities.

In linear colliders, where each bunch collides only once, with typically much smaller beam size and experiencing much stronger forces, the strength of the collision is measured by the ratio of the rms bunch length σ_z to the beam-beam focal length. This ratio, called *disruption parameter* D_y , is related to ξ_y via $D_y = 4\pi\sigma_z\xi_y/\beta_y^*$. Stronger beam-beam focusing and higher D_y lead to effectively smaller beam sizes and a resulting luminosity enhancement; but it also makes the collisions more sensitive to small offsets, resulting in a *kink instability*. Additional beam-beam effects arising in the collisions at linear colliders are the emission of *beamstrahlung* (synchrotron radiation in the field of the opposing beam), along with e^+e^- pair creation, and depolarization by various mechanisms.

Beamstrahlung is relevant for both linear colliders, where it may significantly degrade the luminosity spectrum, and for future highest-energy circular colliders, where it may limit the beam lifetime, and also increases the energy spread and bunch length of the stored beam. For both types of colliders, the beamstrahlung is mitigated by making the colliding beams as flat as possible at the interaction point ($\sigma_x^* \gg \sigma_y^*$) to lower the beams’ EM fields. The photon energy spectrum of the beamstrahlung is characterized by the parameter $\Upsilon = (2/3)\hbar\omega_c/E_b$ [19], with $\hbar\omega_c$ denoting the critical photon energy. The spectrum strongly deviates from the classical synchrotron radiation spectrum for Υ approaching or exceeding 1.

For hadron colliders, two fundamental luminosity limits are the beam lifetime, determined by burn-off in the collisions due to inelastic pp interaction $dN/dt = -\mathcal{L}\sigma_{in}$, and the radiation from the collision debris, which may induce “quenches” (transitions to the normal-conducting state) of the superconducting final quadrupole magnets, and, in the long term, affect the equipment lifetime. Another limit on the achievable integrated luminosity in circular colliders is set by the minimum or average turnaround time (the time between the beam abort at the end of a physics fill and the start of the next physics collisions). Achieving practical filling times with many bunches in the collider requires either fast cycling injector machines and/or the top-up injection operation. The latter makes the average luminosity of circular electron-positron colliders approximately equal to the peak luminosity.

31.3 Recent High Energy Colliders

In this and the following section, elaboration is made on various issues associated with some of the recently operating colliders, particularly factors which impact integrated luminosity. Only general references are provided, where further information can be obtained. A more complete list of recent colliders and their parameters can be found in the High-Energy Collider Parameters tables.

31.3.1 Tevatron

The first superconducting synchrotron in history, the Tevatron [20] was converted into a proton-antiproton collider in 1985. Its 4.4 T dipole magnets employed Nb-Ti superconducting cable op-

also demonstrated an about 2-fold increase of luminosity from *disruption enhancement* due to the mutual focusing of the colliding electron and positron bunches at the interaction point,

