Alma Mater Studiorum · University of Bologna

School of Science
Department of Physics and Astronomy
Master Degree in Physics

Optimisation of LHC Integrated Luminosity

Supervisor:
Prof. Armando Bazzani

Co-supervisor:
Prof. Massimo Giovannozzi

Submitted by:
Giulia Faletti

Academic Year 2020/2021

Tu ne cede malis, sed contra audentior ito.

-Virgilio-

Abstract

Luminosity is the key figure of merit of a collider, its integrated value being proportional to the data collected by the physics detectors. The goal of a collider run is to collect the largest possible integrated luminosity over the fixed time allocated for physics in order to maximise the probability of physics discovery. This study aims at devising a strategy to maximise the integrated luminosity of a circular collider. The research is carried out in two steps: the first one is devoted to optimisation a posteriori of what was done at the CERN Large Hadron Collider during Run 2, whereas the second step consists in the performance analysis of strategies to optimise the integrated luminosity applied in conditions that simulate the control room situation and the evolution of the collider's run.

Introduction

Luminosity is the key figure of merit of a collider, its integrated value being proportional to the data collected by the physics detectors. The goal of a collider run is to collect the largest possible integrated luminosity over the fixed time allocated for physics to maximise the probability of physics discovery.

So far, several ways have been considered to maximise the luminosity delivered into a collider, like keeping small the amplitude function in the interaction point (as planned for High Luminosity LHC) or increasing the bunches overlap regions (e.g. thanks to Crab Cavities). This work examines what can be done in the control room, choosing the appropriate timing in operations by referring specifically to the CERN Large Hadron Collider, and proposes an algorithm that could help the operators make the optimal choices.

The luminosity is not constant over time. Indeed, since the beams degrade over various collisions, also the instantaneous luminosity decays. However, discarding the current beams and refilling a large collider with new beams takes several hours under normal working conditions. To optimise the integrated luminosity yield, it is hence important to model the luminosity evolution and to predict the optimum fill time.

In this work, after a theoretical introduction (see Chap. 1 and Chap. 2), the evolution of luminosity is studied, and it is proposed a way to model it (see Chap. 3). Subsequently, a preliminary *a posteriori* study is exposed in Chapter 4. Here, the idea was to understand if, with complete knowledge of the luminosity evolution of the realised fills, there was room for improvement in terms of integrated luminosity. Based on this study, Chapter 5 devises an on-line optimisation strategy.

Finally, Chapter 6 shows the results obtained applying the proposed strategy and Chapter 6.4 sets out the next steps of this study.

Contents

List of Figures

List of Tables

Chapter 1

Introduction to Accelerator Physics

1.1 Beam Dynamics

In a circular particle accelerator, a charged particle is guided along the circular closed orbit thanks to electromagnetic fields that exert the driving force as a *Lorentz Force*

$$\vec{\mathbf{F}} = e(\vec{\mathbf{E}} + \vec{\mathbf{v}} \times \vec{\mathbf{B}})\,. \tag{1.1}$$

In high-energy accelerators, like the CERN Large Hadron Collider (LHC), electric fields, $\vec{E}$, are used to accelerate particles along the longitudinal direction of the motion, while magnetic fields, $\vec{B}$, are dedicated to focus and bend particles along the transverse plane.

1.1.1 Transverse Dynamics

In Fig. 1.1 is represented the ideal orbit (i.e. the *reference orbit*), along which a charged particle should be guided by the field generated by the synchrotron magnets.

Considering a particle in a local co-ordinate system (x, z, s), moving with a velocity $\vec{v} = (\dot{x}, \dot{z}, \dot{s})$, subjected to a constant magnetic field along the $\hat{z}$ direction,

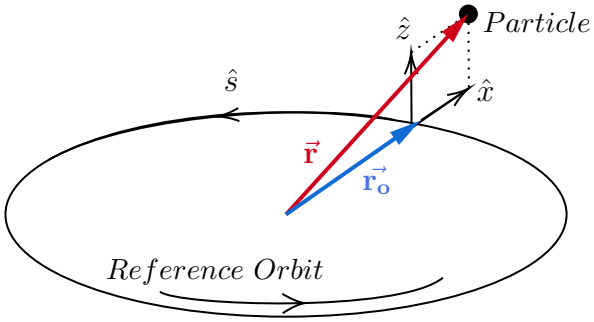

Figure 1.1: The curvilinear co-ordinate system for particle motion in a circular particle accelerator [1].

B_z, the condition that allows the particle to remain on the reference orbit is that the *centripetal force*, acting on it, should be equal to the *Lorentz force*, namely

$$\frac{m\vec{\mathbf{v}}^2}{\vec{\mathbf{r}_0}} = e\vec{\mathbf{v}}B_z. \tag{1.2}$$

From this derives the definition of *beam rigidity*, i.e. the stiffness of the circulating beam under the effect of the magnetic guiding forces

$$B_z\rho = \frac{p}{e}, \tag{1.3}$$

where p is the particle momentum and ρ is the bending radius of the orbit. Eq. (1.3) involves that, for maintaining the radius of the reference orbit constant, B_z must change as the momentum of the charged particle increases.

At this point, it is important to find the equation of motion of a charge particle in a circular accelerator. Given the fact that the transverse dimensions of the beam are small, it is possible to expand the magnetic field around the nominal closed orbit using a *multipolar expansion* [2]

$$B_z + iB_x = \frac{p}{e}\sum_n (ia_n + b_n)(x + iz)^{n-1}, \tag{1.4}$$

where a_n and b_n represent the skew and normal magnetic components, respectively. As a result of this, the magnetic field around the beam is described by a superposition of multipoles, each with a specific effect on the particle's dynamics: b_0 represents the dipolar field that bends the beam, b_1 the quadrupolar field that focuses the beam,

b_2 and b_3 the sextupolar and octupolar fields, respectively, that are generating higher-order effects in the dynamics.

Considering only the dipole and quadrupole contributions to the magnetic field, beam dynamics is linear, and thus it is possible to find the two decoupled equations of transverse motion, an inhomogeneous differential equation and a homogeneous one [3]

$$x''(s) + \left(\frac{1}{\rho^2(s)} - k(s)\right) x(s) = \frac{1}{\rho(s)} \frac{\delta p}{p_0} \tag{1.5}$$

$$z''(s) + k(s) z(s) = 0, \tag{1.6}$$

where k (i.e. the quadrupole strength) and ρ (i.e. the radius of curvature) are naturally periodic with a periodicity equal to the circumference of the accelerator, and $\delta p/p_0$ represents the deviation of the particle's momentum from the nominal one. The two Eqs. (1.5) and (1.6), are called *Hill Equations*[1].

In a generic homogeneous case, i.e.

$$x''(s) + k(s) x(s) = 0, \tag{1.7}$$

the general solution of the Hill equation can be expressed as

$$\begin{aligned} x(s) &= C(s) x_0 + S(s) x_0' \\ x'(s) &= C'(s) x_0 + S'(s) x_0' \,, \end{aligned} \tag{1.8}$$

and they can also be expressed in function of the so-called *transfer matrix* $\mathcal{M}$

$$\begin{bmatrix} x \\ x' \end{bmatrix} = \underbrace{\begin{bmatrix} C(s) & S(s) \\ C'(s) & S'(s) \end{bmatrix}}_{\mathcal{M}} \begin{bmatrix} x_0 \\ x_0' \end{bmatrix} , \tag{1.9}$$

where $C(s)$ is the solution of the homogeneous equation with $C(s) = 1$ and $S(s) = 0$ while $S'(s)$ is the solution with $C'(s) = 0$ and $S'(s) = 1$. For this choice, $C(s)$

[1]These equations were introduced in 1886 by George W. Hill in his studies on periodic differential equations.

and $S(s)$ are called cosine-like and sine-like functions, respectively, or principal trajectories.

To find the solution of the inhomogeneous Hill Eq. (1.5) it is necessary to introduce a dispersion term $D(s)$

$$\begin{aligned} x(s) &= C(s)x_0 + S(s)x_0' + D(s)\frac{\delta p}{p_0} \\ x'(s) &= C'(s)x_0 + S'(s)x_0' + D'(s)\frac{\delta p}{p_0}\,. \end{aligned} \tag{1.10}$$

The solution of $D(s)$ with $D_0 = 0$ and $D_0' = 0$ is

$$D(s) = S(s)\int_0^s d\sigma \frac{1}{\rho(\sigma)}C(\sigma) - C(s)\int_0^s d\sigma \frac{1}{\rho(\sigma)}S(\sigma)\,, \tag{1.11}$$

and in the inhomogeneous case, the form of the transfer matrix is

$$\begin{bmatrix} x \\ x' \\ \delta p/p_0 \end{bmatrix} = \begin{bmatrix} C(s) & S(s) & D(s) \\ C'(s) & S'(s) & D'(s) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_0' \\ \delta p/p_0 \end{bmatrix}\,. \tag{1.12}$$

The solution of the homogeneous Hill Eq. (1.5) can be found also by using Floquet's theorem, which provides the following form

$$u(s) = \sqrt{2J\beta(s)}\cos(\psi(s) + \psi_0) \qquad \text{with } u = \{x, z\}. \tag{1.13}$$

The trajectory function $u(s)$ describes a transverse oscillation about the closed orbit, the so-called betatron oscillation [3]. The quantities J and ψ_0 in Eq. (1.13) are constants determined by the initial conditions, $\sqrt{2J\beta(s)}$ is the amplitude of the oscillation, and $(\psi(s) + \psi_0)$ its phase. β is one of the three *Twiss parameter*[2] called beta function or amplitude function.

Remembering that the quadrupole focusing strength $k(s)$ is periodic with the same period of $\beta(s)$, it is possible to define the betatron tune:

$$Q_u = \frac{\text{Nr. Oscillation Periods}}{\text{Accelerator Circumference}} = \frac{1}{2\pi}\int_0^L \frac{ds}{\beta_u(s)} \qquad \text{where} \qquad u = x, z\,. \tag{1.14}$$

[2] β is the Twiss parameter that defines the other two: $\alpha = -\beta'(s)/2$ and $\gamma = (1+\alpha)^2/\beta$.

The betatron tune is defined by the quadrupole arrangement in the machine. The horizontal Q_x and vertical Q_z tunes define the so-called working point of an accelerator and they values should be chosen to avoid resonances [3].

Emittance

The general solutions of the Hill equations, written as a function of $\beta(s)$, describe the fact that a particle with stable motion in an accelerator follows an ellipse in the phase space (u, u'):

$$\begin{cases} u(s) = \sqrt{2J\beta(s)}\cos(\psi(s) + \psi_0) \\ u'(s) = \sqrt{\frac{2J}{\beta(s)}}(\sin(\psi(s) + \psi_0)) \end{cases} \tag{1.15}$$

By squaring the two previous equations we obtain:

$$2J = \gamma(s)u^2(s) + 2\alpha(s)u(s)u'(s) + \beta(s)u'^2(s) = \varepsilon \tag{1.16}$$

where $\alpha = -\beta'(s)/2$, $\gamma = (1+\alpha)^2/\beta$ and $2J$ is the Courant-Snyder invariant, J being the action, which is equal to the emittance ε. The Eq. (1.16) represents an ellipse in the phase space (u, u'), whose area is $A = 2\pi J = \pi\varepsilon$ (see Fig. 1.2).

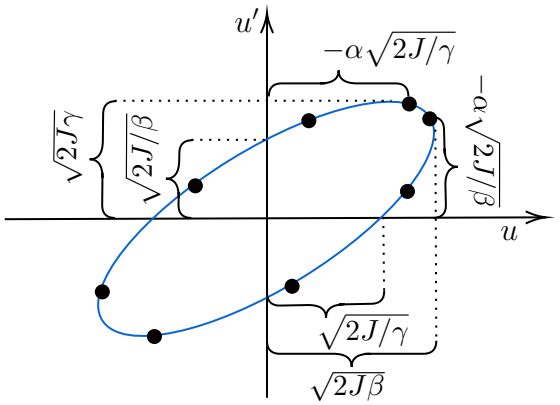

Figure 1.2: The Phase-space ellipse described by the particle motion in (u, u').

[3] For an nth order resonance the particle's trajectory will be closed after n turns, amplifying the effect of the magnetic imperfections.

Thus, the emittance can be defined as the area of the phase-space ellipse divided by π. Considering an ensemble of particles in the beam, the beam emittances is obtained by averaging the emittance of the individual particles. Given that the particles are generically Gaussian distributed, it is customary to consider values of the beam emittance obtained by considering a representative particle with an amplitude corresponding to one rms of the beam distribution.

1.1.2 Longitudinal Dynamics

The first particle accelerators arose by exploiting the ability of the electric field to accelerate or decelerate charged particles. The initial idea was to insert the particle in an area subject to a electric potential difference that would allow the particle to increase its speed [4].

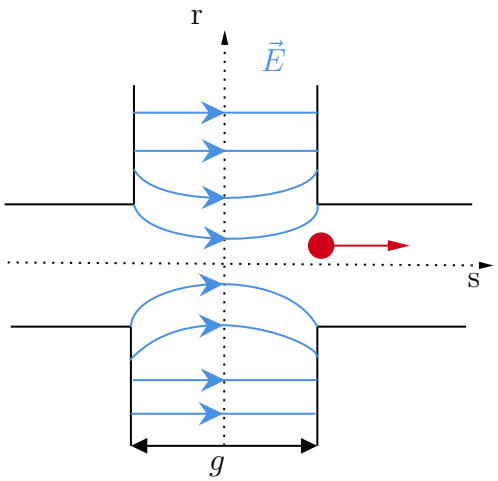

Figure 1.3: Schematic representation of a Radio Frequency cavity.

Figure 1.3 shows a typical acceleration cavity called *Radio Frequency* cavity in which a variable electric field is present. Across the gap, g, the amplitude of the RF voltage is V_{RF} and the particle crosses the gap at a distance r. In this system, the energy gain is

$$\Delta E = e \int_{-g/2}^{g/2} \vec{E}(s, r, t) \cdot d\vec{s}, \tag{1.17}$$

where $E(s, r, t) = E_1(s, r) \cdot E_2(t)$ is the electric field in the cavity gap. If we consider

a simplified model in which

$$E_1(s,r) = \frac{V_{RF}}{g} = constant \tag{1.18}$$

$$E_2(t) = \sin(\omega_{\mathrm{RF}} t + \phi_0), \tag{1.19}$$

for $t = 0$, $s = 0$ and $v \neq 0$ parallel to the electric field, the energy gain becomes

$$\Delta E = eV_{\mathrm{RF}} T_{\mathrm{a}} \sin\phi_0, \tag{1.20}$$

where T_{a} is called the *transit time factor*. The latter is related to the time required to fill the cavity with the electric field, and it is defined as

$$T_{\mathrm{a}} = \frac{\sin\left(\omega_{\mathrm{RF}} g/2v\right)}{\omega_{\mathrm{RF}} g/2v}\,. \tag{1.21}$$

Thus, the transit-time factor conveys the effective acceleration voltage available, and for $g \to 0$, $T_{\mathrm{a}} \to 1$, highlighting the importance of sufficiently narrow RF cavities. Now, it is evident how the longitudinal motion is strongly correlated to the synchronisation between the particles and the acceleration system. Thus, for the latter to be synchronised, the following synchronism condition must occur

$$f_{\mathrm{rev}} = f_{\mathrm{RF}}/h. \tag{1.22}$$

The ratio between the RF frequency and the revolution frequency must be equivalent to the harmonic number, which indicates the maximum number of beam bunches it is possible to load into the accelerator.

At this point, we consider that a particle beam is never mono-energetic, but rather distributed around the nominal value of the momentum. Thus, we define a factor that binds the momentum variation, Δp, to the trajectory variation of the particles, ΔC, the momentum compaction factor α_{p}

$$\frac{\Delta p}{p}\alpha_{\mathrm{p}} = \frac{\Delta C}{C}\,. \tag{1.23}$$

Hence, for momenta greater than the nominal one also the particle trajectory will be longer. The revolution frequency is defined as $f_{\mathrm{rev}} = v/C$. Keeping in mind that for lower energy the velocity will increase faster than the trajectory, while at

higher energy the velocity will be almost constant ($v \simeq c$), we can affirm that there is an energy for which the variation of speed is balanced by the variation of the trajectory called transition energy and defined as

$$\gamma_{\mathrm{tr}} = \sqrt{1/\alpha_{\mathrm{p}}}\,. \tag{1.24}$$

Figure 1.4 represents the longitudinal periodic motion of particles in an accelerating bucket around the nominal phase, for a beam energy below the transition energy. This kind of motion is called synchrotron oscillation and is caused by the distribution of particles' momentum around the nominal value $\Delta p/p = 0$.

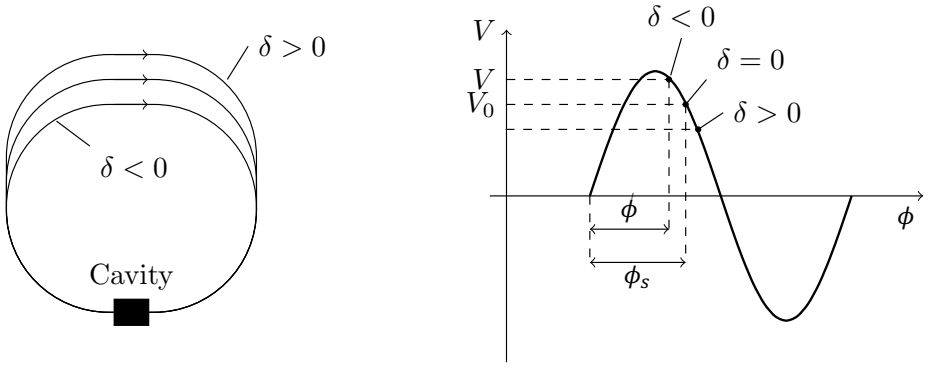

Figure 1.4: The principle of synchrotron oscillations for relativistic particles ($v \approx c$), where $\delta = \Delta p/p$. The effect of a momentum deviation different from zero is illustrated [3].

An ideal on-momentum particle follows the orbit fixed by the construction of the machine, and has a nominal phase ϕ_{s}, called *synchronous phase*. A particle with a lower momentum $\Delta p/p < 0$ will travel a smaller path, arriving earlier in the cavity and thus seeing a larger voltage in the RF cavity. This will then compensate for the lack of energy, and step by step it will approach the ideal particle. Lastly, a particle with a higher momentum $\Delta p/p > 0$ will travel a longer path seeing a smaller voltage in the cavity and again step by step will approach the ideal particle. In conclusion, the effect of an oscillating longitudinal electric field is twofold: on

the one hand, it provides the energy needed to accelerate the beam, and on the other hand, it is also generating a longitudinal focusing of the beam particles.

Figure 1.5 shows a stationary (left) and an accelerating (right) bucket. The separatrix is shown in red and separate the phase space in two regions: a stable one (inside the separatrix and around the origin) and an unstable one (outside the separatrix), The number of buckets along the ring circumference corresponds to the harmonic number h.

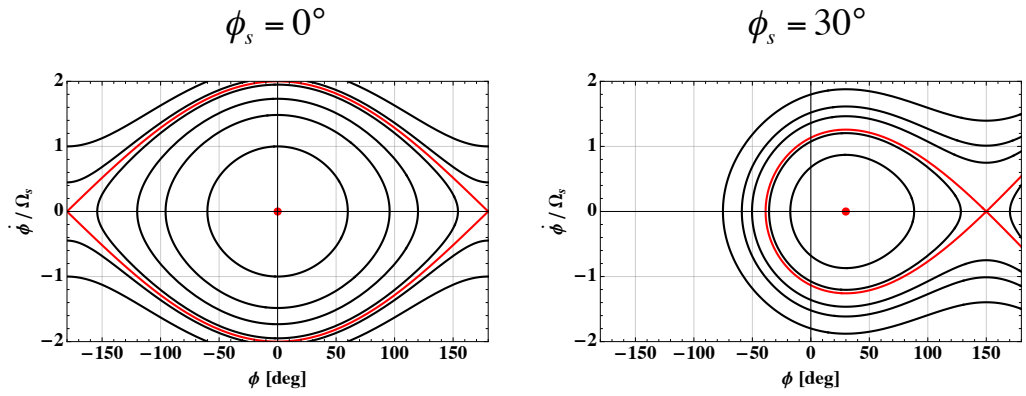

Figure 1.5: Phase-space portrait for stationary (left) or accelerating (right) bucket [5].

Chapter 2

Particle colliders and Luminosity

Since the beginning of the XXth century, particle accelerators have been an indispensable tool for research in high-energy physics. In this field, circular accelerators outperform linear accelerators thanks to their higher acceleration performance, thanks to the multi-pass feature.

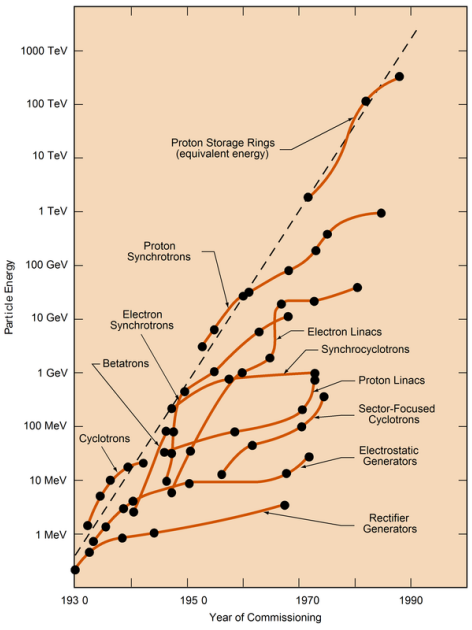

Figure 2.1: A Livingston plot showing the evolution of effective accelerator energy from 1930 to 1990 [6].

Figure 2.1 shows a Livingstone plot reporting the time evolution of accelerators' performance [6].

Today, nuclear and subnuclear physics experiments are performed mainly using large particle accelerators and can be divided into *fixed-target experiment* and *colliding-beam experiment*. The former is mainly used for nuclear physics, while the latter for subnuclear physics, where the energy available for the new processes production plays a fundamental role. The higher centre-of-mass energy will enable the discovery of heavier particles, still unknown, or the determination with high accuracy of the properties of known heavy particles. The highest centre-of-mass energy can be achieved only thanks to particle collider (e.g. the CERN LHC), as can be highlighted thanks to some kinematic calculations [7].

If we consider a collision between two particles of masses m_1 and m_2 the total centre-of-mass energy, E_{cm}, can be expressed as

$$(\mathbf{p}_1 + \mathbf{p}_2)^2 = E_{\mathrm{cm}}^2 = (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2, \tag{2.1}$$

where $\mathbf{p}_i$ is the two ith particle four-momentum, E_i is its energy, and $\vec{p}_i$ its linear momentum. In case of a fixed-target experiment, $\vec{p}_2 = 0$, and thus

$$E_{\mathrm{cm}} = \sqrt{m_1^2 + m_2^2 + 2m_2 E_{1,\mathrm{lab}}}, \tag{2.2}$$

where $E_{1,\mathrm{lab}}$ is the total energy of the first particle in the laboratory reference frame. For a colliding beam experiment, $\vec{p}_1 = -\vec{p}_2$, and thus

$$E_{\mathrm{cm}} = E_1 + E_2. \tag{2.3}$$

Hence, if we consider for instance two protons accelerated up to 7 TeV, in a colliding beam experiment the centre-of-mass energy will be 14 TeV, while in a fixed target one it will be about 115 GeV.

