Chapter 4

The HL-LHC Accelerator Physics Challenges

S. Fartoukh and F. Zimmermann

CERN, BE Department, Genève 23, CH-1211, Switzerland

The conceptual baseline of the HL-LHC project is reviewed, putting into perspective the main beam physics challenges of this new collider in comparison with the existing LHC, and the series of solutions and possible mitigation measures presently envisaged.

1. Introduction and General Description

The HL-LHC is being designed to deliver an integrated luminosity of at least 250 fb$^{-1}$/year in each of the two LHC General-Purpose Experiments ATLAS and CMS for proton operation [1], while operating the other two experiments, ALICE and LHCb, at very low and moderate instantaneous luminosity, of about $10^{31}$ and $1 \sim 2 \times 10^{33}$ cm$^{-2}$s$^{-1}$, respectively [2, 3]. The ambitious performance target for ATLAS and CMS cannot be met without pushing both the optics to their extreme, namely $\beta^*$, and the nominal parameters of the LHC beam. It relies as well on a number of very challenging new equipment and key innovative technologies, such as:

- new larger aperture super-conducting magnets in order to preserve the transverse acceptance of the two high-luminosity insertions at low $\beta^*$, and in particular new inner triplet quadrupoles with a 150 mm coil aperture, more than doubled with respect to the existing Nb-Ti triplet, but still operating at a gradient of 140 T/m (about 30% below the present value) thanks to the Nb$_3$Sn technology,
- crab cavities, which are high-frequency RF transverse deflectors and aim at preserving the luminosity gain with $1/\beta^*$ by ensuring head-on collisions at the interaction point (IP), despite the crossing angle which is needed to separate the two beams after the collision.

Last but not least, the maximum possible peak luminosity is in practice limited by several factors, in particular the number of pile up events per bunch crossing which can rapidly degrade the quality of the data collected for the physics analysis. In this respect, the HL-LHC relies on a constant instantaneous luminosity,
not exceeding $5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, and corresponding to approximately 140 events in average per bunch crossing for operation with 25 ns bunch spacing (and more precisely about 2,750 bunches per beam, see Table 1). This is achieved through challenging luminosity leveling techniques, for instance via a gradual reduction of $\beta^*$ compensating for the proton burn off during the physics coast. In order to sustain such a high luminosity, over typically 10 hours of stable beam, the beam parameters, in particular the total beam current, shall correspond to a so-called virtual luminosity 4 to 5 times higher than the actual (leveled) luminosity. The virtual luminosity would be attained if all the other parameters, for instance $\beta^*$, were pushed to their respective limits at the very beginning of a physics fill. The baseline HL-LHC parameters (25 ns version) are listed in Table 1 (see Chapter 3 for more detail), including some key quantities such as the virtual luminosity introduced above. The leveling time of 9 h has been calculated assuming no emittance growth and a total hadron cross-section of 100 mb. The numbers in parentheses in the right column of Table 1 correspond to an ultimate $\beta^*$ of 10 cm, for which most of the new equipment would be pushed to the limits.

Table 1. Baseline parameters of the HL-LHC (25 ns version) and comparison with the nominal LHC. The numbers in parentheses refer to an ultimate $\beta^*$ of 10 cm.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Nominal LHC</th>
<th>HL-LHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy [TeV]</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Bunch spacing [ns]</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Number of bunches</td>
<td>2,736 (up to 2,808)</td>
<td>2,736 (up to 2,808)</td>
</tr>
<tr>
<td>Bunch charge [$10^{11}$]</td>
<td>1.15</td>
<td>2.2</td>
</tr>
<tr>
<td>Total current [A]</td>
<td>0.58</td>
<td>1.11</td>
</tr>
<tr>
<td>Bunch length [cm]</td>
<td>7.50</td>
<td>7.50</td>
</tr>
<tr>
<td>Energy spread [$10^{-4}$]</td>
<td>1.20</td>
<td>1.20</td>
</tr>
<tr>
<td>Long. emittance [eVs]</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td>$\beta^*$ [cm]</td>
<td>55</td>
<td>15 (10)</td>
</tr>
<tr>
<td>Full crossing angle [$\mu$rad]</td>
<td>300</td>
<td>590 (720)</td>
</tr>
<tr>
<td>Beam separation [$\sigma$]</td>
<td>9.9</td>
<td>12.5</td>
</tr>
<tr>
<td>Normalized transverse emittance [$\mu$m]</td>
<td>3.75</td>
<td>2.5</td>
</tr>
<tr>
<td>Peak luminosity (peak w/o crab cavity) [$10^{34}$ cm$^{-2}$s$^{-1}$]</td>
<td>1.0</td>
<td>7.4 (7.7)</td>
</tr>
<tr>
<td>Virtual luminosity (peak with crab cavity) [$10^{34}$ cm$^{-2}$s$^{-1}$]</td>
<td>1.2</td>
<td>21.9 (30.1)</td>
</tr>
<tr>
<td>Leveled luminosity [$10^{34}$ cm$^{-2}$s$^{-1}$]</td>
<td>NA</td>
<td>5.0</td>
</tr>
<tr>
<td>Leveling time [h]</td>
<td>NA</td>
<td>9.0 (10.2)</td>
</tr>
<tr>
<td>Pile up event/crossing</td>
<td>$\leq 28$</td>
<td>135–138</td>
</tr>
</tbody>
</table>
The matchability of such low $\beta^*$ optics within the specific layout constraints of the LHC, while correcting a series of side effects of chromatic nature (in first instance the linear chromaticity, but also the non-linear chromaticity, the spurious dispersion, and off-momentum $\beta$-beating), has been a long-standing problem which was recently solved by the development of a novel optics concept, the Achromatic Telescopic Squeezing (ATS) scheme [4–6] (see Section 1.1). As targeted in Table 1, doubling the beam current and tripling the beam brightness with respect to the nominal parameters represents a challenge in terms of collective effects (instabilities, intra-beam scattering, beam–beam effects, and electron cloud), which will be reviewed in Section 1.2. Finally, the pros and cons of the various possible options for luminosity leveling will be addressed in Section 1.3, while the last section of the chapter will present a summary and conclusions.

1.1. Optics

1.1.1. Optics constraints and challenges

Reducing the beam sizes at the interaction point, that is acting on $\beta^*$ at constant transverse emittances, is a key ingredient to boost the performance of any collider. Going in this direction, a series of limitations shall however be overcome, driven by the mechanical aperture available in the inner triplet (IT), and also coming from the rest of the ring. Concerning the demand on the mechanical acceptance of the inner triplet, one can always find a solution based on sufficiently large aperture quadrupoles, by weakening their gradient but making them longer (at more or less constant integrated gradient), regardless of $\beta^*$ and of the technology chosen for the triplet [7–9]. Indeed decreasing the operational gradient $G$ of the inner triplet at constant $\beta^*$ and integrated strength, the aperture needed for the beam roughly scales like $1/G^{1/4}$ (or said differently the peak $\beta$-function $\beta_{\text{max}}$ reached in the IT is found to increase with $1/\sqrt{G}$), while the coil aperture can in principle be increased with $1/G$ at constant peak field, that is much faster.

The real optics challenge for low $\beta^*$ is actually elsewhere, i.e. on the “non-triplet” side of the machine [10], where a series of limitations were clearly identified and classified in the framework of the previous upgrade project of the LHC, the so-called Phase I Luminosity Upgrade project [11]. While of very different nature, these limitations can be quantified by the maximum possible peak $\beta$-function which is permitted in the inner triplet, namely $\beta_{\text{max}}$. Indeed, this $\beta_{\text{max}}$ shall then be matched to the regular optics of the arcs within the fixed distance given by the length of the low-$\beta$ insertion, also known as interaction region (IR), and within the aperture and gradient limits of the IR magnets (quadrupoles of the matching section and of the dispersion suppressor). Finally, a clear strategy shall be estab-
lished to ensure a proper control of the chromatic aberrations induced, without exceeding the available strength of the lattice sextupoles. The beam observables to be corrected are not only the linear chromaticity $Q'$, which is increasing linearly with $\beta_{\text{max}}$, but also the non-linear-chromaticities $Q''$, $Q''', \ldots$, the off-momentum $\beta$-beating $\partial \beta / \partial \delta$ (i.e. the chromatic variations of the $\beta$-functions), and the spurious dispersion induced by the crossing angle in the low-$\beta$ insertions. Assuming an upgrade which would essentially only rely on the replacement of the existing inner triplet (see e.g. the former upgrade project of the LHC [11]), and no deep conceptual changes in general beam optics for circular colliders, these limitations can rapidly turn into hard limits driven by the existing hardware in large parts of the ring, and given by:

- the mechanical acceptance of the existing matching section,
- the gradient limits of the matching quadrupoles, and
- the strength limits of the arc sextupoles.

As a result, the minimum possible $\beta^*$ was found to be around 30 cm assuming a Nb-Ti triplet of 120 mm aperture operating at 120 T/m, and only slightly less ($\beta^* \sim 25$ cm), for a triplet of the same aperture but operating with a 50% stronger gradient thanks to the Nb$_3$Sn technology [10]. These minimum values of $\beta^*$ were also considered as hard limits at that time, leaving no operational margin, e.g. for fine tuning the optics, the tune, the chromaticity, etc.

1.1.2. The Achromatic Telescopic Squeezing (ATS) scheme as baseline for the HL-LHC optics

Concerning the first limitation previously mentioned, the only solution is to equip the LHC matching sections with new two-in-one magnets of larger aperture. Concerning the poor optics flexibility observed at low $\beta^*$ in the experimental insertions IR1 and IR5, with some quadrupoles being pushed to very low or very high gradients in the matching section and dispersion suppressor, respectively, one possibility is to allow floating matching conditions at the boundaries of these two insertions. More precisely the idea is to maintain the dispersion matching constraints at the entry and exit of the low-$\beta$ insertions (from Q13.L to Q13.R), but to allow the “auxiliary” insertions on either side (IR8/2 for IR1 and IR4/6 for IR5) to contribute as well to the matching of the $\beta$-functions, at least below a certain value of $\beta^*$. As a result, $\beta$-beating waves are generated in the sectors adjacent to the low-$\beta$ insertions (sectors 45 and 56 for IR5 and sectors 81 and 12 for IR1). Assuming a phase advance per arc cell strictly matched to $\pi/2$ in these sectors, and if correctly phased with respect to the IP, these waves will reach their maximum at every other sextupole, i.e. at the sextupoles belonging to the same electrical circuit (see Fig. 1).
Fig. 1. LHC sextupole powering scheme (a) and zoom of the optics in sector 45: First a pre-squeezed optics (b) with its phasing properties and a typical $\beta^*$ value of 40 cm, followed by two possible telescopic optics further reducing $\beta^*$ in a symmetric (“round” optics in (c) with $\beta^*_x = \beta^*_y$) or asymmetric way (“flat” optics in (d) with $\beta^*_x \neq \beta^*_y$). The dispersion function remains matched in the arcs for the telescopic optics, but the $\beta$-functions are mismatched reaching their maximum at every other sextupole, i.e. at sextupoles belonging to the same electrical circuit. The relative increase of these maxima with respect to the pre-squeezed optics is inversely proportional to the additional reduction of $\beta^*$. As a result, during the telescopic squeeze the chromatic correction of the inner triplet can be achieved at nearly constant sextupole strength.

Consequently, the chromatic correction efficiency of these sextupoles will drastically increase at constant strength which, de facto, will be a definite cure for the third limitation previously mentioned.

