The luminosity of hadron colliders rises with the beam intensity, until some limit is encountered, mostly due to head-on and long-range beam-beam interaction, due to electron cloud, or due to conventional impedance sources. Also beam losses caused by various mechanisms may affect the performance. The limitations can be alleviated, if not overcome, by a proper choice of beam parameters and by dedicated compensation schemes. Examples include alternating crossing at several interaction points, electromagnetic wires, super-bunches, electron lenses, clearing electrodes, and nonlinear collimation. I discuss such mitigating measures and related research efforts, with special emphasis on the LHC and its upgrade.
Beam Dynamics Challenges for Future Circular Colliders

F. Zimmermann, CERN, Geneva, Switzerland

Abstract

The luminosity of hadron colliders rises with the beam intensity, until some limit is encountered, mostly due to head-on and long-range beam-beam interaction, due to electron cloud, or due to conventional impedance sources. Also beam losses caused by various mechanisms may affect the performance. The limitations can be alleviated, if not overcome, by a proper choice of beam parameters and by dedicated compensation schemes. Examples include alternating crossing at several interaction points, electromagnetic wires, super-bunches, electron lenses, clearing electrodes, and nonlinear collimation. I discuss such mitigating measures and related research efforts, with special emphasis on the LHC and its upgrade.

INTRODUCTION

The two primary performance characteristics of a circular collider are its centre-of-mass energy and its luminosity, which together determine the discovery reach. The last half a century has witnessed a factor 6 increase in centre-of-mass energy per decade. With the LHC coming on line in 2007, this trend will continue. In parallel, since 1975, the luminosity has increased on average by a factor 20 every 10 years. Recently, KEKB exceeded the landmark number of $10^{34}$ cm$^{-2}$s$^{-1}$, and both KEKB and PEP-II developed strategies for increasing the luminosity further by at least another factor of 10. The luminosity evolution of hadron colliders is less impressive: after more than 20 years, the first ever-hadron collider, the CERN Intersectioning Storage Rings (ISR), still holds the record. The latter will, however, be surpassed by the LHC in its first weeks of operation.

Extrapolating from past and present experience, beam-dynamics challenges at future circular colliders are: (1) the head-on beam-beam interaction, as encountered, e.g., at the SPS collider and at most lepton machines, (2) the long-range beam-beam collisions, which have become a noticeable limitation in the Tevatron Run-II, (3) the electron cloud, which has been a limiting factor at PEP-II and KEKB, (4) continuous or sudden generation of unbunched beam, as at the Tevatron or HERA, (5) emittance growth due to incoherent effects like intrabeam scattering (IBS), observed, e.g., in RHIC, and (6) collective instabilities driven by conventional impedances, e.g., the laminations of the Lambertson magnets in the Tevatron or LHC collimators. The dynamic aperture can be a concern as well.

As the machines become larger and more complex, storing a higher beam current at increasing energy, new limitations may appear. Examples include machine protection, debris from the collision point, e.g., for LHC ion operation, or increased sensitivity of cold machines to pressure rise. Also, as spot sizes are reduced, maintaining stability of the beam-beam collision offsets at the nanometer level may prove crucial for preserving beam emittance. New concepts will have to be developed to push the energy and luminosity frontiers further. Promising ideas include beam-beam compensation with electron lenses [1] or electro-magnetic lenses [2], hadron superbunch collisions [3], strong rf focusing [4], induction rf [5], high-energy bunched-beam electron cooling [6], high-energy stochastic cooling, and linac-ring collisions.

