Alma Mater Studiorum · University of Bologna

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

Dynamic Aperture Model for LHC Integrated Luminosity Optimisation

Supervisor:
Prof. Massimo Giovannozzi

Co-supervisor:
Dr. Antonio Sbrizzi

Submitted by:
Thiziri AMEZZA

Academic Year 2022/2023

"Chaque génération, sans doute, se croit vouée à refaire le monde. La mienne sait pourtant qu'elle ne le refera pas. Mais sa tâche est peut-être plus grande. Elle consiste à empêcher que le monde se défasse."

Albert Camus.

Abstract

In colliders particle physics experiments, the beam energy and the luminosity are the key figures of merit. To probe new physics with particle colliders requires a high centre-of-mass energy. In addition, to reduce statistical errors, a large number of events should be generated, especially when exploring rare events with a small production cross section.

In this work, the focus is on analysing the evolution of luminosity at LHC during Run 2, which is the basis for providing strategies for integrated luminosity optimisation. To achieve this, we study an efficient model describing the luminosity evolution in a circular collider based on the evolution of the dynamic aperture. This model takes into account various factors such as burn-off and additional pseudo-diffusive effects.

Introduction

In a colliding storage ring, such as the CERN Large Hadron Collider (LHC), luminosity is a crucial indicator for assessing the overall machine performance, as its integrated value is proportional to the data collected by physics detectors.

During a collider run, the primary objective is to maximise the integrated luminosity accumulated over the designated time for physics, which, in turn, leads to an increased number of interactions between particles, thereby enhancing the likelihood of discovering new physics phenomena.

However, these high-energy collisions also give rise to different beam dynamics processes that can lead to particle losses. Minimising these losses is one of the key strategies to boost the delivered luminosity.

The scope of this work is to study a scaling law for luminosity evolution, incorporating both burn-off and additional pseudo-diffusive effects. Notably, the proposed framework can accommodate the time dependence of various beam parameters providing a comprehensive and flexible model.

A dynamic aperture model is used to analyse the data collected by the ATLAS experiment during the LHC Run 2. By employing this sophisticated model, we aim to gain deeper insights into beam dynamics and its evolution in terms of luminosity.

Ultimately, the goal is to enhance the performance of a circular collider by optimising and pushing the limits of integrated luminosity.

Contents

List of Figures

List of Tables

Chapter 1

Principles of Beam Dynamics

Beam dynamics focuses on understanding how charged particles behave inside an accelerator. Indeed, the phase-space region of particle motion within a collider might be affected by the presence of non-linearities in the magnetic field of the superconducting machine. Therefore, it is essential to provide a thorough description of beam dynamics. In this chapter, we present a brief overview of single-particle dynamics with particular emphasis on the transverse beam dynamics and the relativistic nature of the particle's motion.

1.1 Basic Concepts

In a circular collider like the LHC, the particle beams are contained around a reference trajectory, which is a circular closed orbit. On top of this, the motion of a charged particle is guided and controlled by the presence of electromagnetic fields, which give rise to a force known as the Lorentz force, represented as

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

The driving force $\vec{F}$ acting on a charged particle with charge e and velocity $\vec{v}$, as described in Eq. (1.1), is a fundamental framework for understanding how charged particles interact with electromagnetic fields. The first term in Eq. (1.1) represents the effect of the electric field $\vec{E}$ that accelerates the charged particle along the longitudinal direction of the motion (synchrotron motion). The second term is the transverse force needed to focus and bend the particle around the

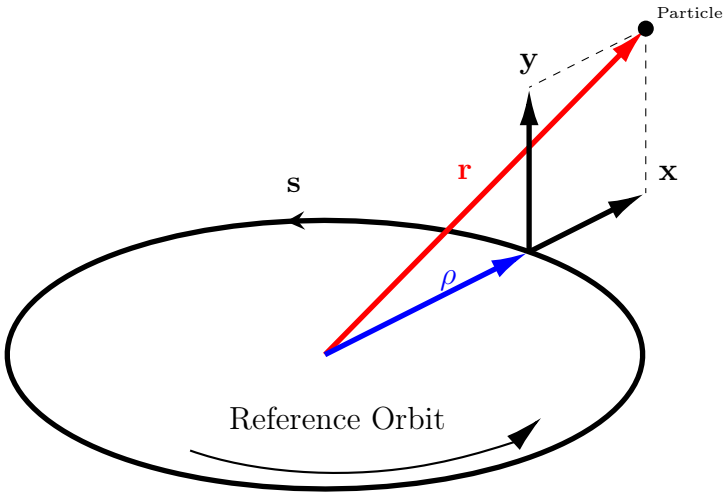


Figure 1.1: The local comoving coordinate system (x, y, s) for particle motion along a circular accelerator, with an ideal closed orbit whose radius of curvature is ρ.

reference orbit (betatron motion); this component is perpendicular to both the particle velocity $\vec{v}$ and the magnetic field direction.

To describe the particle dynamics around a circular, closed orbit, one often introduces a comoving local coordinate system (x, y, s), the so-called *Frenet-Serret coordinate system*[1] for convenience see Fig. 1.1. The s coordinate describes the longitudinal position of the particle along the reference orbit, where x and y denote the transverse dimensions.

1.2 Transverse Dynamics

In a colliding beam storage ring such as the LHC, the beam dynamics are dominated by magnetic fields in the transverse planes. The transverse motion of a particle inside the bunch is complex and can exhibit a non-linear behaviour [14]. To keep the trajectory of a particle within the reference orbit with a curvature radius ρ, a constant magnetic field is applied along the y plane $\vec{B}(0, 0, s) = (0, 0, B_y)$. This bending radius ρ results from the equilibrium between the centrifugal force and the centripetal Lorentz force acting on the particle with momentum $p = mv$,

[1] One curvilinear coordinate system for particle motion along the reference orbit.

where the quantity $B\rho$ is the so-called *beam rigidity*, expressed as

$$B\rho = \frac{p}{e} \tag{1.2}$$

Equation (1.2) means that a particle with higher momentum requires a stronger magnetic field to maintain the circular orbit radius ρ constant, thus ensuring that the particle does not deviate from its path within the accelerator.

In this context, it becomes crucial to evaluate the contribution of the magnetic field to particles close to the design orbit $(0, 0, s)$, and the typical approach is to expand the magnetic field into its multipolar components, namely

$$B_y(x, y) + iB_x(x, y) = \frac{p}{e} \sum_n (b_n + ia_n)(x + iy)^{n-1}, \tag{1.3}$$

where the coefficients a_n and b_n are the skew and upright multipolar components, respectively, and, i.e $n = 1$ corresponds to a dipole, $n = 2$ is quadruple, etc.

Magnetic element designs generate fields with only one multipolar component. Accordingly, multipolar expansion allows for breaking down the magnetic field into upright or skew dipoles, quadrupoles, sextupoles, octupoles, and so on, depending on their effect on the particle's dynamics.

Generally, the dominant forces in an accelerator are dipolar (bending) and quadrupolar (focusing - defocusing) fields. Dipole magnets guide and curve the trajectory of the particles, allowing them to follow the circular path. Simultaneously, quadrupole magnets serve the dual purpose of focusing and defocusing the particle beam within the transverse plane. Higher-order and skew multipoles, on the other hand, are commonly used as correctors to address higher-order effects and imperfections in the magnetic field. They play a crucial role in fine-tuning the particle beam and compensating for deviations from the desired trajectory caused by higher-order magnetic field components.

1.3 Linear Beam Dynamics

The configuration of dipoles and quadrupoles around the whole storage ring in both transverse planes is known as *accelerator optics* or *magnetic lattice*. In this lattice design, the absence of skew and multipoles results in linear beam dynamics and allows the equations of motion to be treated independently in the two transverse

planes. Hence, these equations can be written in terms of some focusing function, often denoted by $k(s)$[2], and using the magnetic rigidity (see Eq. (1.2)), the solution of Eq. (1.1) is given by the following second-order differential equations

$$x''(s) + \left(\frac{1}{\rho^2(s)} - k(s)\right) x(s) = \frac{1}{\rho(s)} \frac{\Delta p}{p_0} \tag{1.4}$$

$$y''(s) + k(s)y(s) = 0 \tag{1.5}$$

as a synthesis of the inhomogeneous equation Eq. (1.4) and the homogeneous one Eq. (1.5) with s-dependent coefficients. The quadrupole strength $k(s)$ and the dipolar radius ρ satisfy a periodic condition, and the quantity $\Delta p/p_0$ is the fractional momentum offset (p_0 is the design momentum of the reference particle).

In the absence of any coupled motions, Eqs.(1.4), and (1.5) are linear, and have the characteristics of *Hill's equation*, defined as

$$u''(s) + k_u(s)u(s) = 0 \tag{1.6}$$

where u stands either for x or y, while $k_u(s) \equiv \{k_x(s) \text{ or } k_y(s)\}$, with $k_x(s) = -\left(k(s) - \frac{1}{\rho^2(s)}\right)$ and $k_y(s) = k(s)$. As we proceed, the concept of phase-space, which corresponds to the coordinate space (u, u'), is introduced. The coordinate u represents the local transverse positions (x or y as illustrated in Fig.1.1), while u' is the angle (particle slope). Later in this chapter, the momentum component p_u will also become relevant.