31.4 Presently Operating Colliders

31.4.1 LHC

With a beam energy of 6.8 TeV (to be raised to the design value of 7 TeV through consolidation and magnet training), the superconducting Large Hadron Collider [27] presently is the world's highest energy collider. In the latest runs, peak luminosities of up to $2.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ have been achieved at 13 TeV cme - more than twice the design value (the current status is best checked at the Web site [28]). To meet its luminosity goals, the LHC operates with a high beam current of approximately 0.5 A, leading to stored energies of about 330 MJ per beam. Controlled energy deposition and component protection are given a high priority and a sophisticated highly efficient system of more than 100 beam collimators is employed [29]. At the energy of about 7 TeV per particle, synchrotron radiation poses a challenge, as the cryogenic system must remove roughly 7 kW due to synchrotron radiation, intercepted with a specially designed *beamscreen* inside the vacuum chamber, at a temperature of about 5–20 K, to be compared with a temperature of 1.9 K for the magnet cold bore. The elevated temperature allows for a more energy-efficient removal of beam-induced heat. The beamscreen also provides an effective cryo-pump for the vacuum system. When synchrotron-radiation photons hit the beamscreen, they can generate photoelectrons. These photoelectrons, and also any other electrons generated in the vacuum system, e.g. by residual-gas ionization, are accelerated in the electric field of the beam and may multiply via secondary-electron emission, with consequent electron cloud development. To mitigate this issue, the beamscreen is regularly subjected to beam-induced surface conditioning (*scrubbing*), thereby lowering the secondary emission yield. The two proton beams of 2556 bunches spaced by 25 ns are contained in separate pipes throughout most of the circumference and are brought together into a single 130 m long beam pipe at the interaction points. To avoid approximately 30 head-on collisions a small crossing angle of about 0.3 mrad is employed, which reduces the luminosity by about 15%. Still, the bunches moving in one direction experience multiple long-range encounters with the counter-rotating bunches and the resulting perturbations of the particle motion substantially contribute to the beam lifetime reduction. The dominant source of approximately 8 hour characteristic luminosity decay time is proton burn-off due to inelastic pp interaction with $\sigma_{\text{in}} \approx 81 \text{ mbarn}$, corresponding to *pile-up* of up to 50 (number of events per individual bunch crossing). In special physics runs with a few bunches and large β^* , the LHC achieved a head-on beam-beam tune shift of $n_{\text{IP}}|\xi| \approx 0.02$ with $n_{\text{IP}} = 2$, about twice as high as in regular operation.

The Tables also show the LHC luminosity performance in Pb-Pb collisions, which for the ATLAS and CMS experiments well exceeded the design value, while for the ALICE experiment, the luminosity is *levelled* near the Pb-Pb design value of $10^{27} \text{ cm}^{-2}\text{s}^{-1}$. The LHC can also provide Pb-p collisions as it did in 2013 and 2016, and other ion-ion or ion-proton collisions, at different energies.

In the coming years, an ambitious upgrade program, HL-LHC [30], with the accompanying LHC Injectors Upgrade [31], has as its target an order-of-magnitude increase in integrated luminosity through doubling the proton beam current, the utilization of new larger aperture Nb₃Sn superconducting final quadrupoles to allow squeezing the β^* to as low as 10 cm, superconducting compact *crab cavities* and luminosity leveling also for ATLAS and CMS as its key ingredients.

31.4.2 Electron-Positron Rings

Asymmetric energies of the two beams allow for the enhancement of *B*-physics research and for interesting interaction region designs. SuperKEKB operates with 7 GeV electron and 4 GeV positron beams since 2018 and is aiming for luminosities of $8 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ [32]. By summer

2022, a world record luminosity of $4.7 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ has been reached, still at rather low beam currents. Vertical beam-beam tune shifts of $\xi_y \approx 0.05$ for the 4 GeV positron beam, and $\xi_y \approx 0.03$ for the 7 GeV electron beam have been achieved. These values are still about a factor of two lower than at the previous KEKB. Since 2020 SuperKEKB operates with a virtual *crab-waist* collision scheme, first developed for the FCC-ee design [33]. The original crab-waist scheme, based on additional sextupole magnets, was earlier implemented at DAΦNE [34]. The general crab-waist concept combines a large Piwinski angle Φ , and an extremely low β_y^* ($\ll \sigma_z$) with the cancellation of the transverse betatron resonances which occur under conditions of $kQ_x + lQ_y = n$, where k, l, n are integers. The latter is achieved by means of existing or additional electromagnetic sextupoles with special betatron phase advances to the collision point [35]. The crab-waist collision scheme has become a design choice for all proposed future e^+e^- circular colliders.

Beside SuperKEKB and DAΦNE, three other e^+e^- ring colliders currently in operation are VEPP-2000 with \sqrt{s} up to 2.0 GeV, BEPC-II with \sqrt{s} up to 4.6 GeV and VEPP-4M with maximum c.m.e. of 12 GeV [1].