This novel approach is particularly well-suited to the LHC for the two following reasons. First, due to the large dynamic range of machine energy, from 450 GeV to 7 TeV, and the reduction in proportion of the transverse emittances during the ramp, the peak $\beta$-functions in the arcs could in principle be increased by a factor of about 16 at top energy without exceeding any aperture-related limits (in practice a
S. Fartoukh and F. Zimmermann

bit less since it is advisable to increase the margins at higher energy). Then, at flat top energy, the quadrupole magnets of the so-called “auxiliary” insertions are either moderately pushed, which is the case for the experimental insertions IR8 and IR2 assuming a $\beta^*$ of a few meters in p-p collision mode, or not pushed at all, in the case of IR4 and IR6 for which the injection optics is kept unchanged during the whole LHC cycle. Therefore all the ingredients are already available in the existing machine to blow up the $\beta$-functions in the arcs 81/12/45/56 at 7 TeV, and to implement the principle of the ATS scheme.

A comprehensive description of the scheme can be found in Ref. [6], in particular concerning the constraints imposed on the betatron phases over the left and right side of the low-$\beta$ insertions, and the optics squeeze which is achieved in a two-stage telescopic mode:

(i) first of all a so-called pre-squeeze, which is a more or less standard squeeze, with additional matching constraints, but acting only on the matching quadrupoles of the low-$\beta$ insertions proper and on the arc sextupoles, till reaching some strength limitations, either for the IR magnets or for the chromaticity sextupole in the arcs,

(ii) then the telescopic squeeze, using only the matching quadrupoles of the neighbouring insertions (IR2/8 for squeezing IR1, and IR4/6 for squeezing IR5), and at constant strength for the chromaticity sextupoles of the arcs.

With the exception of the Q5 matching quadrupoles of IR6 (which would need to be made 25% longer for 7 TeV operation), and heavier interventions obviously needed in the matching sections and inner triplets of IR1 and IR5 (see later), the ATS scheme has been found to be fully compatible with the existing LHC hardware and layout, in order to produce and ensure the chromatic correction of the collision optics with extremely small $\beta^*$, down to 5–10 cm. ATS optics have been built in practice assuming several possible triplet layouts, e.g. the (120 T/m, 120 mm full aperture) Nb-Ti triplet proposed for the former Phase I LHC upgrade project, two other intermediate triplet layouts with an aperture increased up to 140 mm and compatible with the Nb-Ti or Nb$_3$Sn technology [12] (i.e. with an operating gradient of 100 T/m or 150 T/m, respectively), and more recently with the latest (140 T/m, 150 mm) baseline triplet of the HL-LHC [13]. The main difference between these different cases is the so-called pre-squeezed $\beta^*$, ranging from 50 to 40 cm (again with an approximate scaling like $1/\sqrt{G}$), and below which the telescopic techniques of the ATS need to be deployed to reduce $\beta^*$ further down at constant strength for the chromaticity sextupoles of the lattice. A completely new version of the LHC optics based on the ATS scheme was also developed, strictly compatible with the existing layout of the LHC, in particular with its existing (200 T/m–70 mm) Nb-Ti inner triplets. This allowed testing and successfully
validating, with beam, the basic principles of this novel optics scheme via a series of dedicated machine studies which took place in 2011 at 3.5 TeV/beam [14], and ended up in 2012 with 4 TeV/beam where a $\beta^*$ of nearly 10 cm was reached and measured at IP1 and IP5 [15]. These machine studies have of course been achieved in very specific conditions, which are not suitable for nominal operation, in particular without crossing angle in order to preserve the mechanical aperture of the existing IR magnets at such low $\beta^*$.

![Fig. 2. Typical ATS collision optics for the HL-LHC with $\beta^* = 10$ cm at IP1 and IP5 (assuming a machine with the crossing scheme switched off in the four experimental insertions of the machine). While the horizontal dispersion function remains perfectly matched in the eight sectors of the ring, $\beta$-beating waves are clearly visible in the four sectors on either side of the two low-$\beta$ insertions ATLAS and CMS, which is one of the main signatures of the ATS scheme.](image_url)

A typical ATS optics, with $\beta^*$ pushed to its ultimate value of 10 cm at IP1 and IP5, is shown in Fig. 2 for the latest 150 mm aperture Nb$_3$Sn triplet of the HL-LHC (140 T/m). The peak $\beta$-function $\beta_{\text{max}}$ reached in the inner triplet is impressive, about 31 km. The minimum possible pre-squeezed $\beta^*$ is found to be 44 cm in this case (limited by the strength of the arc sextupoles). This means that a mismatch of the $\beta$-functions by 440% needs to be generated in the arcs 81, 12, 45 and 56 (i.e. with peak $\beta$-functions increased by a factor 4.4) in order to build up an ATS collision optics with $\beta^* = 10$ cm. These $\beta$-beating waves are clearly visible in Fig. 2.

A zoomed view of this collision optics is presented in Figs. 3(a)–(c), for the four experimental insertions of the HL-LHC, namely
Fig. 3. Zoomed view of collision optics in IR8, IR1/5 and IR2 running the LHC in ATS-mode for proton-proton physics (left pictures), and in non-ATS mode for ion or proton-ion physics (right picture). In ATS-mode β-beating waves are initiated on the right side of IR8 and absorbed on the left side of IR2 (see Figs. (a) and (c)) in order to gain a factor of 4 to 5 in the β∗ reduction at IP1.
• LHCb (IR8) squeezed to an intermediate $\beta^*$ of 3 m which is more than enough for sustaining over about 10 h a luminosity leveled to $1-2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ assuming the HL-LHC beam parameters given in Table 1.

• Alice (IR2) with $\beta^* = 10$ m where halo collisions (with a large beam–beam separation at the IP of more than $5\sigma$) might be requested in p-p mode, without exceeding an instantaneous luminosity of $10^{31}$ cm$^{-2}$s$^{-1}$.

• ATLAS and CMS (IR1 and IR5) with $\beta^* = 10$ cm, where a factor of about 4 in $\beta^*$ reduction is actually supported by the $\beta$-beating waves generated and absorbed by the matching quadrupoles located on the right of IP8 (resp. IP4) and on the left of IP2 (resp. IP6) for ATLAS (resp. CMS).

Starting from the same injection optics as the one used to build up the 10 cm ATS optics described above, the nominal functionality of the LHC experimental insertions is still preserved: in particular IR2 and IR8 can still be squeezed in the usual way down to their nominal $\beta^*$, or even slightly below (namely one can achieve $\beta^* = 44$ cm at IP1 and IP5, together with $\beta^* = 50$ cm at IP2 and IP8), assuming ion or proton–ion physics during the HL-LHC runs (see Figs. 3(d)–(f)).

The chromatic properties of ATS optics are particularly interesting to analyze. Pushing $\beta^*$ down to 10 cm, the chromatic variations of the betatron tunes are only moderately perturbed by a second order chromaticity $Q''$ showing up over a momentum range of $\delta_p = \pm 10^{-3}$ corresponding to about three times the momentum acceptance of the LHC RF bucket at 7 TeV (see Fig. 4(a)). The chromatic Monetique functions (giving the amplitude of the first order chromatic derivative of the $\beta$-functions) are nicely vanishing in the collimation insertions IR3 and IR7 and at IP1 and IP5 (see Fig. 4(b)), therefore with no impact on the collimation hierarchy, nor on the machine performance. Another important feature lies in the fact that the off-momentum $\beta$-beating waves induced in the two planes by the lattice sextupoles are exactly out of phase by $\pi/2$ with respect to the $\beta$-functions themselves, in particular in the triplet and its neighboring magnets. Therefore, no further degradation of the off-momentum mechanical aperture is induced in the arcs, the matching section and the new triplet, except the usual one coming from the contribution of the dispersion, which remains perfectly matched in the ATS scheme. Finally, an extremely important quantity to control is the spurious dispersion induced by the crossing scheme in IR1 and IR5. This dispersion can reach up to 20 m in the new triplets when pushing $\beta^*$ down to 10 cm with a full crossing angle of 720 $\mu$rad (see Table 1). The spurious dispersion is produced by feed-down effects in the inner triplets of one of the two high luminosity insertions, and then exported to the other one. However, thanks to the specific phasing conditions imposed by the ATS scheme, modest H or V orbit bumps of the order of 4–5 mm generated in the sectors adjacent to IR1 and IR5 are found to be sufficient to correct it back to a level of $\sim 0.5–1$ m in the inner triplet (see Figs. 4(c) and (d)).
Fig. 4. Chromatic properties of the 10 cm ATS optics: (a) Chromatic variations of the betatron tunes (assuming the linear chromaticity to be matched to 2 units using standard $Q'$ knobs), (b) Montague functions $W_x, W_y$ around the LHC ring with IP3 chosen as the origin, and (c) residual horizontal and vertical dispersion mismatch induced by the horizontal and vertical crossing angles at IP1 and IP5, but minimized thanks to orbit bumps generated in the arcs 81, 12, 45 and 56 (d). Without this correction, the horizontal and vertical dispersion would be fully mismatched around the ring, in particular in the collimation insertions IR3 and IR7, the RF insertions IR4 and the dump insertion IR6, and may reach up to 20 m in the inner triplet at $\beta^* = 10$ cm, depending on the betatron phase advances between IP1 and IP5.

1.1.3. Crossing angle and crab cavities

In order to separate the two beams after the collision, a crossing angle $\Theta_c$ is imposed at the interaction point. When reducing $\beta^*$, this angle shall be increased in order to guarantee a sufficiently large normalized beam–beam separation $d_{bb}$, typically of the order of $10\sigma$ for the nominal LHC, but which shall be increased up to $12.5\sigma$ for the HL-LHC, with longer triplets and therefore a lengthening of the
region where the two beams share the same vacuum beam pipe and continue to interact with each other:

\[ \Theta_c = d_{bb} \times \sqrt{\frac{\epsilon}{\beta^*}} = \frac{d_{bb} \sigma^*}{\beta^*}, \]  

(1)

where \( \epsilon \) and \( \sigma^* = \sqrt{\epsilon \beta^*} \) denote the 1\( \sigma \) physical beam emittance and the r.m.s. spot size at the IP, respectively. Increasing the crossing angle, however, affects directly the so-called Piwinski angle \( \phi_w \) which characterizes the overlap of the two colliding beam distributions, and therefore the luminosity in the presence of a non-zero crossing angle:

\[ \phi_w \equiv \frac{\Theta_c \sigma_z}{2 \sigma^*} = \frac{\beta^*_w}{\beta^*}, \]  

(2)

where \( \sigma_z \) is the r.m.s. bunch length and \( \beta^*_w \) is a characteristic \( \beta^* \) defined by

\[ \beta^*_w \equiv \frac{d_{bb} \sigma_z}{2}. \]  

(3)

The potential gain of luminosity with \( 1/\beta^* \) saturates rapidly below this characteristic \( \beta^* \), of the order of \( \beta^*_w \sim 40\text{–}50 \text{ cm} \) for typical (HL)-LHC parameters:

\[ \mathcal{L}(\beta^*) \propto \frac{1}{\beta^* \sqrt{1 + \phi^2}} = \frac{1}{\beta^*_w \sqrt{1 + (\beta^*_w/\beta^*)^2}} \frac{\beta^* \epsilon \beta^*_w}{\beta^*_w} \frac{1}{\beta^*_w}. \]  

(4)

A possible mitigation measure consists of using flat optics, with a \( \beta^* \) as small as possible in the parallel separation plane, if possible down to \( \beta^*_w \approx 15 \text{ cm} \), and of the order of \( \beta^*_S \approx 5\text{–}4 \text{ km} \) for typical (HL)-LHC parameters:

\[ \mathcal{L}(\beta^*) \propto \frac{1}{\beta^* \sqrt{1 + \phi^2}} = \frac{1}{\beta^*_w \sqrt{1 + (\beta^*_w/\beta^*)^2}} \frac{\beta^* \epsilon \beta^*_w}{\beta^*_w} \frac{1}{\beta^*_w}. \]  

(4)

The so-called compact crab cavities [16], running at the main RF frequency of \( \omega_{CC}/(2\pi) = 400 \text{ MHz} \), are planned to be installed in between the recombination dipole D2 and the first quadrupole Q4 of the matching section [17], at about \( \pm 150 \text{ m} \) on either side of the IP where the beam separation has reached its nominal value of 194 mm. The \( \beta \) functions are still relatively high at this location, of the order of \( \beta^*_{CC} \sim 3.5\text{–}4 \text{ km} \) for \( \beta^* = 15 \text{ cm} \), and the phase advance with respect to the IP is still very close to 90°. This in turn reduces the required voltage of the RF deflecting kick, but also and mainly ensures the quasi-closure of the RF bumps on
Fig. 5. Bunch rotation induced by crab cavities, in order to maximize the overlap of the two beam distributions at the interaction point, and therefore the luminosity, in the presence of a non-zero crossing angle.

either side of the IP. The total RF deflecting voltage required per beam and per IP side, $V_{CC}$, is given by

$$\frac{V_{CC}}{E} = \frac{c}{\Theta_{cc}} \times \frac{\Theta_{c}}{2 \sqrt{\beta_{cc} \beta^*}},$$

(5)

where $E$ denotes the beam energy [eV] and $c$ is the speed of light. As an example, $V_{CC} \sim 12.5$ MV is needed in order to restore head-on collisions in the presence of a crossing angle set to $\Theta_{c} = 590 \mu$rad at $\beta^* = 15$ cm (see Table 1).