1.3.1 Lattice Maps

To clarify the structure of the linear betatron motion, the concept of lattice maps comes into play [16]. By imposing initial conditions at $s = s_0$, Hill's equation for piece-wise constant k_u treated as a harmonic oscillator with solution

$$\begin{aligned} u(s) &= C\left(s \mid s_0\right) u(s_0) + S\left(s \mid s_0\right) u'(s_0) \\ u'(s) &= C'\left(s \mid s_0\right) u(s_0) + S'\left(s \mid s_0\right) u'(s_0) \end{aligned} \tag{1.7}$$

$C\left(s \mid s_0\right)$ and $S\left(s \mid s_0\right)$ hold as the cosine-like and sine-like solutions, respectively, and their derivatives with respect to s are $C'\left(s \mid s_0\right)$ and $S'\left(s \mid s_0\right)$. The boundary

[2] An function of the quadrupole magnetic field strength, dependent on the longitudinal position (s).

conditions for these solutions, determined by the initial conditions, are outlined as follows

$$C\left(s_0 \mid s_0\right)=1,\ \ S\left(s_0 \mid s_0\right)=0,\ \ C'\left(s_0 \mid s_0\right)=0,\ \ S'\left(s_0 \mid s_0\right)=1. \tag{1.8}$$

By representing the phase-space coordinates (u, u') at location s with the state-vector $\mathbf{u}(\mathbf{s})$, i.e.

$$\mathbf{u}(\mathbf{s})=\begin{pmatrix} u \\ u' \end{pmatrix}_s \tag{1.9}$$

The solution to Hill's equation can then be elegantly expressed in a matrix formalism as follows

$$\begin{pmatrix} u \\ u' \end{pmatrix}_s = \mathcal{M}\left(s \mid s_0\right)\begin{pmatrix} u \\ u' \end{pmatrix}_{s_0} = \begin{pmatrix} C\left(s \mid s_0\right) & S\left(s \mid s_0\right) \\ C'\left(s \mid s_0\right) & S'\left(s \mid s_0\right) \end{pmatrix}\begin{pmatrix} u \\ u' \end{pmatrix}_{s_0} \tag{1.10}$$

here, $\mathcal{M}\left(s \mid s_0\right)$ represents the betatron transport matrix from a point s_0 to s, and has unit determinant $det(\mathcal{M}\left(s \mid s_0\right))=1$.

In accelerator design, a sequence of magnetic elements is involved, as sketched in Fig.1.2. Hence, this configuration can be represented through a series of matrix multiplications, notably

$$\mathcal{M}\left(s_n \mid s_0\right)=\mathcal{M}\left(s_n \mid s_{n-1}\right)\ldots\mathcal{M}\left(s_3 \mid s_2\right)\cdot\mathcal{M}\left(s_2 \mid s_1\right)\cdot\mathcal{M}\left(s_1 \mid s_0\right) \tag{1.11}$$

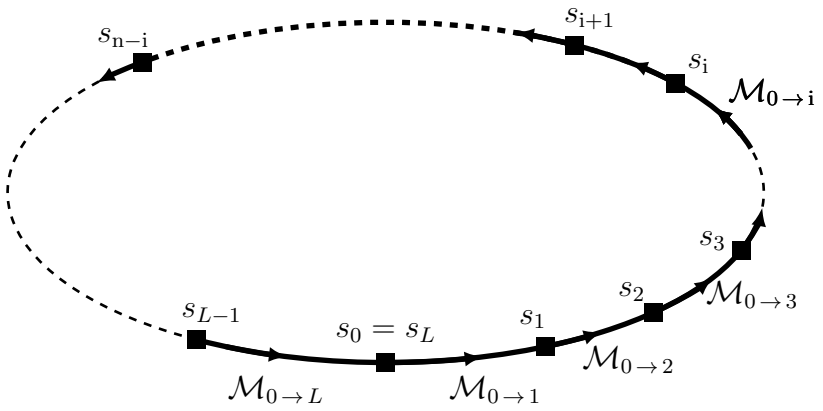


Figure 1.2: Schematic view of a circular accelerator as a series of ith magnetic elements, with i $= 0, 1, 2, .., \mathrm{n}$. s_L represents the length of the accelerator, for a section starting at $s_0 = 0$ and ending at $s_0 = s_L$, it forms a complete one-turn map.

The matrices for transverse planes given as

$$\begin{pmatrix} x(s) \\ x'(s) \end{pmatrix} = \begin{pmatrix} C_x(s) & S_x(s) \\ C'_x(s) & S'_x(s) \end{pmatrix} \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix} \tag{1.12a}$$

$$\begin{pmatrix} y(s) \\ y'(s) \end{pmatrix} = \begin{pmatrix} C_y(s) & S_y(s) \\ C'_y(s) & S'_y(s) \end{pmatrix} \begin{pmatrix} y_0 \\ y'_0 \end{pmatrix} \tag{1.12b}$$

By combining the above matrices (1.12a) and (1.12b), we achieve a comprehensive 4×4 matrix representation, reads

$$\begin{pmatrix} x(s) \\ x'(s) \\ y(s) \\ y'(s) \end{pmatrix} = \begin{pmatrix} C_x(s) & S_x(s) & 0 & 0 \\ C'_x(s) & S'_x(s) & 0 & 0 \\ 0 & 0 & C_y(s) & S_y(s) \\ 0 & 0 & C'_y(s) & S'_y(s) \end{pmatrix} \begin{pmatrix} x_0 \\ x'_0 \\ y_0 \\ y'_0 \end{pmatrix} \tag{1.13}$$

The zeros in the upper and the lower triangular elements of the matrix (1.13) are the signature of an uncoupled motion.

The solution of Eq. (1.6) using the ansatz that incorporates an s-dependent amplitude and phase stated by *Floquet theorem* as

$$u(s) = \sqrt{2J_u\beta_u(s)}cos(\varphi_u(s) + \varphi_{u,0}) \tag{1.14}$$

along the ring, where s is the longitudinal variable, $\sqrt{2J_u\beta_u(s)}$ and $\varphi_u(s)$ are the betatron amplitude and the betatron phase advance, respectively, where J_u and $\varphi_{u,0}$ are constants of integration.

Upon substitution of the Eq. (1.14) into Hill's equation (1.6), a relationship emerging between $\varphi_u(s)$ and $\beta_u(s)$ given by

$$\varphi_u(s) = \oint_0^s \frac{1}{\beta_u(s)} dS \tag{1.15}$$

Additionally, in a machine with a circumference L, the so-called *betatron function* or *β-function* has a periodic condition $\beta_u(s) = \beta_u(s + L)$, and with the constraint $\varphi_u(L) = Q_u$, where Q_u is the transverse betatron tune defined from one turn phase, i.e. number of betatron oscillations around the reference trajectory per revolution, the β-function is related to the machine tune by

$$Q_u = \frac{1}{2\pi} \int_0^L \frac{1}{\beta_u(s)} ds. \tag{1.16}$$

the tune $Q_u \equiv Q_{x,y}$ is imposed by the quadrupole configuration, and its position in the transverse plane Q_x versus Q_y defines the machine's adjustable working point, designed to avoid resonances.

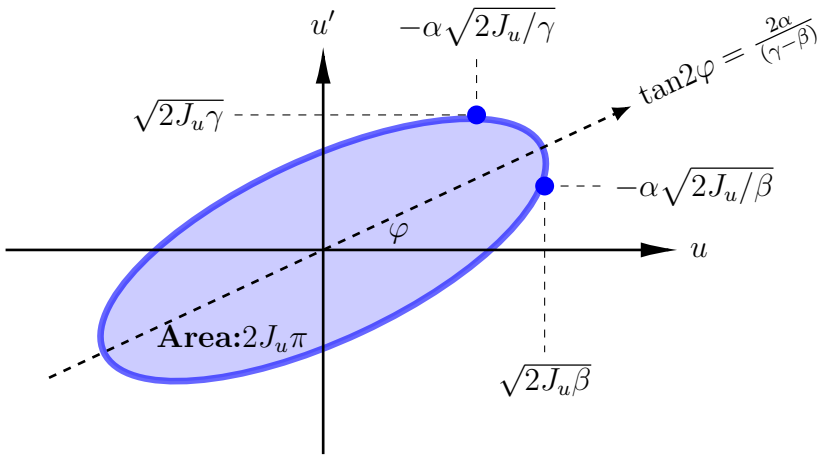


Figure 1.3: The Courant-Snyder ellipse describing the particle motion in a transverse phase-space (u, u').

1.3.2 Courant-Snyder Invariant

The betatron function $\beta_u(s)$ has been identified as a unique function to the configuration of the periodic lattice, and also known as one of *Twiss Parameters*[3] that defines the other two associated parameters $\alpha_u(s)$ and $\gamma_u(s)$, denoted as

$$\alpha_u(s) = -\frac{1}{2}\beta'_u(s), \qquad \gamma_u(s) = \frac{(1+\alpha_u^2(s))}{\beta_u(s)} \tag{1.17}$$

The introduction of twiss parameters yields the definition of the Courant-Snyder Invariant $2J_u$ that maps the betatron motion of a single particle in a phase-space (u, u') by an invariant ellipse, namely

$$2J_u = \gamma_u u^2 + 2\alpha_u u u' + \beta_u u'^2 \tag{1.18}$$

The invariant in Eq. (1.18) remains constant as the particle advances through the accelerator lattice; the ellipse will twist and rotate without changing the area with a constant enclosed area of $2\pi J_u$, as shown in Fig.1.3.