31.4.3 RHIC

The Relativistic Heavy Ion Collider [36] employs 3.45 T Nb-Ti superconducting magnets, and collides combinations of fully-stripped ions such as H-H ($p - p$), p -Al, p -Au, d -Au, h -Au, Cu-Cu, Cu-Au, Zr-Zr, Ru-Ru, Au-Au, and U-U over a wide energy range. The high charge per particle (+79 for gold, for instance) makes intra-beam scattering of particles within the bunch a special concern, even for seemingly moderate bunch intensities. In 2012, 3-D stochastic cooling was successfully implemented in RHIC [37] and is now routinely used. With stochastic cooling, steady increases in the bunch intensity, and numerous other upgrades, RHIC now operates at 44 times the Au-Au design average luminosity. Unique among high energy colliders, RHIC heavy ions beams cross the *transition energy* $\gamma_{rmt} = 1/\sqrt{\alpha_c}$ during acceleration – see Eqs.(31.11, 31.12) – a point where the derivative with respect to momentum of the revolution period is zero. This period of time is kept as short as allowed by the magnet ramp rate and must be dealt with carefully.

RHIC is also unique in its ability to accelerate and collide polarized proton beams. As proton beam polarization must be maintained from its low-energy source, successful acceleration through the myriad of depolarizing resonance conditions in high energy circular accelerators has taken years to accomplish [38]. An energy of 255 GeV per proton with 60% final polarization per beam has been realized. As part of a scheme to compensate the head-on beam-beam effect, electron lenses operated routinely during the record high beam-beam parameter polarized proton operation at 100 GeV energy in 2015 [39].

31.5 Future High Energy Colliders and Prospects

Modern nuclear physics and high energy particle physics face critical questions which require next-generation high-energy and high-intensity experiments using hadron-hadron, lepton-lepton, and lepton-proton colliding-beam facilities. Understanding the structure of the proton and neutron directly from the dynamics of their quarks and gluons governed by the quantum chromodynamics calls for new ion-ion and electron-ion colliders. Two types of colliders are generally aspired by the HEP community [40, 41]: i) Higgs factories with a c.m.e. of 240–250 GeV in e^+e^- collisions for precision studies of the Higgs boson ($m_H = 125$ GeV) and exploration of the Higgs sector in greater detail, including measurements of Higgs couplings to fermions and vector bosons, self-coupling, rare decays, mass and width, that can also deliver other electroweak precision physics, e.g. on the Z -pole (91 GeV c.m.e.), at the W -pair threshold (about 160 GeV), and when run as a top quark factory (365–380 GeV); and ii) colliders with c.m.e. levels significantly beyond those of the LHC to explore the energy frontier for potential discoveries through direct searches in pp , $\mu\mu$, and e^+e^- interactions. In addition, precision physics at future high-luminosity factories operating at

the τ -charm energy also provides sensitivity to new physics at multi-TeV energies and beyond. A comprehensive review of the future colliders' projects, ideas, and R&D activities can be found in Ref. [1]; Ref. [42] presents comparative analysis of implementation challenges of numerous future HEP collider proposals. Below we only briefly summarize leading collider proposals for construction over the next several decades which rely mostly on currently available technologies, such as normal-conducting (NC) or SC RF and/or NC or SC magnets, some of them requiring reasonable scope and duration mission-oriented development programs, as well as advanced schemes based on plasma acceleration and other innovative ideas. Tentative parameters of some of the colliders discussed, or mentioned, in this section are summarized in Table 31.1 and Table 31.2.