While the centre-of-mass energy determines the mass of the products of a collision, the luminosity determines the number of events after a given period of time. The luminosity is defined as

$$L = \frac{dR/dt}{\sigma_{\mathrm{p}}} \quad [\mathrm{cm}^{-2}\,\mathrm{s}^{-1}], \tag{2.4}$$

where σ_{p} is the cross section of the physical process under consideration and dR/dt represents the interaction rate. From that it is possible to define the total number of events collected during a data-taking run as

$$N_{\mathrm{e}} = \sigma_{\mathrm{p}} \int_0^T dt\, L(t) = \sigma_{\mathrm{p}}\, \mathcal{L}_{\mathrm{int}}, \tag{2.5}$$

where T is the total time for physics and $\mathcal{L}_{\mathrm{int}}$ is the integrated luminosity, usually measured in fb^{-1} (i.e. 10^{39} cm^{-2}) [3].

The processes studied in present experiments have smaller and smaller cross-sections, i.e. lower than the nanobarn. For this reason, it is necessary to increase the luminosity to maximise the number of events collected, N_{e}, to achieve an appropriate statistics in a reasonable time.

2.1 The CERN LHC

The CERN LHC is located in the accelerator complex of CERN (Conseil Europé en pour la Recherche Nucléaire) on the Swiss-French border near Geneva. It is the largest and most powerful particle accelerator currently in operation, and it works mainly with proton-proton collisions, but it is also used as ion collider.

The LHC was designed to reach a centre of mass energy of 14 TeV, and for this reason, proton beams must pass through an injector chain (shown in Fig. 2.2) that allows them to gain more and more energy [8].

During Run 2 operations (from 2015 to 2018), the proton source was obtained by subjecting hydrogen to a very intense electric field that ionises the gas, leaving only protons at 50 MeV of total energy at the inlet of the linear accelerator LINAC2 [9]. Here the protons were accelerated up to 160 MeV and subsequently entered the first circular accelerator of the injector chain, the Proton Synchrotron Booster (PSB), which accelerated them up to 1.4 GeV. The PSB and the following Proton Synchrotron (PS) ring determined the bunch train structure of the beams so that it was compatible with the LHC RF system. The PS accelerated protons up to 26 GeV, and then the Super Proton Synchrotron (SPS) accelerated them further up to 450 GeV. In the end, LHC accelerated each beam up to 7 TeV. During the Long

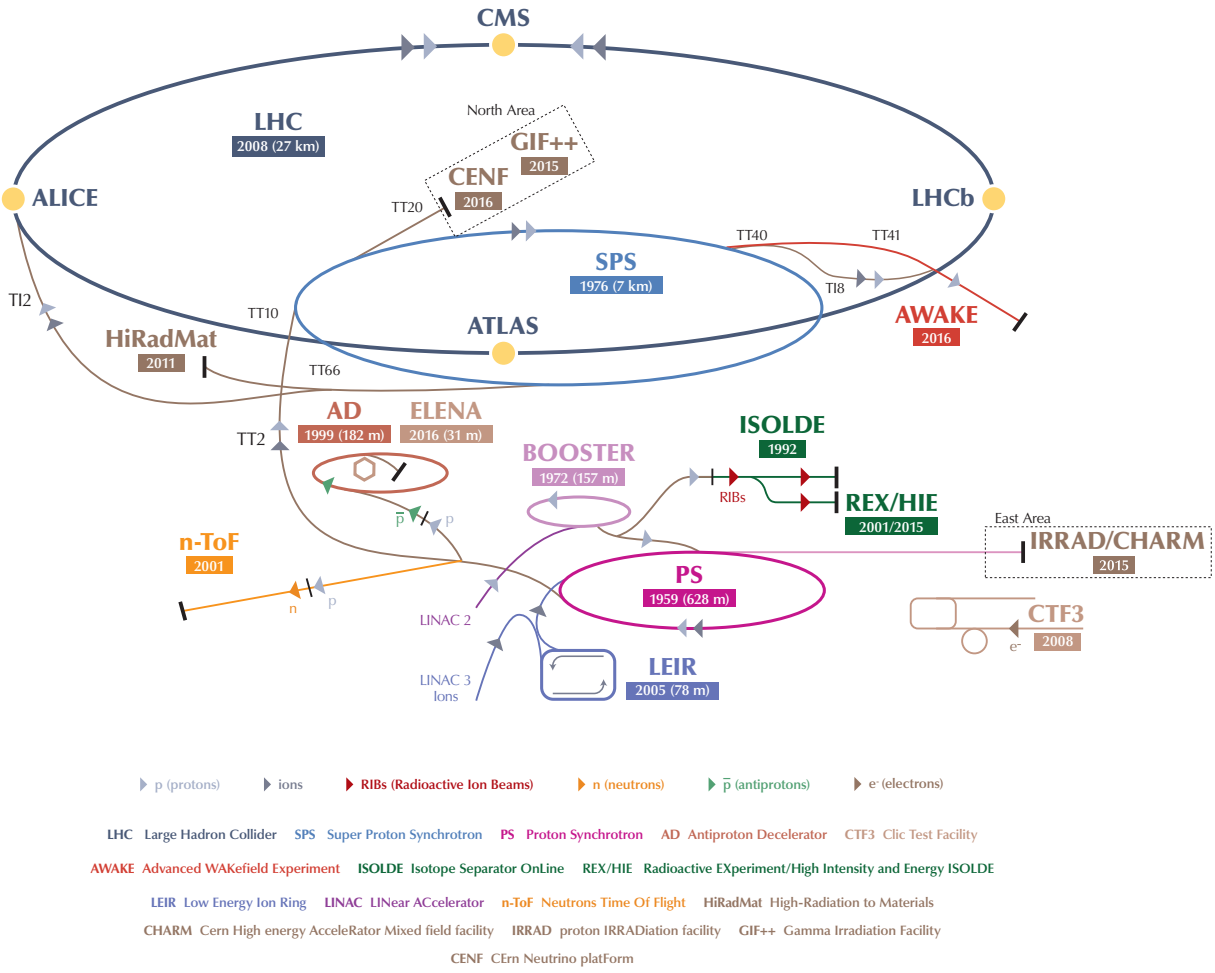

Figure 2.2: The CERN accelerator complex in use during Run 2, showing the path of protons and ions at the injection (the accelerators are not drawn to scale) [2].

Shutdown 2 (LS2), from 2019 to 2020, a massive improvement program called LHC Injectors Upgrade (LIU) was put in place [10]. Hence, the injection chain described above has been modified by replacing LINAC 2 with LINAC 4 and updating PSB, PS and SPS considering the new working conditions expected for high-luminosity LHC.

LHC is built in an underground tunnel of about 27 km in circumference at about 100 meters depth, and, as shown in Fig. 2.3, it consists of 8 arcs interspersed with rectilinear insertion regions (IRs). The insertion regions contain both the

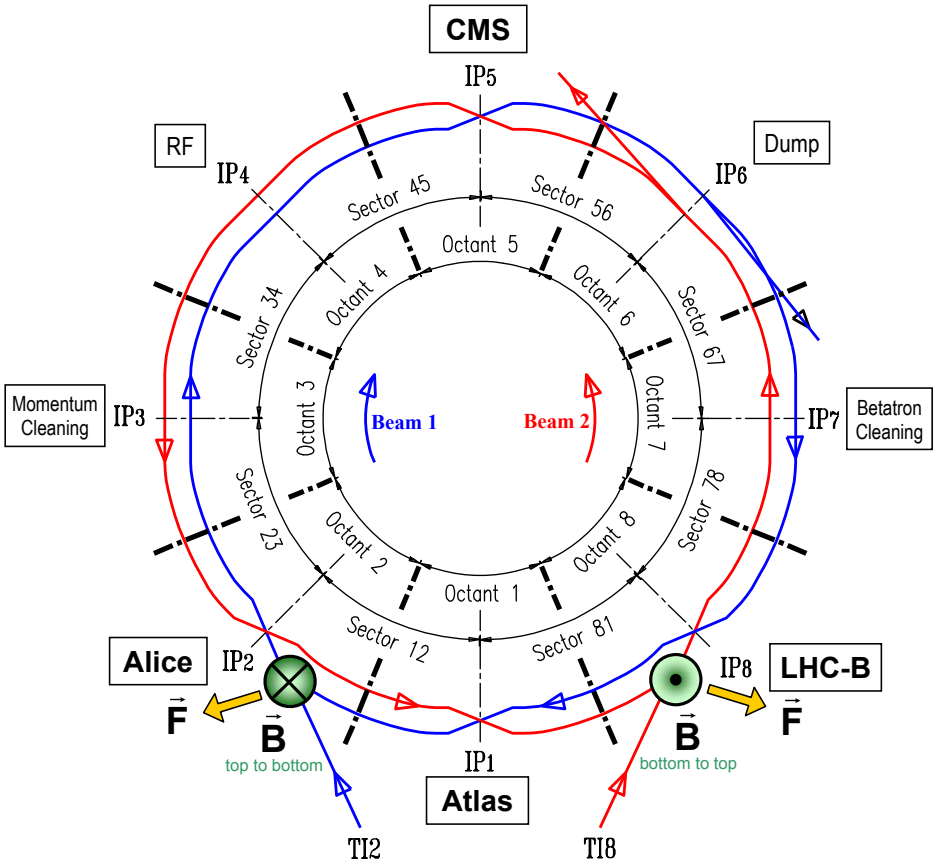

Figure 2.3: Schematic of the LHC rings [11].

interaction points (IPs), where the two LHC beams collide and some accelerator systems [12]. In the interaction points IP1, IP5, IP2 and IP8, we find respectively: ATLAS and CMS, the two LHC detectors that work in conditions of high luminosity, $L_{\text{Peak}} = 2 \cdot 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ [13] (obtained by colliding the beams completely head-on and keeping β^* as small as possible); ALICE and LHCb, which instead work with a much lower luminosity than their colleagues, $L_{\text{Peak}}^{ALICE} = 1 \cdot 10^{31}\,\text{cm}^{-2}\,\text{s}^{-1}$ and $L_{\text{Peak}}^{LHCb} = 2 \cdot 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$ [13].

Each LHC arc consists of 23 cells 106.9 m long, schematically represented in Fig. 2.4. Each cell is composed of six dipoles and two quadrupole magnets, plus a number of ancillary magnets that are supposed to correct the closed orbit, the tune, the linear coupling, the chromaticity and the non-linear field quality of the dipoles.

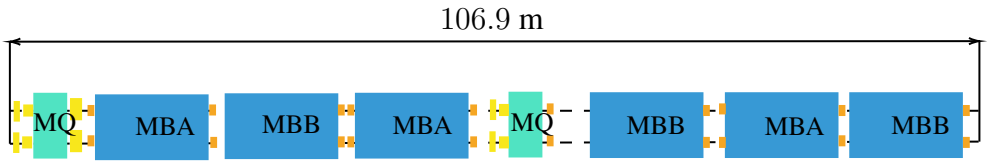

Figure 2.4: Schematic layout of an LHC arc cell. MBA/B are the bending dipoles, MQ are the focusing quadrupoles (both superconducting). The smaller magnets in between are dipolar (orbit), quadrupolar (tune) and higher order correctors [14].

2.2 Luminosity Computation

It is possible to derive the expression of the luminosity in terms of the collider parameters. Colliders typically use bunched beams, thus, Fig. 2.5 sketches the case of two colliding bunches moving through each other [7].

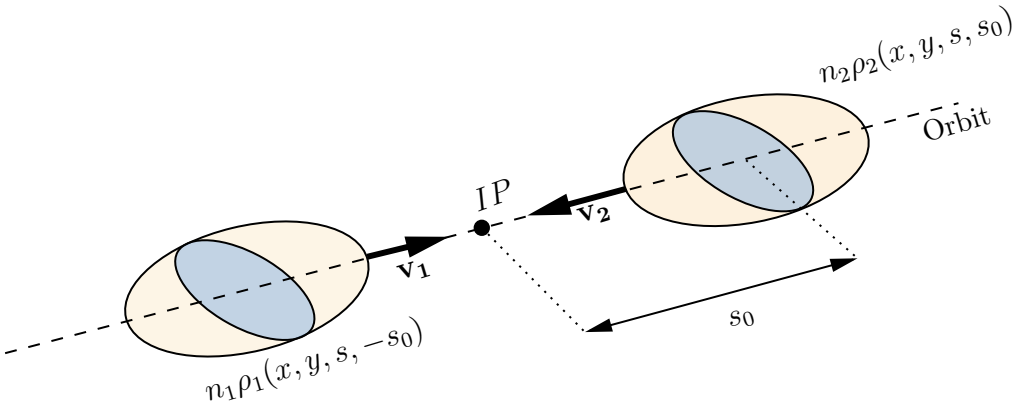

Figure 2.5: Schematic view of the interaction of two colliding beams.

In the considered case, the overlap integral of the two distributions is proportional to the luminosity as follows

$$L = n_1 n_2 K \int\int\int\int_{-\infty}^{+\infty} dx dy ds ds_0 \, \rho_1(x, y, s, -s_0)\rho_2(x, y, s, s_0), \tag{2.6}$$

where $s_0 = ct$, $K = \sqrt{(\mathbf{v_1} - \mathbf{v_2})^2 - (\mathbf{v_1} \times \mathbf{v_2})^2/c^2}$ is the kinematic factor, which is equal to $2c$ for symmetric colliders, i.e. $\mathbf{v}_1 = -\mathbf{v}_2$ and $\mathbf{v}_1 \times \mathbf{v}_2 = 0$ [15], and ρ_1 and ρ_2 are the time-dependent, normalised beam density distribution functions, n_i

being the number of particles in the ith bunch.

At this point, considering only head-on collisions, i.e. $\mathbf{v_1} = -\mathbf{v_2}$, of equal beams, with uncorrelated Gaussian density distributions, we obtain

$$L = \frac{2n_1 n_2 f_{\text{rev}} k_b}{(\sqrt{2\pi})^6 \sigma_s^2 \sigma_x^{*2} \sigma_y^{*2}} \int\int\int\int dx dy ds ds_0 \, e^{-\frac{x^2}{\sigma_x^{*2}}} e^{-\frac{y^2}{\sigma_y^{*2}}} e^{-\frac{s^2}{\sigma_s^2}} e^{-\frac{s_0^2}{\sigma_s^2}}, \tag{2.7}$$

where f_{rev} is the revolution frequency, k_{b} is the number of colliding-bunch pairs, σ_s is the rms bunch length, and $\sigma^*_{x,y}$ represent the transverse rms beam size at the interaction point.

Integrating the previous equation, considering that

$$\int_{-\infty}^{+\infty} e^{-at^2} dt = \sqrt{\pi/a} \tag{2.8}$$

we obtain the following expression for the luminosity

$$L = \frac{n_1 n_2 f_{\text{rev}} k_{\text{b}}}{4\pi \sigma_x^* \sigma_y^*}, \tag{2.9}$$

that can be rewritten in the following form, considering that $\sigma_x^* = \sigma_y^* = \sqrt{\beta^* \varepsilon / \gamma_{\text{r}}}$

$$L = \frac{n_1 n_2 f_{\text{rev}} k_{\text{b}} \gamma_{\text{r}}}{4\pi \beta^* \varepsilon}. \tag{2.10}$$

where γ_r is the relativistic factor, β^* is the amplitude function in the interaction point and ε is the emittance. Thus, to increase luminosity, there are several possibilities: the amplitude function (or β function) could be kept small, the number of particles per bunch or the number of bunches in the machine could be intensified, and finally, the total energy could be increased [16].

2.3 Reduction Factors

Equation (2.10) represents the maximum luminosity value that can provide an ideal collider. This value, however, might be lowered by some effects, which occur in real machines. There are principally three additional effects that we need to consider: the crossing angle, the collision offset, and the hourglass effect.

2.3.1 Crossing Angle

The crossing angle is a voluntarily-inserted effect needed to avoid unwanted collisions outside of the designated interaction point [16].

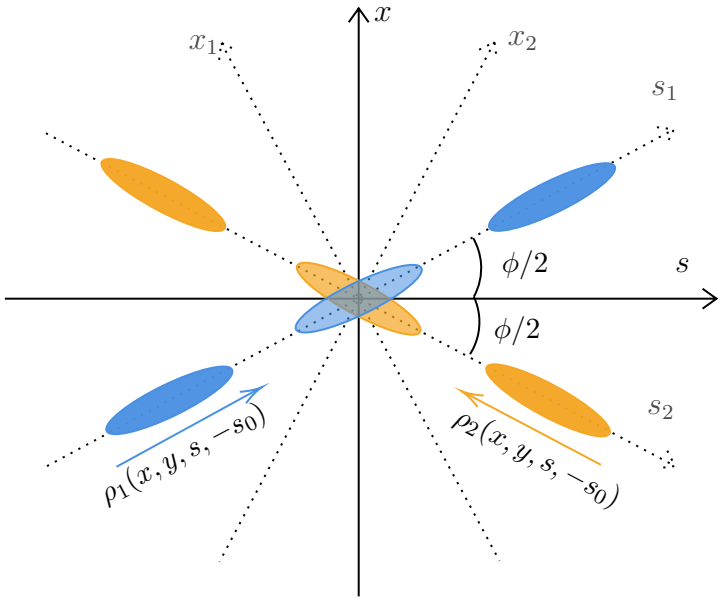

Figure 2.6: Rotated reference system for collision at a finite crossing angle with a schematic view of two beams colliding.

Assuming that the crossing angle is in the horizontal plane (as displayed in Fig. 2.6), we can evaluate the overlap integrals performing a co-ordinates transformation [7]. From $x - s$, we go to $x_1 - s_1$ for Beam 1 and $x_2 - s_2$ for Beam 2, where

$$x_1 = x\cos(\phi/2) - s\sin(\phi/2) \qquad s_1 = s\cos(\phi/2) + x\sin(\phi/2)$$
$$x_2 = x\cos(\phi/2) + s\sin(\phi/2) \qquad s_2 = s\cos(\phi/2) - x\sin(\phi/2).$$

In this situation, the overlap integral becomes

$$L = A\int\int\int\int_{-\infty}^{+\infty} dxdydsds_0\,\rho_{1x}(x_1)\rho_{1y}(y_1)\rho_{1s}(s_1-s_0)\rho_{2x}(x_2)\rho_{2y}(y_2)\rho_{2s}(s_2+s_0), \tag{2.11}$$

where $A = 2\cos^2(\phi/2)n_1n_2f_{\mathrm{rev}}k_{\mathrm{b}}$ and $2\cos^2(\phi/2)$ is the kinematic factor for bunch velocities which are not collinear. Computing the previous integral, considering

that $\int_{-\infty}^{-\infty} e^{-(at^2+bt+c)}dt = \sqrt{\pi/a}e^{\frac{b^2-ac}{a}}$, leads to

$$L = \frac{n_1 n_2 f_{\text{rev}} k_{\text{b}}}{4\pi\sigma_x^*\sigma_y^*} F_{\text{c}} \tag{2.12}$$

where F_{c} is the reduction factor for the crossing angle, which for small crossing angles and $\sigma_s >> \sigma_{x,y}^*$ assumes the form

$$F_{\text{c}} = \frac{1}{\sqrt{1 + (\frac{\sigma_s}{\sigma_x^*}\frac{\phi}{2})^2}}, \tag{2.13}$$

where $\frac{\sigma_s}{\sigma_x}\frac{\phi}{2}$ is called the Piwinski angle.

2.3.2 Collision Offset

The collision offset can be voluntarily introduced in the machine to perform some measurements or to adjust the trajectory of the beams, but it can also arise spontaneously as a result, e.g. of ground motion or beam-beam effects [7].

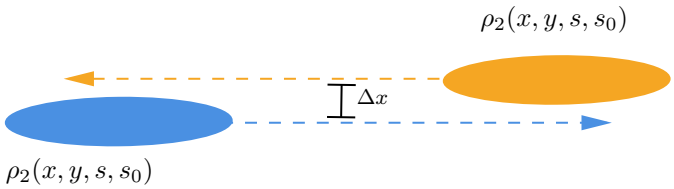

Figure 2.7: Schematic view of two bunches colliding with a transverse offset Δx in the horizontal plane.

In Fig. 2.7 is presented the simplest case of transverse offset of colliding beams, where the reduction factor F_{o} is

$$F_{\text{o}} = \exp\left(-\frac{\Delta x^2}{4\sigma_x^{*2}}\right). \tag{2.14}$$

It could also happen to have both crossing angle and transverse offset. In this case, in addition to the two reduction factors, F_{c} and F_{o}, it is necessary to add to the luminosity expression a relative factor,

$$F_{\text{c-o}} = \exp\left(\frac{(\frac{\Delta x \sin(\phi/2)}{2\sigma_x^{*2}})^2}{\frac{\sin^2(\phi/2)}{\sigma_x^{*2}} + \frac{\cos^2(\phi/2)}{\sigma_s^{*2}}}\right). \tag{2.15}$$

2.3.3 Hourglass effect

Equation (2.10) assumes that the beta function is constant and equal to β^* along the longitudinal direction, s. However, this assumption is correct only if $\beta^* >> \sigma_s$, otherwise, for $\beta^* \sim \sigma_s$, the variation of the beta function along the longitudinal direction cannot be neglected anymore and it is

$$\beta(s) = \beta^* \left[1 + \left(\frac{s}{\beta^*} \right)^2 \right], \tag{2.16}$$

should be taken into account, as it implies a luminosity reduction described by the reduction factor F_{h}. In this case, the luminosity can be expressed as

$$L = \frac{n_1 n_2 f_{\mathrm{rev}} k_{\mathrm{b}}}{4\pi \sigma_x^* \sigma_y^*} \underbrace{\frac{\cos(\phi/2)}{\sqrt{\pi}\sigma_s} \int_{-\infty}^{+\infty} ds \frac{e^{-s^2 A(s)}}{1 + (s/\beta^*)^2}}_{F_{\mathrm{h}}}, \tag{2.17}$$

where

$$A(s) = \frac{\sin^2(\phi/2)}{(\sigma_x^*)^2 [1 + (s/\beta^*)^2]}. \tag{2.18}$$

The hourglass reduction factor cannot be computed analytically, and it has to be computed numerically, and some examples are given in Table 2.1. Considering the relation between the beam size and the beta function,

$$\sigma_{x,y}(s) = \sigma_{x,y}^* \sqrt{1 + \left(\frac{s}{\beta_{x,y}^*} \right)^2},$$

also the beam size is not constant but increases approximately linearly with the distance from the interaction point.

In Fig. 2.8 the beta-function variation with respect to the distance from the interaction point is shown for two values of β^*. From the plot, it is also possible to understand why this effect is called the hourglass effect.