The solution of Hill's equation for a constant k_u takes the following form at a position s within the accelerator after N revolutions:

$$\begin{cases} u(s) = \sqrt{2J_u\beta_u(s)}cos(\varphi_u(s) + N\varphi_{u,0})) \\ u'(s) = -\sqrt{\frac{2J_u}{\beta_u(s)}}[sin(\varphi_u(s) + N\varphi_{u,0})) + \alpha_u(s)cos(\varphi_u(s) + N\varphi_{u,0}))] \end{cases} \tag{1.19}$$

[3]Are the Courant-Snyder parameters $\alpha_u(s)$, β and $\gamma_u(s)$ that are functions of the lattice only and don't depend in the initial condition of the particle.

The Courant-Snyder invariant provides insight into the phase-space evolution of a particle and helps to understand the structure of its linear dynamics within an accelerator. For a beam with many particles, all particles move on nested ellipses. Indeed, knowing how these ellipses transform and evolve is equivalent to understanding the overall dynamics of the particle beam as it travels through the accelerator lattice.

1.3.3 Emittance

Since the twiss parameters that describe the orientation of the phase-space ellipse are independent of the particle's initial conditions, while the location of the particle on the ellipse and the size (area) of the ellipse depends on the initial conditions of the particle, it is convenient to define an alternative, Betatron representation of the orbit formulated as

$$u(s) = \sqrt{\epsilon_u}\sqrt{\beta_u(s)}cos(\varphi_u(s) + \varphi_{u,0}) \tag{1.20}$$

where ϵ_u is the single-particle transverse emittance, and is equal to twice the betatron action J_u, represented as

$$\pi\epsilon_u = const = 2\pi J_u \tag{1.21}$$

Nevertheless, when dealing with a distribution of particles in the beam, the transverse emittance is defined as the average measure of the beam phase-space area (divided by π) that will contain a specified fraction of the beam particles, i.e.

$$\epsilon_{rms} = \sqrt{[\langle u^2\rangle\langle u'^2\rangle - \langle uu'\rangle^2]} \tag{1.22}$$

where ϵ_{rms} is the *rms transverse emittance* and $\langle ...\rangle$ is the distribution Average.

The statistical measure in Eq. (1.22) characterizes the spread of particles in both position and momentum spaces, effectively quantifying their dispersion. The phase-space area of the multiparticle beam is constant, but the particles are randomly distributed in the phase-space volume. If the beam distribution in the action space is Gaussian, which is the case at LHC, the beam size σ_u is linked to the beam emittance ϵ_u as

$$\sigma_u = \sqrt{\beta_u\epsilon_u} \tag{1.23}$$

here, ϵ_u is the full emittance equal to four times the rms emittance, i.e. $\epsilon_u = 4\epsilon_{rms}$ guarantees that the enclosed area phase-space volume is $\pi\epsilon_{rms}$. Hence, the quantity defined in Eq. (1.23) determines the beam's envelope radius.

The effective area of the Courant-Snyder ellipse ϵ_u is of fundamental interest within the realm of accelerator physics. Liouville's theorem states that as the ellipse rotates in the u−u' plane, the emittance of a particle beam subjected to conservative forces remains constant, thereby preserving the beam envelope area. However, a dichotomy arises when the particle beam is accelerated along the circumference of the circular path, resulting in a decrease in emittance with increasing particle momentum, referred to as *adiabatic damping*[4].

To gain a deeper insight into the shrinkage of the physical emittance ϵ_u, we consider the particle momentum $\mathbf{p}$(u,s) before and after the acceleration process, as shown in Fig.1.4.

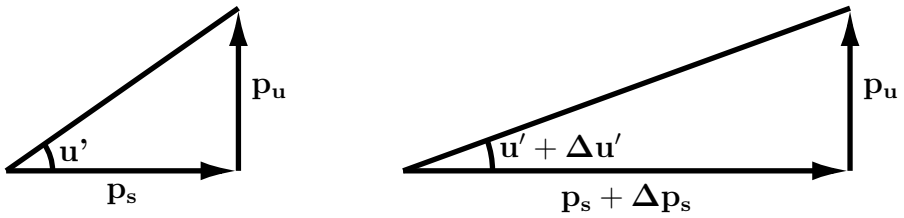


Figure 1.4: Evolution of momenta before (left) and after particle acceleration (right).

In a machine with linear optics, the particle's longitudinal momentum p_s increases to $p_s + \Delta p_s$ while the transverse component p_u remains constant, and the angle u' becomes smaller; the trajectory slope changes because the transverse velocity of a particle doesn't change during acceleration. Therefore, the length of the hypotenuse becomes $u' + \Delta u'$, and the change of trajectory slope corresponds to

$$\Delta u' = -u' \frac{\Delta p}{p_0} \tag{1.24}$$

where $\Delta p/p_0$ is the momentum deviation from the reference momentum. Any change in the slope will reduce the emittance as a function of the momentum. Therefore, to maintain consistency within this dynamic framework, the adiabatic damping becomes explicit by scaling the emittance into a normalized emittance in the following form

$$\epsilon_n = \beta_r \gamma_r \epsilon_u \tag{1.25}$$

Here, β_r and γ_r are the relativistic factors, ϵ_u is the physical emittance, while the normalized emittance ϵ_n independent on the beam energy remains conserved even under acceleration.

The normalized emittance remains constant throughout the entire circumference of the ring independently from the beam energy, adhering to Liouville's

[4] A misnomer there is no damping process involved.

theorem. This property presents a notable benefit for beam diagnostics, enabling the evaluation of beam emittance at collision energy and the estimation of its influence on accelerator luminosity, as detailed in the subsequent sections.

1.3.4 Off-momentum Particles

The discussion above assumes that all particles in the beam have the same momentum p_0. Nevertheless, in practice, this ideal scenario doesn't hold. A spread in particle momentum causes particles within a bunch to follow shifted trajectories through the magnetic elements. Considering a dipolar field and recalling the equation of rigidity[5](see Eq. (1.2)), one obtains that off-momentum particles follow different orbits around the accelerator lattice. The deviation from the design momentum p_0 with non-zero momentum offset Δp is quantified as

$$\delta = \frac{\Delta p}{p_0} \tag{1.26}$$

where δ is the spread in momenta. This deviation can be associated with chromatic effects that arise from the dipolar field, i.e. chromaticity is a measure of how the focusing properties change for particles that are off their design momentum.

The spread in momentum manifests as dispersion, a phenomenon in which particles with different momentum deviate from the design trajectory. This dispersion affects the dynamics of the particle and can introduce chromatic aberrations in the beam dynamics, particularly in the presence of a dipolar field. The inhomogeneous equation of motion, previously provided by Eq. (1.5), can be extended to account for these effects by considering the effects of dispersion and chromaticity on off-momentum particles, thus

$$u''(s) + k_u(s)u(s) = \frac{1}{\rho(s)}\frac{\Delta p}{p_0} \tag{1.27}$$

The solution of Eq. (1.27) can be written as the sum of its homogeneous and inhomogeneous solutions, namely

$$u(s) = u_H(s) + u_I(s) = u_H(s) + D_u(s)\frac{\Delta p}{p_0} \tag{1.28}$$

[5]Considering a constant effective length of the dipole $(\rho\Delta u') \equiv constant$, off-momentum particles rigidity: $B(\rho + \Delta\rho) = (p + \Delta p)/e \implies \frac{\Delta\rho}{\rho} = \frac{\Delta p}{p_0} \implies \Delta u' = -u'\frac{\Delta p}{p_0}$ (different deflection $\implies$ different orbit).

where $u_H(s)$ represents the homogeneous solution given by Eq. (1.14), while the inhomogeneous part $u_I(s)$ given in terms of dispersion function $D_u(s)$, and can be defined as

$$\Delta u(s) = u_I(s) = \delta D_u(s). \tag{1.29}$$

Thus, the dispersion equation satisfies

$$D_u''(s) + k_u(s) D_u(s) = \frac{1}{\rho(s)} \tag{1.30}$$

Using the matrix formalism in Eq. (1.10), the solution for the dispersion

$$D_u(s \mid s_0) = S(s \mid s_0) \int_{s_0}^{s} \frac{C(\bar{s})}{\rho(\bar{s})} d\bar{s} + C(s \mid s_0) \int_{s_0}^{s} \frac{S(\bar{s})}{\rho(\bar{s})} d\bar{s} \tag{1.31}$$

This solution indeed satisfies the differential equation of the dispersion. The general betatron solutions can be obtained from applying the 3×3 transfer matrices including dispersion propagation from s to s_0, yielding

$$\mathcal{M}_{3\times3} = \begin{pmatrix} C(s \mid s_0) & S(s \mid s_0) & D(s \mid s_0) \\ C'(s \mid s_0) & S'(s \mid s_0) & D'(s \mid s_0) \\ 0 & 0 & 1 \end{pmatrix} \tag{1.32}$$

Recalling Eq. (1.30), a particular solution for the inhomogeneous equation is given by setting $D_u(s) = constant = D_0$, takes the form

$$\begin{pmatrix} u \\ u' \\ \Delta p/p \end{pmatrix}_s = \mathcal{M}_{3\times3} \begin{pmatrix} u \\ u' \\ \Delta p/p \end{pmatrix}_{s_0} \quad \text{and} \quad \begin{pmatrix} D(s) \\ D'(s) \\ 1 \end{pmatrix} = \mathcal{M}_{3\times3} \begin{pmatrix} D_0 \\ D_0' \\ 1 \end{pmatrix} \tag{1.33}$$

In a periodic lattice, the dispersion function has to fulfill the periodic boundary conditions, this implies $D(s_0) = D(s_0 + L)$, where L presents the length of one period of the lattice.