31.5.1 Ion-Ion and Electron-Ion Colliders

NICA (Nuclotron-based Ion Collider fAcility) is a new accelerator complex under construction at the Joint Institute for Nuclear Research (JINR, Dubna, Russia) to study properties of hot and dense baryonic matter, strong interactions between quarks and gluons, and spin physics [43]. NICA will provide a variety of beam species, ranging from protons and polarized deuterons to massive gold ions. The 500 m circumference SC magnet based collider is designed for average luminosity in heavy ion and light ion interactions at $\sqrt{s_{NN}}=4\text{--}11$ GeV of $1\times 10^{27}\text{cm}^{-2}\text{s}^{-1}$ for a variety of nuclei up to $^{197}\text{Au}^{79+}$, and for polarized proton and deuteron collisions at $\sqrt{s}=12\text{--}27$ GeV with $\mathcal{L}=(1\text{--}10)\times 10^{31}\text{cm}^{-2}\text{s}^{-1}$. NICA major accelerator challenges include strong intrabeam scattering and space-charge effects which will be mitigated by extensive use of electron and stochastic cooling systems.

The recently announced Electron-Ion Collider (EIC) for nuclear physics research will be built at Brookhaven National Laboratory in the US and arrange collisions between an the reconfigured RHIC with a 41–275 GeV proton beam and a 5–18 GeV electron beam stored in a new ring (eRHIC) [44]. The EIC physics requirements [45], include highly polarized ($P_{e,n} \sim 70\%$) electron and nucleon beams (as the precision of measurements of interest scales as $\mathcal{L}P_e^2 P_n^2$), a spectrum of ion beams from deuterons to the heaviest nuclei (U or Pb), variable c.m.e. values from $\sqrt{s}=20$ GeV to 140 GeV, high luminosities of up to $10^{33\text{--}34}\text{ cm}^{-2}\text{s}^{-1}$, as well as possibilities of having more than one interaction region. Main accelerator design challenges on the path to the required energy, luminosity, and polarization, include the development of SRF crab-cavities and advanced SC magnets for interaction region focusing, energy-recovery linac (ERL) based electron cooling of hadron beams, essential to attain luminosities two orders of magnitude beyond the predecessor HERA ep collider, and high intensity polarized particle sources, augmented by the development of special magnets and operational techniques to preserve the polarization through the acceleration process to the collisions, including swap-out injection.

31.5.2 Higgs/Electroweak Factories

Higgs factory proposals generally aim at improving the precision of coupling measurements of Higgs boson, top quark, W and Z by an order of magnitude or more compared with previous studies. Two proposals for ~ 100 km circumference circular e^+e^- colliders have recently gained momentum: the Future Circular Collider (FCC-ee) at CERN [33] and the Circular Electron-Positron Collider (CEPC) in China [46]. Design philosophy of these machines assumes use of the maximum RF power available to compensate $O(100$ MW) synchrotron radiation losses $P_{SR} = 2I \cdot \Delta E_{SR}$ and operation at the beam-beam limit ξ_y that yields peak luminosity:

$$\mathcal{L} = \frac{3}{16\pi r_0^2(m_e c^2)} \frac{P_{SR} \xi_y \rho}{\beta_y^* \gamma^3}, \quad (31.17)$$

that scales approximately as $1/E_b^{3.5}$ for practical limits on P , ξ_y and β_y^* . The short beam lifetime at the high target luminosity, due to radiative Bhabha scattering, requires these machines to be

e^\pm ($\alpha \approx 1/137$ denotes the fine-structure constant), and the last factor H_D , typically between 1 and 2, represents the enhancement of luminosity due to the *pinch effect*, the additional focusing occurring during the collision of oppositely charged bunches. The management of P_{wall} leads to an upward push on the bunch population N with an attendant rise in the energy radiated due to the electromagnetic field of one bunch acting on the particles of the other (beamstrahlung). Keeping a significant fraction of the luminosity close to the nominal energy represents a design goal, which is met if N_γ does not exceed a value of about 1. A consequence is the use of flat beams, where N_γ is managed by the beam width, and luminosity adjusted by the beam height, thus the explicit appearance of the vertical beam size σ_y^* .