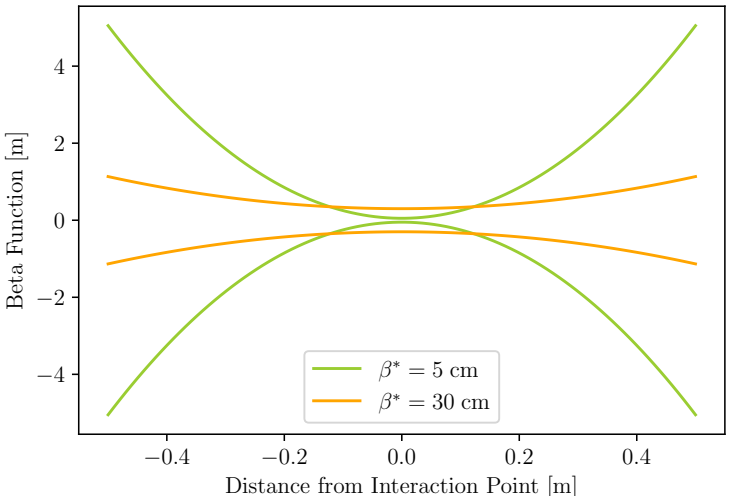

Figure 2.8: Schematic illustration of the hourglass effect. The beta function is plotted for two values of β^*.

Table 2.1 shows how the previously described effects impact the LHC luminosity production using the nominal beam parameters [16].

Table 2.1: Typical values of LHC parameters and the corresponding reduction factors [14].

Effect	**Reduction Factor F**		
Crossing Angle	$F_{\mathrm{c}} = 0.84$	$\phi = 285\,\mu\mathrm{rad}$	$\sigma_s = 7.5\,\mathrm{cm}$
Collision Offset	$F_{\mathrm{o}} = 1$	$\Delta x = 0$	
	$F_{\mathrm{o}} = 0.779$	$\Delta x = 1\sigma$	
	$F_{\mathrm{o}} = 0.368$	$\Delta x = 2\sigma$	
	$F_{\mathrm{o}} = 0.105$	$\Delta x = 3\sigma$	
	$F_{\mathrm{o}} = 0.018$	$\Delta x = 4\sigma$	
	$F_{\mathrm{o}} = 0.002$	$\Delta x = 5\sigma$	
Hourglass	$F_{\mathrm{h}} = 1$	$\beta^*/\sigma_s > 7$	LHC
	$F_{\mathrm{h}} = 0.9$	$\beta^*/\sigma_s \sim 2$	HL-LHC

For typical LHC parameters, the crossing angle effect consists of a reduction

in luminosity of about 16%. However, as planned for High Luminosity LHC, crab cavities can reduce the impact of crossing angle on integrated luminosity, and Fig. 2.9 shows their working scheme.

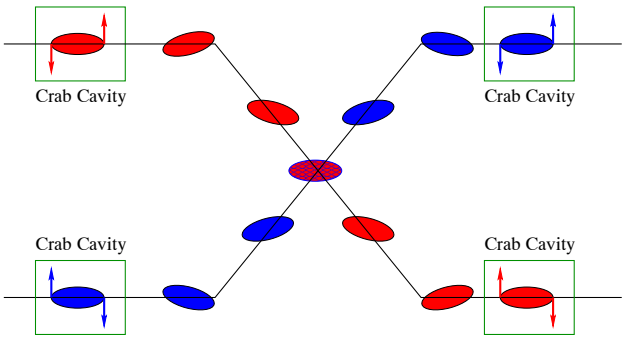

Figure 2.9: Concept of crab crossing scheme using RF cavities to maximise the bunch overlap at the collision points [17].

Moreover, looking at the table, it is evident that the collision offset effect becomes more and more intense as the distance between the beams increases. Lastly, the hourglass effect comes into play only when working in high luminosity conditions, i.e. with a very small β^*.

2.4 Beam-Beam Effects

Another phenomenon that impacts on the luminosity is the beam-beam interaction, as already mentioned in section 2.3.2.

A beam consists of a large number of charged particles and therefore each bunch exerts a force on the charges of the other beam, and the particles in each bunch are also interacting between them [18], although this effect is supposed to be negligible at the collision energy of modern colliders. To study the beam-beam force, we consider a bunch with Gaussian charge distribution of the type

$$\rho(x,y,z) = \frac{nq}{4\pi\epsilon_0} e^{-(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2}+\frac{z^2}{2\sigma_z^2})}, \tag{2.19}$$

and the electric potential in the beam's rest frame is given by

$$\phi(x,y,z) = \frac{nq}{\sqrt{\pi}} \int_0^\infty dt \, \frac{e^{-(\frac{x^2}{2\sigma_x^2+t} + \frac{y^2}{2\sigma_y^2+t} + \frac{z^2}{2\sigma_z^2+t})}}{\sqrt{(2\sigma_x^2+t)(2\sigma_y^2+t)(2\sigma_z^2+t)}}. \tag{2.20}$$

From the last equation, we can derive the transverse electric field considering that

$$\mathbf{E}(x,y,z) = -\nabla\phi(x,y,z). \tag{2.21}$$

To simplify the calculations, from this moment on we will consider round beams, i.e. with $\sigma_x = \sigma_y = \sigma$. In this approximation, the charge distribution becomes

$$\rho(x,y) = \frac{nq}{2\pi\sigma^2} e^{-(\frac{x^2+y^2}{2\sigma^2})}, \tag{2.22}$$

where n is the charge line density in the lab frame, and the radial electric field is given by the Gauss theorem

$$2\pi r E_r = \frac{1}{\epsilon_0} \int_0^r dr' \frac{2\pi n q r'}{2\pi\sigma^2} e^{-(r'^2/2\sigma^2)} = \frac{nq}{\epsilon_0}[1 - e^{-(r^2/2\sigma^2)}]. \tag{2.23}$$

Similarly, for the magnetic field

$$2\pi B_\phi = \mu_0 \int_0^r dr' 2\pi r' \beta c \rho(r'). \tag{2.24}$$

At this point, the force of the bunch on a single test particle in the other bunch is given by

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = q(E_r + \beta c B_\phi)\hat{\mathbf{r}}, \tag{2.25}$$

which, using $\epsilon_0\mu_0 = c^{-2}$, corresponds to

$$F_r = \frac{nq^2(1+\beta^2)}{4\pi\epsilon_0\sigma^2 r}[1 - e^{-r^2/2\sigma^2}]. \tag{2.26}$$

The graph in Fig. 2.10 represents the trend of the radial force as a function of the ratio r/σ. In the orange shaded area we have the situation of particles nearby the beam centre, i.e. $r << \sigma$. In this case, the particle will experience a linear defocusing effect with a gradient

$$k_x = k_y = \frac{nq^2(1+\beta^2)}{4\pi\epsilon_0\sigma^2}, \tag{2.27}$$

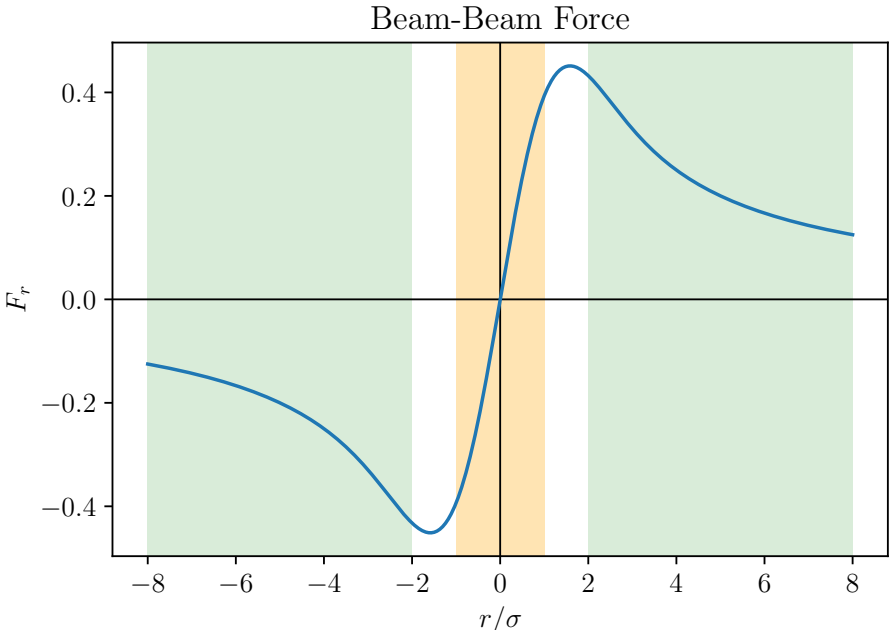

Figure 2.10: The blue line in the graph represents the trend of the beam-beam force, $y(r/\sigma) = (1 - e^{-r^2/2\sigma^2})/r$. The force is represented in arbitrary units. The orange shaded area indicates the linear behaviour of the force, while the green ones the deviation from the linear behaviour.

that generates a linear tune shift. Otherwise, the green shaded areas represent particles far from the centre, i.e. $r >> \sigma$. In this case, the effect is non-linear and it manifests as an amplitude-dependent tune shift, and for a bunch with many particles as a tune spread. This might bring to a poor beam lifetime and a growth of the emittance that leads to poor luminosity lifetime.

Particles that experience the linear effect are subjected to a quadrupole-like kick. This kick can be evaluated considering the two-dimensional force and multiplying it with the longitudinal distribution

$$F_r(r, s, t) = \frac{nq^2(1+\beta^2)}{4\pi\epsilon_0\sigma_s r}[1 - e^{-r^2/2\sigma^2}][e^{-(s+vt)^2/2\sigma^2}]. \tag{2.28}$$

Now, integrating $F_r(r, s, t)$, we find the radial deflection

$$\Delta r' = \frac{1}{mc\beta\gamma}\int_{-\infty}^{+\infty} dt\, F_r(r, s, t) = \frac{2nr_0}{\gamma r}[1 - e^{-r^2/2\sigma^2}] \tag{2.29}$$

where $r_0 = e^2/4\pi\epsilon_0 mc^2$ is the classical particle radius. For small amplitudes, the asymptotic limit is

$$\Delta r'|_{r\to 0} = \frac{nr_0}{\gamma\sigma^2} r = kr, \tag{2.30}$$

where k is the gradient induced by the beam-beam force, which is equivalent to the action of a thin lens element of gradient k, which has the effect of changing the tune of a quantity

$$\Delta Q = \frac{1}{4\pi}\frac{\beta^* n r_0}{\gamma\sigma^2} = \frac{1}{4\pi}\frac{n r_0}{\gamma\varepsilon}, \tag{2.31}$$

that corresponds to the so-called *beam-beam parameter* ξ.

In conclusion, we can write the luminosity in terms of the beam-beam parameter so to make it explicit the impact of this interaction on the luminosity

$$L = \frac{k f_{\text{rev}} n}{r_o \beta^*}\xi\gamma, \tag{2.32}$$

where, for beam-stability reasons, the beam-beam parameter ξ should be much smaller than 0.5.

What we have seen so far is related to a single-particle in a bunch, i.e. we have considered the incoherent effect of the beam-beam force, but beam-beam interaction acts also coherently on the entire bunch. In this case, if opposite bunches interact at a certain distance, each of them experiences a dipolar kick generated by the other bunch.

In the LHC, the configuration of the collisions (see Fig. 2.6) provide both head-on and long-range beam-beam interactions. In the latter case, the beams are affected by a dipolar kick and by the non-linear effects of the beam-beam force previously described.

2.5 Luminosity Measurements

The knowledge of luminosity is fundamental to provide a means of calibrating a physics detector and to optimise the machine operations. To reach this knowledge, different measurement methods exist. We can start by considering that

$$L = \frac{R_{\text{ref}}}{\sigma_{\text{ref}}} \tag{2.33}$$

where R_{ref} is the interaction rate of a reference collision process and σ_{ref} is the reference process' cross-section [19].

Therefore, if we consider a known process, whose cross-section is known with a good precision (as can be any QED process for lepton colliders or inelastic pp collision for hadron ones), we can derive the luminosity indirectly by measuring the event rate R_{ref}.

The simplest method to measure the event rate is the Event Counting method, which implies the determination of the fraction of bunch crossings during which a determined detector records a reference process. This type of measurement refers to the relative luminosity, i.e. the relative variation of the instantaneous luminosity over a certain time interval.

It is also useful to measure the absolute luminosity, which is the luminosity expressed on an absolute scale at a certain point in time. The *absolute luminosity* is linked to the *relative luminosity* by means of a calibration constant, which in the case of Eq. (2.33) is the cross-section σ_{ref}. However, the absolute luminosity calibration may also depend on other parameters, such as beam profiles. The latter can be evaluated using the wire scan technique, in which a thin wire is moved through the beam and its interactions with the beam are detected by some calorimeters and give the needed signal [12].

Another possibility to measure both the luminosity L, and the total cross-section σ_{tot} is to consider an elastic scattering of protons and apply the Optical Theorem [7]. The optical theorem states that

$$\lim_{t\to 0}\frac{d\sigma_{\text{el}}}{dt} = (1+\rho^2)\frac{\sigma_{\text{tot}}^2}{16\pi} = \frac{1}{L}\frac{dN_{\text{el}}}{dt}\Big|_{t=0}, \tag{2.34}$$

where σ_{el} is the elastic cross-section, ρ is the ratio of the real to imaginary parts of the scattering amplitude evaluated at the vanishing momentum transfer, $(dN_{\text{el}}/dt)_{t=0}$ is the differential elastic counting rate at small momentum transfer and $\sigma_{\text{tot}} = \sigma_{\text{el}} + \sigma_{\text{inel}} = 4\pi \text{Im} f(t=0)$ is the total cross section with f the scattering amplitude. From the previous equation, we have that the luminosity is equivalent to

$$L = \frac{(1+\rho^2)}{16\pi}\frac{(N_{\text{inel}}+N_{\text{el}})^2}{(dN_{\text{el}}/dt)_{t=0}}, \tag{2.35}$$

where $N_{\text{inel}} + N_{\text{el}}$ is the total number of events, and the total cross section assumes the form

$$\sigma_{\text{tot}} = \frac{N_{\text{inel}} + N_{\text{el}}}{L}. \tag{2.36}$$

2.6 Luminosity Evolution

During a fill, i.e. the time when the beams are made to collide, the luminosity does not remain constant. The collisions cause the decay of the beam intensity, which results in luminosity decay. To derive the expression of the evolution of luminosity as a function of the machine parameters, we can consider round Gaussian beams, i.e. with equal horizontal and vertical r.m.s. emittances, $\varepsilon = \varepsilon_x = \varepsilon_y$, and equal beta functions at the collision point, $\beta^* = \beta^*_x = \beta^*_y$, neglecting reduction factors such as hourglass effect [20]. In this case, the luminosity is defined as

$$L = \frac{n_b^2 f_{\text{rev}} k_{\text{b}} \gamma_{\text{r}}}{4\pi\beta^*\varepsilon} \tag{2.37}$$

where we can define the total beam intensities as $N_{\text{b}} = n_{\text{b}} k_{\text{b}}$.

Considering the burn-off as the only source of beam loss, it is possible to define the rate of change of the beam intensity as

$$\frac{dN_{\text{b}}}{dt} = -\sigma_{\text{tot}} n_{\text{IP}} L = -\kappa N_{\text{b}}^2(t), \tag{2.38}$$

where σ_{tot} represents the total cross-section of the process, n_{IP} the number of collision points and $\kappa = \sigma_{\text{tot}} f_{\text{rev}} n_{\text{IP}} \gamma_{\text{r}} / (4\pi\beta^*\varepsilon)$. Assuming a *constant emittance*, the solution of Eq. (2.38) is

$$N_{\text{b}}(t) = \frac{N_0}{1 + t/\tau_{\text{b}}}, \tag{2.39}$$

where $N_0 = N_{\text{b}}(t = 0)$ is the initial intensity and

$$\tau_{\text{b}} = \frac{1}{\kappa N_0} = \frac{4\pi\beta^*\varepsilon}{\sigma_{\text{tot}} f_{\text{rev}} n_{\text{IP}} \gamma_{\text{r}} N_0} = \frac{N_0}{L(0)\sigma_{\text{tot}} n_{\text{IP}}} \tag{2.40}$$

is the beam lifetime. Thus, it is possible to describe the behaviour of the luminosity decay as

$$L(t) = \frac{L(0)}{(1 + t/\tau_{\text{b}})^2}\,, \tag{2.41}$$

and the integrated luminosity after a certain time t can be obtained as

$$\mathcal{L}_{\text{int}} = \int_0^t dt' L(t') = \tau_{\text{b}} L(0) \frac{t/\tau_{\text{b}}}{1 + t/\tau_{\text{b}}}. \tag{2.42}$$

In practice, the situation is much more complex, and effective models can be devised. Several options are available: one is a generalised form of the solution of Eq. (2.38) [21]

$$L = \frac{L(0)}{(1 + t/\tau)^{\mu}}, \tag{2.43}$$

where $L(0)$, τ and μ are the fit parameters, and a second one is a double-exponential model [22] of the type

$$L(t) = a \exp(-t/b) + c \exp(-t/d), \tag{2.44}$$

where a, b, c, d are the fit parameters. In the following, only the second model will be used.

2.7 Luminosity Optimisation

As previously highlighted, luminosity is a key quantity to define the experimental potential of a collider. In particular, to reach the required accuracy in the determination of the new particles' parameters, it is essential to increase the number of collected events, Eq. (2.5), which is equivalent to maximising the *integrated luminosity*. To this end, several non-excluding approaches are possible. For instance, the machine parameters such as β^* and the crossing angle, can be varied during the fill. Additionally, the length of the fill can be judiciously chosen taking into account several parameters, such as the turn-around time, the probability of a fill failure etc. The first approach was implemented during the LHC Run 2 operations, and it was one among several improvements that led to the considerable progress, as shown in Fig. 2.11 [23].

The following is a brief review of the current method used at LHC to pick up an optimal fill time. In Fig. 2.12, a typical LHC operational cycle is represented, and, here, it is possible to distinguish two different time intervals: one in which

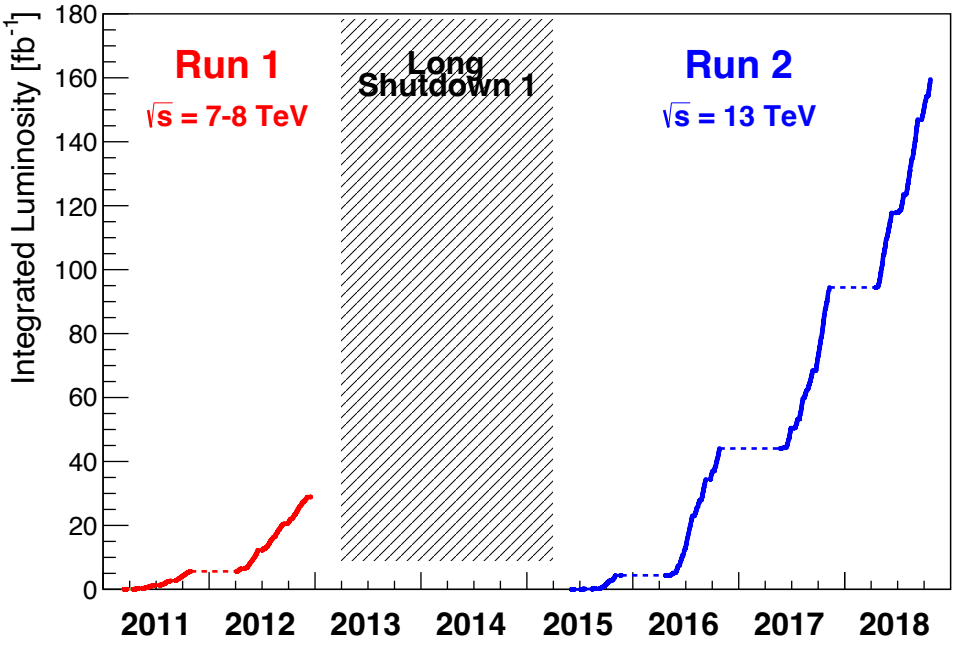

Figure 2.11: Integrated luminosity between 2011 and 2018 for proton operation during Run 1 and Run 2 [23].

there is luminosity production, called fill time (t_{f}), and one in which the luminosity is absent, called turnaround time (t_{ta}).

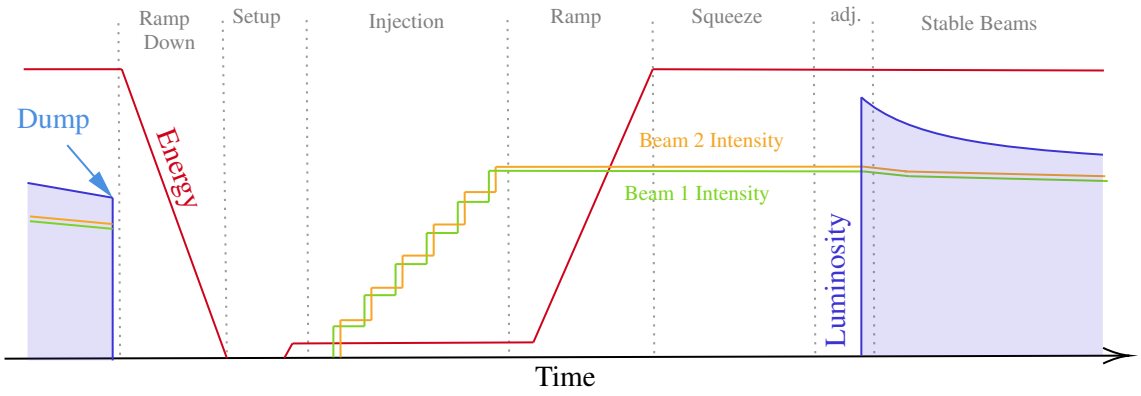

Figure 2.12: Example of the LHC operational cycle from the dumping of the previous fill to the declaration of "Stable Beams" for the subsequent one [2].

The turnaround time includes the ramp-down phase, injection setup, injection

phase, acceleration, collimation setup and β^* squeeze, until the declaration of "Stable Beams"[20]. Moreover, the turnaround time comprises also the time needed to solve technical difficulties and the waiting periods due to nonavailability of the injectors or of the LHC. With the goal of optimising the overall integrated luminosity, it is necessary to maximise the average luminosity $\langle L \rangle$ of a complete operational cycle of the collider [7]. Considering the luminosity evolution described by (2.41) we have that

$$\langle L \rangle = \frac{\int_0^{N_\mathrm{f} t_\mathrm{f}} dt L(t)}{t_\mathrm{f} + t_\mathrm{ta}} = L(0) \frac{t_\mathrm{f}}{(1 + t_\mathrm{f}/\tau_\mathrm{b})(t_\mathrm{f} + t_\mathrm{ta})}, \tag{2.45}$$

where N_f is the number of fills in the run. The fill time can be chosen by the operators in the control room, so we can consider it as a free parameter. At this point, the optimisation of (2.45) give us the optimum fill time,

$$t_\mathrm{f}^\mathrm{opt} = \sqrt{\tau_\mathrm{b} t_\mathrm{ta}}. \tag{2.46}$$

For the LHC case, $t_\mathrm{ta} \sim 10$ h, $\tau_\mathrm{b} \sim 15$ h and so $t_\mathrm{f}^\mathrm{opt} \sim 15$ h [7]. Typically, the operators in the control room choose the fill length, taking into account the estimated optimal fill time (2.46).

This thesis work focuses on enhancing the strategies used to determine the optimal fill length taking into account realistic models for the luminosity evolution, the turn around time, and other statistical factors that affect the simple considerations used to derive Eq. (2.46).

Chapter 3

Luminosity Evolution Data

Our analysis of possible optimisation strategies for the integrated luminosity produced by the LHC is based on the data obtained by the ATLAS experiment [24], as explained in this chapter.