In order to minimize the beam loading in the crab cavities, a strict control of the closed orbit is also demanded at the position of the crab cavities. At this location, however, the crossing and parallel separation bumps are not yet closed in the LHC, inducing in particular a non-zero closed orbit which can substantially vary depending on the operation mode of the machine (injection, ramp, squeeze, Vernier scans,...). It was however found that the closure of these bumps upstream of the crab cavities could actually be achieved by means of new strong orbit correctors which would be installed on the non-IP side of D2, and by reinforcing the triplet orbit corrector MCBX located on the non-IP side of Q3 (see Fig. 6 and Ref. [17] for more details).

1.1.4. Layout and mechanical aperture

After having introduced the two main ingredients, namely the ATS optics for producing and properly controlling the chromatic aberrations of very low $\beta^*$ collision optics, and the crab cavities to maximize the efficiency of a reduced $\beta^*$ in terms of luminosity performance, the need and the basic parameters (in particular the mechanical aperture) of new HL-LHC magnets and various equipments can be defined. First of all, it is clear that for the target $\beta^*$ values ranging from 15 to 10 cm (see Table 1), that is a factor of 4 to 5 below the nominal collision $\beta^*$ of the LHC, the peak $\beta$-functions reached in the inner triplet increase in proportion. Since the beam sizes go with the square root of the $\beta$-functions, the aperture of the
Fig. 6. Nominal (left) and new (right) crossing (red) and parallel separation (black) bumps. In both cases, the recombination dipole D2 has been displaced towards the IP in order to accommodate crab cavities on the IP-side of Q4. In the first case, these bumps are closed at Q6, as in the existing LHC. In the second case, the bumps can be closed before the crab cavities using the triplet orbit correctors (in particular the one on the non-IP side of Q3, see Section 1.1.4), but also additional orbit correctors installed on the non-IP side of D2 [17] (Courtesy of Riccardo de Maria).

new triplet shall be more or less doubled (with respect to the 70 mm coil aperture of the existing low-β quadrupoles), similarly for the TAS (which is an absorber in front of the first low-β-quadrupole Q1), and the separation dipole D1. Detailed aperture calculations demonstrated as well the need to replace the recombination dipole D2 and the first two quadrupoles of the matching section, namely Q4 and Q5, with magnets of larger aperture [4, 5, 12, 13]. In this respect, the coil aperture of the new triplet was fixed to 150 mm, which is compatible with an operating gradient of 140 T/m for the Nb₃Sn technology (see Chapter 6). While in the existing machine the separation dipole D1 is a normal conducting dipole made of six consecutive modules in IR1 and IR5, the new D1 will be superconducting in the high-luminosity insertions, with an aperture of 150 mm as for the new triplets. The aperture of the recombination dipole D2 will also be increased, from 80 mm to 105 mm, that of Q4 from 70 mm to 90 mm, and that of Q5 from 56 mm to 70 mm (see Table 2).

A sketch of the layout of the new triplet up to D1 is shown in Fig. 7, with a zoomed view of the triplet corrector package on the non-IP side of Q3. The main features of this layout, and in particular its modifications and similarities with respect to the existing one, are emphasized below:

- The triplet remains symmetric, that is with the same magnetic length for Q1 and Q3 ($L_{Q1} \equiv L_{Q3} \approx 8$ m, to be compared with 6.4 m for the existing magnets). Q1 and Q3 are however split into two 4-m-long magnets, for technical reasons. Also, contrary to the existing triplet layout, Q2a and Q2b are hosted in two sep-
Table 2. Aperture and performance specifications for the main HL-LHC equipments foreseen to be installed in IR1 and IR5. The numbers in parentheses refer to the nominal LHC.

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Coil aperture [mm] (gap for the existing D1)</th>
<th>Aperture separation [mm]</th>
<th>Performance for 2-in-1 equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAS</td>
<td>60 (34)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Triplet (Q1-2-3)</td>
<td>150 (70)</td>
<td>-</td>
<td>140 T/m (205 T/m)</td>
</tr>
<tr>
<td>D1</td>
<td>150 (63)</td>
<td>70 (63)</td>
<td>35 T · m (27 T · m)</td>
</tr>
<tr>
<td>D2</td>
<td>105 (80)</td>
<td>186 (188)</td>
<td>35 T · m (35 T · m)</td>
</tr>
<tr>
<td>Crab cavity</td>
<td>84</td>
<td>194</td>
<td>12.5 MV/beam/IP side</td>
</tr>
<tr>
<td>Q4</td>
<td>90 (70)</td>
<td>194 (194)</td>
<td>550 T/m × m (550 T/m × m)</td>
</tr>
<tr>
<td>Q5</td>
<td>70 (56)</td>
<td>194 (194)</td>
<td>770 T/m × m (770 T/m × m)</td>
</tr>
</tbody>
</table>

Fig. 7. Sketch of the layout of the new high-luminosity insertions from the IP up to D1 (top picture, courtesy of Ezio Todesco) and zoom into the triplet corrector package installed on the non-IP side of Q3 (bottom picture, courtesy of R. de Maria).

Arate cryostats and have a magnetic length $L_{Q2a} \equiv L_{Q2b} \approx 6.8$ m, to be compared with 5.5 m for the existing Q2’s.

- Nested horizontal and vertical orbit correctors, so-called MCBX, are installed on the non-IP side of Q2b and IP-side of Q2a, with an integrated strength which is rather similar to the one available in the existing MCBX magnets, but of course with a coil aperture of 150 mm. The integrated strength (or more precisely the length) of the MCBX installed in a corrector package on the non-IP side of Q3
The HL-LHC Accelerator Physics Challenges

(see below) is however almost doubled in order to cope with the requirement of the new crossing scheme. The main specifications of these orbit correctors are given in Table 3, together with the new orbit correctors which are needed on the non-IP side of D2 in order to close the crossing and separation bumps before the crab cavities.

- A series of multipole correction coils, for all orders from $n = 1$ to $n = 6$, normal and skew (except $b_2$), are hosted in a separated cryostat, the so-called triplet corrector package (CP) installed on the non-IP side of Q3. Compared to the nominal LHC, only the normal and skew decapole (namely $a_5$ and $b_5$), and the skew dodecapole (namely $a_6$) magnets are new types of triplet correctors. They were indeed found to be fundamental in order to preserve the dynamic aperture of the machine in the presence of field imperfections in the new triplet and D1 (see Section 1.1.5 and specification in Table 4).
- Finally, as already mentioned, the separation dipole D1 will be superconducting in the HL-LHC, with a coil aperture of 150 mm, a length of 6.7 m, and an integrated strength of about 35 T·m.

### Table 3. Performance specification for the new triplet and D2 orbit correctors [13].

<table>
<thead>
<tr>
<th>Corrector</th>
<th>Location</th>
<th>Aperture type</th>
<th>Integrated strength [T·m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCBX2a</td>
<td>IP-side of Q2a</td>
<td>Triplet</td>
<td>2.5 (in both planes)</td>
</tr>
<tr>
<td>MCBX2b</td>
<td>Non-IP side of Q2b</td>
<td>Triplet</td>
<td>2.5 (in both planes)</td>
</tr>
<tr>
<td>MCBX3</td>
<td>Non-IP side of Q3</td>
<td>Triplet</td>
<td>4.5 (in the crossing plane)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.5 (in the parallel sep. plane)</td>
</tr>
<tr>
<td>MCBRD</td>
<td>Non-IP side of D2</td>
<td>D2</td>
<td>7.0 (in the crossing plane)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.0 (in the parallel sep. plane)</td>
</tr>
</tbody>
</table>

### Table 4. Performance specification for the triplet multipole correctors [18].

<table>
<thead>
<tr>
<th>Corrector</th>
<th>Multipole</th>
<th>Aperture type</th>
<th>Integrated strength [mT·m] at 50 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>MQSX</td>
<td>$a_2$</td>
<td>Triplet</td>
<td>1000</td>
</tr>
<tr>
<td>MCSSX</td>
<td>$a_3$</td>
<td>Triplet</td>
<td>63</td>
</tr>
<tr>
<td>MCSX</td>
<td>$b_3$</td>
<td>Triplet</td>
<td>63</td>
</tr>
<tr>
<td>MCOX</td>
<td>$b_5$</td>
<td>Triplet</td>
<td>46</td>
</tr>
<tr>
<td>MCDX</td>
<td>$b_5$</td>
<td>Triplet</td>
<td>25</td>
</tr>
<tr>
<td>MCTSX</td>
<td>$a_6$</td>
<td>Triplet</td>
<td>25</td>
</tr>
<tr>
<td>MCTX</td>
<td>$b_6$</td>
<td>Triplet</td>
<td>86</td>
</tr>
</tbody>
</table>
Concerning the layout of the matching section, D2, Q4 and Q5 are planned to be replaced by magnets of larger aperture. D2 will in addition be shifted by about 15 m towards the IP in order to liberate sufficient room for the integration of the crab cavities, and Q5 displaced by about 10 m towards the arcs, which was found to be optimal in terms of optics flexibility with the new longer HL-LHC triplet. On the other hand, from a strictly qualitative point of view, the conceptual layout of the matching section does not need to be further modified for the HL-LHC, thanks to the telescopic squeeze possibilities offered by the ATS.

The normalized aperture of the new high luminosity insertions (from Q7.L to Q7.R) is shown in Fig. 8, including some mechanical tolerances and a tolerance budget for the beam based on the LHC operational experience so far (i.e. 2 mm closed orbit and 10% β-beating). A minimum of 12σ is reached in the inner triplet at β∗ = 15 cm, which is considered to be sufficient for machine protection and collimation (see Chapters 12 and 13).
1.1.5. Dynamic aperture

In the presence of non-linear field imperfections, both in the existing LHC magnets and in the new ones foreseen for the HL-LHC, the preservation of the dynamic aperture (DA), that is the region of the transverse phase space where the particle motion remains stable after a very large number of turns, is a fundamental input driving both the design of the optics, of the layout (in terms of need for correctors) and, of course, of the magnets themselves. The target dynamic aperture has been fixed to $10.5\sigma$ in collision for the baseline $\beta^*$ of 15 cm, with a scaling like $\beta^*$ for smaller $\beta^*$ [19], i.e. reaching the $7\sigma$ extension of the secondary halo (see Chapter 13) for the ultimate $\beta^*$ of 10 cm. More precisely, although the new HL-LHC magnets do not seem to degrade the DA for the injection optics at 450 GeV [20], with still quite relaxed assumptions concerning the field quality which is expected for these new magnets, the situation is much more critical in collision when the optics is squeezed. This fact is both related to the features of the collision optics, with a net increase of the $\beta$-functions in the arcs imposed by the ATS scheme, and of course the very large beam sizes reached in the inner triplet and downstream magnets at very small $\beta^*$.