The ILC [47, 48] is based on 1.3 GHz standing-wave superconducting accelerating structures operating at 2 K with 31.5 MV/m average gradient, up to 8 nm vertical beam size at the IP, and luminosity comparable to the LHC. Progress toward higher field gradients and Q values continues to be made, with nitrogen-doping techniques being a recent example [49]. CLIC is based on a novel two-beam acceleration scheme [50]. Here, room-temperature NC copper high-gradient 12 GHz accelerating structures are powered by a high-current 1.9 GeV drive beam to efficiently enable accelerating gradients of up to 100 MV/m (though optimal $G=70$ MV/m for the first CLIC stage at $\sqrt{s}=380$ GeV, and for this stage an alternative RF power drive option with 12 GHz klystrons powering is also being considered). To reach their design luminosities, both CLIC and ILC require unprecedented rates of positron production about 40 times the world record set by the SLC positron source, and very tight control of imperfections, such as $O(10 \mu\text{m})$ accuracy of pre-alignment of the main linac and beam delivery system components at the level, suppression of fast vibrations of the quadrupoles due to ground motion to $O(1 \text{ nm})$ level at frequencies above 1 Hz, advanced beam-based trajectory tuning, and mitigation of the effect of wakefields [51].

Recent developments of the RF technologies show promise of more compact linear e^+e^- Higgs/Electroweak factories based on either ~ 70 MV/m travelling-wave 1.3 GHz SRF structures [52] or on 70–120 MV/m high-efficiency, distributed RF coupling, normal-conducting cold copper 5.7 GHz accelerating cavities at liquid nitrogen temperature [53].

There are a number of alternative ideas proposed for studies of the Higgs/Electroweak physics, such as high-energy, high-luminosity e^+e^- collider in a 100 km tunnel using ERLs to accelerate particles to collision energy in 4 to 6 passes and return up to 81% of the energy back into the SRF cavities on deceleration turns, thus, lowering the required facility power several-fold [54]; similar power recovery in one pass can greatly improve efficiency of linear colliders [55–57]; an arrangement of $\gamma\gamma$ collision through near-IP conversion of high energy electron beams into intense photon beams by backward Compton scattering off a high-power laser [58,59] or off an FEL photon pulse [60,61]; $\mu\mu$ Higgs factory with unprecedented 0.004% energy resolution [62], and a high-energy lepton-hadron collider bringing into collision a 60-GeV electron beam from an ERL with the 7 TeV protons circulating in the LHC (LHeC) [63]. At lower energies, Super Tau-Charm Factory proposals aim at the production and precise study of charmonium states and of the tau lepton [64,65].

31.5.3 Energy Frontier Circular Colliders

Several hadron and lepton colliders have been proposed to extend the energy reach beyond the LHC. As noted above, ambitious plans have been proposed to upgrade the FCC and CEPC to hadron colliders – FCC-hh at CERN and Super Proton Proton Collider (SPPC) in China, respectively – by means of next- or next-next generation SC magnets installed in the arc sections of the 100 km rings, so as to enable \sqrt{s} of the order 100 TeV or above [66,67]. Comparable discovery reach is expected for a circular 10–14 TeV muon collider [68], significantly beyond that of practical e^+e^- linear colliders.

The maximum beam energy Eq.(31.1) is directly proportional to the magnetic field and to the

and compressor rings, in which high intensity 1-3 ns long proton bunches are formed, (iii) a liquid-mercury or other high-power target for converting the proton beam into a tertiary muon beam with energy of about 200 MeV, (iv) a multi-stage ionization cooling section that reduces transverse and longitudinal emittances by several orders of magnitude and creates a low emittance beam, similarly to that recently demonstrated [77], (v) a multistage acceleration system, possibly employing either rapid cycling synchrotrons or RLAs to accelerate muons in a modest number of turns up to the final energy using superconducting RF technology, and, finally, (vi) a 3–14 km diameter collider ring, where counter-propagating muon beams are stored and collide over the roughly 3000 turns corresponding to the muon lifetime.