3.1 The Dataset

From the ATLAS data directory [25] we extracted the luminosity data files, each containing multiple entries of luminosity integrated over a short time interval defined by the experiment. In Fig. 3.1 is presented a segment of one of these files for clarity.

```
1461365329 1 4.93401241302 0. 4.32383963915e-22 0.
1461365390 1 4.92993354797 0. 4.32501542323e-22 0.
1461365450 1 4.92800760269 0. 4.32793945664e-22 0.
1461365510 1 4.92684078217 0. 4.33191696087e-22 0.
1461365570 1 4.91852092743 0. 4.32738936732e-22 0.
1461365619 1 4.91928911209 0. 4.33022990553e-22 0.
1461365668 1 4.91526603699 0. 4.33076866742e-22 0.
1461365728 1 4.90638399124 0. 4.32616542505e-22 0.
1461365788 1 4.90047359467 0. 4.3252130049e-22 0.
1461365823 1 4.88737773895 0. 4.31714424346e-22 0.
```

Figure 3.1: Example of a part of a luminosity data file produced by the ATLAS experiment.

Each file contains in order the *UNIX time*[1] in which the luminosity measurement took place, a stable-beam flag (i.e. a float value between 0 and 1 which corresponds to the fraction of time spent in stable beams for this time bin), the luminosity measured in $\mathrm{Hz}\,\mu\mathrm{b}^{-1}$, a point-to-point error on luminosity in $\mathrm{Hz}\,\mu\mathrm{b}^{-1}$, specific luminosity in $\mathrm{Hz}\,\mu\mathrm{b}^{-1}$ and point-to-point error on specific luminosity.

By extracting from these files the UNIX times and luminosities, we obtain a plot for each fill like the one shown in Fig. 3.2.

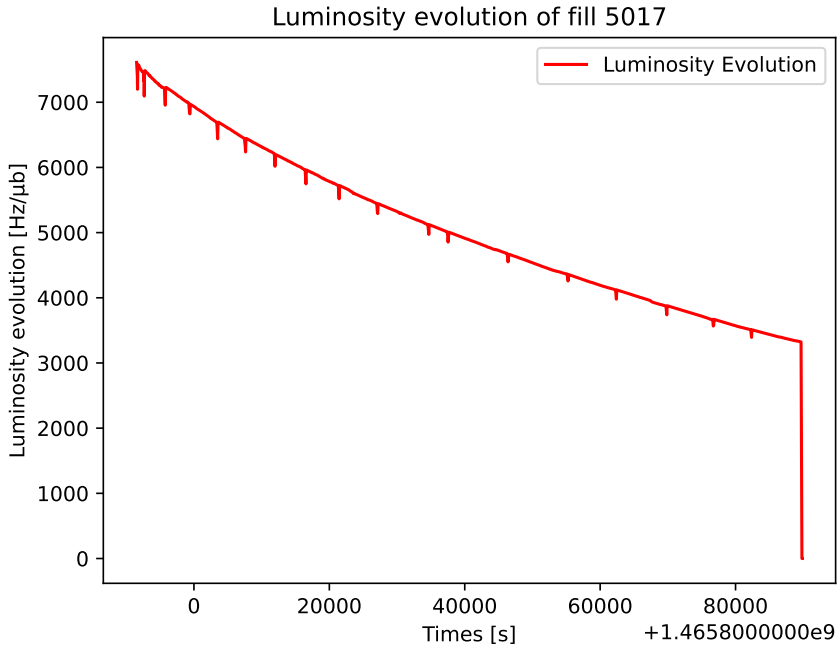

Figure 3.2: Example of luminosity evolution plot.

3.1.1 Fills selection

In [26] it is presented a report about LHC Run 2 fills divided per year (2016, 2017, 2018). Each fill is defined with particular types, e.g. physics, ions, special, rump-up, and so on. For our study, we were just interested in physics fills, which, as the name suggests, are the fills dedicated to the standard proton-physics experiments.

[1]UNIX time, i.e. in seconds since UTC Jan 1, 1970, 00:00:00, not counting leap seconds.

In the following, only those fills that are defined as physics, and whose preceding turnaround time is not affected by special events, such as faults or other issues are retained.

3.2 Discriminating Data

Observing the various plots produced (as described in the previous section), irregularities show up. For clarity, some significant examples are given in Fig. 3.3. These irregularities can be caused by several reasons, but the principal one is the

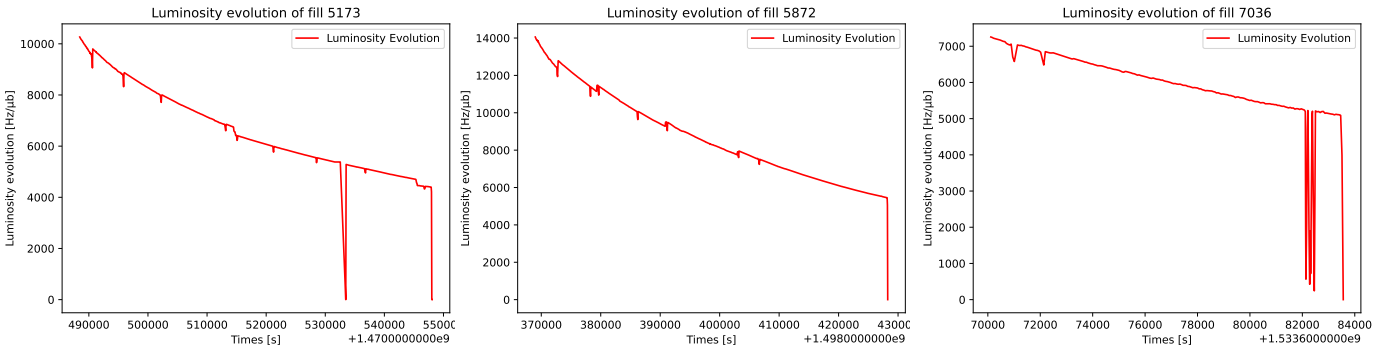

Figure 3.3: Example of three fills (5173, 5872, 7036) luminosity evolution.

re-positioning of beams so to re-optimise the luminosity.

The goal of our preliminary analysis is to obtain a realistic and accurate model of the evolution of luminosity that can be used to generate synthetic fills or extrapolate existing fills beyond the actual fill time. The first step was to remove some of the data from the measured luminosity curves to ensure that the final model was not affected by artefacts in the luminosity measurement.

We performed the data selection by eliminating the null-luminosity values (caused by the re-positioning of the beams or measurement issues) and the points with a too high derivative (in absolute value). The results of this selection are visible in Fig. 3.4.

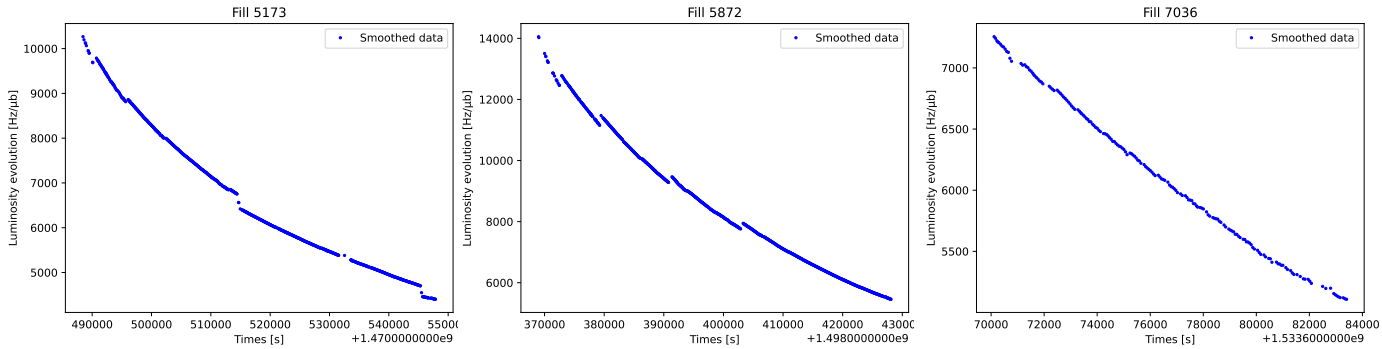

Figure 3.4: Example of three fills (5173, 5872, 7036) luminosity evolution after the data discrimination.

3.3 Fitting Data

Initially, the model selected to fit the data was the following:

$$f(t) = a\exp(-bt) + c\exp(-dt). \tag{3.1}$$

This model allows a good empirical description of the luminosity evolution. $a + c$ represents the peak luminosity, while b and d are the time constants of the luminosity decay [22]. At this point, it was necessary to normalise the time intervals due to the considerable orders of magnitude included in the UNIX times, which are not easy to manage for an exponential fit. Considering for each fill an interval of time that goes from the minimum UNIX time, $t_{\mathrm{U}_{\mathrm{min}}}$, to the maximum one, $t_{\mathrm{U}_{\mathrm{max}}}$, we first re-scaled the UNIX times as follows:

$$t_{\mathrm{s}} = t_{\mathrm{U}} - t_{\mathrm{U}_{\mathrm{min}}}\,, \tag{3.2}$$

where t_{s} is the rescaled time that goes form 0 to $t_{\mathrm{U}_{\mathrm{max}}} - t_{\mathrm{U}_{\mathrm{min}}}$, and t_{U} is the UNIX time. Thus, the fitting model is

$$f(t_{\mathrm{s}}) = a_{\mathrm{s}}\exp(-b_{\mathrm{s}}t_{\mathrm{s}}) + c_{\mathrm{s}}\exp(-d_{\mathrm{s}}t_{\mathrm{s}})\,. \tag{3.3}$$

We can define the normalisation of this rescaled quantity as $t_{\mathrm{n}} = t_{\mathrm{s}}/t_{\mathrm{s}_{\mathrm{max}}}$ obtaining

$$f(t_{\mathrm{n}}) = a_{\mathrm{s}}\exp(-b_{\mathrm{s}}t_{\mathrm{n}}t_{\mathrm{s}_{\mathrm{max}}}) + c_{\mathrm{s}}\exp(-d_{\mathrm{s}}t_{\mathrm{n}}t_{\mathrm{s}_{\mathrm{max}}}) \tag{3.4}$$

$$= a_{\mathrm{s}}\exp(-b_{\mathrm{n}}t_{\mathrm{n}}) + c_{\mathrm{s}}\exp(-d_{\mathrm{n}}t_{\mathrm{n}}), \tag{3.5}$$

where $b_n = b_s t_{s_{max}}$ and $d_n = d_s t_{s_{max}}$.

In Fig. 3.5 three examples (one for each year) of the double exponential fits are shown.

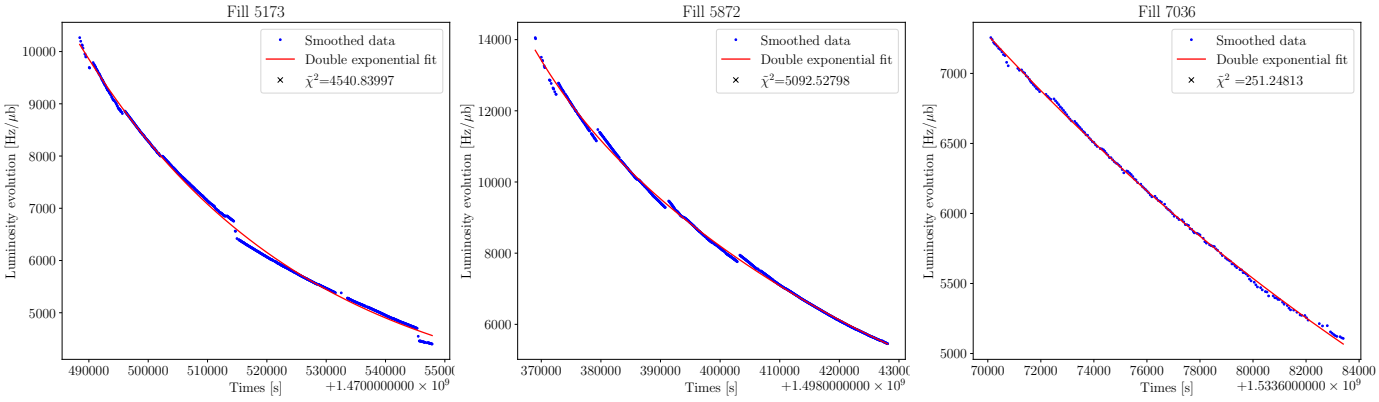

Figure 3.5: Example of double exponential fit of three fills (5173, 5872, 7036) luminosity evolution.

To understand if it was possible to uniquely describe each year thanks to an annual model, the distributions of the fit parameters for each year were studied, as shown in Figs. 3.6, 3.7 and 3.8.

It is therefore evident that it is not possible to describe the set of fills studied with a single model due to so different realisations of the same.

3.4 Other Fitting Models

During the analysis we considered also a three-parameter model and a two-parameter model. The former has the following shape

$$f(t) = a(\exp(-bt) + \exp(-dt)). \tag{3.6}$$

The two-parameter model has the following shape

$$f(t) = a\exp(-bt). \tag{3.7}$$

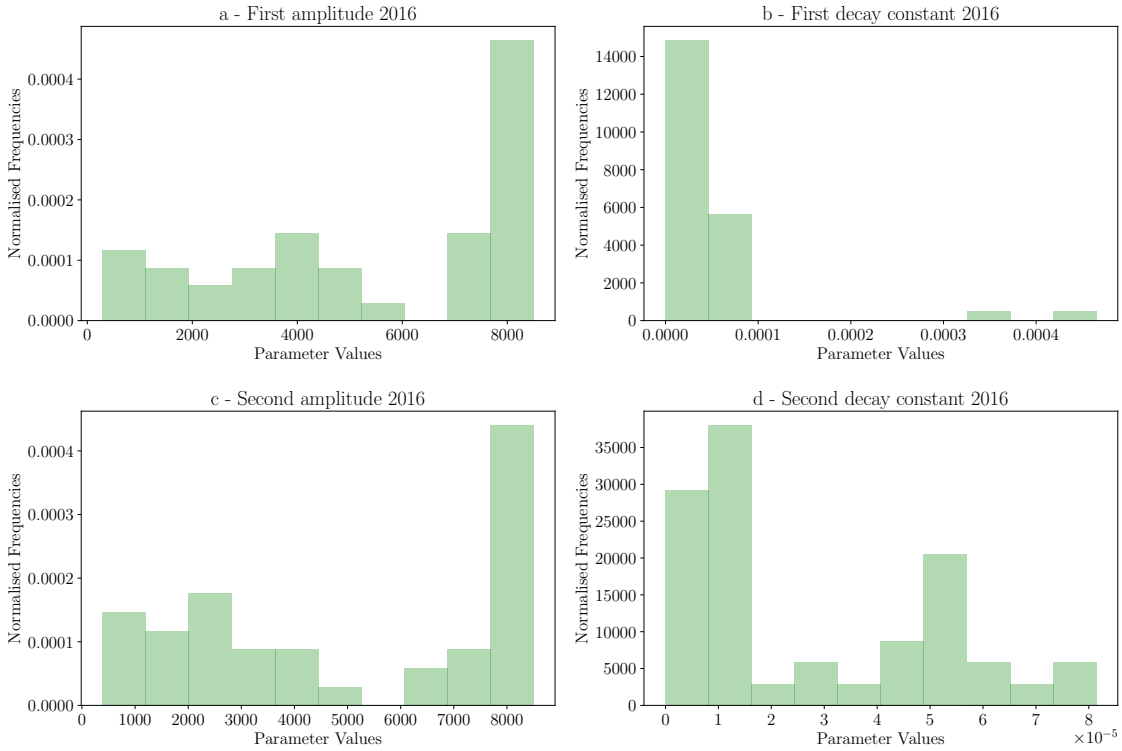

Figure 3.6: Distributions of the fit parameters for the year 2016. On the x-axis there are the parameter values while on the y-axis there are the normalised frequencies.

The four-parameter model choice was taken observing the distributions of the reduced chi square of the different models, as shown in Fig. 3.9.

Figure 3.10 shows a further comparison of the three fit models chosen. On the x-axis is shown the integrated luminosity measured at LHC, and on the y-axis is shown the one obtained using the three fit models. In this case, if we look at points perfectly aligned along the bisector of the graphs we can conclude that the models interpolate the data very well and no bias is visible for most of the fit models (the only exception is the two-parameter case in 2016).

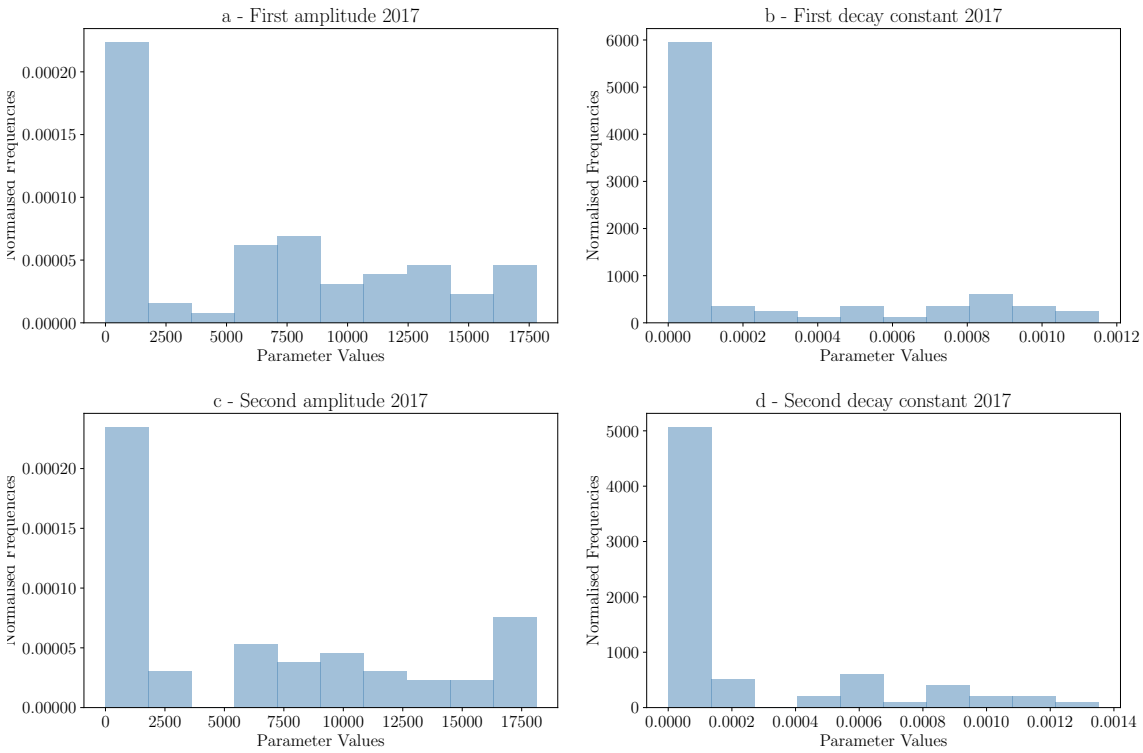

Figure 3.7: Distributions of the fit parameters for the year 2017. On the x-axis there are the parmeter values while on the y-axis there are the normalised frequencies.

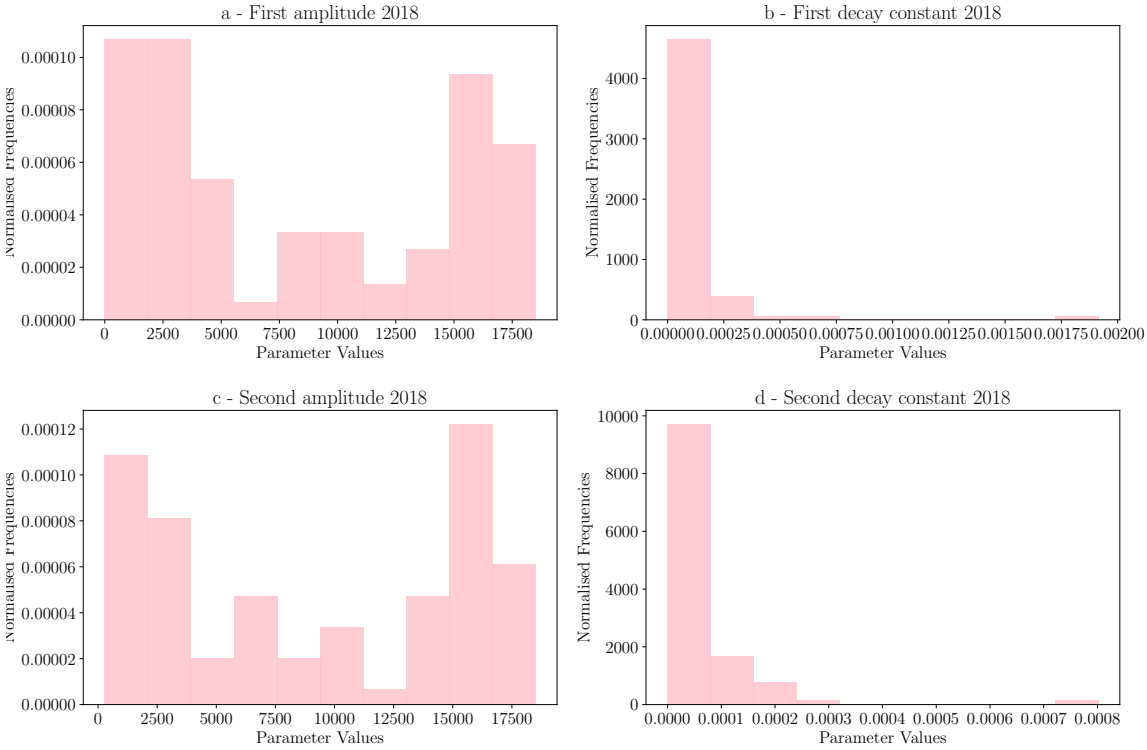

Figure 3.8: Distributions of the fit parameters for the year 2018. On the x-axis there are the parmeter values while on the y-axis there are the normalised frequencies.