BEAM-BEAM INTERACTION

In a circular collider, the beams collide at one or several interaction points (IPs) on every turn. For lepton colliders, the diffusion limiting the achievable beam-beam tune shift is driven by the interplay of beam-beam forces and synchrotron radiation, which can be reliably modeled [7]. On the other hand, it is difficult to predict the exact beam-beam limit for hadron colliders. Computer simulations suggest that, if the collisions are head on and without any errors, hadron colliders should be more stable and reach higher ‘head-on’ beam-beam tune shifts $\xi_{HO}$ than lepton colliders. In reality, however, beam-beam tune shifts achieved at hadron colliders are roughly a factor 10 lower than those at lepton machines. This may be attributed to the absence of radiation damping, and the thereby increased sensitivity to noise, errors and small imperfections. E.g., simulations [7] show that the introduction of a crossing angle induces a substantial drop in beam-beam tune shift. Also a random beam-beam offset, $\Delta x$, varying from turn to turn generates emittance growth, which for $n_{IP}$ IPs is

$$\Delta \epsilon / (\Delta t) \approx n_{IP} 4 \pi^2 (\Delta x)^2_{rms} \xi_{HO} f_{rev} / \beta^*.$$  

Demanding less than 10% emittance growth per hour, for a SuperLHC [8] with $\beta^* = 0.25$ m, $n_{IP} = 2$, $\xi_{HO} = 0.005$, $\gamma \approx 7500$, and $\epsilon \approx 3.75 \mu$m, the tolerance on the turn-to-turn variation is $(\Delta x)_{rms} \leq 12$ nm, about 0.1% of the IP beam size. One source of beam-beam offsets that can be important, especially for large machines, is natural ground motion. It is amplified either dynamically by the optical lattice (e.g., for wave lengths comparable to, or shorter than, the betatron period) or by mechanical resonances of magnet supports. Ground motion results in a varying beam-beam offset, which fortunately is not random, but highly correlated between successive turns. Extending the formalism in [9] to estimate the associated emittance growth we weight the offset $\Delta x > 0$ with the perturbation frequency $f$ and obtain

$$\Delta \epsilon / (\Delta t) \approx n_{IP} 8 \pi^2 x_{rms} \xi_{HO}^2 R^2 P(f_0) f_0^2 / \beta^*,$$

where $R$ is the lattice response function, $P(f)$ the ground motion power spectrum, $f_0 \approx v / C$ a limiting frequency above which the lattice amplifies ground waves, $C$ the circumference, and $v$ the speed of the ground waves (about 2500 m/s). As an example, we assume a power spectrum measured in a quiet LEP tunnel [10] $P(f) \approx 10^{-16} / f^3$ m$^2$/Hz where $f$ is given in Hz. This corresponds to 1 nm integrated rms ground motion above 6 Hz. Inserting nominal LHC parameters we then estimate a relative growth rate of only $3 \times 10^{-3}$ per hour, roughly consistent with the much more detailed analysis in [11]. Considering instead the natural ground motion of the HERA site [12], the
LHC emittance growth would be several percent per hour. For an asymmetric interaction-region optics, as in the LHC, also lower-frequency ground motion may contribute to the variation in beam-beam offset [9].

Long-range collisions unavoidably appear when colliding beams with close bunch spacing. As the bunches approach the primary interaction point (IP) or depart from it after collision, they encounter bunches of the opposing beam at intervals equals to half the bunch spacing. If both beams followed the same orbit, each of these parasitic collisions would introduce a beam-beam tune shift equal to that at the primary collision point. Therefore, to reduce the effect of the parasitic collisions, the beams are separated either by dipoles close to the IP or by means of a crossing angle $\theta_c$. In the latter case the transverse beam-beam distance in units of rms beam sizes at the long-range collision points is $d = \theta_c \sqrt{\beta^*/(\gamma e)}$. For the same crossing angle, a smaller $\beta^*$ implies a smaller normalized separation. If $\beta^*$ is small, all the long-range collisions next to one primary IP occur either at about equal betatron phase, or, on the other side of the IP, with opposite sign of separation at a phase advance of about $\pi$, yielding the same net effect. Thus, in a good approximation, the deflections from all $n_{LR}$ long-range collisions on both sides of the IP can be added.