Concerning the first item, a potential strong limitation of the ATS scheme is indeed its detrimental impact on the DA, due to the net increase of the peak $\beta$-functions in the arcs during the telescopic squeeze, combined with the non-linear field imperfections of the main dipoles and quadrupoles, and with the strength imposed on some sextupole families for the chromatic correction. On the other hand, thanks to the tremendous efforts which were deployed for the LHC main magnets during the construction and installation phases of the machine, both in terms of field quality specifications [21] and monitoring [22, 23], and in terms of sorting strategy [24], the existing machine was found compatible with the ATS scheme for this aspect, although not designed for it. More precisely,

- within an increase of the peak $\beta$-functions in the arcs by a factor of 4 in both planes (corresponding to a $\beta^*$ of about 10 cm starting from an optics pre-squeezed to $\beta^* \approx 40$ cm), or up to 8 in one plane, but only 2 in the other plane (corresponding to a flat optics with at typical $\beta^*$ aspect ratio of 4, e.g. with $\beta^* \approx 20/5$ cm in the crossing/separation planes, alternated between IR1 and IR5, when starting from a 40 cm pre-squeezed optics [4]),
- considering only in a first step the field imperfections of the arc magnets as measured and installed in the existing LHC ring (i.e. assuming new IR magnets for the HL-LHC which are ideal from the field quality point of view), and
- ensuring that the sextupole families participating in the chromatic correction do contain an even number of magnets (for a two-by-two compensation at $\pi$ of their contribution to the third order resonance driving terms), the (HL)-LHC
dynamic aperture was found to be about 40 beam $\sigma$ for the pre-squeezed optics, but dropping to about 15$\sigma$ and 11$\sigma$ for the ATS round and flat optics, described above [17, 25]. This net degradation determines the limit of the ATS scheme in terms of $\beta^*$ reach. Fortunately, the $\beta^*$ value achievable lies within (and even beyond) the HL-LHC target range.

Concerning the impact of the new matching section quadrupoles Q4 and Q5, and separation-recombination dipoles D1 and D2, and based on first estimates of their expected field quality, the situation is still relatively comfortable for the baseline $\beta^*$ of 15 cm of the HL-LHC [26]. Preliminary designs of the new D2 magnet, however, seem to indicate that its field quality could be much worse than initially thought. Therefore, dedicated studies are ongoing to improve the D2 design, in order to decide whether or not this magnet needs to be equipped with dedicated spool-piece correctors. The latter would definitely solve the problem.

Implementing the triplet corrector package described in Section 1.1.4, and assuming the field quality presently expected for the inner triplet [27] (see Table 5), the target DA of 10.5$\sigma$ seems to be within reach for $\beta^* = 15$ cm in collision (see Fig. 9(a)). However, dedicated studies and iterations with the magnet builders are still ongoing in order to identify, and if necessary to act on, the most dangerous multipoles [28], so as to clearly meet the target DA at $\beta^* = 15$ cm, and to further improve the situation in order to gain margin in view of the ultimate $\beta^*$ of 10 cm (see Fig. 9(b)).

Table 5. Field imperfections expected at high current in the new 150 mm aperture inner triplet quadrupoles [27], expressed in terms of relative normal and skew multipole errors in units of $10^{-4}$ at a reference radius of 50 mm. A reduction (by at most 50%) is recommended for some multipoles [28] (see numbers in parenthesis). The superscripts $S$, $U$ and $R$ stands for systematic, uncertainty (maximum possible average deviation with respect to the expected systematic over the production), and random (giving the r.m.s fluctuation from magnets to magnets over the production).

<table>
<thead>
<tr>
<th>n</th>
<th>$a_n^S$</th>
<th>$a_n^U$</th>
<th>$b_n^S$</th>
<th>$b_n^U$</th>
<th>$b_n^R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.800</td>
<td>0.800</td>
<td>0</td>
<td>0.820</td>
<td>0.820</td>
</tr>
<tr>
<td>4</td>
<td>0.650</td>
<td>0.650</td>
<td>0</td>
<td>0.570</td>
<td>0.570</td>
</tr>
<tr>
<td>5</td>
<td>0.430</td>
<td>0.430</td>
<td>0</td>
<td>0.420</td>
<td>0.420</td>
</tr>
<tr>
<td>6</td>
<td>0.310</td>
<td>0.310</td>
<td>0.800</td>
<td>1.100</td>
<td>1.100</td>
</tr>
<tr>
<td>7</td>
<td>0.190 (0.152)</td>
<td>0.190 (0.095)</td>
<td>0</td>
<td>0.190 (0.095)</td>
<td>0.190 (0.095)</td>
</tr>
<tr>
<td>8</td>
<td>0.110 (0.088)</td>
<td>0.110 (0.055)</td>
<td>0</td>
<td>0.130 (0.065)</td>
<td>0.130 (0.065)</td>
</tr>
<tr>
<td>9</td>
<td>0.080 (0.064)</td>
<td>0.080 (0.040)</td>
<td>0</td>
<td>0.070 (0.035)</td>
<td>0.070 (0.035)</td>
</tr>
<tr>
<td>10</td>
<td>0.040</td>
<td>0.040 (0.032)</td>
<td>0.150 (0.075)</td>
<td>0.200 (0.100)</td>
<td>0.200 (0.100)</td>
</tr>
<tr>
<td>11</td>
<td>0.026</td>
<td>0.026 (0.021)</td>
<td>0</td>
<td>0.026 (0.021)</td>
<td>0.026 (0.021)</td>
</tr>
<tr>
<td>12</td>
<td>0.014</td>
<td>0.014</td>
<td>0</td>
<td>0.018 (0.014)</td>
<td>0.018 (0.014)</td>
</tr>
<tr>
<td>13</td>
<td>0.010</td>
<td>0.010</td>
<td>0</td>
<td>0.009 (0.007)</td>
<td>0.009 (0.007)</td>
</tr>
<tr>
<td>14</td>
<td>0.005</td>
<td>0.005</td>
<td>-0.040 (-0.020)</td>
<td>0.023 (0.012)</td>
<td>0.023 (0.012)</td>
</tr>
</tbody>
</table>
The HL-LHC Accelerator Physics Challenges

**Collective effects**

Collective beam effects arise from the electromagnetic interaction of the beam particles among themselves, with their environment (including electron cloud) and with the other beam. The two times higher than nominal beam current of the
HL-LHC (see Table 1) will enhance such interactions, so that collective effects may ultimately limit the performance of the HL-LHC. Depending on the charge per bunch, on the bunch length, and on the bunch filling pattern around the machine, electromagnetic interactions with the beam surroundings give rise to parasitic losses (overheating of sensitive components due to the excitation of trapped modes), can cause beam instabilities, or may degrade the beam quality by inducing emittance growth and poor lifetime of all or of some specific bunches. For the nominal LHC a systematic review of single beam collective effects was given in Ref. [29], and a comprehensive updated summary in Ref. [30].

The next section summarizes the current knowledge of the longitudinal and transverse impedance in the actual LHC, while Section 1.2.2 presents an impedance forecast for the HL-LHC, and Section 1.2.3 discusses some consequences of these impedances for Landau damping. Section 1.2.4 then reviews synchrotron radiation and electron cloud, Section 1.2.5 intrabeam scattering, and Section 1.2.6 Touschek scattering. Lastly, Section 1.2.7 addresses beam-beam effects at the HL-LHC.

1.2.1. Present LHC Impedance

The longitudinal impedance of the present LHC is known to be compatible with the LHC design specifications, at least the inductive part [31]. Measurements of the frequency shift of bunch-length oscillations with bunch intensity have yielded a limit consistent with the design impedance budget of \( |Z/n| \sim 0.1 \Omega \). Observations of the loss of Landau damping during acceleration (without controlled emittance blow up) have confirmed this estimate [32].

Concerning the transverse impedance of the LHC as built, the total transverse effective impedances can be estimated from the measured tune slopes and bunch parameters [33] to be between 4 and 12 M\( \Omega \)/m at injection (larger in the vertical plane than in the horizontal), and between 30 and 50 M\( \Omega \)/m at 4 TeV (here larger in the horizontal plane than in the vertical), assuming average \( \beta \) functions of \( \langle \beta_x \rangle \approx 66 \text{ m} \) and \( \langle \beta_y \rangle \approx 72 \text{ m} \) around the LHC ring [34].

The measurements at 4 TeV were performed with the so-called ‘tight collimator settings’ of 2012 (collimators in IR3 and 7 closed in towards the end of the ramp) prior to the squeeze [35]. Therefore, the collimator settings at the time of the 4-TeV measurement were almost exactly identical to those used during physics operation, the only significant difference being the settings of the special absorbers (TCLs) located downstream of the high luminosity IPs. The latter were retracted during these beam measurements, but this should not make any significant difference for impedance since, firstly, the TCLs are made from copper and, secondly, the \( \beta^* \) in IP1 and IP5 was 11 m.
Within the measurement accuracy ($\pm 2 \times 10^{-4}$ per $10^{11}$ protons for the tune variation with intensity), the horizontal and vertical betatron tunes were found insensitive to the rms bunch length at injection, when the latter was varied between 8.6 and 11.6 cm. The rms bunch lengths at 4 TeV were then adjusted in between 7.1 and 8.1 cm, and, also here, no noticeable change of tune was observed within the range of the bunch-length variations.

Both at injection and at 4 TeV beam energy the measured tune shifts — and hence also the associated effective impedances — are up to three times larger than expected from the LHC impedance model [36]. The LHC transverse broadband impedance is thought to be dominated by the collimation system.

1.2.2. HL-LHC impedance

**Longitudinal impedance**

The HL-LHC longitudinal impedance is composed of several significant contributions, the largest of which, over a wide frequency range, up to about 100 MHz, is the resistive wall of the arc beam screen, as is illustrated in Fig. 10.

Another important part of the machine impedance, both longitudinally and transversely, is high-Q resonances, normally related to higher-order modes (HOMs) in cavity-like objects. As for the LHC, such modes, either damped or un-damped, will be present in the HL-LHC, from the 400 MHz main RF cavities, the transverse damper system, and/or the vacuum chambers of the four main experiments. New for the HL-LHC will be an extensive crab cavity system, with up to 16 novel compact RF cavities per beam, with a total transverse RF voltage of about 20 MV (similar to the main RF system), and possibly as well additional 200 MHz (6 MV) or 800 MHz (8 MV) lower or higher harmonic RF cavities. Higher- or lower-order modes in these new cavities will significantly add to the set of higher-order modes in the LHC. Also the inductive impedance is expected to at least double due to these new equipments.

The large contribution of the crab cavities to the overall longitudinal impedance is evident from Fig. 10 (the orange portions). The longitudinal effective inductive impedance for 16 crab cavities is $(Z/n)_{el} \approx 0.04 \Omega$, which corresponds to an increase by about 50%, in the total longitudinal inductive impedance, compared with the present LHC, with possible implications for bunch lengthening and loss of longitudinal Landau damping.