1.4 Non-linear Beam Dynamics

The transverse dynamics discussed above considers only the linear terms, while the non-linear components are conveniently ignored. In the linear regime, the beam dynamics is described by the linear differential equations (1.4) and (1.5)). However, in high-energy physics, the particle beam lines or accelerators do not operate under such ideal and linear circumstances. Non-linear magnetic fields are

often incorporated within an accelerator lattice, either by design purpose or due to intrinsic lattice imperfections.

Explicitly, non-linear components can be integrated into the accelerator lattice to correct specific effects in lattice functions. As already mentioned, effects such as dispersion, determined by δD_u, occur within the beam transport system due to the deviations in the particle momentum from the ideal design momentum. The dispersion effect introduces perturbations in the focusing magnets, leading to a shift in the tune, a phenomenon known as chromaticity, i.e. a dominant source of the lattice non-linearity. For the sake of beam stability and to prevent beam loss due to tune shifts into resonances, sextupole magnets are needed for compensating the chromatic aberrations.

On the other hand, the inclusion of non-linear elements through the powering of non-linear magnets, or through magnetic errors, can have a deep impact on the beam dynamics. It may cause particles to diffuse to large amplitudes, or be pushed onto resonances. Indeed, non-linearities in the beam transport system contribute significantly to the Dynamic Aperture (DA), which manifests in the form of perturbations in the betatron function or through resonances. These effects can be detrimental to both the beam lifetime and luminosity production.

Non-linear effects stand as a primary factor limiting performance within the accelerator. Therefore, the linear formalism used so far has to account for non-linear phenomena. In the following section, the *Beam-Beam Interaction*, is introduced as a main source of non-linear effects within the lattice.

1.4.1 Beam-Beam Force

The beam-beam interaction is likely one of the main limitations to collider performance, leading to emittance blow-up, beam losses, and a reduction of the beam lifetime. These interactions generate electromagnetic potentials with a non-linear characteristic, resulting in significant perturbations in high-density beams, thus limiting the machine's luminosity production.

Since we are interested in the transverse fields, adopting the approximate approach presented in [25, 35], assuming the charge density within a bunch of particles follows a bi-Gaussian distribution in the transverse frame, for a round beam ($\sigma_x = \sigma_y = \sigma_r$), expressed as

$$\rho(r) = \frac{n}{\sigma_r\sqrt{2\pi}} e^{-\frac{-r^2}{2\sigma_r^2}}, \text{ where } r = (x, y) \tag{1.34}$$

The potential function arising from the system therefore is

$$\varphi(r) = \frac{ne}{\sqrt{4\epsilon_0}} \frac{1}{\sqrt{\pi}} \int_0^\infty dq \frac{e^{-\left(\frac{r^2}{q_r}\right)}}{\sqrt{q_r}}, \text{ with } q_r = 2\sigma_r^2 + q \tag{1.35}$$

here, n denotes the number of particles in the bunch, e represents the elementary charge, and ϵ_0 is the permittivity of free space.

The associated potential in Eq. (1.35) satisfies the equation of the Poisson equation [17], which links the potential to the charge-density distribution ρ

$$\nabla^2 \varphi(r) = \frac{-\rho(r)}{\epsilon_0} \tag{1.36}$$

and relates the potential to the transverse electric field of the bunch by the gradient of the scalar potential, accordingly

$$\mathbf{E}(r) = -\nabla \varphi(r) \tag{1.37}$$

In the bunch's rest frame ($v = 0$), the electric field lines radiate uniformly in all directions,

$$E_r(r) = -\frac{ne}{4\pi\epsilon_0} \frac{1}{r} \left[1 - e^{-\left(r^2/2\sigma_r^2\right)}\right] \text{ with } r^2 = x^2 + y^2 \tag{1.38}$$

For details of the calculation the reader is referred to Refs. [25, 35, 14]

To obtain the force exerted on particles in the counter-rotating beams, one must first compute the force that the electric fields generate. This involves boosting the electric field into a co-moving frame ($v \approx c$), using a Lorentz boost, the field lines confined almost in the transverse planes. The Lorentz force that acts on a particle with charge q as presented in Eq. (1.1) can be reformulated as

$$\vec{F} = q\left(E_r + \beta_r c B_\phi\right)\vec{r} \tag{1.39}$$

From Eq. (1.35) and with $r^2 = x^2 + y^2$, one can get the electric and magnetic fields in the collision, namely

$$E_r(r) = -\frac{\gamma nq}{4\pi\epsilon_0} \frac{1}{r} \left[1 - e^{-\left(r^2/2\sigma_r^2\right)}\right], \quad B_\phi(r) = -\frac{nq\beta_r c\mu_0}{4\pi} \left[1 - e^{-\left(r^2/2\sigma_r^2\right)}\right] \tag{1.40}$$

Using $\epsilon_0\mu_0 = c^{-2}$, the radial force, obtained in a closed form

$$F_{\mathrm{r}}(r) = -\frac{ne^2\left(1 + \beta_r^2\right)}{2\pi\epsilon_0} \frac{1}{r} \left[1 - e^{-r^2/2\sigma_r^2}\right] \tag{1.41}$$

the force in Eq. (1.41), depends only on distance r from bunch centre, hence, for a component in the action space $u = \{x, y\}$, we get

$$F_u(r) = -\frac{ne^2\,(1+\beta_r^2)}{2\pi\epsilon_0}\frac{u}{r^2}\left[1 - e^{-r^2/2\sigma_r^2}\right] \tag{1.42}$$

The forces described in Eq1.41 and Eq1.42 are computed under the assumption that the charges of the test particle and the opposing beam are of opposite charges. Indeed, the particles passing through or nearby the Gaussian bunch are deflected by the beam-beam force, i.e. the change of the slope of the particle trajectory imparted by the momentum kick. The kick can be respectively, focusing or defocusing if the particle have opposite or same sign sign with respect to the counter-rotating beam producing the force.

For small amplitude r (i.e. $r \ll \sigma_r$), the single particle experiences a linear beam-beam force and as the radial amplitude grows, larger than $1\sigma_r$ (i.e. $r \gg \sigma_r$), the force deviates sharply from its linear trend leading to a non-linear behavior in the lattice.

The kick profile as a function of the radial amplitude r is illustrated in Fig.1.5. The linear force component, relevant for short-range interactions as presented within the pink-highlighted segment maintains a linear pattern for $r \lesssim 1\sigma_r$, results in a steady *beam-beam tune shift*, acting like a quadrupole. In contrast, the nonlinear component, acting over large distances as showcased by the purple region, induces an *amplitude-dependent tune shift* (*detuning*), analogous with the non-linear dynamics seen with higher-order multipoles. The particle's radial distance from its counteracting beam center has an important impact on the beam dynamics. In short-range interactions when two bunches overlap (i.e. $\langle r \rangle = 0$), this is known as *head-on (HO) beam-beam force*. When the non-linear force acts from large distances, it's referred to as the *long-range (LR) beam-beam interaction*. These two regimes are elaborated upon in Chapter 2.

The Beam-Beam Parameter

Introducing the time-like variable, i.e. multiply the two-dimensional force by the longitudinal Gaussian distribution with a width of σ_s, which depends on both position s and time t.

$$F_r(r,s,t) = -\frac{ne^2\,(1+\beta_r^2)}{\sqrt{(2\pi)^3}\epsilon_0\sigma_s}\frac{1}{r}\left[1 - e^{\frac{r^2}{2\sigma_r^2}}\right]\left[e^{-\frac{(s+vt)^2}{2\sigma_s^2}}\right]. \tag{1.43}$$

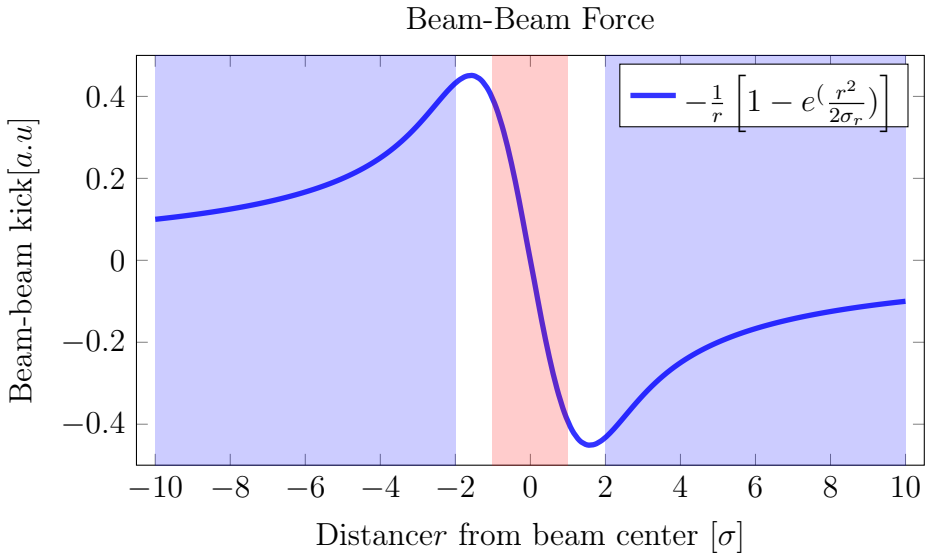


Figure 1.5: The beam–beam force (blue line) for round Gaussian beam in arbitrary units versus the radial component in units of rms beam size σ.