Figure 1 illustrates the simulated effect of long-range collisions in the LHC. Details of the employed ‘weak-strong’ simulation are described in [13, 14]. The blue curves, labelled ‘$x - y$ crossing’, refer to the nominal situation. There is a steep increase in diffusion rate at an amplitude of about 6-6.5 $\sigma$. The other two curves represent alternative schemes of beam crossing at the two primary IPs. Here, a similarly high diffusion level is reached only for about 2$\sigma$ larger amplitudes, not far from the center of the opposing beam, at 9.5$\sigma$. At amplitudes below 6$\sigma$ the simulated diffusion is smaller for the blue nominal curves, but here the overall values are low and they might correspond to the resolution limit of the simulation and not reflect true chaotic behavior.

$$\xi_{LR} \approx 2nLR(2\xi_{HO})/d^2.$$

$$d = \theta_c \sqrt{\beta^*/(\gamma e)}.$$
ternating crossing case, the trajectories at large amplitudes \((x \approx y \approx 6\sigma)\) are unstable and lost, while, for the equal-plane crossing, the motion remains bounded and no losses occur. It is clear that the tune footprint alone provides an incomplete characterization of the dynamics. Experimental simulations of the two crossing schemes will be compared at the SPS [2].

![Sample transverse trajectories from LHC long-range simulation projected onto the amplitude plane](image)

Figure 2: Contour plot of the determinant of the Jacobian matrix for long-range beam-beam collisions as a function of amplitudes \(x\) and \(y\) in \(\sigma\) for horizontal-horizontal crossing (left) and horizontal-vertical crossing (right).

![Sample transverse trajectories from LHC long-range simulation projected onto the amplitude plane](image)

Figure 3: Sample transverse trajectories from LHC long-range simulation projected onto the amplitude plane \(y = \sqrt{\frac{2}{\alpha}} \beta x\) vs. \(x = \sqrt{\frac{2}{\alpha}} \beta x\) in units of \(\sigma\). The color distinguishes different trajectories.

To combine all the advantages of a crossing angle with those of a head-on collision, crab cavities were proposed for linear colliders [20] and soon after for storage rings [21]. In 2005 for the first time a crab cavity will be installed at an operating collider, KEKB [22]. The crab cavity is an rf cavity operating in the transverse dipole mode, which deflects the head and tail of a bunch in opposite direction so that in the beam-beam centre-of-mass frame the collision becomes head on. Since there is neither geometric luminosity reduction nor excitation of synchrotron resonances, the crab-crossing scheme allows for large crossing angles, and, thereby, avoids the problem of the long-range collisions. Ideally, there should be one crab cavity on each side of the IP for either of the two beams. The required crab-cavity rf voltage is \(V_{\text{crab}} = cE_b \tan(\theta_c/2)/(\epsilon \omega t \beta \epsilon_{\text{crab}})\). If the rf phases of the crab cavities for the two beams drift with respect to each other, a net deflection to the beam center results which is different for the two beams, so that they will collide with a transverse offset \(\Delta x\). If \((\Delta x)_{\text{max}}\) denotes the acceptable random offset, the left-right phase stability tolerance is \(\Delta \phi_{\text{crab}} \leq (\Delta x)_{\text{max}}2\pi/(\lambda t \theta_c)\). Example parameters are listed in Table 1. The tight limit at SuperLHC [8, 23] corresponds to the 12-nm tolerance on random offsets derived from Eq. (1). If crab cavities are not available or apertures limited, and long-range collisions at harmful distance cannot be avoided, various types of compensation are possible, in addition to the alternating crossing at two IPs. The Tevatron Electron Lens collides a pulsed electron beam of close to 1 A current with the antiproton bunches, which cancels part of the fields and tune shifts induced by the proton collisions [1]. With a second lens coming into operation in 2004, both transverse tunes can be corrected for each bunch. By shaping the transverse profile of the electron beam even a compensation of the nonlinear beam-beam force is conceivable. For the LHC a compensation of the long-range collisions around each IP was proposed in [24]. It consists of thin water-cooled about 1-m long conducting tubes parallel to beam at about the same normalized transverse distance and the same betatron phase as the long-range encounters. The shape of the field mimics that of the other beam. With the opposite direction of current a nearly perfect cancellation can be accomplished. Several prototype devices are installed in the SPS. Machine experiments over the last two years so far confirmed simulations [2]. With a pulsed wire current, also the tune shifts of individual bunches could be corrected.