**Transverse impedance**

Figure 11 compares the transverse impedance models for the LHC and the HL-LHC. As it is already the case for the LHC, the impedance of the HL-LHC will be dominated by the collimation system. The effective impedance is sensitive to
the normalized aperture of the secondary collimators, i.e. expressed in units of the transverse rms beam size, as well as to the collimator material and its possible coatings. The transverse impedance scales inversely with (slightly less than) the third power of the normalized aperture of the collimator jaws, and, therefore, the effective impedance, i.e. the impedance weighted by the local beta function, depends only weakly on the beta value at the collimator.

Figure 12 illustrates that over a wide frequency range the collimators contribute about 90% of the total transverse impedance in the HL-LHC.

Figure 13 (top picture) reveals that more than half of this transverse impedance is due to the secondary collimators ("TCSG"), which have fairly long jaws placed
Fig. 11. Comparison of the transverse impedance models for the LHC and HL-LHC as a function of frequency. The LHC and HL-LHC beam power spectrum at top energy extends up to about 1 GHz (more precisely $6 \times 10^8$ Hz for a Gaussian beam, where the beam spectrum would drop by $1/e$) [38] (Courtesy N. Mounet).

close to the beam. As part of the LHC upgrade it is considered to replace 22 secondary collimators in the betatron cleaning insertion of IR7, and 8 secondary collimators in the momentum cleaning section of IR3, by new collimators using jaws made from molybdenum-graphite (Mo-Gr) instead of the presently employed carbon-fibre reinforced carbon (CFC). The CFC had been chosen for its robustness in case of a beam impact, but it exhibits a rather low conductivity. The bottom picture of Fig. 13 shows the expected overall impedance reduction thanks to the replacement of the secondary collimators in IR7. For frequencies between a few 100 kHz (a few tens of revolution harmonics) and a few GHz (beyond the cutoff of the bunch spectrum) the impedance will be reduced by about a factor of two, which should ensure beam stability at the (twice the nominal) bunch current of the HL-LHC, provided the normalized aperture of the collimators remains unchanged. The latter, in turn, depends on the apertures of the final triplet and of the matching-section magnets, as well as on the $\beta^*_\text{eff}$ values.

As in the longitudinal plane, high-Q resonances can also be a significant part of the transverse machine impedance, with, again, many of these resonances related to higher-order modes (HOMs) in cavity-like objects. Such modes will be present in the 400 MHz main RF cavities, the transverse damper system, and/or the vacuum chambers of the four main experiments. Similar to the longitudinal plane, the HL-LHC crab-cavity system will significantly increase the set of higher-order modes. Namely, the 16 crab cavities per beam will add about 25% to the total effective
Fig. 12. Composition of the real (top) and imaginary transverse impedance (bottom) for the HL-LHC at 7 TeV as a function of frequency. The resistive-wall impedance of the collimators is shown in gray, the geometric collimator impedance in green [38] (Courtesy N. Mounet).

The impedance of other additional devices, such a long-range wire compensators and, especially, (hollow) electron lenses could potentially be large and harmful, and requires a careful investigation and strict impedance control during design and fabrication. Other possibly harmful devices are the kickers and monitors currently under study as part of a stochastic cooling implementation for the HL-LHC ion operation.
The HL-LHC Accelerator Physics Challenges

The transverse resistive wall impedance is another potential concern. At twice the nominal LHC bunch charge the resistive wall instability rise time for the most unstable rigid coupled-bunch mode will shrink to about 150 turns or 13 ms. About half of this growth rate comes from the Cu-coated arc beam screen, and the other half from the collimator impedance at 8 kHz (which is not much affected by the
planned collimator modifications). A growth rate of 150 turns should however be easily damped by the transverse damper system.

Space-charge

Considering a Gaussian longitudinal bunch distribution, the incoherent space-charge tune shift is given by

$$\Delta Q_{\text{sc}}^{\text{inc}} = -\frac{N_b r_p C}{4\pi \beta \gamma^2 \varepsilon_n \sqrt{2\pi} \sigma_z}.$$  \hspace{1cm} (6)

Here $N_b$ is the number of protons per bunch, $C$ the ring circumference, $\sigma_z$ the rms bunch length, $r_p$ the classical proton radius, $\beta = v/c$ the beam velocity normalized to the speed of light, $\gamma = (1 - \beta^2)^{-1/2}$ the Lorentz factor, and $\varepsilon_n$ the normalized transverse beam emittance. At injection energy, assuming the same bunch length as for the nominal LHC, the incoherent space charge tune shift for HL-LHC bunch intensity and emittance is $-3.6 \times 10^{-3}$ and the direct space charge impedance $i8.9 \text{ M}\Omega/m$. These values are significantly larger than for the nominal LHC, with possible consequences including (enhanced) Landau damping [41] and (reduced) dynamic aperture [42]. Reassuringly in beam studies at the SPS much larger space-charge tune shifts (less than $-0.1$) did not significantly affect beam lifetime or stability [43]. However, any effect of direct space charge arises from its interplay with the nonlinear magnetic fields of a storage ring, which are different in the superconducting LHC and the warm SPS. At 7 TeV the incoherent HL-LHC space-charge tune shift and direct space-charge impedance become $-2 \times 10^{-5}$ and $i0.62 \text{ M}\Omega/m$, respectively.

1.2.3. Landau damping

Loss of longitudinal Landau damping has been observed in the present LHC on different parts of the cycle (flat bottom, ramp and flat top) for bunches with small longitudinal emittances [32]. For bunch intensities of around $1.5 \times 10^{11}$ protons, injection phase oscillations are not damped on the flat bottom with longitudinal emittances (defined as $4\pi \sigma_z \sigma_E$, with $\sigma_E$ the rms energy spread), of less than 0.5 eVs. Quadrupole (or non-rigid dipole) instability has been observed during acceleration for emittances below 0.4 eVs (ramp in 2010) and on the flat top below about 0.7 eVs. As a mitigation measure, in normal operation the beam is stabilized by controlled emittance blow-up during the ramp [44]. A possible additional approach considered for the HL-LHC is to use a second RF system, operating at the second or third harmonic frequency of the fundamental system (i.e. at 800 MHz or 1.2 GHz), which could more than double the threshold for the loss of longitudinal Landau damping [45].
Landau damping in the transverse plane is provided by existing dedicated octupole magnets located in the arcs (MO), as well as, during physics, primarily by the large tune spread resulting from the head-on collisions. On the other hand, during the optics squeeze the tune spread induced by the long-range (LR) beam–beam interactions increases with $1/\beta^*$ (assuming the crossing angle to be kept constant during this process, as in 2012). Therefore, depending on the MO polarity, the LR beam–beam tune spread can add up or, on the contrary, compensate the amplitude detuning induced by the Landau octupoles themselves. In addition, during the collapsing of the parallel separation bumps for bringing the two beams into collision, the aforementioned effect is further amplified because the degree of parallel separation at the main IP also affects the beam–beam distance at the long-range beam–beam encounters. Finally at small parallel separation, the tune spread due to the beam–beam forces at the main IP changes sign, and Landau damping can be lost completely if the tune spread provided by the octupoles is insufficient, or if (for the wrong sign of the octupole polarity) at a certain transverse separation the tune shift due to the beam–beam force cancels the effect of the octupoles, resulting in essentially zero tune spread and onset of instability [46]. All together, for the negative MO polarity, and in the presence of long-range beam–beam interactions, the overall tune spread may first vanish on the anti-diagonal of the tune diagram towards the end of the squeeze, and on the diagonal during the collapsing process of the separation bumps [46] (see Fig. 14). For this reason the LHC octupole polarity was changed from negative to positive in the middle of the 2012 run [46], although the positive polarity is known to be non-optimal with regard to single beam stability [47] as, in the LHC, the impedance induced tune shifts have negative real parts.

This effect can also be analyzed from the point of view of Landau stability diagrams, as is illustrated in Fig. 15, which shows stability limits in complex tune plane computed separately for each of the following components for the nominal LHC parameters [48]: (1) the Landau octupoles, (2) the long-range beam–beam interactions, and (3) head-on collisions at the interaction points 1 and 5.

For the 2012 LHC parameters, combining the above three contributions, and analyzing the situation for both octupole polarities, Fig. 16 shows the negative imaginary part of the tune shift, $-\text{Im}(\Delta Q)$, related to the Landau damping rate through $1/\tau_{\text{damp}} = -2\pi f_{\text{res}} \text{Im}(\Delta Q)$, as a function of the real part of the tune shift and of the beam–beam separation at the two main collision points IP 1 and 5, but ignoring any collisions in IPs 2 and 8 [48]. This figure highlights the lack of stability for a transverse beam–beam separation around $2\sigma^*$, especially for the case where the Landau octupole magnets are powered at negative polarity (left picture). Comparing the left and the right pictures, this figure illustrates as well the aforementioned effect during the last part of the collision making.
Fig. 14. Tune footprint up to transverse amplitude of $6\sigma$ for the 2012 LHC parameters (4 TeV/beam, $N_b = 1.5 \times 10^{11}$, $\gamma e = 2.5 \mu m$) for bunches experiencing the maximum number of long-range interactions in IR1 and IR5. The Landau octupoles are assumed to be powered with a negative polarity (i.e. $\partial Q_x/\partial J_x \equiv \partial Q_y/\partial J_y < 0$ and $\partial Q_y/\partial J_x \equiv \partial Q_x/\partial J_y > 0$). The different pictures correspond to: (a) the end of the ramp; the end of the squeeze with the full (b) and an intermediate parallel separation (c), corresponding to $\delta x^* \sim 70\sigma^*$ and $\delta x^* \sim 15\sigma^*$, respectively; the very end of the collision making process with a small (d) and or very small (e) parallel separation of $\delta x^* \sim 5\sigma^*$ and $\delta x^* \sim 1.8\sigma^*$, respectively; (f) the beginning of the physics coast with head-on collisions at IP1 and IP5.
Fig. 15. Comparison of stability diagrams from either octupoles powered with at maximum negative excitation (500 A magnet current), long-range collisions in IPs 1 and 5, or head-on collisions in IPs 1 and 5, for the nominal LHC parameters [48] (Courtesy X. Buffat).

Fig. 16. Stability diagram as a function of beam separation in IPs 1 and 5 for a bunch experiencing the maximum number of long-range interactions assuming LHC parameters from 2012, and considering both polarities of the octupoles [48] (the left and the right picture; Courtesy X. Buffat).

The Landau damping scales with the transverse geometric emittance, and with the normalized octupole strength, both of which decrease by a factor $4/7 \approx 0.57$ at the design beam energy of 7 TeV, compared with the 4 TeV from 2012. In
addition the bunch charge will be about 40% higher for the HL-LHC beam ($2.2 \times 10^{11}$ p/bunch) than what it was for the 50 ns LHC beam in 2012 ($1.5–1.6 \times 10^{11}$ p/bunch). In total without collision and prior to the final $\beta^*$ squeeze, the HL-LHC beam may be about four times more unstable transversely than it has been the case in 2012 (where, for the negative octupole polarity, already many fills were aborted due to transverse instabilities). A possible countermeasure could be to ramp and squeeze with colliding beams, thereby stabilizing the beam thanks to the larger beam–beam tune spread. Also, after the squeeze, the ATS optics greatly increases the effective strength of the Landau octupoles, by almost a factor of 4 for $\beta^*_{x,y} = 15$ cm, thanks to the beta wave introduced in the arcs around IP1 and IP5. Finally long-range beam–beam compensators based on electro-magnetic wires [49] could locally correct the tune spread induced by the long-range beam–beam interactions, and therefore restore the full freedom concerning the choice of the octupole polarity.