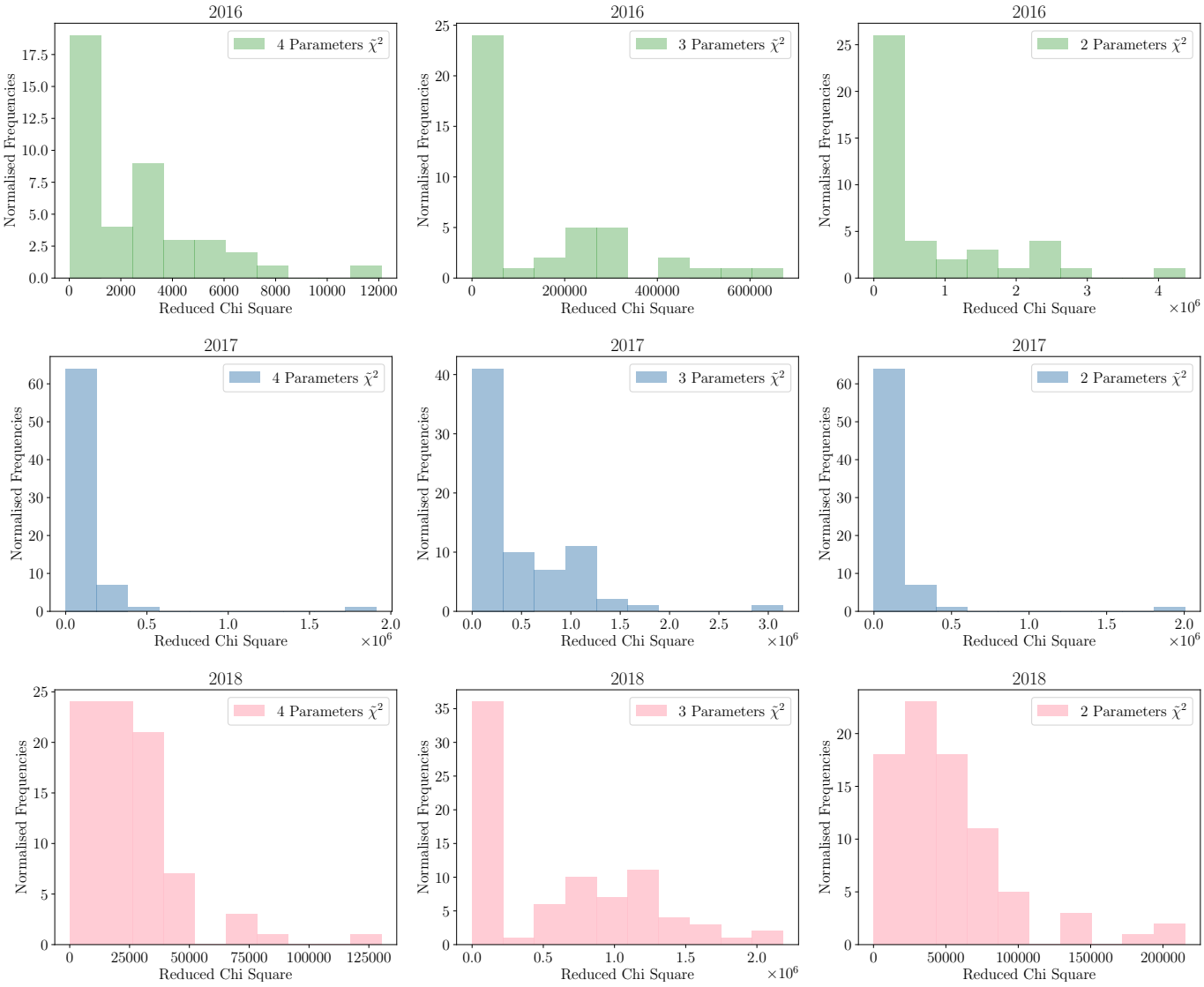

Figure 3.9: $\tilde{\chi}^2$ distributions of 2016 (green plots), 2017 (blue plots) and 2018 (pink plots) for the three different fitting models: (left) four-parameter model; (centre) three-parameter model; (right) two-parameter model.

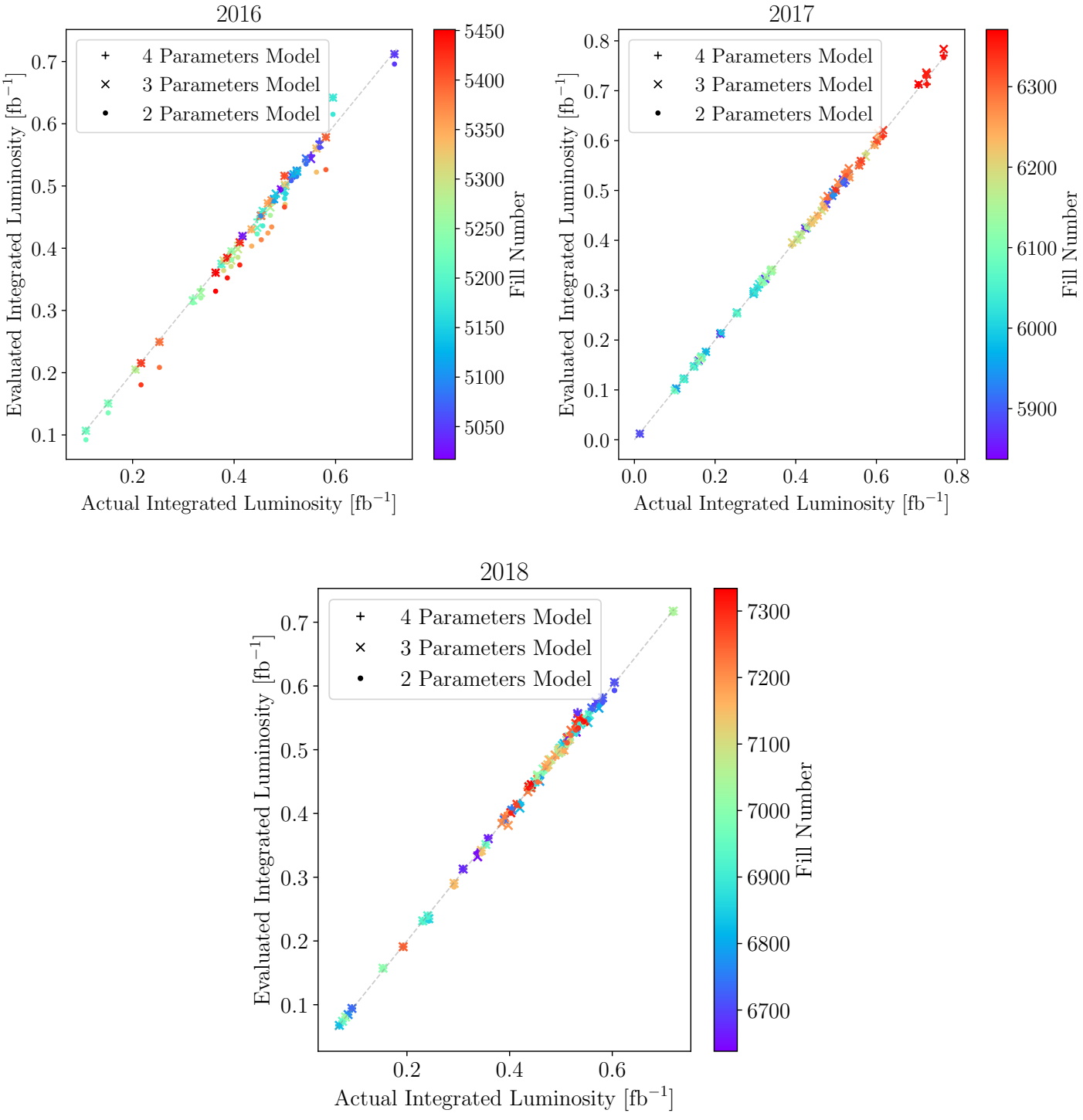

Figure 3.10: Qualitative analysis of the fitting models for the three years of LHC Run 2. On the x-axis there is the integrated luminosity measured by ATLAS at LHC operation, while on the y-axis there is the integrated luminosity evaluated with the three different fitting models. On the secondary y-axis (on the right) there is a chromatic scale that represents the different fill numbers.

Chapter 4

Numerical Optimisation of the Integrated Luminosity

Taking into account what we did in the previous chapter, we implemented a numerical optimiser to probe what is the possible gain in terms of integrated luminosity with respect to the collected integrated luminosity for each of the three years of Run 2.

To this aim, we selected a number of fills for 2016, 2017, and 2018, describeing how many and how they have been selected and tried to determine fill lengths that maximise the total integrated luminosity for the year under consideration. The optimisation is performed with the constraint that the sum of the optimised fill lengths equals that of the actual fill lengths.

4.1 Optimiser Implementation

The algorithm is implemented using Python's `minimize` routine. Since there is no unique model able to describe all fills, to obtain the total luminosity $\mathcal{L}_{\text{tot}}$ it is necessary to sum all single integrated luminosity. The latter is given by the

integral of the fit function of each fill. The function to be optimised[1] is therefore

$$\mathcal{L}_{\text{tot}} = \sum_{i}^{N_{\text{f}}} \int_{0}^{t_i} dx \left[a_{n_i} \exp\left(-\frac{b_{n_i}}{k_i} x \right) + c_{n_i} \exp\left(-\frac{d_{n_i}}{k_i} x \right) \right], \tag{4.1}$$

where N_{f} is the number of fills considered in our study, $k_i = (t_{\text{U}_{\max}} - t_{\text{U}_{\min}})_i$, $t_i \in [0, k_i]$ is the optimum re-scaled time, determined by the optimiser, and the coefficients of the fits are the ones described in the previous chapter. The following constraint

$$\sum_{i}^{N_{\text{f}}} t_i - \sum_{i}^{N_{\text{f}}} t_{r_i} = 0, \tag{4.2}$$

has been imposed to the optimiser, where t_{r_i} are the times that actually took place at LHC.

4.2 Optimisation Algorithm

In general, with optimisation, we mean the choice of the optimal element (depending on the purpose of the optimisation) belonging to a set of possible candidates. Mathematically, this corresponds to determining the extreme value of the function that we want to optimise, i.e. the objective function $f(x)$, in a given domain [27]:

$$\nabla f(x) = 0. \tag{4.3}$$

The solutions of this system are defined *stationary points* of the function, and can be *local minima*, *local maxima*, or *saddle points*. In a case, like ours, of multivariate optimisation, to check the nature of stationary points we need to analyse the Hessian Matrix,

$$\left[H_f(x) \right]_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}. \tag{4.4}$$

If the Hessian is *positive definite* the stationary point will be a local minimum, if it is *negative definite* a local maximum and if it is *indefinite* a saddle point.
In a multivariate case, the analytical approach of finding the roots of the gradient is

[1] As explained in section 4.2 maximising $\mathcal{L}_{\text{tot}}$ is equivalent to minimise $-\mathcal{L}_{\text{tot}}$. This actually is what we have done considering that `minimise` is a minimising routine.

hardly achievable. Instead, techniques, that start at some point of the co-ordinate space and use different methods to move toward the function searching for minima, must be considered.
For our study, we are interested in maximising the objective function. Note that maximising $f(x)$ is equivalent to minimising $-f(x)$, this is why from this moment on we will talk about minimisation.

Many multivariate optimisation algorithms have their roots in the steepest descent method. If we consider the gradient of the objective function $\nabla f(x)$ in a given point x, the negative gradient $-\nabla f(x)$ always points in the direction where the function decreases the most. Thus, the steepest descent method provides a strategy to move along the direction of the negative gradient by a certain distance l_k, and iterate this strategy until reaching the local minimum:

$$x_{k+1} = x_k - l_k \nabla f(x_k), \tag{4.5}$$

where l_k is called *line search parameter.*

In this case, the multivariate optimisation algorithm chosen is the Sequential Least Squares Programming (SLSQP) Algorithm, which deals with constrained optimisation problems of the form

$$\begin{aligned} &\min_x f(x) \\ \text{subject to} : \; &c_j(x) = 0, \quad j \in \mathcal{E} \\ &c_j(x) \geq 0, \quad j \in \mathcal{I} \\ &lb_i \leq x_i \leq ub_i, \quad i = 1, \ldots, N. \end{aligned}$$

where $\mathcal{E}$ and $\mathcal{I}$ are sets of indices containing equality and inequality constraints[2], and lb_i and ub_i are the lower and upper boundaries for each optimal element.

Quadratic constrained optimisation problems with equality constraints are defined as

$$\begin{aligned} \min \; f(x) &= \frac{1}{2} x^T H x + x^T d \\ \text{subject to} : \; Jx &= b, \end{aligned} \tag{4.6}$$

[2]For the studied case, we considered equality constraints.

where H is the hessian matrix of the function and J is the Jacobian matrix of the constraints [28] . The Lagrangian of the optimisation problem is

$$L(x, \lambda) = \frac{1}{2} x^T H x + x^T d + \lambda^T (Jx - b), \tag{4.7}$$

where λ is the Lagrange multiplier.

The necessary condition, i.e. the *Karush-Kuhn-Tucker condition* or KKT, such that x^* is the solution of the optimisation problem, implies the existence of a vector λ^* capable of satisfying the following system

$$\begin{bmatrix} H & -J^T \\ J & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} -d \\ b \end{bmatrix}. \tag{4.8}$$

If the Hessian matrix H is a positive definite matrix, then the unique solution of Eq. (4.7) is (x^*, λ^*).

In the sequential quadratic programming model, an optimisation sub-problem is set to find the ideal step to move toward the function

$$x_{k+1} = x_k + a_k l_k \tag{4.9}$$

where $a_k \in (0, 1]$ and l_k is the solution of the optimal sub-problem [29]. The latter can be defined as follows

$$\begin{aligned} &\min_x \nabla f^T(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T H(x_k)(x - x_k) \\ &\text{subject to : } s(x_k) + \nabla s^T(x_k)(x - x_k) = 0, \end{aligned}$$

where $s(x) = Jx - b$, $\nabla f(x_k) = (\partial f/\partial x_1, \ldots, \partial f/\partial x_n)^T|_{x=x_k}$,
$\nabla s(x_k) = (\partial s/\partial x_1, \ldots, \partial s/\partial x_n)^T|_{x=x_k}$ and $H(x_k)$ is the Hessian matrix of $f(x)$.

In this kind of algorithms, penalty functions have the goal of simplifying the optimisation problem. In fact, these functions introduce an artificial penalty for violating the constraint. In our particular case (i.e. Eq. (4.6)) the penalty function is

$$P(x, \pi) = \frac{1}{2} x^T H x + x^T d + \pi (Jx + b) \tag{4.10}$$

where π is the penalty coefficient.

4.3 Optimisation Results

The Fig. 4.1 shows the results of the optimisation in which we had considered the same number of fills and only let the optimiser change the times according to the optimal ones.

In Table 4.1 are shown the results of the optimisation for the total luminosity of each year.

Table 4.1: Total Luminosity Results of Numerical Optimisation.

	2016	**2017**	**2018**
LHC actual total luminosity [fb^{-1}]	18.48	28.67	35.72
LHC optimised total luminosity [fb^{-1}]	19.07	29.23	37.11
Relative increase	3.2%	2.0%	3.9%

We remark, that although the gain in integrated luminosity is not extremely large, still it could be achieved at zero cost, simply developing the appropriate strategy to determine on-line the optimal fill length (see next chapters).

The analysis presented so far assumed that the actual value of N_f was indeed optimal. This assumption has been scrutinised by eliminating in succession one, two or three fills (considering all possible combinations) and adding the times of the deleted fills and the respective delivery times to the physical time. Thus, we modified the Eq. (4.2) as follows:

$$\sum_{i}^{N_\text{f}-j} t_i - \sum_{i}^{N_\text{f}-j} t_{r_i} - \sum_{k}^{j} t_{\text{ta},k} = 0, \quad \text{where} \quad j = 1, 2, 3 \tag{4.11}$$

where t_{r_i} are the actual fill times and $t_{\text{ta},k}$ is the turnaround time of the removed fills. In Figs. 4.2 and 4.3 are shown the results of this second optimisation.

From this analysis, it is possible to conclude that the initial number of fills was the optimal one and what we have to change is the length of each fill according to the optimal one.

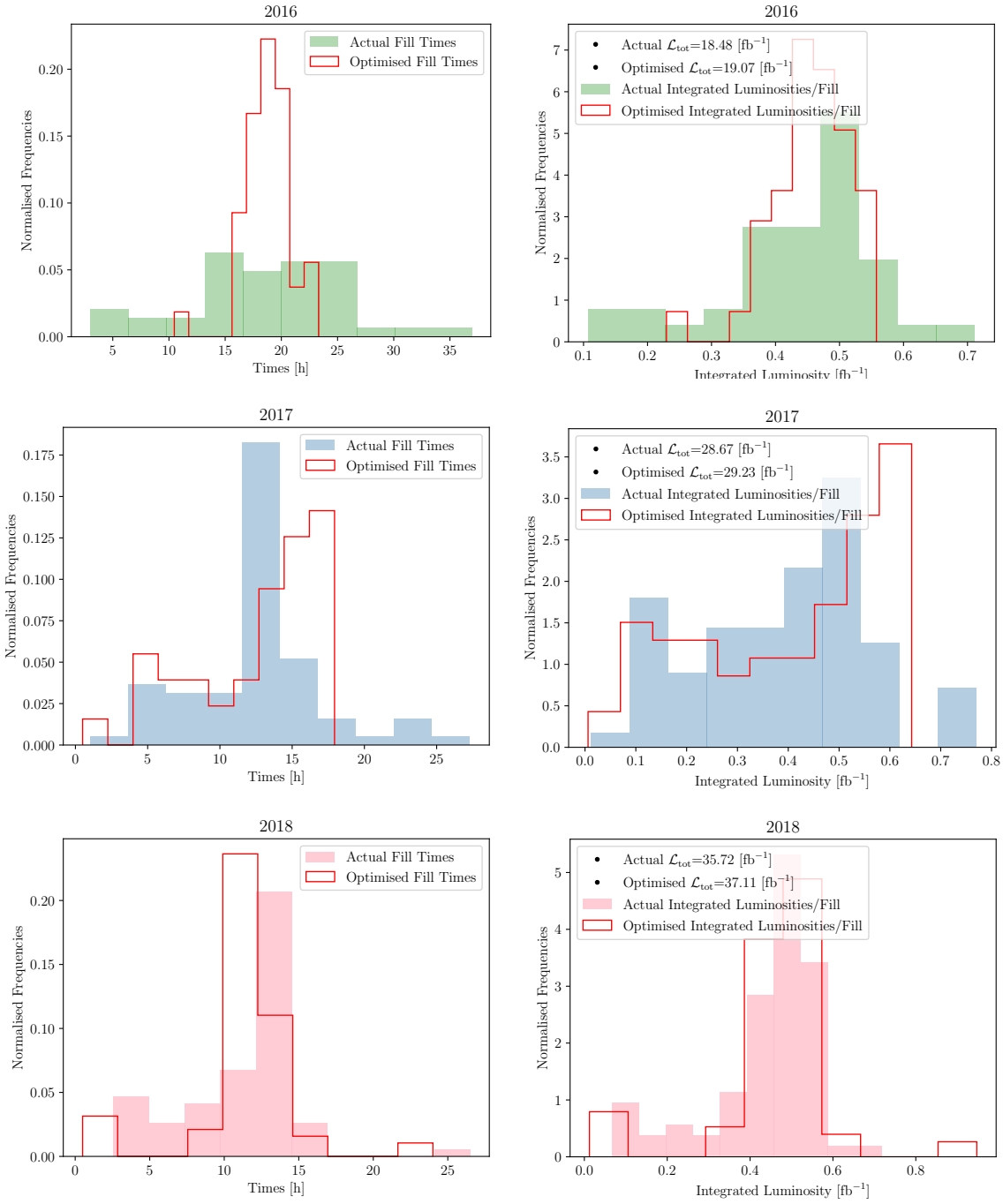

Figure 4.1: (left) Comparisons between the fill times that took place at LHC during Run 2 (green plots for 2016, blue plots for 2017 and pink plots for 2018) and the optimal fill times chosen by our algorithm; (right) Comparison between the integrated luminosities/ of LHC Run 2 (green plots for 2016, blue plots for 2017 and pink plots for 2018) and the optimal ones obtained by our algorithm.

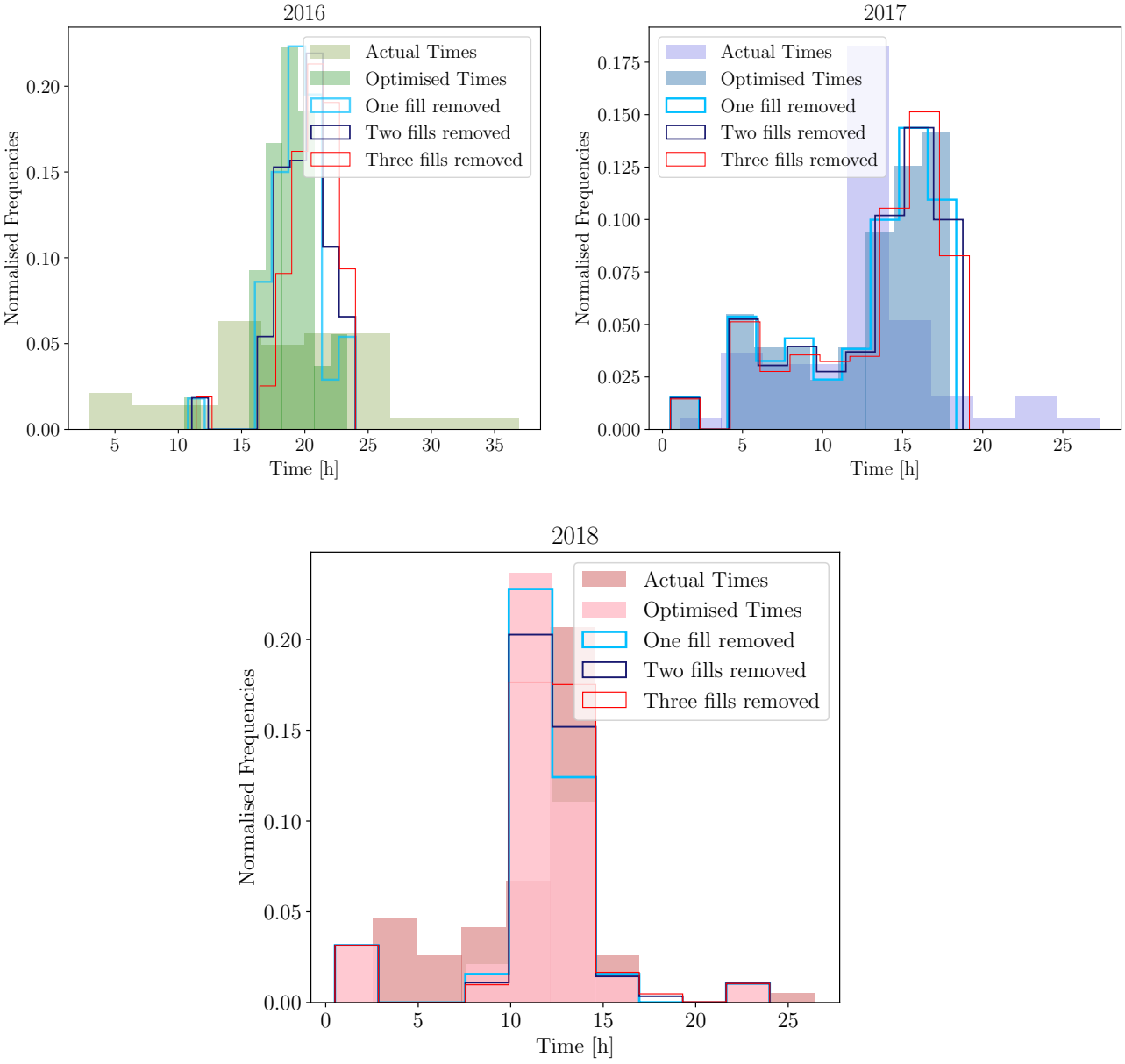

Figure 4.2: Comparison between the fill times that took place at LHC during 2016, 2017 and 2018, the optimal fill times chosen by our algorithm and the optimal times where one, two and three fills have been removed.