| Table 1: Crab cavity parameters for KEKB and SuperLHC |
|-------------------------------|------------|-------------|-------------|
| variable                      | symbol     | KEKB HER    | SuperLHC    |
| beam energy                   | \(E_b\)    | 8 GeV       | 7 TeV       |
| rf frequency                  | \(f_{\text{crab}}\) | 509.9 MHz   | 1.3 GHz     |
| crossing angle                | \(\theta_c\) | 11 mrad     | 8 mrad      |
| IP beta function              | \(\beta^+\) | 0.33 m      | 0.25 m      |
| cavity beta funct.            | \(\beta_{\text{av}}\) | 100 m      | 2 km       |
| kick voltage                  | \(V_{\text{crab}}\) | 1.44 MV     | 46 MV       |
| phase tolerance               | \(\Delta \phi_{\text{crab}}\) | 0.06 mrad   |             |

**TOUSCHEK AND IBS**

The operation of HERA and the Tevatron is complicated by unbunched beam, which emerges outside the normal bunch-train structure, and occasionally causes spikes in the physics detectors. The dilution of the abort gap compromises the machine protection and can lead to magnet quenches during beam dump. Several origins of beam loss from the rf bucket have been identified, the most prominent ones being rf noise, intrabeam scattering and Touschek effect. The first two processes lead to particle losses from the rf bucket, only if the longitudinal emittance of the bunch is blown up sufficiently to fill the bucket area. In case the bunch occupies a small portion of the bucket, unbunched beam can still be created by the Touschek effect.

The name Touschek effect [25] refers to a particle-particle collision within a bunch, by which so much energy is transferred from transverse into longitudinal phase space, that the scattered particles leave the stable rf bucket. The loss rate due to Touschek scattering is quadratic in the bunch population, \(dN_b/\alpha N_0 t = -N_b^2 t/(1 + \alpha N_0 t)\). Hence, the number of particles outside the rf bucket increases as \(N_{\text{umb}} = \alpha N_0^2 t/(1 + \alpha N_0 t)\), with \(N_0 = N_b(0)\). For round beams and considering a bunch small compared with the bucket size, the Touschek scattering rate is [26, 27] \(\alpha_{\text{rd}} = \pi \alpha V \sigma_{\text{T}}^3 \sigma_\eta V \eta D(\eta/\sigma_q)\), where \(V = \frac{8\pi^3}{3} \sigma_x \sigma_y \sigma_z\) denotes the bunch volume, \(\eta \equiv (\Delta E/E)_{\text{max}}\) the maximum energy deviation ac-
cepted by the rf system, $\delta q \equiv \gamma \sigma_x/\beta_x$, and, at $\epsilon > 20$, $D(\epsilon) \approx e^{-\epsilon}$. For the ultimate bunch intensity, $N_b = 1.7 \times 10^{11}$, in the LHC, unbunched beam is produced at a rate per proton of $3 \times 10^{-4} \text{ hr}^{-1}$ during injection and $1.3 \times 10^{-3} \text{ hr}^{-1}$ at top energy [28]. Once protons are outside of the bucket, they lose energy due to synchrotron radiation. At top energy in the LHC, this energy loss amounts to $d\delta /dt \approx -1.0 \times 10^{-5} \text{ s}^{-1}$ at 7 TeV. The LHC collimators defining an energy aperture of about $4 \times 10^{-3}$, a scattered proton is lost after $\gamma_{\text{loss}} \approx 6.5$ minutes and one expects to observe a steady-state coasting beam fraction of $\alpha_{\text{coast}} \approx 1.4 \times 10^{-5}$. The Touschek rates at HERA and the Tevatron are $10^{-3}$ [28] and $2 \times 10^{-4}$ [29], respectively, per hour and per proton.