1.2.4. Synchrotron radiation and electron cloud

The LHC is the first proton storage ring for which synchrotron radiation is a noticeable effect. At top energy, the synchrotron radiation gives rise to a significant heat load, which for the HL-LHC amounts to about 3886 W in total, and which is intercepted by a beam screen at a temperature of 5–20 K. The critical photon energy is 43 eV and the ring-average photon flux $3.4 \times 10^{21}$ cm$^{-2}$s$^{-1}$ per beam. The synchrotron radiation also leads to a shrinkage of the beam emittance during physics stores, with an emittance damping time of about 13 h longitudinally, and twice longer in the transverse plane. When operating at 3.5–4 TeV in 2011–12 the damping has been about ten times weaker. At injection energy, the synchrotron radiation stays negligible in terms of heat load (though still usable for beam diagnostics purposes), with a total radiated power of order 0.1 W.

Table 6 summarizes the various heat loads induced by the circulating beam on the arc beam screen for the HL-LHC at 7 TeV. The corresponding values for the nominal LHC are also shown for comparison. The total value of these conventional heat loads determines the cooling capacity remaining for any additional source of heat, such as due to incident electrons discussed in the following.

Namely, seed electrons created by ionization of the residual gas at injection or photo-electrons liberated by the large number of hard U.V. synchrotron radiation photons at 7 TeV are pulled towards the positively charged LHC proton bunches. When they hit the opposite wall, they generate secondary electrons which can in turn be accelerated by the next bunch if they are slow enough to remain in the vacuum chamber until the next bunch arrives.
Table 6. Summary of beam-induced heat loads on the arc beam screen at 7 TeV beam energy for the nominal LHC and for the HL-LHC (assuming the same bunch length). The three columns show the various sources together with the corresponding heat loads for the two machines in units of mW/m.

<table>
<thead>
<tr>
<th>Heat Source</th>
<th>Power [mW/m] LHC at 7 TeV</th>
<th>Power [mW/m] HL-LHC at 7 TeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synchrotron Radiation [50]</td>
<td>170</td>
<td>330</td>
</tr>
<tr>
<td>Ohmic Losses [51]</td>
<td>110</td>
<td>400</td>
</tr>
<tr>
<td>Pumping Slots [52]</td>
<td>~10</td>
<td>~35</td>
</tr>
<tr>
<td>Welds [29]</td>
<td>~10</td>
<td>~35</td>
</tr>
<tr>
<td>Shielded Bellows [29]</td>
<td>≤ 30</td>
<td>≤ 100</td>
</tr>
<tr>
<td>Total</td>
<td>~330</td>
<td>~900</td>
</tr>
</tbody>
</table>

Fig. 17. Illustration of the beam induced multipacting process leading to the build up of an electron cloud in the LHC beam pipe. The horizontal axis is time, illustrating the bunch spacing. Blue curvy lines represent synchrotron radiation photons hitting the chamber wall at the moment of a bunch arrival. Emitted photoelectrons (red) gain energy in the electric field of the passing bunch. When hitting the opposite side of the beam pipe, these are either reflected (still red) or generate lower-energy secondary electrons (green), which are then again accelerated transversely during the subsequent bunch passage (Courtesy F. Ruggiero).

This mechanism is illustrated in Fig. 17. It can lead to the fast build up of an electron cloud, with potential implications for beam stability, emittance growth, and heat load on the cold beam screen in the arcs. Electron cloud effects have been actively investigated at CERN since 1997 by analytic estimates, simulations, and experimental tests [53–55]. The electron cloud can ultimately cure itself since the secondary emission yield of the chamber surface decreases as a result of the continuing bombardment with electrons of sufficiently high energy (e.g. above about 30 eV). The variation of the maximum secondary emission yield, $\delta_{\text{max}}$, with electron dose measured in the laboratory for the colaminated copper surface of the LHC beam screen (at perpendicular incidence of the primary electrons) is illustrated in Fig. 18. The LHC strategy to overcome electron-cloud related limitations...
Fig. 18. Variation of the secondary emission yield of copper with the incident electron dose for different energies of the impinging electrons at normal incidence on colaminated Cu of the LHC beam screen. The squares represent the $\delta_{\text{max}}$ after an additional dose of $10^{-2}$ C mm$^{-2}$ at 200 eV [56] (Courtesy R. Cimino).

Electron-cloud effects and a reduction of the secondary emission yield with time have indeed been observed in the first years of LHC operation, where operation with 50-ns spacing has become possible thanks to rapid surface improvement (lowering of the secondary emission yield) in dedicated “scrubbing” runs with 25-ns beams [57–59]. Figure 19 presents the LHC surface conditioning observed during the 2011–12 LHC run for an uncoated field-free region, Fig. 20 the one for the cold arcs in 2011.

Arc heat-load observations in 2012, together with associated ‘benchmarking’ simulations, have, however, given rise to a concern that further scrubbing at 25-ns spacing may not decrease the secondary emission to a level sufficiently low to completely avoid electron-cloud build up during physics operation with 25-ns beams. Particularly critical are the arc quadrupole magnets, for which the electron-cloud threshold is lower than for dipoles or field-free regions (see Fig. 21) [59]. In addition at beam energies of 6.5 or 7 TeV the amount of photoelectrons will be much higher than at the 3.5–4 TeV energies of 2011–13. Also the generation of “UFOs” (micro-size dust particles falling into the beam, sometimes leading to very fast losses and consequently triggering a beam abort [60]) could be related to the presence of an electron cloud, as both occur more strongly, or more frequently, at the 25-ns bunch spacing. One conjecture is that the electron-cloud build-up could
lead to a charging of macro-particles residing inside the vacuum system, thereby making them more likely to move under the influence of the electric field of the proton beam. For all these reasons, HL-LHC physics operation with 50-ns spacing (implying lower luminosity or higher pile up) is still retained as a back-up.

Fig. 19. LHC surface conditioning in an uncoated field-free region around the LHC vacuum gauge VGI.141.61.4.B, from 2010 to 2013, inferred by benchmarking of simulations with pressure-rise observations. The symbol $\delta_{\text{max}}$ denotes the maximum secondary emission yield, measurements of which were shown in Fig. 18. The calculated multipacting thresholds for bunch spacings of 50 and 25 ns are marked for reference (Courtesy O. Dominguez).

Fig. 20. LHC surface conditioning in the cold arcs during 2011: beam intensity (top) and estimated evolution of the maximum secondary emission yield (bottom) as inferred by benchmarking of simulations with cryogenic heat-load measurements (Courtesy G. Iadarola and G. Rumolo).
Fig. 21. Simulated electron-cloud heat load in dipoles, field-free regions and quadrupole magnets of the LHC cold arcs, for a single beam of 2,808 bunches (25 ns bunch spacing) as a function of maximum secondary emission yield [59] (Courtesy G. Iadarola, G. Rumolo).

Fig. 22. Heat load per unit length inside an arc quadrupole for a single beam of 2,808 bunches as a function of bunch intensity simulated by the PyECLOUD code; the different curves refer to various values for the maximum secondary emission yield as indicated [62] (Courtesy G. Rumolo, G. Iadarola).

option. With 50-ns electron cloud is predicted not to be a problem at the HL-LHC either [61]. Fortunately, at the higher bunch intensity of the HL-LHC the electron cloud build up and related heat load inside the quadrupole magnets are expected to improve [62], as is illustrated in Fig. 22, so that for the HL-LHC the 25-ns operation should be more easily feasible than for the present LHC.
The HL-LHC Accelerator Physics Challenges

Fig. 23. Simulated heat load per unit length inside an arc dipole at 7 TeV beam energy as a function of the maximum secondary emission yield $\delta_{\text{max}}$; the different curves refer to various values of the bunch spacing as indicated; the total beam current is held constant at 1.12 A; a low-energy electron reflectivity of $R \approx 0.7$ and a characteristic incident electron energy $\epsilon_{\text{max}} = 332$ eV (specifically, $\epsilon_{\text{max}}$ refers to the energy at which the maximum yield $\delta_{\text{max}}$ is attained) have been assumed [58] (Courtesy O. Dominguez).

The HL-LHC could conceivably also operate with shorter bunch spacing. The variation of the electron-cloud heat load for shorter spacing is illustrated in Fig. 23, where the total beam current is held constant. This figure indicates that a bunch spacing equal to half the nominal, $\sim 12.5$ ns is about the worst case in terms of the onset of multipacting, while the heat load decreases again for even shorter distances between bunches, approaching the limit of a coasting continuous beam with static electric potential (in which the electrons cannot gain any energy).

1.2.5. Intra-beam scattering

In the dispersive regions of the ring, small-angle particle–particle collisions within a bunch couple the horizontal and longitudinal particle oscillations and, above transition energy, give rise to an irreversible emittance growth in both planes. The horizontal emittance growth rate caused by this intra-beam scattering (IBS) is roughly proportional to $N_b/\left(\epsilon_{\perp}^2 \epsilon_{\text{e}}^{0.75}\right)$, which increases by a factor 4.7 when transiting from the LHC to the HL-LHC beam parameters. The longitudinal IBS growth rate roughly scales as $N_b/\left(\epsilon_{\perp}^2 \epsilon_{\text{e}}^{1.75}\right)$, which rises only by a factor 3.1 (assuming the same longitudinal emittance, bunch profile, and RF voltage). In addition, the IBS growth rates also depend on the optics. Especially for the ATS collision optics [63] deviations from the simple scaling are expected due to the large $\beta$-beating waves induced in four of the arcs, which are not part of the LHC design optics. At 7 TeV...
beam energy, for the nominal LHC the longitudinal IBS emittance growth time is 57 h and the horizontal growth time is 102 h. For the HL-LHC, with the ATS optics, these growth times become 19 and 18 h, i.e., they decrease by a factor of about 6 horizontally (the difference from the aforementioned factor of 4.7 is attributed mainly to the change in optics) and by a factor of about 3.0 longitudinally. For constant normalized emittances the IBS growth times are approximately independent of the beam energy. However, the longitudinal emittance is a factor 2.5 smaller at injection than in collision (as a result of a controlled blow up on the LHC ramp) which leads to larger IBS growth rates at lower energy. As a result, for the HL-LHC the IBS rise times at injection are 5 h and 7.5 h, respectively, in the longitudinal and in the horizontal plane. Therefore, an emittance growth of about 10% and 7% is expected to occur longitudinally and transversely, respectively, during 30 minutes on the injection plateau of the HL-LHC. This emittance growth could be reduced by means of a lower frequency capture RF system (e.g. 200 MHz), allowing for a larger longitudinal emittance at injection, which enters roughly quadratically or linearly into the longitudinal and transverse IBS rise times, respectively. Another possible mitigation could be longitudinal bunch shaping using two RF systems (e.g., 200 and 400 MHz, or 400 and 800 MHz): A flat bunch of total length \(2\sqrt{\pi}\sigma_z\) has a factor \(\sqrt{2}\) lower IBS rate than a Gaussian bunch with rms length \(\sigma_z\) of the same charge [64].

The IBS growth rates cited above were computed with the latest version of MAD-X [65] considering either the design optics, for the nominal LHC, or the ATS optics (SLHCV3.1b [12]), with \(\beta^*_{x,y} = 15\) cm in IPs 1 and 5 for the HL-LHC collision scenarios and with \(\beta^*_{x,y} = 11\) m at injection energy.