and using the Newton's law, the radial kick $\Delta r'$ that a particle receives from the opposing beam is computed from the integral of the force in Eq1.43 over the collision (i.e. time of passage) as

$$\Delta r' = \frac{\Delta r}{r} = \frac{1}{mc\beta_r\gamma_r}\int_{-\infty}^{+\infty} F_{\mathrm{r}}(r,s,t)dt \tag{1.44}$$

Hence, the radial kick is given by

$$\Delta r' = -\frac{2nr_0}{\gamma_r}\frac{1}{r}\left[1-e^{-\frac{r^2}{2\sigma_r^2}}\right] \tag{1.45}$$

where r_0 designates the classical particle radius, i.e. $r_0 = e^2/4\pi\epsilon_0 mc^2$, m is the mass of the particle. The expressions for the kick, in the transverse planes with $u = \{x, y\}$, given as

$$\Delta u' = -\frac{2nr_0}{\gamma_r}\frac{u}{r^2}\left[1-e^{-\frac{r^2}{2\sigma_r^2}}\right] \tag{1.46}$$

For small amplitudes r, one can derive the linear tune shift. The bunch behaves like a focusing lens for a particle traveling in the opposite direction and is equivalent to a lens with focal lengths $\frac{1}{f}$, in a linear approximation

$$\Delta u'|_{r\to 0} = -\frac{nr_0}{\gamma_r\sigma_r^2}u = -\frac{1}{f}u \tag{1.47}$$

where the slope of the force at zero amplitude (i.e. the asymptotic limit at $r = 0$), is the linear kick with a focal length $\frac{1}{f}$ given as proportionality factor which relates to the tune shift ΔQ as

$$\Delta Q = \frac{1}{4\pi}\frac{\beta^*}{f} = \frac{nr_0\beta^*}{4\pi\gamma_r\sigma_r^2} \tag{1.48}$$

where β^* is the amplitude function (β^*-function) at the interaction point.

The quantity in Eq. (1.48) represents the linear beam-beam tune shift, also known as the *linear beam–beam parameter* ξ.

The beam-beam parameter ξ quantifies the strength of the beam-beam force but does not reflect the nonlinear nature of the force.

Chapter 2

The CERN Large Hadron Collider Performance

In a circular collider such as the LHC, the centre-of-mass energy enables the discovery of heavy particles and limits the mass of the collision products, and their branching ratios. The LHC foremost aims to probe the unknown realms beyond the Standard Model, targeting a centre-of-mass collision energy of up to 14 TeV. However, achieving optimal performance is a complex endeavour with challenges such as minimising beam losses and maintaining beam quality.

To this end, it is essential to consider the number of effective collisions (events) to gauge the collider's performance. Hence, the ability of a particle physics experiment to generate events during the collision of two particle beams is quantified by the so-called luminosity, which determines the collision rate assigned to the available statistics for a given running time.

In this chapter, the LHC machine layout, performance, and operations challenges are addressed. In addition, the concept of luminosity as a figure of merit of a particle collider performance is introduced. We will delve deeper into this concept in the following chapters.

2.1 The CERN Accelerator Complex

CERN, the European Organization for Nuclear Research, is located on the Swiss-French border near Geneva and hosts the world's largest accelerator complex. This facility houses multiple particle accelerators, with several dedicated to supplying to the Large Hadron Collider (LHC) proton nad ion beams. Figure 2.1 shows a schematic of the accelerator complex. The LHC is a circular collider, designed to

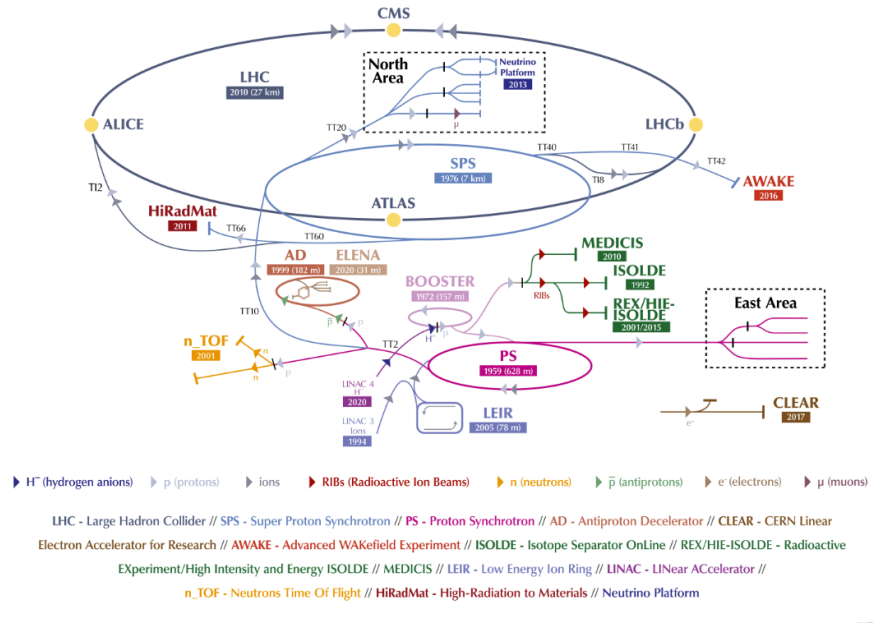


Figure 2.1: The CERN accelerator complex [28].

achieve a centre-of-mass energy of 14 TeV; so far up to 13.6 TeV has been achieved during Run 3. It operates mainly with proton beams but also functions as an ion collider. Therefore, to reach the nominal collision energies at LHC interaction points, proton beams are guided through a series of injectors that systematically boost their energy. The LHC proton injectors consist of a linear accelerator (LINAC4), the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS), the Super Proton Synchrotron (SPS), culminating in the LHC itself as illustrated in Fig. 2.1.

In this section, a brief overview of the different stages of proton beam acceleration at the LHC injector chain during Run 2 together with the machine layout is discussed.

2.1.1 LHC Injector Chain

In Run 2, the main physics program is to operate LHC with proton beams accelerated to 6.5 TeV. Protons are produced by the ionization of hydrogen gas in a plasma chamber, leaving only protons at 100 KeV to enter the linear accelerator (LINAC2) [32].

LINAC2 compresses the protons into packets, referred to as *bunches* and pre-accelerates them to 50 MeV of kinetic energy. LINAC2 then delivers the high-intensity proton bunches to the Proton Synchrotron Booster (PSB) to be accelerated to 2 GeV (kinetic energy) before transferring them to the Proton Synchrotron (PS) ring. The PS Booster and PS rings establish the bunch train structure compatible with the LHC RF system, accelerating the beams up to 26 GeV. Once the beam reaches the extraction energy, it is transferred to the SPS for a further increase in energy.

The Super Proton Synchrotron (SPS) requires up to 4 PS injections of 72 bunches at 25 ns spacing to accumulate up to 288 bunches in the ring, each injection corresponds to a batch (group of bunches). Once completed, the 26 GeV beams are then accelerated to the LHC injection energy of 450 GeV. Depending on the desired filling scheme, it takes up to 39 SPS batch injections per ring of 72 bunches each (a total of about 2500 bunches per ring) to fill the LHC with 25 ns separation bunches.

Afterward, the LHC itself provides the last energy boost, accelerating the beam in each ring up to nominal collision energies at LHC interaction points.

2.1.2 LHC Lattice Geometry

LHC is a 26.6 km twin-ring collider that produces luminosity to study the physics of the Standard Model, as well as, potential physics beyond the Standard Model. Structurally, it is segmented into 8 circular arcs with rectilinear insertion regions (IRs) in between, as shown in Fig. 2.2. The design of the LHC allows for the storage of two beams that counter-rotate in separate vacuum chambers (rings). It is only in proximity to the predetermined Interaction Points (IPs) that the beams merge into a common vacuum chamber, where they collide.

Four interaction regions have been chosen for housing the experiments, namely the Interaction Points, i.e. IP1, IP5, IP2, and IP8.

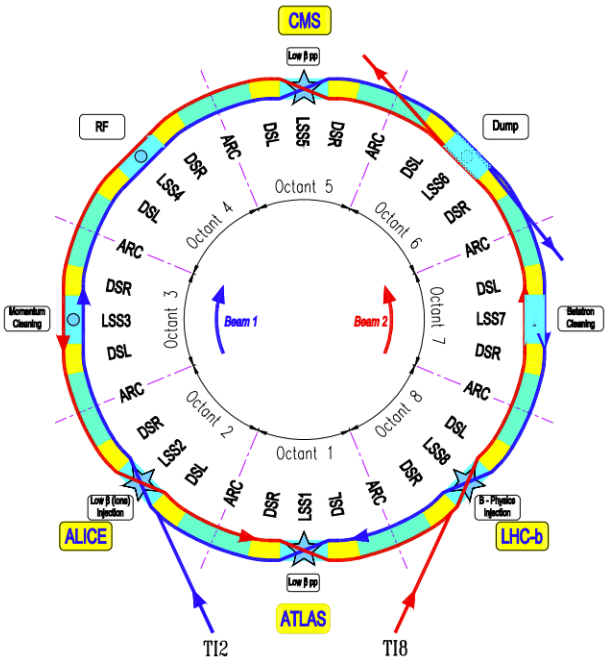


Figure 2.2: Schematic of the LHC rings [5].