Intrabeam scattering (IBS) is a similar phenomenon. The term refers to multiple scattering inside the bunch, which, above transition, causes a blow up of the longitudinal and transverse emittances. The general theories by Bjorken-Mtingwa and Piwinski consist of intricate expressions. However, considerable simplifications are possible [30]. Further, considering a smooth lattice with constant beta function and dispersion, and a round Gaussian beam with $D_x \approx \sqrt{\beta_x \epsilon_x}$, the longitudinal IBS amplitude growth rate can be written

$$\frac{1}{T_\delta} \frac{d\sigma_\delta}{dt} \approx \frac{r_x^2 c N_b (\log)}{16\gamma (\epsilon_x^{3/2}) (\gamma \sigma_x \sigma_\delta) \beta_\delta^{3/2}},$$

where $(\log) \approx 24$ denotes the Coulomb logarithm, and $1/T_\delta \approx \sigma_x^2 D_x^2 / (\beta_x \epsilon_x) / T_\delta$, which, by our assumption, is larger than the longitudinal rate. In [31] this approximation was compared with the exact computation of MAD. For fixed bunch length, the longitudinal IBS growth rate scales with the inverse square of the momentum spread. Thus, it can be made negligible by increasing the energy spread, e.g., for a superbunch [31]. If the increased energy spread requires a larger momentum bandwidth for the squeezed collision optics, a solution would be to adapt the Raimondi-Seryi final focus [32] with local chromatic correction and dispersion across the low-$\beta$ quadrupoles to a storage ring.

**ELECTRON CLOUD**

Unlike for past colliders where particles and anti-particles circulated in opposite directions around the same ring, with about equal current, in the present or next generation of colliders (RHIC, LHC, B factories,...) each beam has its own separate beam pipe. This leads to a pulsed electric field which can create and sustain an electron cloud of substantial density inside the beam pipe. There are four sources of electrons: (1) gas ionization, (2) beam loss at the wall, (3) photo-emission from synchrotron radiation (in positron rings and in the LHC at top energy), and (4) beam-induced multipacting, which is an amplification process driven by acceleration of electrons in the beam field and subsequent secondary emission at the chamber wall.

Electron clouds have been observed at almost all storage rings operating with single beams of positrons or protons, except for at large bunch spacings. Concerning proton bunch trains, since 1999 electron-cloud effects are observed with the LHC beam in the CERN SPS at bunch spacings of 25 ns and, much weaker, 75 ns [34]. In 2002, electron-cloud phenomena were noticed at the Tevatron, when operated with uncoalesced proton beam (19 ns spacing). At RHIC, pressure rise due to electron cloud is seen both with 55 ns and 110 ns bunch spacing, initially in the warm straight sections only. In 2004, RHIC also reported evidence of electron-cloud induced pressure rise in a cold section of the ring. Figure 4 summarizes observed intensity thresholds for electron-cloud effects as a function of bunch spacing for a number of storage rings. The data for the CERN SPS, with three different bunch spacings, for the CERN PS, for the APS, and the Tevatron all lie approximately on a straight line with a slope of about 1 on this log-log plot. At RHIC the threshold appears to be lower, while at DAFNE and KEKB it is higher. The different thresholds might reflect differences in the surface parameters, e.g., different levels of conditioning from ‘scrubbing’ (by the electron cloud itself). Also shown in the figure are design working points for the LHC and, as a reference, for the damping rings of several proposed linear colliders.

Figure 4: Electron-cloud threshold bunch intensity vs. bunch spacing observed at existing storage rings (blue) and nominal working points of several future projects (red).

The electron cloud may have various harmful effects. It can induce a pressure rise, as in the SPS, at RHIC, PEP-II and the Tevatron. It can cause fast single-bunch or coupled-bunch instabilities, as seen at the SPS, KEKB and PEP-II. Or it may lead to a heat load on the beam screen inside the cold magnets of the LHC. Indeed, measurable heat load due to electron bombardment of the chamber wall was detected in warm and cold calorimeters at the SPS. In the LHC, an additional, potentially larger contribution to the heat load can arise indirectly from scattering off the residual gas, if the pressure increases as a result of the electron cloud. The electron cloud can also perturb beam diagnostics, e.g., SEM profile monitors or electro-static position pick ups.