1.2.6. **Touschek scattering**

Touschek scattering refers to particle–particle collisions within a bunch, by which enough energy is transferred from the transverse into the longitudinal oscillations such that the scattered particles leave the stable RF bucket. The Touschek loss rate for the nominal LHC has been estimated in [66]. According to this study, for the nominal LHC coasting beam is produced at a rate per proton of \(2 \times 10^{-4}\ h^{-1}\) during injection and of \(8 \times 10^{-5}\ h^{-1}\) at 7 TeV. Since the loss rate per proton due to Touschek scattering is linear in the bunch population and roughly inversely proportional to the square of the transverse emittance, it will be about four times higher at the HL-LHC, with a coasting beam production rate per proton of \(8 \times 10^{-4}\ h^{-1}\) during injection and of \(2 \times 10^{-4}\ h^{-1}\) at 7 TeV. In addition, the total number of protons circulating in the ring is doubled. Therefore, the absolute Touschek scattering rate is expected to be about one order of magnitude higher at the HL-LHC compared to the nominal LHC.
Once the protons are outside the RF bucket, they lose energy due to synchrotron radiation. If the collimators provide an energy aperture of $3.9 \times 10^{-3}$, a scattered proton is lost after about 390 h at injection or after 6.5 min at top energy, respectively. While the energy drift due to synchrotron radiation is not noticeable at injection (which means that most protons coasting outside of the RF bucket are lost during a short time interval at the start of the energy ramp), at 7 TeV it gives rise to a steady-state coasting beam component of about $10^{-5}$ for the present LHC and of $10^{-4}$ for the HL-LHC. This unavoidable coasting beam component, which will also populate the beam abort gap, could be reduced (linearly) by collimating more tightly in energy.

1.2.7. Beam–beam effects

Beam–beam effects include incoherent effects, such as betatron tune spreads associated with the nonlinear head-on and long-range collisions as well as a reduction of dynamic aperture in case of insufficient beam separation at the parasitic encounters, and coherent effects affecting orbit, tunes, and chromaticities of the different bunches (depending on their different collision schedules) or coherent oscillation modes.

For the HL-LHC both long-range and head-on beam–beam effects can be more severe than for the nominal LHC.

The effect of the long-range collisions is enhanced compared with the nominal LHC, since both the charge per bunch is almost doubled and the number of relevant long-range encounters (till the entrance of the separation dipole D1) is increased, due to the longer final triplet, from about 16 to 19 encounters for the latest layout of new insertions ATLAS and CMS [13]. Figure 24 illustrates that a crossing angle of $590 \mu$rad (12.5σ) is required to obtain a short-term dynamic aperture around $7\sigma$ for a $\beta^*$ of 15 cm (round). This choice of crossing angle is confirmed by Fig. 25, which compares the tune footprints up to $6\sigma$ computed for crossing angles of 475 $\mu$rad (10σ normalized separation) and for 590 $\mu$rad (12.5σ), the former case appearing “pathological”. These results are also consistent with a scaling law for the dynamic aperture induced by the long-range beam–beam interaction, as first observed by Irwin [67].

The residual effect of long range beam–beam encounters can be mitigated by electromagnetic wires [49], which are being investigated for a possible deployment at the HL-LHC.

The head-on beam–beam tune shift at one IP with crab crossing and round colliding beams is given by

$$\Delta Q_{bb} = \frac{r_p N_b}{4 \pi \gamma \varepsilon_{x,y}},$$

(7)
where $r_p$ denotes the “classical proton radius,” $N_b$ the bunch population and $(\gamma e_{x,y})$ the normalized transverse emittance. For the HL-LHC beam parameters (see Table 1) and full crabbing at IP1 and IP5 (or collision at small Piwinski angle as in IR8), $\Delta Q_{bb} \approx 0.011$ per IP. With collisions at three IPs, the total tune shift becomes $\Delta Q_{tot} \approx 0.033$, which is more than three times higher than for the nominal LHC. Some simulation results indicate that this value may be close to the beam–beam limit of the LHC [69]. At the beam–beam limit, once it is reached, the luminosity can however be further optimized by increasing bunch length and crossing angle [71], essentially at the beginning of the store when the bunch charge is maximal. Another means for mitigating the total head-on beam–beam tune shift is to assume a luminosity leveling through a variation of the parallel separation at IP8 rather than with $\beta^*$, as successfully deployed in 2012. Finally an alternative which emerged very recently but has not yet been approved for the HL-LHC baseline is the so-called “crab-kissing” scheme [72], which consists in leveling the luminosity (and the peak line density of pile-up events, see Subsection 1.3) by acting on...
additional (rotated) crab cavities providing an RF dipole field in the separation plane of IR1 and IR5, so as to generate a time-dependent parallel separation at the IP and, therefore, to simultaneously reduce the head-on beam–beam tune shift at IP1 and IP5 [73].

Coming back to the present baseline ($\beta^*$ leveling at IP1, IP5 and IP8) and for the HL-LHC beam parameters given in Table 1, Fig. 26 presents a beam–beam tune footprint and frequency map for a worst case scenario [74], assuming collisions with a very small Piwinski angle at IP8 (relatively small luminosity and large $\beta^*$), and head-on fully crabbed collisions at IP1 and IP5 with $\beta^*_{x,y}$ equal to 40 cm and a crossing angle of 360 $\mu$rad (12$\sigma$ separation at this value of $\beta^*$, compared to 590 $\mu$rad for the nominal HL-LHC $\beta^*$ of 15 cm). The tune footprint straddles resonances of 7th, 10th and 13th order, all of which are known to degrade the beam lifetime in collision, according to past experience at the Sp$\bar{p}$S collider [75]. Figure 27 shows the corresponding plots for a later time in the store, including changes due to luminosity leveling, namely with a smaller $\beta^*$ of 15 cm, a lower bunch population of $1.06 \times 10^{11}$ and a larger (nominal HL-LHC) crossing angle of 590 $\mu$rad (12.5$\sigma$ separation) [74]. Resonance effects have all but disappeared from the frequency maps.

Other important considerations related to beam–beam effects are the bunch to bunch variation of closed orbit, betatron tune and chromaticity, arising from the fact that different bunches do not experience the same number of long-range (LR) beam–beam interactions depending on their relative position in the trains. Neglecting the LR interactions in IR2 and IR8, this effect is vanishing in terms of tune shift assuming an alternated horizontal-vertical crossing angle in IR1 and IR5 [76].
Fig. 26. Tune footprint (left) and frequency map (right) for the HL-LHC with fully crabbed collisions in 3 IPs, considering $\beta^*_x = 40 \, \text{cm}$, $N_p = 2.2 \times 10^{11}$ protons per bunch, $\theta_c = 360 \, \mu\text{rad}$ full crossing angle, normalized long-range separation of $12\sigma$, crab voltage of $7.7 \, \text{MV}$ per side of IP, and head-on beam–beam tune shifts of $\xi_x = 0.0306$ and $\xi_y = 0.031$ in total for the 3 IPs, as obtained in weak-strong simulations with the code LIFETRAC [74] (Courtesy A. Valishev).

Fig. 27. Tune footprint (left) and frequency map (right) for the HL-LHC with fully crabbed collisions in 3 IPs, considering $\beta^*_x = 15 \, \text{cm}$, $N_p = 1.1 \times 10^{11}$ protons per bunch, $\theta_c = 590 \, \mu\text{rad}$ full crossing angle, normalized long-range separation of $12.5\sigma$, crab voltage of $12.5 \, \text{MV}$ per side of IP, and head-on beam–beam tune shifts of $\xi_x = 0.015$ and $\xi_y = 0.016$ in total for the 3 IPs, as obtained in weak-strong simulations with the code LIFETRAC [74] (Courtesy A. Valishev).

bunch to bunch variations of chromaticity are also marginal for the ATS optics, and essentially driven by the sextupole-like component of the LR beam–beam interaction combined with an almost vanishing spurious dispersion in the low-$\beta$ insertion (see Subsection 1.1.2). The most visible effect concerns the bunch to bunch variation in terms of closed orbit, scaling as follows with the bunch charge $N_b$, the
normalized emittance $\gamma e$, the normalized crossing angle $d_{bb}$, and the total number of LR beam–beam interactions: $N_{LR}$:

$$\delta x[\sigma] \propto \frac{N_{LR} N_b}{\gamma e d_{bb}}.$$  \hspace{1cm} (8)

In a worst case scenario without $\beta^*$ leveling, in particular when starting the store with a minimal beam–beam separation of $12.5\sigma$ (see Table 1), this effect would be almost tripled for the HL-LHC, reaching about $\pm 0.3\sigma$, to be compared with $\pm 0.1\sigma$ for the nominal LHC [77]. The first and last bunches of each train may then collide with the other beam with a relative offset of about $\delta x_{b1} - \delta x_{b2} \sim 0.5\sigma^*$ at the IP depending on the optics of both beams. This offset will be visible in terms of luminosity (5% luminosity loss for these specific collisions), but should remain marginal in terms of integrated performance over the overall collision schedule. Both simulations for the HL-LHC [70] and LHC operational experience so far [78] indicate that the emittance growth due to static collision offsets at the level of a few tenths of the rms beam size is insignificant compared with the emittance growth caused by the nonzero crossing angle. Even this small effect could be further mitigated by an appropriate adjustment of the phase advance between IP1 and IP5 for the clockwise rotating beam, and between IP5 and IP1 for the other beam.

1.3. Dealing with pile up limits

1.3.1. Detector limitations

The detector technology sets limits on the total number of events per crossing (e.g. for calorimetry), as well as on the longitudinal event line density (for tracking of the primary vertices) and, possibly, also on the number of events per unit time during the collision. The nominal LHC parameters correspond to a peak pile up of about 20 events per crossing. During the 2012 LHC run the average pile up already was about 21 for ATLAS and CMS, and the maximum pile up in physics runs about 40 (i.e. twice the design). Higher peak values, close to 80 events per crossing, were reached in dedicated machine studies with a few bunches. The new ATLAS and CMS detectors are being designed for an average pile up of $\mu_{tot} = 140$ events per bunch crossing, with tails up to 200. At 25 ns bunch spacing, with about 2,800 bunches per beam, and considering a total inelastic cross-section of 85 mbarn an average pile up of 140 events per crossing corresponds to a luminosity of $5 \times 10^{34}$ cm$^{-2}$s$^{-1}$ (more precisely $5.2 \times 10^{34}$ cm$^{-2}$s$^{-1}$ for 2,808 bunches per beam). By contrast, the new LHCB detector is designed for an instantaneous luminosity not exceeding $2 \times 10^{33}$ cm$^{-2}$s$^{-1}$, corresponding to about 5.2 visible pile up events.
on average per bunch crossing [3] (for an inelastic cross-section of 70 mb, which is visible by this detector, and assuming 2,400 collisions at IP8, compared with about 2,800 at IP1 and IP5).

In addition to the limit of \( \mu_{\text{tot}} = 140 \) in the new ATLAS and CMS detectors, the average peak pile-up line density should not exceed 1.3 events per mm, with tails up to 1.8 events per mm [80]. The latter condition determines the minimum extent of the longitudinal “luminous region”, which is related to the bunch length, the value of \( \beta^* \), the normalised crossing angle and the voltage assumed in the crab cavities (maximum voltage for full crabbing as in the baseline, or partial crabbing). Assuming transverse and longitudinal bunch distributions to all be of Gaussian shape, and denoting the effective full crossing angle (including the possible effect of crab cavities) by \( \theta_c \), the rms length of the luminous region can be written as

\[
\sigma_{\text{lum}} = \sigma_z \left( \frac{1}{\sqrt{\left( \frac{\sigma_z}{\sigma_{\theta z}} \right)^2 + 1}} + \phi_w \right),
\]

where \( \phi_w \) signifies the Piwinski angle. For the baseline parameters of Table 1 at \( \beta^* = 15 \) cm, the extension of the luminous region is expected to be 4.4 cm r.m.s., and the peak line density of pile-up is calculated as 1.27 event/mm at \( \mu_{\text{tot}} = 140 \), including the degradation due to the hourglass effect and due to the RF curvature of the crab-cavity field [72].