Specifically, the two high-luminosity experiments ATLAS [11] and CMS [12] are located at IP1 and IP5, operating to reach a Peak luminosity of $L^{\text{PEAK}}_{ATLAS\&CMS} \approx 2{\cdot}10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. On the contrary, the ALICE [10] and LHCb [13] detectors, located at IP2 and IP8, respectively, operate at significantly lower luminosity levels than the former detectors, $L^{\text{PEAK}}_{ALICE} \approx 1.10^{31}\,\text{cm}^{-2}\,\text{s}^{-1}$ and $L^{\text{PEAK}}_{LHCb} \approx 2.10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$. The experiment's main operational parameters required to reach the target luminosity during Run 2 (e.g. transverse emittance, intensity, batch spacing,etc.) are specified in [13].

The LHC arcs are constructed from FODO cells, and each arc is made up of 23 cells 106.9 m long, with 6 dipoles located between the focusing quadrupoles of every cell as depicted in Fig. 2.3. Note that an upgrade project is ongoing to increase the luminosity at the LHC (HL-LHC) [1] to be implemented during the long shutdown 3 (LS3, 2026-28) . This is also combined with the LHC Injector Upgrade (LIU) [15] that occurred during the long shutdown 2 (LS2, 2019-21).

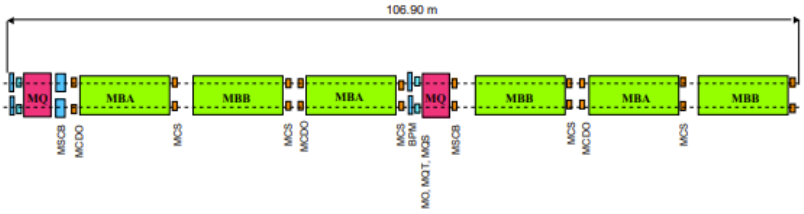


Figure 2.3: Schematic layout of the LHC arc cell including the corrector magnets [5]. MBA/B are the dipoles and MQ are the quadrupoles. The smaller magnets in between are dipolar (orbit), quadrupolar (tune), and higher-order correctors.

2.2 Luminosity of a Circular Collider

The key parameter describing the performance of a circular collider is its luminosity (integrated over time). The instantaneous luminosity $L(t)$, defined as the ratio between the number of events per second dR/dt and the corresponding interaction cross section σ_{int}.

$$\frac{dR}{dt} = L(t)\sigma_{int}, \tag{2.1}$$

usually measured in $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Moreover, its integrated value is related to the total number of events collected during a physics run [19], and

$$R = \sigma_{int} \int_0^T L(t)\,\mathrm{d}t \; = \sigma_{int}\,\mathcal{L} \tag{2.2}$$

where R is the total number of events during the run, T is the total time allocated for physics and $\mathcal{L}$ is the integrated luminosity delivered to the physics experiments, and its unit is femtobarns fb^{-1}.

A higher integrated luminosity indicates a larger dataset, which can improve the statistical accuracy of measurements and increase the potential for making new discoveries, especially when looking for rare events.

At LHC, the luminosity is produced by colliding bunches of particles and is closely related to the beam density distribution. In the case of two colliding bunches containing n_1 and n_2 particles, one can derive the expression of the instantaneous luminosity can be derived following the procedure in [14, 19, 26, 27, 30].

$$L \propto \kappa \int_{-\infty}^{+\infty} \rho_1\left(x, y, s, -s_0\right) \rho_2\left(x, y, s, s_0\right) dr^4, \tag{2.3}$$

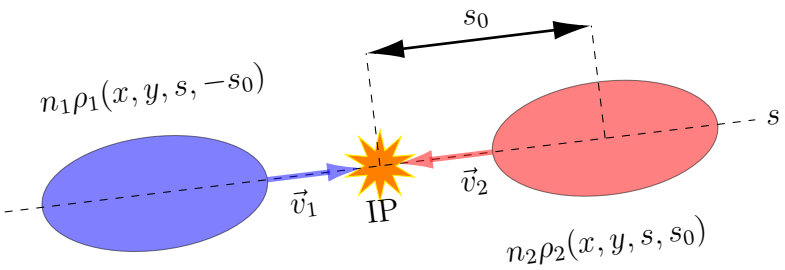


Figure 2.4: Sketch of a head-on collision, the two bunches moving against each other with increasing overlap.

where ρ_1 and ρ_2 represent the charge density distribution in the phase-space $\mathbf{r} = (x, y, s, s_0)$ and $s_0 = c \cdot t$ is the time-like component, while κ is the kinematic relativistic factor expressed as

$$\kappa = \sqrt{(\vec{v}_1 - \vec{v}_2)^2 - (\vec{v}_1 \times \vec{v}_2)^2 / c^2}. \tag{2.4}$$

Assuming bunches with uncorrelated Gaussian density distributions,

$$\rho(x, y, s, s_0) = \frac{1}{(\sqrt{2\pi})^3 \sigma_s \sigma_x^* \sigma_y^*} e^{-\frac{x^2}{\sigma_x^{*2}}} e^{-\frac{y^2}{\sigma_y^{*2}}} e^{-\frac{(s-s_0)^2}{2\sigma_s^2}}, \tag{2.5}$$

where σ_s denotes the rms bunch length; assuming equal bunches ($\sigma_{1s} \approx \sigma_{2s}$), while σ_x^* and σ_y^* characterize the horizontal and vertical transverse rms bunch dimension at the interaction point.

Considering only head-on colliding bunches travelling towards each other (as shown in Fig. 2.4), $\|\vec{v}_1\| = \|\vec{v}_2\| \approx c$, the combined integral of both distributions correlates directly to the luminosity, described as

$$L = \frac{2 n_1 n_2 f_{\text{rev}} k_b}{(\sqrt{2\pi})^6 \sigma_s^2 \sigma_x^{*2} \sigma_y^{*2}} \iiiint e^{-\frac{x^2}{\sigma_x^{*2}}} e^{-\frac{y^2}{\sigma_y^{*2}}} e^{-\frac{s^2}{\sigma_s^2}} e^{-\frac{s_0^2}{\sigma_s^2}} dx dy ds ds_0, \tag{2.6}$$

where the kinematic factor becomes $\kappa \approx 2$, since the crossing angle at collision point is null. the parameter f_{rev} is the collider revolution frequency, and k_b is the number of colliding bunches.

In the scenario under discussion, integrating Eq. (2.6) and using:

$$\int_{-\infty}^{+\infty} e^{-At^2} dt = \sqrt{\pi / A} \tag{2.7}$$

the fundamental expression of L is given by

$$L = \frac{f_{\mathrm{rev}}\, k_{\mathrm{b}}\, n_1\, n_2}{4\,\pi\sigma_x^*\sigma_y^*} \tag{2.8}$$

where f_{rev} is the collider revolution frequency and k_b is the number of colliding bunches. σ_x^* and σ_y^* characterise the horizontal and vertical transverse rms bunch dimension at the interaction point.

Equation (2.8) shows that the luminosity depends on the number of particles per bunch and the size of the beam. Due to the impossibility of measuring the beam size at the collision point [32], the beam width at the IP is inferred from the calculated beam emittance, which is provided by beam profile monitors.

Assuming round optics[1], with equal beam parameters for both beams $\sigma_x^* = \sigma_y^* = \sigma^* = \sqrt{\beta^*\epsilon^*/(\gamma_{\mathrm{r}}\beta_{\mathrm{r}})}$, the luminosity can be recast in terms of the rms normalised transverse emittances ϵ^* and the amplitude function β^* at the IP, namely

$$L = \frac{\gamma_{\mathrm{r}}\, f_{\mathrm{rev}}\, k_{\mathrm{b}}\, n_1\, n_2}{4\,\pi\epsilon^*\beta^*} \tag{2.9}$$

where γ_r and β_{r} are the relativistic factor at the IP, and β^* commonly referred to as the β-function.

For optimal performance, the LHC injector chain must generate beams with high brightness, i.e. beam of high intensity with minimal normalised emittance. Then, the LHC should generate very small β-function at the IPs [36].

2.3 Performance Limitation

The previous discussion centered on the ideal scenario of head-on collisions between bunches with uncorrelated Gaussian profiles in all planes, and the solution of Eq. (2.9) neglects any complexities that might impact the luminosity computation. In practise, a *crossing angle* is introduced to prevent unwanted collisions on either side of the interaction point, which would spoil the quality of the beam. However, the crossing angle results in a reduction in luminosity.

[1]This assumes that the β-function at the collision point is equal for the two beams and for both planes, i.e. $\beta_{1x}^* = \beta_{2x}^* = \beta_{1y}^* = \beta_{2y}^*$ [27].