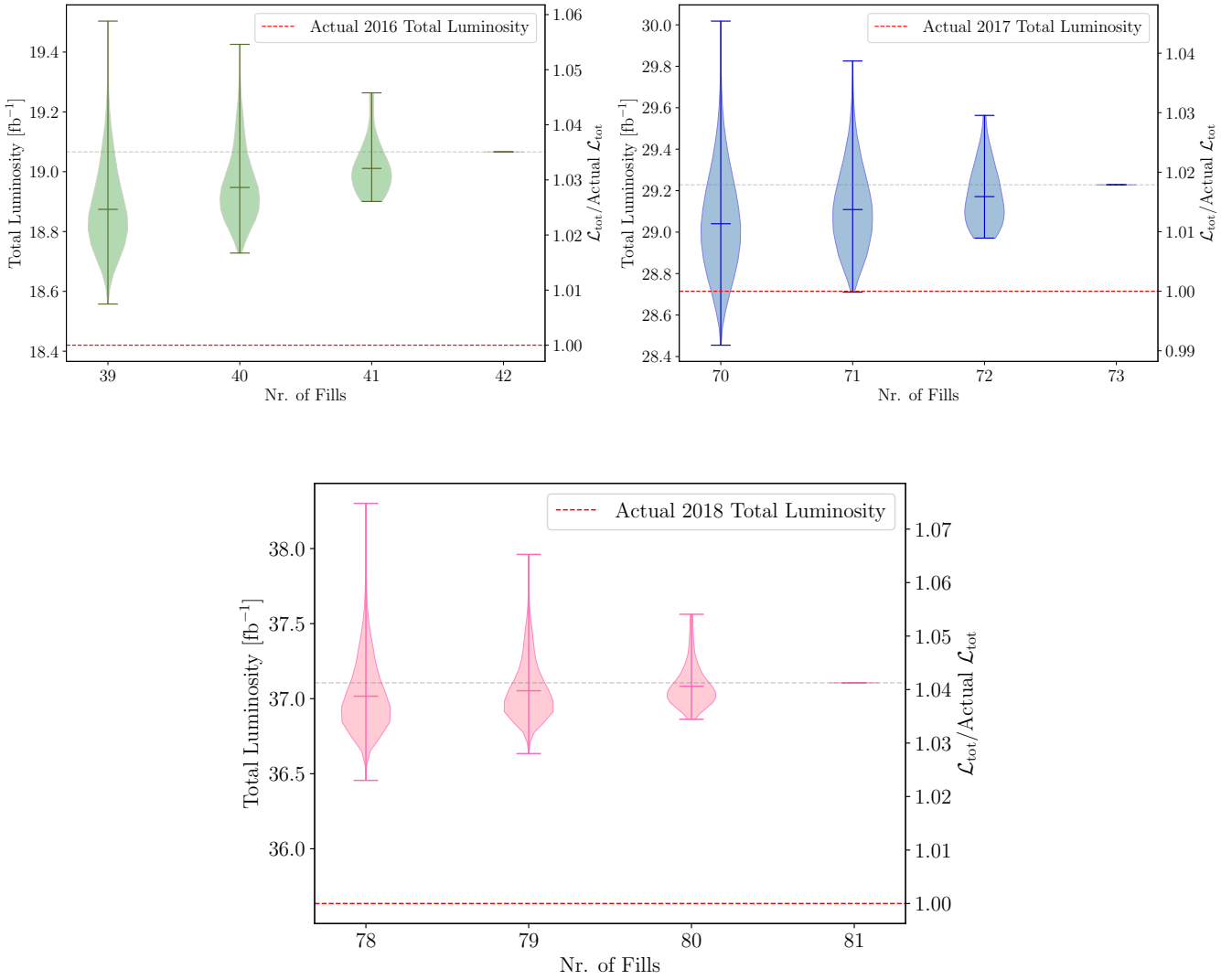

Figure 4.3: Violin plots in which we can observe the distributions of total luminosity in all the studied cases, form right to left: the optimisation of time lengths, the removal of one fill, the removal of two fills and the removal of three fills. On the secondary y-axis the ratio between the optimised luminosity and the measured one is shown.

Chapter 5

On-line Optimisation Strategies

This chapter presents the optimisation strategies that could be realised at the LHC to increase the integrated luminosity.

The process of luminosity production is sketched in Fig. 5.1 where the magnetic cycle is shown and the main quantities are introduced.

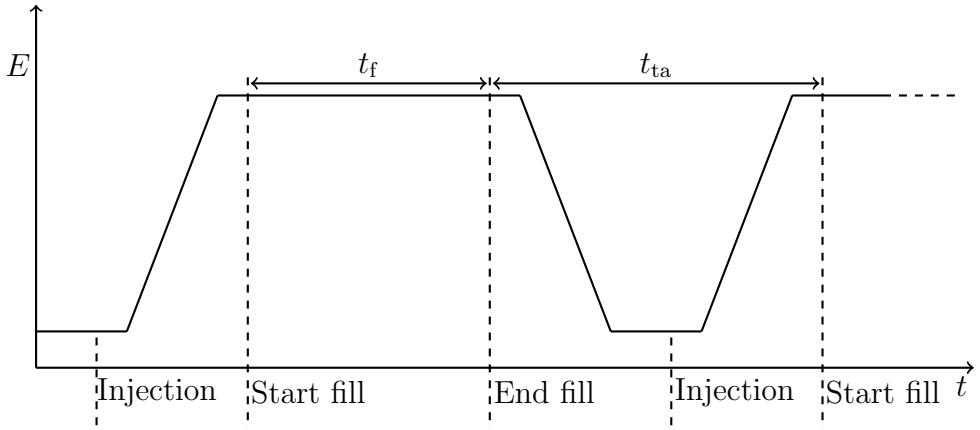

Figure 5.1: Sketch of the luminosity production process in a circular collider, where E is the beam energy.

In this study, $L(t)$ represents the evolution of the luminosity during a fill. Any model can be considered, as the goal is to provide an optimised strategy to collect luminosity in a circular collider once a model for describing the function L is given.

The goal is to maximise the integrated luminosity, and, in a purely deterministic case, this corresponds to maximising

$$\begin{aligned}\mathcal{L}_{\mathrm{tot}}(t_{\mathrm{f}}) &= N_{\mathrm{f}} \int_0^{t_{\mathrm{f}}} \mathrm{d}t\, L(t) \\ &= \frac{T}{t_{\mathrm{ta}} + t_{\mathrm{f}}} \int_0^{t_{\mathrm{f}}} \mathrm{d}t\, L(t)\,,\end{aligned} \tag{5.1}$$

where T is the total time for physics, t_{ta} is the so-called turn around, i.e. the time elapsed from the end a fill for physics and the start of the next one, t_{f} is the length of the physics fill. We remark that the second step in Eq. (5.1) deserves some care. In principle, one should write

$$\mathcal{L}_{\mathrm{tot}}(t_{\mathrm{f}}) = \left\lfloor \frac{T}{t_{\mathrm{ta}} + t_{\mathrm{f}}} \right\rceil \int_0^{t_{\mathrm{f}}} \mathrm{d}t\, L(t)\,, \tag{5.2}$$

where $\lfloor \cdot \rceil$ stands for the nearest integer. However, under the assumption that $T \gg t_{\mathrm{ta}} + t_{\mathrm{f}}$, which is always the case in reality, one can simply consider the form (5.1) that will be correct with a high degree of accuracy.

Under these assumptions, the optimisation is performed by noting that $\mathcal{L}_{\mathrm{tot}} = \mathcal{L}_{\mathrm{tot}}(t_{\mathrm{f}})$ and the maximum can be found by solving the equation

$$\frac{\mathrm{d}\mathcal{L}_{\mathrm{tot}}}{\mathrm{d}t_{\mathrm{f}}} = 0\,. \tag{5.3}$$

5.1 The $L(t)$ model

We now define a model of the evolution of the luminosity to replace it in the following calculations to derive an optimal fill time scale-law with respect to the turnaround time. As presented in [30], the luminosity can be derived form the only time-dependent beam parameter: the intensity N[1]. Considering the burn off as the only relevant mechanism for a time-variation, in case of round beams ($\epsilon_x^* = \epsilon_y^* = \epsilon^*$) and round optics ($\beta_x^* = \beta_y^* = \beta^*$), it is possible to find

$$L(t) = \frac{\Xi N_{\mathrm{i}}^2}{(1 + \sigma_{\mathrm{int}} n_{\mathrm{c}} \Xi N_{\mathrm{i}} t)^2}, \tag{5.4}$$

[1] $N_i = k_{\mathrm{b}} n_i$, where n_i represents the bunch population and k_{b} is the number of colliding bunches.

where σ_{int} is the cross section for interaction of charged particles, n_{c} stands for the number of collision points and the two colliding beams have been assumed to be of equal intensity (which is the best scenario as far as the luminosity is concerned). The term Ξ stands for

$$\Xi = \frac{\gamma_{\mathrm{r}} f_{\mathrm{rev}}}{4\pi\epsilon^* \beta^* k_{\mathrm{b}}} F(\theta_{\mathrm{c}}, \sigma_{\mathrm{z}}, \sigma^*),$$

where γ_{r} is the relativistic γ-factor, f_{rev} is the revolution frequency, k_{b} the number of colliding bunches, ϵ^* is the RMS normalised transverse emittance, β^* is the value of the beta-function at the collision point and the factor F accounts for the reduction in volume overlap between the colliding bunches due to the presence of a crossing angle

$$F(\theta_{\mathrm{c}}, \sigma_{\mathrm{z}}, \sigma^*) = 1/\sqrt{1 + (\frac{\theta_{\mathrm{c}}}{2}\frac{\sigma_{\mathrm{z}}}{\sigma_*})}, \qquad \text{with } \sigma^* = \sqrt{\beta^* \epsilon^* / (\beta_{\mathrm{r}} \gamma_{\mathrm{r}})},$$

where σ_{z} is the longitudinal RMS dimension, β_{r} is the relativistic β and $\theta_{\mathrm{c}}/2$ is the half crossing angle. At this point, assuming the simple case of equal intensities for both beams, it is possible to obtain for the burn off part

$$\mathcal{L}_{\mathrm{int}}(t) = \int_0^t dt L(t) = \frac{N_{\mathrm{i}}^2 t_{\mathrm{f}} \Xi}{(N_{\mathrm{i}} t_{\mathrm{f}} \Xi n_{\mathrm{c}} \sigma_{\mathrm{int}} + 1)} = \frac{N_{\mathrm{i}} \Xi}{\varepsilon f_{\mathrm{rev}}} \frac{\varepsilon N_{\mathrm{i}} f_{\mathrm{rev}} t}{1 + \varepsilon N_{\mathrm{i}} f_{\mathrm{rev}} t}, \tag{5.5}$$

where $\varepsilon = \sigma_{\mathrm{int}} n_{\mathrm{c}} \Xi / f_{\mathrm{rev}}$. The next table 5.1 shows the values of the above-mentioned parameters for Run 2.

Table 5.1: LHC Run 2 Parameters for the Luminosity Model [14] [23].

Parameters		**2016**	**2017**	**2018**
L_{Peak}	$[10^{34}\ \text{cm}^{-2}\,\text{s}^{-1}]$	1.4	2.1	2.1
E_{beam}	[TeV]	6.5	6.5	6.5
f_{rev}	[kHz]	11.2	11.2	11.2
N	$[10^{11}]$	$1.0 - 1.25$	$1.0 - 1.25$	$1.0 - 1.25$
β^*	[cm]	0.4	$0.4 - 0.3$	$0.3 - 0.25$
ϵ^*	[µm]	~ 2.2	~ 2.2	~ 1.9
n_{c}		2	2	2
σ_{z}	[m]	0.102	0.102	0.102
σ_{int}	$[10^{-30}\ \text{m}^2]$	7.95	7.95	7.95
k_{b}		2220	2556/1868	2556
$\theta_{\text{c}}/2$	[µrad]	$185 - 140$	$150 - 120$	$160 - 130$
γ_{r}		6929.64	6929.64	6929.64
β_{r}		~ 1	~ 1	~ 1
No. days of physics operations		146	140	145

5.2 Optimisation with a fixed t_{ta}

A very first attempt of integrated luminosity optimisation can then be carried out by solving the Eq. (5.3) or going through the Lagrange Multipliers method, which is based on the *Lagrange Multipliers Theorem*.

If $f, g \in \mathcal{C}^1(\mathbb{R}^2)$ and $(\hat{x}, \hat{y})$ is a constrained critical point of f, with $g(\hat{x}, \hat{y}) = c$ and $\nabla g(\hat{x}, \hat{y}) \neq (0, 0)$, then it exists $\lambda \in \mathbb{R}$ (called Lagrange Multiplier) such that [31]:

$$\nabla f(\hat{x}, \hat{y}) = \lambda \nabla g(\hat{x}, \hat{y})$$

In first approximation, it is possible to consider t_{ta} as given, and so the function

to be optimised is

$$f(t_{\mathrm{f}}, N_{\mathrm{f}}) = N_{\mathrm{f}} \int_0^{t_{\mathrm{f}}} dt L(t), \tag{5.6}$$

and the constraint function is

$$g(t_{\mathrm{f}}, N_{\mathrm{f}}) = N_{\mathrm{f}} t_{\mathrm{f}} + N_{\mathrm{f}} t_{\mathrm{ta}}. \tag{5.7}$$

Thus, the system to be solved will be:

$$\begin{cases} \left(\frac{\partial f(t_{\mathrm{f}}, N_{\mathrm{f}})}{\partial t_{\mathrm{f}}}\right) = \lambda \left(\frac{\partial g(t_{\mathrm{f}}, N_{\mathrm{f}})}{\partial t_{\mathrm{f}}}\right) \\ \left(\frac{\partial f(t_{\mathrm{f}}, N_{\mathrm{f}})}{\partial N_{\mathrm{f}}}\right) = \lambda \left(\frac{\partial g(t_{\mathrm{f}}, N_{\mathrm{f}})}{\partial N_{\mathrm{f}}}\right) \\ N_{\mathrm{f}} t_{\mathrm{f}} + N_{\mathrm{f}} t_{\mathrm{ta}} = T \end{cases} \Rightarrow \begin{cases} N_{\mathrm{f}} L(t_{\mathrm{f}}) = \lambda N_{\mathrm{f}} \\ \int_0^{t_{\mathrm{f}}} dt L(t) = \lambda (t_{\mathrm{f}} + t_{\mathrm{ta}}) \\ N_{\mathrm{f}} t_{\mathrm{f}} + N_{\mathrm{f}} t_{\mathrm{ta}} = T \end{cases} \tag{5.8}$$

where $L(t_{\mathrm{f}})$ is the integrable function of (5.6) evaluated in t_{f}. At this point, solving the system, it is possible to find that:

$$\begin{cases} \lambda = L(t_{\mathrm{f}}) \qquad \text{(considering that } N_{\mathrm{f}} \neq 0) \\ \int_0^{t_{\mathrm{f}}} dt L(t) = L(t_{\mathrm{f}})(t_{\mathrm{f}} + t_{\mathrm{ta}}) \\ N_{\mathrm{f}} = \frac{T}{t_{\mathrm{f}} + t_{\mathrm{ta}}} \end{cases} \tag{5.9}$$

Substituting equations (5.4) and (5.5) in the system (5.9):

$$\begin{cases} \lambda = \frac{\Xi N_{\mathrm{i}}^2}{(1+\sigma_{\mathrm{int}} n_{\mathrm{c}} \Xi N_{\mathrm{i}} t_{\mathrm{f}})^2} \\ \frac{N_{\mathrm{i}}^2 t_{\mathrm{f}} \Xi}{(N_{\mathrm{i}} t_{\mathrm{f}} \Xi n_{\mathrm{c}} \sigma_{\mathrm{int}} + 1)} = \frac{\Xi N_{\mathrm{i}}^2}{(1+\sigma_{\mathrm{int}} n_{\mathrm{c}} \Xi N_{\mathrm{i}} t_{\mathrm{f}})^2} (t_{\mathrm{f}} + t_{\mathrm{ta}}) \\ N_{\mathrm{f}} = \frac{T}{t_{\mathrm{f}} + t_{\mathrm{ta}}} \end{cases} \tag{5.10}$$

Solving the previous system is equivalent to solve the following equation:

$$\frac{N_{\mathrm{i}}^2 \Xi t_{\mathrm{f}}}{N_{\mathrm{i}} n_{\mathrm{c}} \Xi \sigma_{\mathrm{int}} t_{\mathrm{f}} + 1} = \frac{N_{\mathrm{i}}^2 \Xi (t_{\mathrm{f}} + t_{\mathrm{ta}})}{(N_{\mathrm{i}} n_{\mathrm{c}} \Xi \sigma_{\mathrm{int}} t_{\mathrm{f}} + 1)^2}, \tag{5.11}$$

whose real solutions for the optimal fill time t_{f} are:

$$t_{\mathrm{opt}} = \pm \frac{\sqrt{t_{\mathrm{ta}}}}{\sqrt{N_{\mathrm{i}} n_{\mathrm{c}} \Xi \sigma_{\mathrm{int}}}}. \tag{5.12}$$

where only the positive result has a physical meaning, and it is the one that maximises our initial function (5.6). It is possible to conclude that, for relatively consistent variations in the turn-around time, there will be rather small variations in the optimal fill time.

5.3 Optimisation with a distribution of t_{ta}

At this point, it is possible to consider n values t_i, distributing according to a certain probability density function (p.d.f.), representing n realisations of the turn-around time. In this case the function to be maximised is:

$$\mathcal{L}_{\mathrm{tot}}(\hat{t}) = n \int_0^{\hat{t}} \mathrm{d}t \, L(t) \, = n \frac{N_{\mathrm{i}}^2 \hat{t} \Xi}{(N_{\mathrm{i}} \hat{t} \Xi n_{\mathrm{c}} \sigma_{\mathrm{int}} + 1)}, \tag{5.13}$$

which assumes that the fills should be of equal length although the turn-around times are not, with the constraint

$$\sum_{i=1}^{n} t_i + n \, \hat{t} = T \, . \tag{5.14}$$

One can assume to replace the sum of the numbers by a term $n \, \tau$ so that

$$n = \frac{T}{\tau + \hat{t}} \tag{5.15}$$

and the optimisation of Eq. (5.13) becomes of the same type as the problem (5.1). In this case, indeed, the equation to be optimised is

$$f(n, \hat{t}) = n \int_0^{\hat{t}} dt L(t), \tag{5.16}$$

and the constraint function is

$$g(n,\hat{t}) = n\tau + n\hat{t}. \tag{5.17}$$

Thus, the system to be solved will be:

$$\begin{cases} \left(\frac{\partial f(n,\hat{t})}{\partial \hat{t}}\right) = \lambda\left(\frac{\partial g(n,\hat{t})}{\partial \hat{t}}\right) \\ \\ \left(\frac{\partial f(n,\hat{t})}{\partial n}\right) = \lambda\left(\frac{\partial g(n,\hat{t})}{\partial n}\right) \\ \\ n\tau + n\hat{t} = T \end{cases} \Rightarrow \begin{cases} \lambda = L(\hat{t}) \qquad \text{(considering that } n \neq 0) \\ \\ \int_0^{\hat{t}} dt L(t) = \lambda(\hat{t}+\tau) \\ \\ n = T/(\tau+\hat{t}) \end{cases} \tag{5.18}$$

which leads to a solution similar to the previous case (see Eq. (5.12)):

$$t_{\rm opt} = \pm \frac{\sqrt{\tau}}{\sqrt{N_{\rm i} n_{\rm c} \Xi \sigma_{\rm int}}}, \tag{5.19}$$

.

It can be checked *a posteriori* that the assumption of optimising the integrated luminosity by using equal fills lengths is the correct one. In fact, in this case

$$\begin{aligned} \mathcal{L}_{\rm tot}(\hat{t}_1,\hat{t}_2) &= (n-1)\int_0^{\hat{t}_1} {\rm d}t\, L(t) + \int_0^{\hat{t}_2} {\rm d}t\, L(t) \\ &= \frac{T-\hat{t}_2-\tau}{\tau+\hat{t}_1}\int_0^{\hat{t}_1} {\rm d}t\, L(t) + \int_0^{\hat{t}_2} {\rm d}t\, L(t)\,, \end{aligned} \tag{5.20}$$

where the constraint (5.14) has been adapted to this new case. The maximisation of $\mathcal{L}_{\rm tot}(\hat{t}_1,\hat{t}_2)$ is obtained by considering as the function to be optimised

$$f(n,\hat{t}_1,\hat{t}_2) = (n-1)\int_0^{\hat{t}_1} dt L(t) + \int_0^{\hat{t}_2} dt L(t), \tag{5.21}$$

and as the constraint function

$$g(n,\hat{t}_1,\hat{t}_2) = n\tau + n\hat{t}_1 - \hat{t}_1 + \hat{t}_2. \tag{5.22}$$

Thus, the system to be solved will be:

$$\begin{cases} \left(\frac{\partial f(n,\hat{t}_1,\hat{t}_2)}{\partial n}\right) = \lambda\left(\frac{\partial g(n,\hat{t}_1,\hat{t}_2)}{\partial n}\right) \\ \left(\frac{\partial f(n,\hat{t}_1,\hat{t}_2)}{\partial \hat{t}_1}\right) = \lambda\left(\frac{\partial g(n,\hat{t}_1,\hat{t}_2)}{\partial \hat{t}_1}\right) \\ \left(\frac{\partial f(n,\hat{t}_1,\hat{t}_2)}{\partial \hat{t}_2}\right) = \lambda\left(\frac{\partial g(n,\hat{t}_1,\hat{t}_2)}{\partial \hat{t}_2}\right) \\ n\tau + (n-1)\hat{t}_1 = T \end{cases} \Rightarrow \begin{cases} \int_0^{\hat{t}_1} dt L(t) = \lambda(\tau - \hat{t}_1) \\ nL(\hat{t}_1) - L(\hat{t}_1) = \lambda(n-1) \\ L(\hat{t}_2) = \lambda \end{cases} \hat{t}_1 = \hat{t}_2 \equiv \hat{t} \\ n = T/(\tau + \hat{t}) \tag{5.23}$$

where is possible to verify that $\hat{t}_1 = \hat{t}_2$.