The intense neutrino flux originating from the multi-TeV muon beams decaying in the collider poses another challenge — the need to minimize the environmental impact. The collider complex is usually located underground and when the produced neutrinos emerge at the surface, a small fraction interacts with the rock (and other material) and produces ionizing radiation dose scaling as γ^3 . The impact of this neutrino-induced radiation can be mitigated, for example, by continually adjusting the orbits of the beams to spread them out on a wider area, by deeper collider tunnels or with a further reduced emittance of the muon beam so that the required luminosity could be obtained using a substantially smaller number of muons.

31.5.4 Plasma Acceleration and Other Advanced Concepts

Since about the mid-1950s, it has been understood that collective plasma-based accelerators promise extremely large accelerating gradients, approximately three orders of magnitude greater than ~ 100 MV/m obtained in conventional breakdown limited RF structures [78]. Ionized plasmas can sustain electron plasma density waves with electric fields in excess of $E_0 = cm_e\omega_p/e$ or

$$E_0 \approx 96 \text{ [V/m]} \sqrt{n_0 \text{[cm}^{-3}]}, \quad (31.20)$$

where n_0 denotes the ambient electron number density and $\omega_p = \sqrt{e^2 n_0 / (m_e \epsilon_0)}$ is the electron plasma frequency [79].

Such gradients can be effectively excited by either powerful external pulses of laser light or by electron bunches if they are shorter than the plasma wavelength $\lambda_p = c/\omega_p \approx 1 \text{ mm} \times \sqrt{10^{15} \text{ cm}^{-3}/n_0}$, or by longer beams of protons if their charge density is modulated with the period of λ_p . In the past decade, we have seen impressive progress in the plasma wakefield acceleration of high-quality beams. Laser-driven electron energy gain of about 8 GeV over 20 cm of plasma with density $3 \times 10^{17} \text{ cm}^{-3}$ has been demonstrated at the BELLA facility at the Lawrence Berkeley National Laboratory (LBNL) [80]. Short electron bunches were used to boost the energy of externally injected electron bunches by 9 GeV over 1.3 m of $\sim 10^{17} \text{ cm}^{-3}$ plasma at the FACET facility in SLAC [81]. The AWAKE experiment at CERN used self-modulating long 450 GeV proton bunches to accelerate electrons to 2 GeV over 10 m of 10^{15} cm^{-3} plasma [82].

Whether plasma acceleration will find application in an HEP facility is not yet clear, given the necessity of staging and phase-locking acceleration in multiple plasma chambers [83]. Another critical issue is the power efficiency η for a collider based on plasma acceleration, whose luminosity would still be described by Eq. (31.18). Positrons can be accelerated as well but, contrary to electrons, plasma-wakefields will naturally defocus positively charged particles. Such fundamental deficiency can be addressed, e.g., in more complicated plasma acceleration schemes for e^+ [84]; or by acceleration of electrons in both linacs with conversion into a $\gamma\gamma$ collider at the IP. In addition, many novel approaches have been proposed to extend the energy or physics reach of future particle colliders, reduce their cost, and improve their luminosity or energy efficiency. Those include: i) *dielectric wakefield accelerators* in which resonant dielectric accelerating structures are fed by ultra-short RF pulses of wakefields driven by either collinear or preceding high charge

- [86] E. Peralta *et al.*, Nature **503**, 7474, 91 (2013).
- [87] V. D. Shiltsev, Physics-Uspekhi **55**, 10, 965 (2012).
- [88] T. Tajima and M. Cavenago, Physical Review Letters **59**, 13, 1440 (1987).
- [89] D. Budker *et al.*, Annalen der Physik **532**, 8, 2000204 (2020).