2.3.1 Geometric Luminosity Reduction

Following the approach used in [30], a new coordinate system has been adopted to highlight the geometry of the crossing region where the beams collide (see Fig. 2.5). The two beams are related by the following transformations

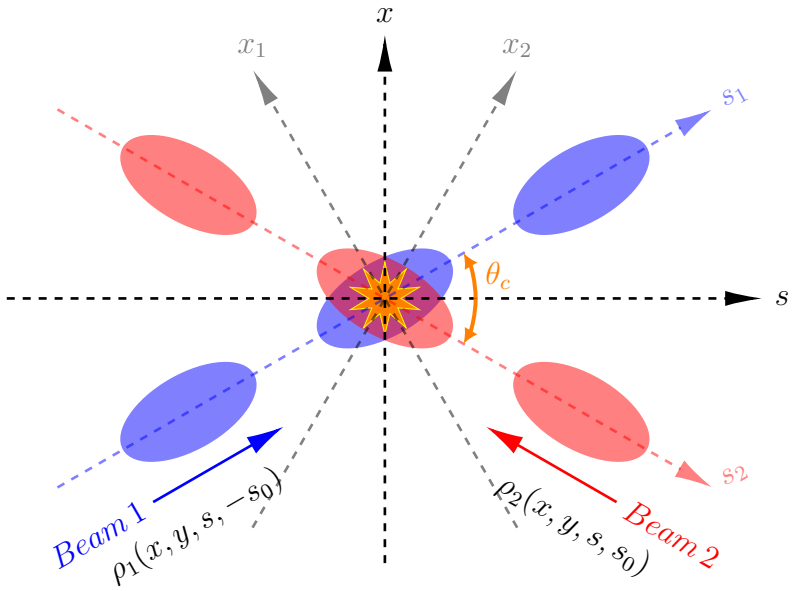


Figure 2.5: Sketch of the coordinate system of two bunches colliding with a crossing angle θ_c at the IP.

$$\begin{aligned} x_1 &= x\cos(\theta_c/2) - s\sin(\theta_c/2) \quad s_1 = s\cos(\theta_c/2) + x\sin(\theta_c/2) \\ x_2 &= x\cos(\theta_c/2) + s\sin(\theta_c/2) \quad s_2 = s\cos(\theta_c/2) - x\sin(\theta_c/2). \end{aligned} \tag{2.10}$$

with θ_c is the full crossing angle at IP.

The value of the overlap integral of Eq. (2.3), is reduced when beams are colliding with a crossing angle at the IP according to

$$\begin{aligned} L = 2\cos^2(\theta_c/2) n_1 n_2 f_{rev} \kappa_b \iiiint_{-\infty}^{+\infty} & \rho_{1x}(x_1)\,\rho_{1y}(y_1)\,\rho_{1s}(s_1 - s_0) \\ & \times \rho_{2x}(x_2)\,\rho_{2y}(y_2)\,\rho_{2s}(s_2 + s_0)\, d\hat{r}^4 \end{aligned} \tag{2.11}$$

where $2\cos^2(\theta_c/2) \equiv \kappa$, is the kinematic factor and and $\hat{\mathbf{r}} = (x_{1,2}, y_{1,2}, s_{1,2}, s_0)$.

Using the identity $\int_{-\infty}^{-\infty} e^{-\left(at^2+bt+c\right)} dt = \sqrt{\pi/a}\, e^{\frac{b^2-ac}{a}}$, the reduction in lumin-

osity is quantified by the *reduction factor*, namely

$$L = \frac{n_1 n_2 f_{\text{rev}} k_{\text{b}}}{4\pi\sigma_x^*\sigma_y^*}\mathcal{F}(\theta_c) \tag{2.12}$$

where $\mathcal{F}(\theta_c)$ is the luminosity geometric reduction factor due to the crossing angle at the IP. In particular, for small values of x and crossing angles $tan(\theta_c) \approx \theta_c$, after integrating in y and s_0, and using the approximation $\sigma_s \gg \sigma_{x,y}^*$, the reduction factor $\mathcal{F}(\theta_c)$ can be expressed as

$$\mathcal{F}(\theta_c) = \left(1 + \left(\frac{\sigma_s}{\sigma_x^*}\frac{\theta_c}{2}\right)^2\right)^{-\frac{1}{2}} \tag{2.13}$$

The calculation above takes into account only the effect of the crossing angle. However, in reality, there is an additional effect that arises from the transverse beam separation when the two beams are not colliding head-on.

Figure 2.6: Sketch of two beams that do not overlap in the transverse plane and collide with an offest δu at the IP.

As presented in Fig. 2.6, the two bunches collide with an offset δu in the transverse plane, with $u = \{x, y\}$. Collision offset can arise either parasitically, due to orbit imperfections, or intentionally for luminosity control or beam separation scans [18]. To address the reduction in luminosity owing to such transverse separation, a specific metric is used, known as the *separation factor*

$$\mathcal{F}(\sigma_u^*) = \exp\left(-\frac{\delta u^2}{4\sigma_u^{*2}}\right). \tag{2.14}$$

In the overall geometry, the luminosity reduction counts as including a correction factor $e^{\frac{\mathcal{B}^2}{\mathcal{A}}}$ when the crossing angle and collision offset occur simultaneously. Taking into account the geometric factors only in the horizontal plane, i.e. $\mathcal{F}(\sigma_u^*) = \mathcal{F}(\sigma_x^*) = \mathcal{F}(\sigma^*)$ and $\mathcal{F}(\theta_c)$ given as in Eq. (2.13), and introducing the variables $\mathcal{A}$ and $\mathcal{B}$ as [26, 14]

$$\mathcal{A} = \frac{\sin^2\frac{\theta_c}{2}}{\sigma^{*2}} + \frac{\cos^2\frac{\theta_c}{2}}{\sigma_s^2}, \quad \mathcal{B} = \frac{(\delta)\sin\left(\frac{\theta_c}{2}\right)}{2\sigma^{*2}} \tag{2.15}$$

the luminosity expression in Eq. (2.9) is recalculated by including the three reduction terms $\mathcal{F}(\sigma^*)$, $e^{\frac{\mathcal{B}^2}{\mathcal{A}}}$, $\mathcal{F}(\theta_c)$ [26], namely

$$L = \frac{\gamma_{\rm r}\, f_{\rm rev}\, k_{\rm b}\, n_1\, n_2}{4\,\pi\epsilon^*\beta^*}\ \cdot \mathcal{F}(\sigma^*) \cdot e^{\frac{\mathcal{B}^2}{\mathcal{A}}} \cdot \mathcal{F}(\theta_c). \tag{2.16}$$

In general, when discussing symmetric interaction region optics and removing any unwanted effects at the IP, it is often assumed that the betatron function $\beta(s)$ is constant throughout the length of the bunch and equal to its smallest value, β^* [27].

Nevertheless, the β-functions have their minima at the collision point and vary away from the IP. In a collider, when β^* is comparable to the rms bunch length or even smaller (i.e. $\beta^* \lesssim \sigma_s$), the variation of the β-function becomes significant and couples the transverse bunch size with the longitudinal position (i.e. $\sigma_u \sim \sigma_s$) [14]. The β-functions depend only on the longitudinal coordinate s as

$$\beta(s) = \beta^* \left[1 + \left(\frac{s}{\beta^*}\right)^2\right], \tag{2.17}$$

causing a parabolic behavior in the transverse bunch size σ_u, as a function of s

$$\sigma_u(s) = \sigma_u^* \sqrt{1 + \left(\frac{s}{\beta^*}\right)^2}. \tag{2.18}$$

Equation (2.18) implies that the beam size σ_u is not constant due to the growth of the betatron function away from the IP. In the literature this parabolic effect is known as the *hourglass effect*; beam size has the shape of an hourglass, as illustrated in Fig. 2.7, the hourglass effect due to the change in $\beta(s)$ along the bunch length is more important when $\beta^*(s)$ is very small.

In a limiting case without crossing angle and for a head-on symmetric, round Gaussian beams, after changing the integration variable into $\hat{S} = s/\sigma_s$, and using the ratio $h = \beta^*/\sigma_s$ [33], we obtain the *hourglass reduction factor luminosity* as

$$\mathcal{H} = \frac{L(\sigma_s)}{L(0)} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{e^{-\hat{S}^2}}{\left[1 + \left(\hat{S}/h\right)^2\right]} d\hat{S}, \tag{2.19}$$

where $\mathcal{H}$ is the *hourglass correction factor*

$$\mathcal{H} = \sqrt{\pi} \cdot h \cdot e^{h^2} \cdot erfc(h) \tag{2.20}$$

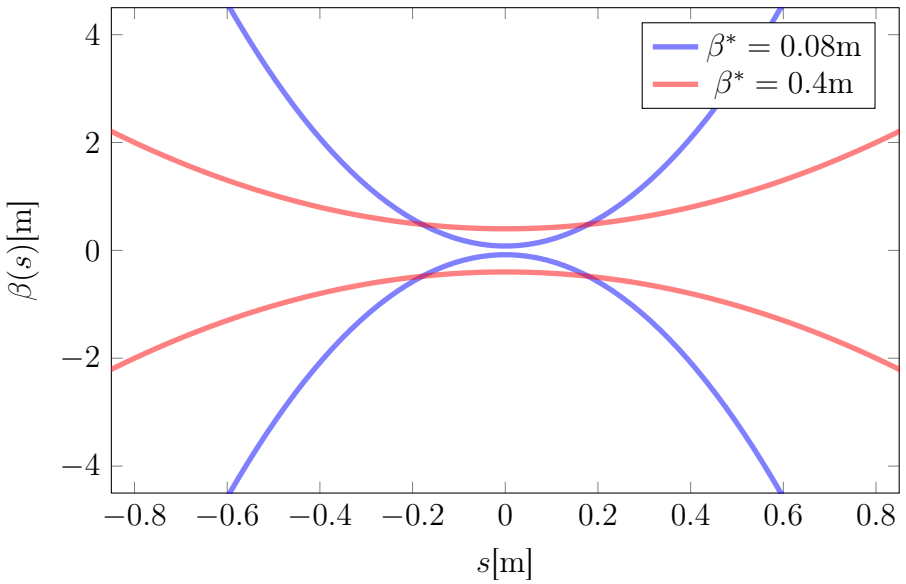


Figure 2.7: Schematic visualization reflecting the parabolic behavior caused by the hourglass effect for two different values of β^* at IP (the origin of the s-coordinate is such that $s = 0$ corresponds to the IP).

where $erfc(h)$ is the complementary error function [20].