Obviously, in the realistic case, the n values of the turn-around are not known *a priori*, and the luminosity evolution model is not the same for all fills (as shown in Chap. 3), which modifies the scheme previously described. Let us assume that t_j, $1 \leq j \leq i$ and $\hat{t}_j$, $1 \leq j \leq i-1$ are the turn-around times and the optimal fill times, respectively for all the fills from 1 to i. The problem is to determine $\hat{t}_i$ so to maximise

$$\mathcal{L}_{\rm tot}(\hat{t}_i, \hat{t}) = \sum_{j=1}^{i-1} \mathcal{L}_{\rm int}(\hat{t}_j) + \int_0^{\hat{t}_i} {\rm d}t\, L(t) + \frac{T - \sum_{j=1}^{i} \left(t_j + \hat{t}_j\right)}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \int_0^{\hat{t}} {\rm d}t\, L_{\rm mp}(t)\,, \tag{5.24}$$

where and additional optimisation parameter has been introduced, namely $\hat{t}$, which represents the optimum fill time of future fills. Here, $\mathcal{L}_{\rm int}$ stands for the integrated luminosity in a single fill. The third term in Eq. (5.24) introduces a relationship between $\hat{t}_i$ and $\hat{t}$. We remark that $L_{\rm mp}(t)$ stands for the most probable value of the function representing the luminosity evolution. It is clear that in case the luminosity function is Gaussian distributed, the most probable value corresponds to the average value. Such a value should be determined on the basis of the collection of fills already occurred. Moreover the term

$$\frac{1}{i}\sum_{j=1}^{i} t_j \tag{5.25}$$

is intended to provide the average turn-around time. The optimisation is performed as follows:

$$\begin{cases} \dfrac{\partial \mathcal{L}_{\mathrm{tot}}}{\partial \hat{t}_i} & = L(\hat{t}_i) - \dfrac{1}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \displaystyle\int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t) = 0 \\ \\ \dfrac{\partial \mathcal{L}_{\mathrm{tot}}}{\partial \hat{t}} & = \dfrac{T - \sum_{j=1}^{i}(t_j + \hat{t}_j)}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \left(L_{\mathrm{mp}}(\hat{t}) - \dfrac{1}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \displaystyle\int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t) \right) = 0\,. \end{cases} \tag{5.26}$$

Chapter 6

Application of On-line Optimisation Strategies

6.1 Preliminary Studies

Before applying the strategy proposed in Chap. 5, we carried out some preliminary studies. The first one is about the quality of the description of the luminosity evolution reached by our fitting models. We did this by comparing the integrated luminosity, as obtained from the models, with the actual value (present in the summary files of the ATLAS data directory). We considered, from that moment on, only those fills that satisfy the following condition:

$$\frac{|\mathcal{L}_{\mathrm{Fit}} - \mathcal{L}_{\mathrm{LHC}}|}{\mathcal{L}_{\mathrm{LHC}}} \leq 0.02\,, \tag{6.1}$$

where $\mathcal{L}_{\mathrm{Fit}}$ is the fill integrated luminosity evaluated by the fitting models, while $\mathcal{L}_{\mathrm{LHC}}$ is the actual LHC integrated luminosity. To determine the value for the selection criterion reported in Eq. (6.1), we performed the study shown in Fig. 6.1. It is clearly seen that for a selection threshold larger than 0.02 the amount of discarded or retained fills varies only slightly, which explains the threshold selected. Fig. 6.2 shows how the models of the retained fills provide a good description of luminosity evolution.

We also carried out a study on the stability of the fit parameters. The test is

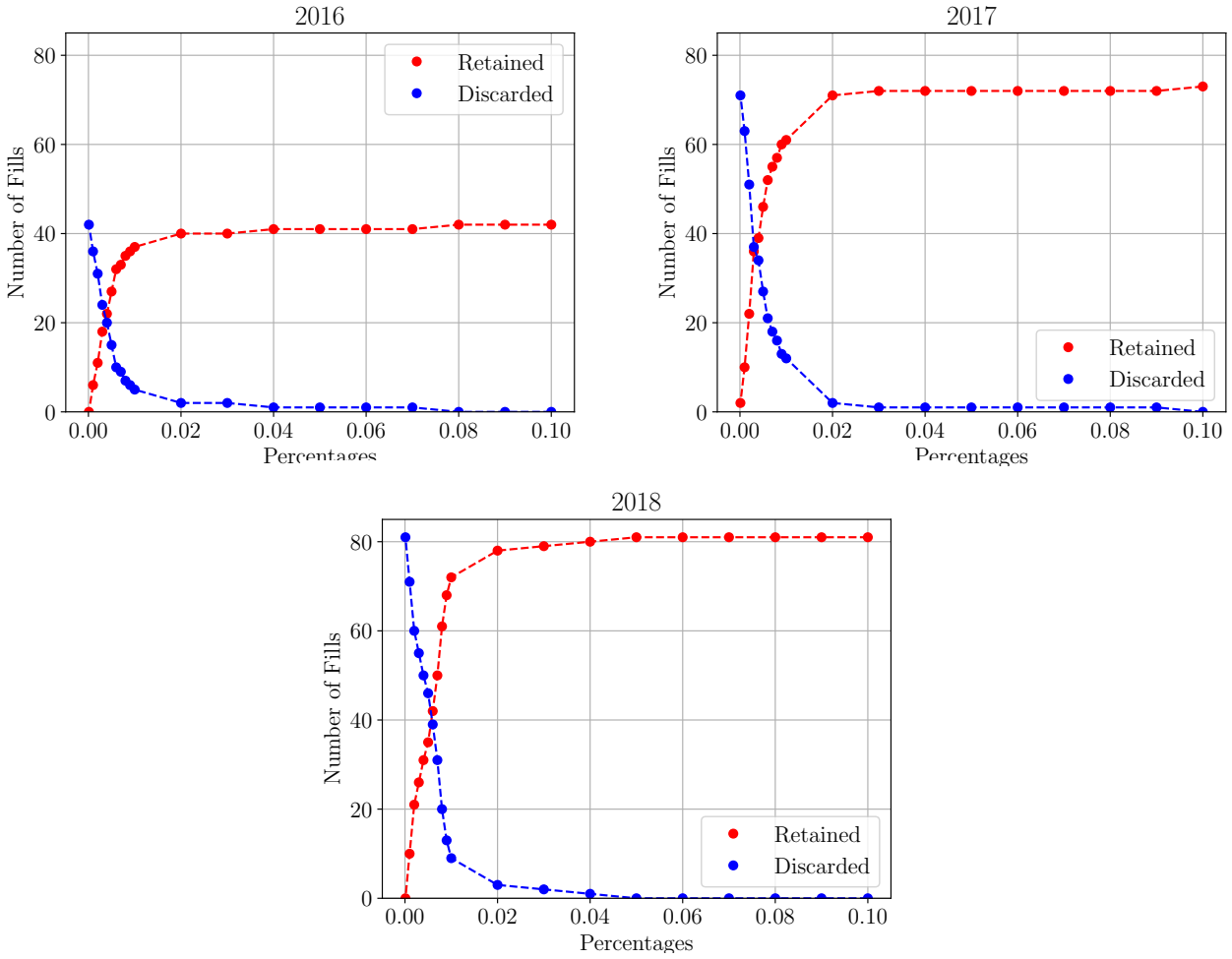

Figure 6.1: Plots of the convergence study for the selection criterion for the three years of Run 2 (upper left 2016, upper right 2017 and at the bottom 2018).

carried out by performing the fit of the fill luminosity with the selected model by changing the amount of measured data used in the fit procedure. The results are shown in Figs. 6.3.

The on-line optimisation strategy, based on the solutions of Eq. (5.26), requires the current-fill model, $L(\hat{t}_i)$, and thus, the decision on its length should be taken using the available set of data representing the luminosity evolution. The purpose of this study is to assess whether there is a strong of the fit parameters on time

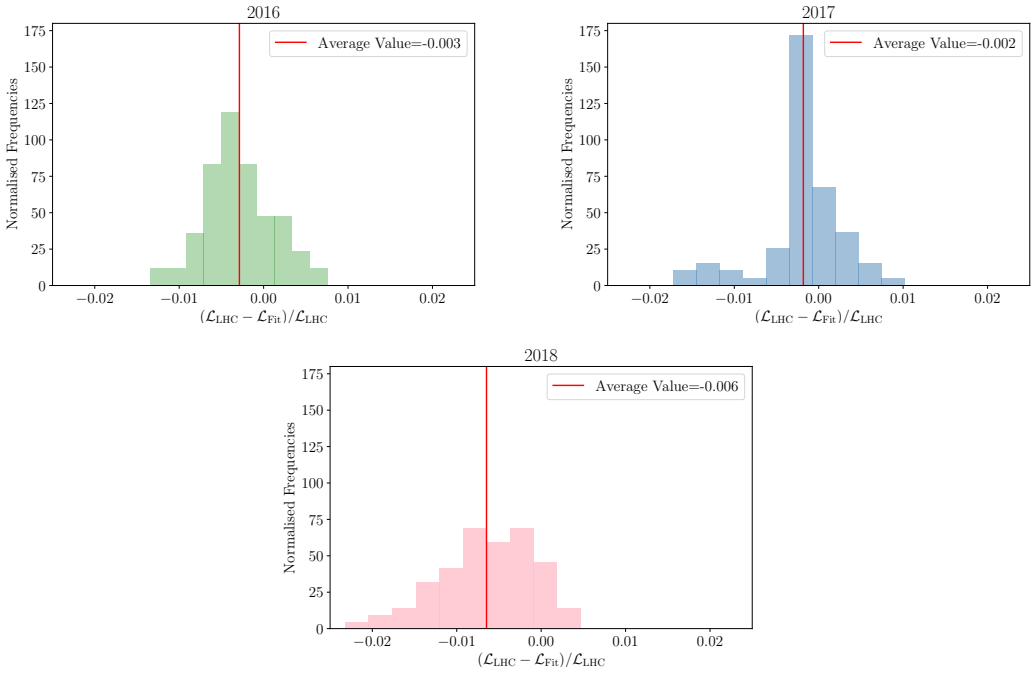

Figure 6.2: Plots showing the distributions of the relative differences between $\mathcal{L}_{\rm Fit}$ and $\mathcal{L}_{\rm LHC}$, for the years 2016 (top-left), 2017 (top-right) and 2018 (bottom). A good agreement is clearly visible.

elapsed since the beginning of the fill. The results of this analysis are shown in Fig. 6.3 where the dependence of the fit parameters b and d as a function of time during the fill is shown for the three years. The different length of the various curves is due to the difference in $t_{\rm fill}$. The observed behaviour is qualitatively the same for the three years. The clear dependence on time of b and d is a sign that the double-exponential model does not capture the full physics of the luminosity evolution. This aspect will be investigated in the future by attempting different models, e.g. the one discussed in Ref. [30].

We remark that the fit parameters a and c do not feature any dependence on time, as is expected, representing the initial value of the luminosity.

6.2 Optimiser Implementation

The optimisation algorithm is implemented by solving Eq. 5.24, which corresponds to solve the system shown in Eq. (5.26). The first step is the derivation of an analytical model for the most probable luminosity, L_{mp}. To this aim, the code performs a scan over time on all the preceding-year fills and evaluates the mode, i.e. the most probable value, of the luminosity distribution at each time value defined in the scan. For each new fill realised during the current year, the algorithm updates the most probable luminosity, and performs a fit of it using the double-exponential model, producing the analytical expression sought. Before starting a new fill, the algorithm checks that the constraint is respected, which is carried out by verifying that the sum of all turn-around times and fill times that have been realised up to that moment is less than the total time set for physics. It is worth stressing that the total time for physics in our simulations is not the actual total time of the LHC physics run for a given year, but rather the sum of all retained fills and the corresponding turn-around time.

The final step consists in finding the solution $\hat{t}$ of the following equation:

$$L_{\mathrm{mp}}(\hat{t}) - \frac{1}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t) = 0\,, \tag{6.2}$$

which is then used to solve

$$L(\hat{t}_i) - \frac{1}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t) = 0\,, \tag{6.3}$$

finding in the end the optimal fill time, $\hat{t}_i$.

6.3 Optimisation Results

We performed the simulation of the on-line optimisation on 2018 data, for different knowledge levels of the luminosity fit models. We started considering a case in which the fit parameters are determined using the full length of the fill, and the results are shown in Fig. 6.4 (top) and Table 6.1.

Table 6.1: Comparison between the actual LHC data and the two optimisations performed.

	Complete Knowledge	**One hour**	**Five hours**
T_{LHC} [days]	57.3		
T_{Online} [days]	53.6	57.2	55.0
T_{Num} [days]	57.3		
$\mathcal{L}_{\mathrm{Tot}}^{\mathrm{LHC}}/day$ [fb^{-1}]	0.61		
$\mathcal{L}_{\mathrm{Tot}}^{\mathrm{Online}}/day$ [fb^{-1}]	0.63	0.59	0.62
$\mathcal{L}_{\mathrm{Tot}}^{\mathrm{Num}}/day$ [fb^{-1}]	0.63		
Relative increase in $\mathcal{L}_{\mathrm{Tot}}^{\mathrm{Online}}/day$	3.8%	−3.0%	2.3%
Relative increase in $\mathcal{L}_{\mathrm{Tot}}^{\mathrm{Num}}/day$	3.6%	-	-

The optimisation acts on the fill length and given that the trend is to reduce it with respect to the actual length, the total time for physics is shorter than the actual one. The top-left plot of Fig. 6.4 shows that the distribution of the fill length for the on-line optimisation is much more similar to that provided by the numerical optimisation, compared to what occurred at LHC. Furthermore, the spread of the distribution is also narrower for the on-line and numerically-optimised cases than for the actual LHC case. Given the difference in effective length of the physics run, the results in terms of integrated luminosity will be normalised providing the collected luminosity per day. The overall gain of integrated luminosity per day is about 3.8% for the on-line optimised case.

The results concerning the case in which the on-line optimisation is carried out using the fill luminosity model derived from the data of the first hour are shown in Fig. 6.4 (centre) and Table 6.1). In contrast to the previous case, the distribution of the fill times has an average value that is larger than that of the actual LHC case and is considerably broader. This feature is introduced by the values of the luminosity models that, being evaluated over one hour, only, are not accurate enough. This induces a net loss of integrated luminosity of 3%.

Finally, we considered the case corresponding to the on-line optimisation performed using the luminosity models obtained with five hours of data, and the results are shown in Fig. 6.4 (bottom) and Table 6.1). As expected, the results of this last optimisation are rather similar to those of the case of complete knowledge, although the gain in integrated luminosity/day is 2.3%. In conclusion, we can say that the on-line optimisation proposed provides a possible daily integrated luminosity gain up to 3.8%, similarly to what happens in the numerical optimisation case (see Chap. 4). The key to this encouraging result is the accuracy of the luminosity model.

The actual implementation of the proposed on-line optimisation would consist of an *ad hoc* software that continues to interpolate online the evolution of the luminosity of the current fill. Combined with this, the on-line estimate of the optimal fill time should be available for the decision on whether or not use it to end the fill. This should allow us to achieve optimisation results compatible with those of the case called complete knowledge.

6.4 Future Developments

The conclusions reached in this thesis are certainly encouraging, although further investigations are needed before robust conclusions about the efficiency of the proposed method can be drawn.

Future investigations have been planned, such as:

1. As mentioned in section 6.1, for years in which the machine parameters change drastically (e.g. 2016 and 2017), we need to find a more complete way to compute the most probable luminosity. Similarly, a strategy should be devised for the case in which a bootstrap is needed, as there are no previous fills to be used to predict the behaviour of future fills. This case corresponds to the situation of a collider after the start up phase.

2. Using the preliminary studies on turn-around time statistics presented in Appendix A, generate runs with Monte Carlo algorithms for cases in which

the optimised physics time is shorter than the expected (as in the case of on-line optimisation presented in the previous chapter).

3. Devising more complex optimisation strategies like the one presented in Appendix B, which take into account the failure probability of a fill in the optimisation of its duration.

4. Improve the modelling of luminosity evolution by using more refined physics models. After that, machine learning techniques might be considered, too.

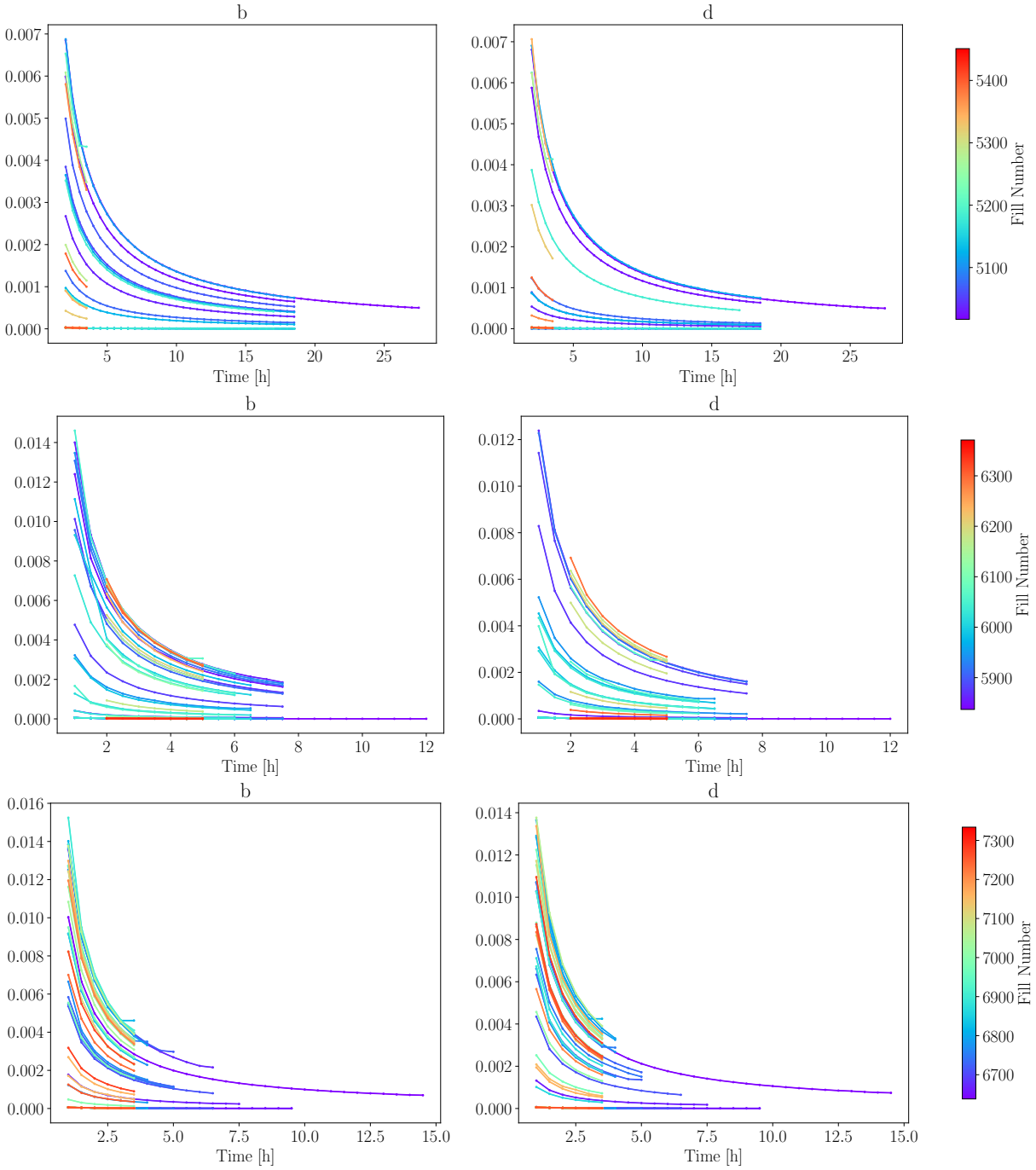

Figure 6.3: Variation of the fit parameters (see Eq. (3.3)) as a function of the amount of data used in the fit, which corresponds to the time elapsed during the fill. The parameters b (left) and d (right) are shown for the 2016 (top), 2017 (centre), and 2018 (bottom), respectively. The curves cover the actual fill length. The observed behaviour is qualitatively the same for the three years.

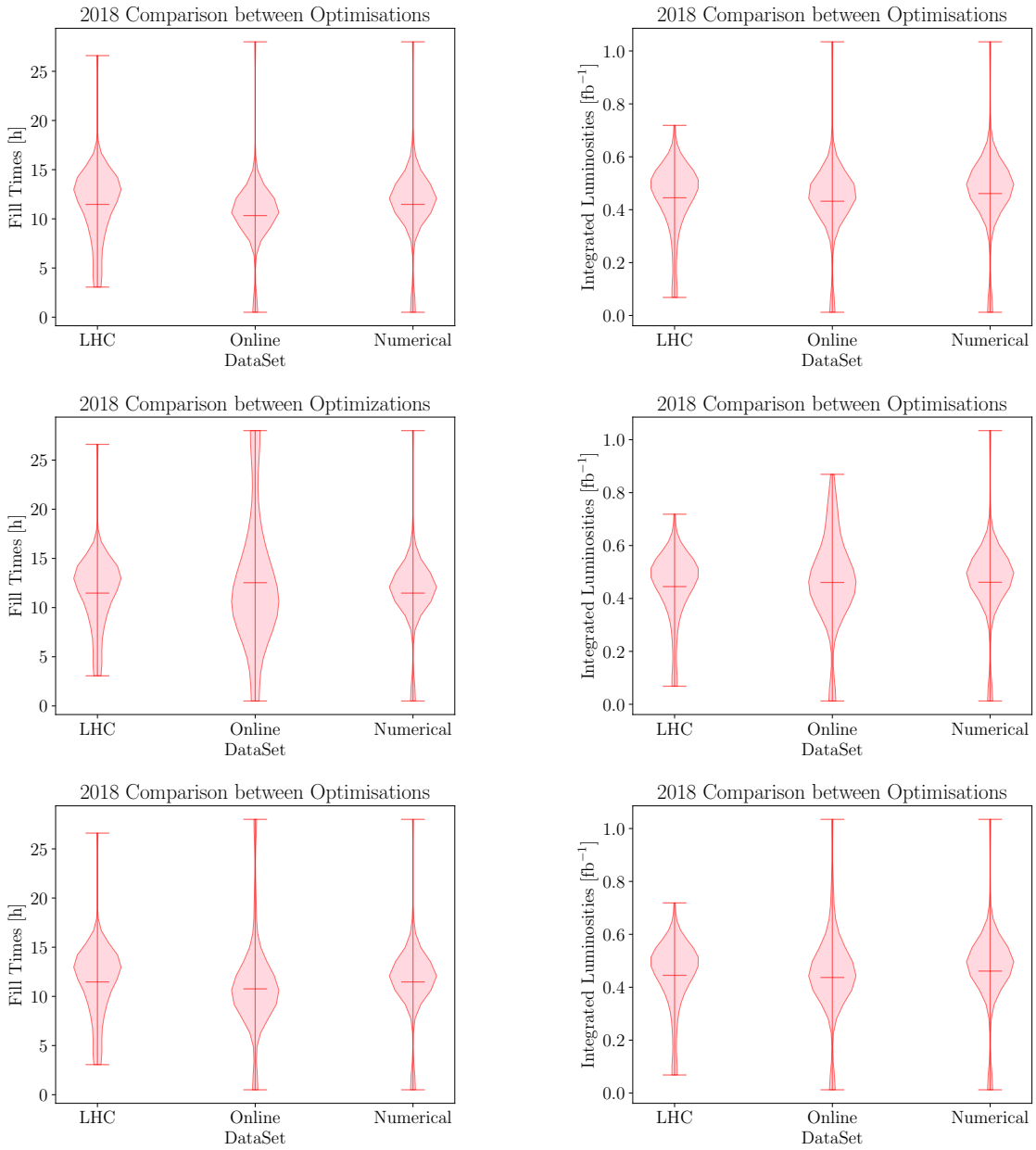

Figure 6.4: The graphs show the on-line optimisation results in the case of complete knowledge of the fitting models (top), partial knowledge, i.e. with the model determined with one hour of luminosity data (centre), and with five hours of luminosity data (bottom). The left plots show the distribution of the actual LHC fill times, of the on-line-optimised fill times, and of the numerically-optimised fill times. The right plots show the distribution of the integrated luminosity in each fill for the LHC run, for the on-line-optimised fill times, and of the numerically-optimised fill times.

Chapter 7

Conclusions

In this thesis, we have shown the it is potentially possible to optimise the collection of the LHC integrated luminosity only by implementing an appropriate algorithm aimed at determining the optimal fill length (see Chapter 4 and 6). The results reported in this thesis are not final, and a number investigations have been outlined and will be part of future activities (see Section 6.4). These new aspects should make the proposed approach more realistic and hence, the possible improvements in the collection of integrated luminosity more reliable.

All in all, we believe that this line of research might bring some useful and interesting outcomes with respect to the efficiency of the luminosity collection by the CERN LHC.