A limit on the number of events per unit time during a bunch collision may also exist, essentially for forward physics, but has not yet been specified.

Finally, it is expected that bunch spacings shorter than 25 ns cannot be handled by either the present or the upgraded LHC detectors.

1.3.2. Luminosity leveling

The HL-LHC upgrade project aims at achieving a ‘virtual’ peak luminosity that is considerably higher than the maximum value imposed by the acceptable event pile up and to deploy a controlled reduction of the peak luminosity during operation (‘luminosity leveling’) so that the operational luminosity can be sustained over a significant length of time.

This luminosity leveling during a physics store can be accomplished in a number of ways: (1) dynamic \( \beta^* \) squeeze, (2) crossing angle variation, (3) changes in the crab RF voltage, (4) dynamic bunch-length reduction, or (5) controlled variation of the transverse distance between the two colliding beams\(^a\).

\(^a\)Historically, leveling with \( \beta^* \) variation was mentioned for the LHC ion programme around 2000 [81]. For ions the history of this proposal goes back further than this, probably to about 1995 [82]. Leveling for \( pp \) collisions in the context of the LHC luminosity upgrade was first proposed in February 2007
Due to proton consumption in the collisions, the total beam intensity, \( N_{\text{tot}} \), decays as 
\[
\frac{dN_{\text{tot}}}{dt} = -n_{IP} \sigma_{\text{tot}} L, 
\]
where \( n_{IP} \) denotes the number of high-luminosity IPs (\( n_{IP} = 2 \) for HL-LHC), \( \sigma_{\text{tot}} \) the total hadron cross-section (\( \sigma_{\text{tot}} \approx 100 \text{ mbarn} \)), and \( L \) the luminosity at IP1 and IP5. Setting \( L \) equal to the leveled luminosity, \( L_{\text{lev}} \), the effective beam lifetime is
\[
\tau_{\text{eff}} = \frac{N_{\text{tot}}(t \equiv 0)}{n_{IP} \sigma_{\text{tot}} L_{\text{lev}}}. \tag{10}
\]
Next, introducing the ratio of virtual peak luminosity and leveled luminosity, \( k = \hat{L}/L_{\text{lev}} \), we can express the maximum leveling time as
\[
t_{\text{lev}} = \tau_{\text{eff}} \left( 1 - \frac{1}{\sqrt{k}} \right) \equiv \tau_{\text{eff}} K, \tag{11}
\]
where \( K \equiv (1 - 1/\sqrt{k}) \) designates the ratio of leveling time and effective beam lifetime. For the general case, where the physics run is extended beyond the end of the leveling period by a certain decay time \( t_{\text{dec}} \) (see Fig. 28), the time-averaged luminosity becomes [86]
\[
L_{\text{ave}} = \frac{t_{\text{lev}} + t_{\text{dec}} \tau_{\text{eff}} / (t_{\text{dec}} + \tau_{\text{eff}})}{t_{\text{dec}} + t_{\text{lev}} + t_{\text{ta}}}, \tag{12}
\]
with \( t_{\text{ta}} \) denoting the average turnaround time, i.e. the time between the end of one physics run and the start of the next one (time needed for magnet ramp down, injection, acceleration, \( \beta^* \) squeeze, collimation set up, etc.).

The average luminosity assumes a maximum value, \( L_{\text{ave, opt}} \), for \( t_{\text{dec}} \) equal to the ‘optimum decay time’ [86]:
\[
t_{\text{dec, opt}} = \frac{\tau_{\text{eff}}}{1 + K} \left(-K + \sqrt{(K^2 + (1 + K)t_{\text{ta}} / \tau_{\text{eff}})} \right). \tag{13}
\]
The larger the turnaround time \( t_{\text{ta}} \) is, compared with the effective lifetime \( \tau_{\text{eff}} \), the longer is the optimum decay time. The resulting optimum total length of a single run is \( t_{\text{run, opt}} = (t_{\text{lev}} + t_{\text{dec, opt}}) \).

Both dynamic \( \beta^* \) squeeze and bunch-length variation were analysed as examples. A subsequent discussion concluded that “Luminosity leveling in IP1 and 5 is a very attractive solution for constant luminosity throughout the runs, achievable by i.e. a continuous \( \beta^* \) squeeze. Emittance blow-up effects remain to be assessed. Difficulties of changing the bunch length with RF systems is stressed with very high RF voltage required. Changing the \( \beta^* \) is the most promising option, although it is probably difficult to use in practice because of the behaviour of superconducting magnets.” A few months later leveling by crossing angle variation in the so-called “early separation scheme” was proposed as another attractive option [84]. Leveling using the crab RF voltage was suggested as well [85].

The former two options have however the drawback of inducing very high line pile-up density at the beginning of the physics fill (about 5 event/mm). Very recently, the so-called crab-kissing scheme was introduced [72], assuming crab cavities in the parallel separation plane in anti-phase between Beam 1 and Beam 2, and offering a powerful luminosity leveling tool while reducing simultaneously the peak pile up line density and the head-on beam–beam tune shift. This scheme, however, is still under discussion, essentially due to its additional need in terms of hardware, and therefore will not be discussed any further here.
For comparison, at the nominal LHC without leveling the optimum run time is
\[
t_{\text{run,nol}} = \sqrt{t_{\text{ta}}/\tau_{\text{eff}}},
\]
yielding the average luminosity of
\[
L_{\text{ave,nol}} = \hat{L}/\left(1 + \sqrt{t_{\text{ta}}/\tau_{\text{eff}}}\right)^2,
\]
where \(\hat{L}\) denotes the initial peak luminosity.

The variation of the beam–beam tune shift during a physics store depends on
the leveling scheme [87]. In case of \(\beta^*\) variation, the beam–beam tune shift is max-
imized at the beginning, but then decreases during the store. When leveling via the
bunch length, crossing angle or crab voltage the tune shift is minimized at the be-
inning but then increases, by a factor of 2 to 3. When leveling with the transverse
offset the beam–beam tune shift changes sign and its modulus can increase even
more strongly during the leveling process, by up to an order of magnitude [87].

The integrated annual luminosity for LHC and HL-LHC can simplistically be
estimated by multiplying the total time scheduled for physics production \(T\), the
machine availability \(A\) (time without hardware failures divided by total time sched-
uled), and the average luminosity over the time periods without any hardware fail-
ure, as
\[
L_{\text{int}} \equiv \int_{\text{year}} L(t)\,dt = T_{\text{tot}}AL_{\text{ave}}.
\]
(14)

For the HL-LHC target parameters we require \(L_{\text{int}}\) to be 250 fb\(^{-1}\), consider \(T_{\text{tot}} = 160\) days (per year). We can then use the above equation to deduce the minimum availability required to be about 45%, which may be compared with an actual LHC availability of 73% in the year 2012.

Defining the machine efficiency, \(E\), as the time spent in physics divided by the
total allocated calendar time, we can also estimate the minimum needed efficiency,
as
\[
E \approx A t_{\text{run}}/(t_{\text{run}} + t_{\text{ta}}).
\]
(15)

For example, assuming the estimated minimum machine availability \(A\) of 45%
deduced previously, an optimum HL-LHC run (leveling) time \(t_{\text{run}}\) of 9 h, and an
average turnaround time of 5 h, the minimum needed efficiency \(E\) becomes 29%,
which is lower than the efficiency of 36.5% obtained during LHC 4-TeV operation
in 2012, and, therefore, appears an achievable target.

More refined estimates of integrated luminosities or necessary efficiencies
might be obtained by considering a realistic random run-time distribution of (pre-
maturely aborted) physics stores.

Figure 28 illustrates the HL-LHC luminosity time evolution for a single fill
with and without leveling. The curve for the nominal LHC is also included for
comparison. Figure 29 displays the luminosity evolution with and without leveling
over several successive fills, where the fill length for either case has been optimized
for maximum luminosity, assuming a turnaround time of 5 h. Without leveling the
Fig. 28. HL-LHC luminosity evolution as a function of time, for a single fill, without (red), and with leveling at a pile up of 140 events per crossing (blue), compared with the LHC design (black). An inelastic pile-up cross-section of 85 mbarn is assumed for the mapping between number of pile-up events and luminosity, while a total cross-section of 100 mbarn is adopted for evaluating the proton burn off rate during the store.

Fig. 29. HL-LHC luminosity evolution as a function of time, during several length-optimized fills over 30 h, without (dashed red), and with leveling at a pile up of 140 events per crossing (dashed blue). A turnaround time of 5 h has been assumed. The corresponding time-averaged luminosity values are indicated by the solid red and blue lines.

The optimized fill length would be less than half the fill length of the leveled case, implying a significant increase in the fraction of time spent for turnaround without any luminosity. The solid lines indicate the time-averaged luminosities with and without leveling. The difference is only some 25% while the peak luminosity differs by 400%. Finally, Fig. 30 shows the integrated luminosity as a function of time. It is about 4 fb⁻¹ per day. With an efficiency of 40% and considering 160 days of days scheduled for physics per calendar year, the luminosity delivered per year will exceed 250 fb⁻¹.
1.4. Summary and conclusions

The yearly performance targeted by the HL-LHC (\(\sim 250 \text{fb}^{-1}\)) is a factor of 5 to 6 larger than the one expected for the nominal LHC. In order to reach this goal, both the optics (\(\beta^*\)) and the beam parameters shall be pushed by substantial factors in order to achieve a so-called virtual luminosity of about \(2 \times 10^{35} \text{cm}^{-2}\text{s}^{-1}\), which will enable long fills at a leveled luminosity of \(5 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}\) (limited by the number of pile up events per bunch crossing), to be compared with the LHC design peak luminosity of \(10^{34} \text{cm}^{-2}\text{s}^{-1}\) (see Table 1). The generation and chromatic correction of the collision optics with extremely small \(\beta^*\) requires the deployment of a novel scheme, the so-called Achromatic Telescopic Squeezing (ATS) scheme, but also a series of new magnets of larger aperture including a challenging 150 mm aperture inner triplet (but not only). Novel RF devices, the so-called compact crab cavities, are also mandatory in order to maximize the luminosity at small \(\beta^*\) despite the crossing angle. Due to the increased bunch charge and smaller emittance, collective effects may ultimately limit the performance of the HL-LHC. Some of them are however still expected to be well under control (e.g. IBS), and/or with clear mitigation measures already identified (e.g. head-on and long-range beam-beam effects). On the other hand, as for the nominal LHC and pending more understanding and operational experience at 7 TeV per beam, some uncertainties remain concerning the beam stability in the transverse plane, electron-cloud effects for 25-ns bunch spacing, UFOs, the attainable machine availability and efficiency, as well as machine protection (e.g. crab-cavity failure modes).
Acknowledgments

The authors would like to warmly thank many collaborators of the HiLumi Design Study for various inputs and materials, in particular X. Buffat, O. Dominguez, J. Esteban Muller, M. Giovannozzi, G. Iadarola, R. de Maria, N. Mounet, Y. Nosochkov, G. Papotti, G. Rumolo, B. Salvant, E. Shaposhnikova, E. Todesco, A. Valishev, and S. White.

We also would like to express our gratitude to L. Rossi and O. Brüning for inviting us to address this very interesting aspect of the HL-LHC project, and for their diligent proofreading of an earlier draft version.

References


[34] B. Salvant, private communication (2014).


[38] N. Mounet, First Estimates of Intensity Limitations from HL-LHC Transverse Impedance, HiLumi WP2 Task 2.4 meeting, CERN, 22 January 2014.