Figure 5 shows the simulated heat load in the LHC arcs as a function of bunch spacing for the design bunch population of $1.15 \times 10^{11}$ and two values of the maximum secondary emission yield. The nominal bunch spacing is 25 ns, and the cryogenic cooling capacity available for electron-induced heating is of the order 1 W/m.

A multitude of countermeasures and strategies have been developed and adopted for suppressing electron-cloud effects. Among these are multi-bunch and intra-bunch feed-
LOW ENERGY, to the electron-cloud space-charge field and to the initial
tensions at higher bunch intensity indicate a saturation of the
multiple single short bunches of the same total charge [36].

The bunch structure is critical: in simulations a few flat su-
pression that can be intercepted by a local absorber, while

For the nominal LHC bunch spacing of 25 ns, simula-
tions at higher bunch intensity indicate a saturation of the
total number of electrons. This saturation can be attributed to the
electron-cloud space-charge field and to the initial
low energy, \( E_0 \) (\( E_0 \approx 1.8 \) eV), of secondary electrons,
which limit the electron line density building up between
bunches to [37] \( \lambda_e \approx E_0/(r_e m_e e^2) \approx 1.3 \times 10^9 \) m\(^{-3}\),
independent of beam current.

OUTLOOK

There are many other challenges facing the next generation
of high-energy colliders. One of these is collimation and machine protection. Simultaneously fulfilling the requirements for high cleaning efficiency, acceptable
impedance, and collimator survival in case of kicker failure or poor beam lifetime is a highly non-trivial task for the LHC [38]. A possible solution for future projects may be nonlinear collimation, as has been proposed for linear colliders [39, 40]. For circular colliders, there are two possible approaches: (1) single-turn collimation with a pair of
nonlinear elements, and (2) multi-turn collimations with a
single element. In (1), a pair of nonlinear elements is separated by an optical – I transform. The first blows up the
beam size and deflects halo particles to large amplitudes
where they can be intercepted by a local absorber, while
afterwards the second cancels the nonlinear aberration intro-
duced by the first. In (2), a single nonlinear element is
used to create an artificial controllable dynamic aperture.

Cooling colliding hadron beams would counteract emi-
tance growth from beam-beam interaction and IBS, and, in
addition, reduce the emittances during the store. For ex-
ample, using a 100-mA 54-MeV 5-MW bunched electron
beam accelerated by an energy-recovery linac is projected
to yield about a factor 10 increase of average luminosity for
RHIC-II [6]. Bunched beam stochastic cooling or optical
stochastic cooling are complementary techniques, that may
prove essential for controlling beam tails. Synchrotron ra-
diation will be helpful at the LHC, the first hadron collider
where radiation damping exceeds the IBS growth rate.

Several design studies for ring-linac colliders were car-
ried out in recent years, e.g., for ERHIC [6], and ELIC. At
high energy, a 7-TeV proton superbunch of the LHC could
be collided with a CLIC bunch train of, e.g., 75 GeV [42].

I thank K. Akai, R. Assmann, O. Bruning, A. Burov,
S. Fartoukh, A. Faus-Golfe, W. Fischer, J. Jowett, J.-P.
Koutchouk, K. Ohmi, Y. Papaphilippou, and F. Ruggiero
for helpful discussions.

REFERENCES

[1] V. Shiltsev, these proceedings; also HALO’03 Montauk.
[29] Y. Papaphilippou et al., LHC-Project-Note-244 (2000).
[30] F. Zimmermann et al., CERN-LHC-Project-Note 223 (2000);
[35] A. Krasnov, LHC Project Report 671 (2003); G. Stupakov,
M. Pivi, ECLoud’04 Napa (2004).
[48] D. Schulte, F. Zimmermann, these proceedings.