Appendix A

Statistics of the Turn-around Time

In this appendix, we analysed the data of LHC Run 2 concerning the turn-around times of the years 2016, 2017 and 2018, excluding those that represent exceptional conditions, and therefore negligible for statistical studies. [23] presents an overview of Run 2 data.

A.1 Kolmogorov-Smirnov Test

The first step of data analysis was the execution of a statistical test of Kolmogorov-Smirnov on pairs of samples [32]. These test statistics are used to determine whether two distributions (e.g. F(x), G(x)) differ or whether an underlying probability distribution differs from a hypothesised distribution.

Considering two independent samples, there will be two possible hypothesis:

$$\begin{cases} H_0 : \ F(x) = G(x) \\ H_1 : \ F(x) \neq G(x) \end{cases} \tag{A.1}$$

In this case, there are three datasets to compare: the 2016 sample, the 2017 sample, and the 2018 sample. The results of the test are shown in Table A.1.

For this test, a significance level, $1 - \alpha$, of 95% was imposed, which entails α= 0.05.

Table A.1: Results of the Kolmogorov-Smirnov test.

Samples	**2016 - 2017**	**2016 - 2018**	**2017 - 2018**
KS Statistics ($D_{\#,n}$)	0.11	0.16	0.06
p-value	0.21	0.02	0.75

To interpret the results, first we observe the p-value: if the p-value is bigger than α it is not possible to reject H_0, while if it is smaller the rejection is possible.

Looking at the obtained p-values, we have that:

1. p-value(2016-2017)$> \alpha$: The distribution of the 2016 sample is the same as that of the 2017 one;

2. p-value(2016-2018)$< \alpha$: The distribution of the 2016 sample is not the same as that of the 2018 one;

3. p-value(2017-2018)$> \alpha$: The distribution of the 2017 sample is the same as that of the 2018 one.

Secondly, we look at the statistics. A 95% confidence level critical value is defined as follows:

$$D_{\text{crit},0.05} = 1.36\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\,, \tag{A.2}$$

where n_1 and n_2 are the number of elements of the two studied samples. In this case, if the critical value D_{crit} is bigger than the obtained statistics D_n it is possible to accept H_0, while if it is smaller H_0 has to be rejected. Considering that $n_1 = 156$, $n_2 = 194$ and $n_3 = 218$, we have that:

$$\begin{cases} D_{1,\text{crit},0.05} = 0.15 \\ D_{2,\text{crit},0.05} = 0.14 \\ D_{3,\text{crit},0.05} = 0.13 \end{cases} \tag{A.3}$$

This means that:

1. $\underline{D_{1,\mathrm{crit},0.05} > D_{1,n}}$: The distribution of the 2016 sample is the same of the 2017 one;

2. $\underline{D_{2,\mathrm{crit},0.05} < D_{2,n}}$: The distribution of the 2016 sample is not the same of the 2018 one;

3. $\underline{D_{3,\mathrm{crit},0.05} > D_{3,n}}$: The distribution of the 2017 sample is the same of the 2018 one.

In conclusion, it is possible to state that 2016 data seem to be distributed similarly to the 2017 and similarly to the 2017 and 2018 data. However, for the 2016 and 2018 samples, it is not possible to accept the hypothesis of equal distributions. The latter conclusion can be linked to the consistent difference in the number of objects in the two samples, and to the changes that are made to the machine (i.e. LHC) from year to year.

A.2 Evaluation of the t_{ta} distribution

Knowledge of the statistical distribution of the turn-around time is essential for building an optimisation strategy for the integrated luminosity, as pointed out in section 5.3. For this reason, we analysed the data of the LHC Run 2 by producing histograms that describe the distribution of the turn-around time, and by looking for a model that would better fit them (see Fig. A.1). The first model chosen for the fit of the experimental data is the exponential one with an offset, which is able to represent rather well the tail of the data but not the peak:

$$f(x) = \lambda \exp\big(-\lambda(x - B)\big), \tag{A.4}$$

The expectation value for this model is $E[x] = e^{\lambda b}/\lambda$. For the bad representation of the peak, we have also considered a power law, which acts contrary to the exponential, representing quite well the peak of the data and badly the tail:

$$f(x) = \frac{A}{(x - B)^n}. \tag{A.5}$$

Finally, combining the two previous models, we have come to the following truncated power-law model:

$$f(x) = \frac{A}{(x-B)^n} \exp(-\lambda x), \tag{A.6}$$

whose first momentum (i.e. the expectation value) is $E[t] = \int_{t_{\min}}^{t_{\max}} dt' f(t') t'$. In Fig. A.1, it is possible to view all fits performed with the respective reduced chi-square values, which are now also shown in the next table for greater clarity.

Table A.2: Results of the Reduced Chi-Square $\tilde{\chi}^2$ test.

Samples	**2016**	**2017**	**2018**
Power Law $[\times 10^{-4}]$	1.4	5	7
Exponential Law $[\times 10^{-4}]$	7	9	6
Truncated Power Law $[\times 10^{-4}]$	2	5	6

Observing the Table A.2, it is possible to state that the best fit is the one obtained with the truncated power-law model. In Fig. A.2, we have reported the model chosen fit with the parameters (presented also in Table A.3) obtained for the various years analysed.

Table A.3: Parameters from the fits of data.

Parameters	**2016**	**2017**	**2018**
$\boldsymbol{A}$ $[\mathrm{h}^{1+n}]$	0.4 ± 0.1	0.30 ± 0.08	0.3 ± 0.2
$\boldsymbol{B}$ [h]	2.000 ± 0.005	2.0 ± 1.9	1.5 ± 0.1
n	0.7 ± 0.7	0.5 ± 0.6	0.3 ± 0.9
λ $[\mathrm{h}^{-1}]$	0.10 ± 0.04	0.1 ± 0.1	0.2 ± 0.2

Once the truncated power-law fitting model was established, we evaluated the averages as shown in Table A.4.

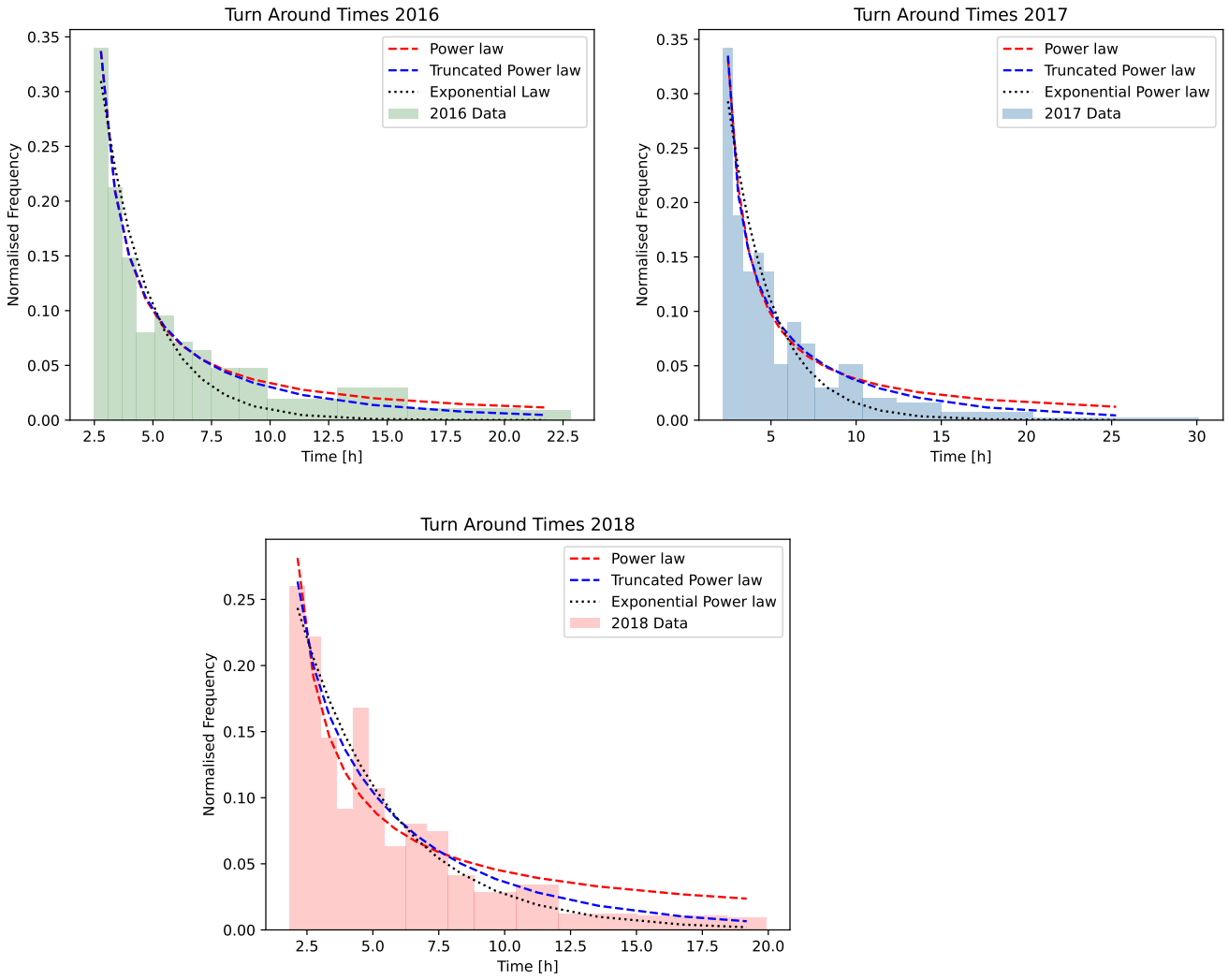

Figure A.1: The fits for the 2016, 2017 and 2018 with the previously described functions. All the plots present in the figure have been obtained with the `lmfit` python function.

Table A.4: Averages turn-around time for the years 2016, 2017, 2018.

	2016	**2017**	**2018**
Average Turn-around Time [h]	8.9	8.9	7.2

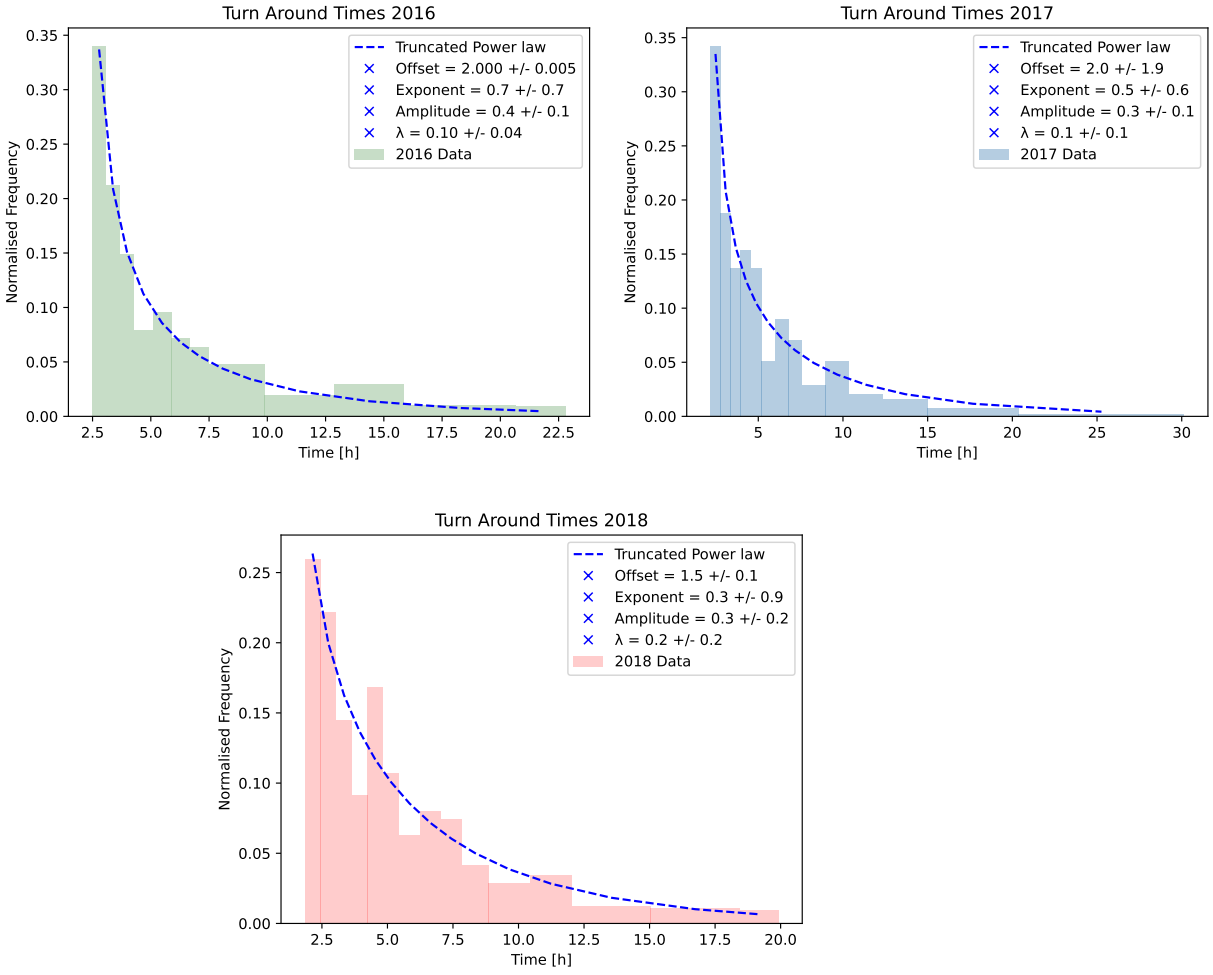

Figure A.2: Plots of the data with the truncated power-law model for the turn-around times distributions. All the plots present in the figure have been obtained with the `lmfit` python function.

Appendix B

Failure Probability of a Fill Optimisation Strategy

In this appendix we present the case in which we take into account the failure probability of a fill. There are two options to proceed: assuming a constant failure probability p, or assume a probability that depends on the fill duration $p(t_{\mathrm{f}})$, and of course it is assumed that each fill is statistically independent.

In the first case, we can consider

$$\mathcal{L}_{\mathrm{tot}}(\hat{t}_i, \hat{t}) = \sum_{j=1}^{i-1} \mathcal{L}_{\mathrm{tot}}(\hat{t}_j) + (1-p) \int_0^{\hat{t}_i} \mathrm{d}t\, L(t) + \kappa\,(1-p)^{\kappa} \int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t)\,, \qquad \text{(B.1)}$$

where

$$\kappa(\hat{t}_i, \hat{t}) = \frac{T - \sum_{j=1}^{i-1} \left(t_j + \hat{t}_j\right) - t_i - (1-p)\hat{t}_i}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \qquad \text{(B.2)}$$

for simplifying the notation. Note in the estimate of the luminosity contribution from future fills, we neglect the fact that in case of failure of the current fill, some integrated luminosity might have been collected, and $\hat{t}_j$ should be intended as the actual fill length even if it might be shorter than the optimal value due to a failure.

The optimal values t_i and $\hat{t}_i$ are the solutions of

$$\begin{cases} \dfrac{\partial \mathcal{L}_{\mathrm{tot}}}{\partial \hat{t}_i} & = (1-p)\, L(\hat{t}_i) + \dfrac{\partial \left[\kappa\,(1-p)^{\kappa}\right]}{\partial \hat{t}_i} \displaystyle\int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t) = 0 \\ \\ \dfrac{\partial \mathcal{L}_{\mathrm{tot}}}{\partial \hat{t}} & = \kappa\,(1-p)^{\kappa}\, L_{\mathrm{mp}}(\hat{t}) + \dfrac{\partial \left[\kappa\,(1-p)^{\kappa}\right]}{\partial \hat{t}} \displaystyle\int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t) = 0\,, \end{cases} \tag{B.3}$$

which gives

$$\begin{cases} (1-p)\, L(\hat{t}_i) - (1-p)^{\kappa} \dfrac{(1-p)}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} [\ln(1-p)+1] = 0 \\ \\ L_{\mathrm{mp}}(\hat{t}) - \dfrac{[\ln(1-p)+1]}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \displaystyle\int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t) = 0\,. \end{cases} \tag{B.4}$$

In the second case, we should consider

$$\mathcal{L}_{\mathrm{tot}}(\hat{t}_i, \hat{t}) = \sum_{j=1}^{i-1} \mathcal{L}_{\mathrm{int}}(\hat{t}_j) + (1-p(\hat{t}_i)) \int_0^{\hat{t}_i} \mathrm{d}t\, L(t) + \kappa \left(1-p(\hat{t})\right)^{\kappa} \int_0^{\hat{t}} \mathrm{d}t\, L_{\mathrm{mp}}(t)\,, \tag{B.5}$$

where

$$\kappa(\hat{t}_i, \hat{t}) = \frac{T - \sum_{j=1}^{i-1} \left(t_j + \hat{t}_j\right) - t_i - (1-p(\hat{t}_i))\hat{t}_i}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}}\,, \tag{B.6}$$

and the solution can be found by taking the derivatives of Eq. (B.5).

Appendix C

Developed Software

All codes implemented during the thesis work are available in the following GitHub repositories:

1. **Numerical Optimisation:**
   github repository clone GiuliaFaletti/NumericalOptimization
   `https://github.com/GiuliaFaletti/NumericalOptimization.git`;
2. **Optimisation Strategies:**
   github repository clone GiuliaFaletti/OptimizationStrategies
   `https://github.com/GiuliaFaletti/OptimizationStrategies.git`;
3. **Turn-Around Statistics:**
   github repository clone GiuliaFaletti/LuminosityOptimization
   `https://github.com/GiuliaFaletti/LuminosityOptimization.git`.

Bibliography

[1] S.Y. Lee, **Accelerator Physics**, *2nd ed.* World Scientific, URL: `https://cds.cern.ch/record/939203` (London, 2004).

[2] M. Hostettler, **LHC Luminosity Performance**, PhD Thesis, in CERN-THESIS-2018-051 (Bern, 2018).

[3] K. Willie, **The physics of particle accelerators: an introduction**, Oxford Univ. Press, URL: `https://cds.cern.ch/record/560708` (Oxford, 2000).

[4] B. J. Holzer, **Introduction to Longitudinal Beam Dynamics**, in CERN Accelerator Schools, pp. 47-61, DOI: 10.5170/CERN-2014-005.41 (2013).

[5] M. Giovannozzi, **Accelerator Physics**, Lecture Notes (AA.2019-2020).

[6] A.C. Mueller, **Review of Accelerators for Accelerator-Driven Systems (ADS)**, in Thorium Energy for the World, pp. 243-248 (2016).

[7] W. Herr and B. Muratori, **Concept of Luminosity**, in CAS - CERN Accelerator School: Intermediate Accelerator Physics, pp. 361-378, DOI: 10.5170/CERN-2006-002.361 (2006).

[8] L. Evans and P. Bryant, **LHC Machine**, in JINT, Vol. **3**, p. S08001, DOI: 10.1088/1748-0221/3/08/S08001 (2008).

[9] J.L. Sirvent Blasco, **Beam secondary shower acquisition design for the CERN high accuracy wire scanner**, PHD Thesis, Barcelona University (September 2018).

[10] E. Shaposhnikova et al., **LHC Injectors Upgrade (LIU) project at CERN**, in Proceedings of IPAC2016, pp. 992-995, URL: https://cds.cern.ch/record/2207337/files/mopoy059.pdf (Busan, 2016).

[11] M. Gyr, **Polarities of LHC septum magnets MSI and MSD** (2008).

[12] P. Hopchev, **Absolute luminosity measurements at LHCb**, PHD Thesis, Grenoble University (2006).

[13] H. Burkhardt and I. Efthymiopoulos, **Chapter 8: Interface with Experiments. Interface with Experiments**, in CERN Yellow Report, DOI: 10.5170/CERN-2015-005.157 (May 2017).

[14] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole and P. Proudlock, **LHC Design Report Vol.1: The LHC Main Ring**, DOI: 10.5170/CERN-2004-003-V-1.

[15] W. Herr, **Kinematic of Particle Beams I - Relativity**, in URL: https://indico.cern.ch/event/226938/contributions/478804/attachments/373731/519906/rel1.pdf.

[16] H. Schmickler, **Luminosity and (future) colliders**, in CERN Accelerator School 2019, URL: https://indico.cern.ch/event/808940/contributions/3553685/attachments/1909425/3154647/luminosity_colliders.pdf.

[17] R. Calaga, **Crab Crossing for the LHC Upgrade**, in Proceedings of SRF2011, URL: https://accelconf.web.cern.ch/SRF2011/papers/friob05.pdf (Chicago, 2011)

[18] W. Herr and T. Pieloni, **Beam-Beam Effects**, URL: https://cds.cern.ch/record/941319/files/p379.pdf (2014).

[19] P. Grafström and W. Kozanecki, **Luminosity determination at proton colliders**, in Prog. Part. Nucl. Phys., Vol. **81**, pp. 97-148, DOI: 10.1016/j.ppnp.2014.11.002 (2015).

[20] M. Benedikt, D. Schulte and F. Zimmermann, **Optimising integrated luminosity of future hadron colliders**, in Physical Review Special Topics - Accelerators and Beams, DOI: 10.1103/PhysRevSTAB.18.101002 (October 2015).

[21] V. Shiltsev and E. McCrory, **Characterizing Luminosity Evolution in the Tevatron**, in Proceedings of the 2005 Particle Accelerator Conference, pp. 2536-2537, DOI: 10.1109/PAC.2005.1591170 (2005).

[22] G. Papotti, **Observations on Luminosity curves**, presented at the 38th Meeting of LHC Beam Operation Committee, URL: `https://lhc-beam-operation-committee.web.cern.ch/lhc-beam-operation-committee/minutes/Meeting38-22_05_2012/lumi.GP.22May12.pdf` (May 2012).

[23] J. Wenninger, **Operation and Configuration of the LHC in Run 2**, in CERN-ACC-NOTE-2019-0007 (March 2019).

[24] CERN, **ATLAS Experiment**, URL: `https://atlas.cern/`.

[25] CERN, **"Massi File" Specifications**, in LHC Program Coordination URL: `https://lpc.web.cern.ch/MassiFileDefinition_v2.htm`.

[26] CERN, **Fill Table**,
URL: `https://lpc.web.cern.ch/cgi-bin/showFillTable.py`.

[27] R. Johansson, **Numerical Python: Scientific Computing and Data Science Applications with Numpy, SciPy and Matplotlib**, *Apress* (2018).

[28] J. Nocedal and S.J. Wright, **Numerical Optimisation**, *Springer* (New York, 2006).

[29] F. Zhengqing, L. Goulin and G. Lanlan, **Sequential Quadratic Programming Method for Nonlinear Least Squares Estimation and Its Ap-**

plication, in Mathematical Problems in Engineering, DOI: https://doi.org/10.1155/2019/3087949 (2019).

[30] M. Giovannozzi and F.F. Van der Veken, **Description of the luminosity evolution for CERN LHC including dynamic aperture effects, Part I: The model**, in Nuclear Inst. and Methods in Physics Research, Vol. **A**, 905, pp. 171-179 (2018).

[31] M. Eleutieri, **Analisi II - Ottimizzazione Vincolata**, Lecture Notes.

[32] Yadolah Dodge, **The Concise Encyclopedia of Statistics**, DOI: https://doi.org/10.1007/978-0-387-32833-1, *Springer* (New York, NY).