At this point, given the various geometric factors that may limit the luminosity evolution, one can perform a global correction to the nominal luminosity value, including those factors. Putting all the geometric factors together, the nominal luminosity expression in Eq. (2.9) modified to get the design luminosity as

$$L = \frac{\gamma_r \, f_{rev} \, k_b \, n_1 \, n_2}{4 \, \pi \epsilon^* \beta^*} \cdot \mathcal{F}(\sigma^*) \cdot e^{\frac{\mathcal{B}^2}{\mathcal{A}}} \cdot \mathcal{F}(\theta_c) \cdot \mathcal{H}. \tag{2.21}$$

2.3.2 Beam-Beam Limit

In colliding-beam facilities, the electromagnetic forces generated by the counter-rotating beams that interact with each other are the most noteworthy factor limiting the accelerator's performance. As already detailed in Section 1.4.1 in Chapter 1, this force has two distinct regimes: a linear component for head-on (HOI) beam-beam interactions at very small distances, and a strongly non-linear behaviour when the Coulomb forces act at large distances, a feature known as long-range (LRI) beam-beam interactions.

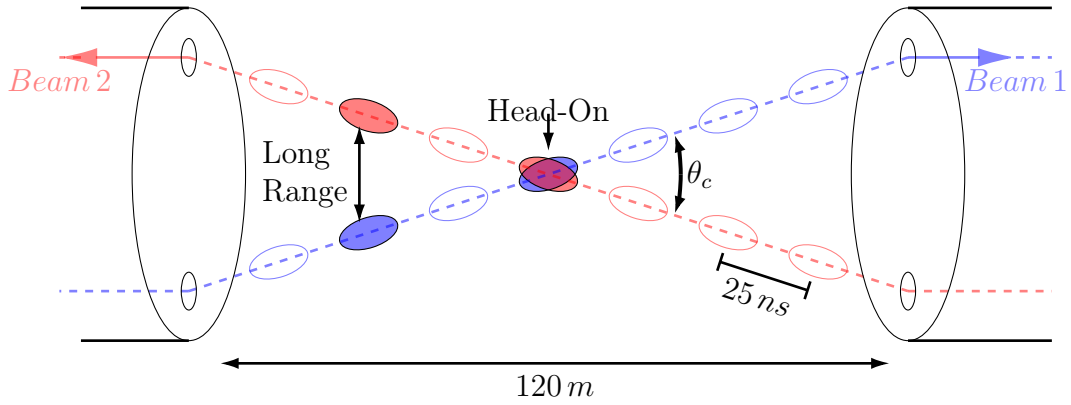


Figure 2.8: A schematic of head-on and long-range beam-beam interactions between two beams colliding at a crossing angle θ_c.

The core concept is shown in Fig. 2.8, the two counter-rotating beams share the same region around the IP for more than 120 m and spatially maintain a crossing angle. In particular, at LHC, the bunches are spaced 25 ns apart, and when two bunches collide in the centre or pass close to each other, several long-range encounters are made left and right of head-on collisions. As a consequence, the bunches continuously experience the non-linear electromagnetic fields from the opposite beam's bunches. This, in turn, limits the maximum particle density per bunch, and given the definition of luminosity and its direct relation to the number of collision points, the machine luminosity is inherently reduced.

Beam-Separation Dependence

In order to mitigate the undesirable effects of LRI [29], the bunches are separated and configured to collide at a finite crossing angle (quasi-head-on). Although widening the separation between beams weakens the LRI, this comes at the cost of lowering luminosity, since operating with a large crossing angle reduces the overlapping region between the two colliding bunches. Therefore, a minimum separation is essential to ensure that the bunches collide at a very small crossing angle. Conversely, a smaller beam-beam separation increases the force exerted during LRI, which in turn shrinks the stable area, causing particle instability and ultimately particle losses.

Traditionally, the strength of the LRI is quantified through the linear beam-beam parameter ξ (see Eq. (1.48)), and this parameter represents the largest tune shift attributable to the HOI in a bunch. As for the specific tune shift arising from

LRI, it is worth addressing the beam-beam force as outlined in Eq. (1.41)

$$F_{\mathrm{r}}(r) \propto -\frac{n}{\sigma} \cdot \frac{1}{r}\left[1 - e^{-r^2/2\sigma^2}\right]. \tag{2.22}$$

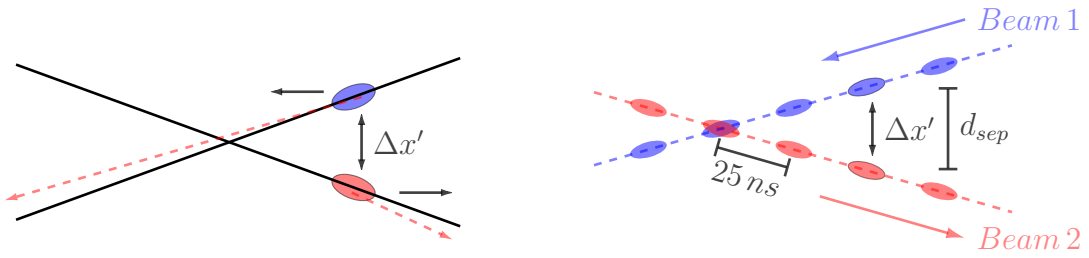
Given a beam-beam separation distance d_{sep} between the two beams in the horizontal plane and considering $r^2 = (x+d_{sep})^2+y^2$, the kick experienced in both transverse planes can be represented as

$$\Delta x' = -\frac{2nr_0}{\gamma_r}\frac{(x+d_{sep})}{r^2}\left[1 - e^{-\frac{r^2}{2\sigma_r^2}}\right], \quad \Delta y' = -\frac{2nr_0}{\gamma_r}\frac{(y)}{r^2}\left[1 - e^{-\frac{r^2}{2\sigma_r^2}}\right]. \tag{2.23}$$

For large enough beam-beam separation distance d_{sep}, the exponential term becomes negligible, i.e. $e^{-\frac{r^2}{2\sigma^2}} \ll 1$, making the tune shift from LRI to first order, inversely proportional to the square of the beam-beam separation distance d_{sep} as

$$\Delta Q_{LRI} = -\frac{nr_0\beta^*}{2\pi\gamma_r(d_{sep}^2)}. \tag{2.24}$$

In this framework as shown in Fig. 2.9b, the tune shift $\Delta Q_{LR} \propto -\frac{n}{d_{sep}^2}$ is calculated as a function of the beam separation d_{sep}.



(a) **(b)**

Figure 2.9: (a): Schematic view of the orbit change due to the beam-beam deflection, (b): Sketch of the geometry of beam-beam interactions at an IP.

This implicitly states that, for a given beam separation, all particles in the bunch experience the same beam-beam kick, equivalent to that affecting a zero-amplitude particle for that specific value of d_{sep}, i.e. its effect depends on beam-beam separation d_{sep}, and the beam-beam separation d_{sep} at first LRI is proportional to the crossing angle, written as

$$d_{sep} = \frac{\theta_c}{2}\sqrt{\frac{\gamma_r\beta^*}{\epsilon}}. \tag{2.25}$$

Thus, to ensure that LRI remains within tolerable limits by providing a minimum beam separation d_{sep}, a sufficiently large crossing angle is vital.

More importantly, the exponential component retains its significance for small separations, leading to an amplitude-dependent tune shift, i.e. detuning with amplitude or tune spread. Additionally, when the two beams collide with an offset and in the presence of non-linearities, both bunches undergo a dipole-like collective force or coherent beam-beam kick. This force causes the beam to deviate from its intended path, as shown in Fig. 2.9a, leading to orbit changes from its predefined closed orbit by introducing an amplitude-independent contribution. Mathematically, this is seen as a constant contribution in the kick formula in Eq. (2.23) when it is developed in series [25].

$$\Delta x' = \frac{const}{d} \cdot \left[1 - \frac{x}{d} + \mathcal{O}(\frac{x^2}{d^2}) + ...\right]. \tag{2.26}$$

This particular shift affects the dynamical stability of the beam differently than the HOI. All particles oscillate around the orbit at different tunes, deviating from the coherent tune and forming the tune footprint in tune space as presented in Fig. 2.10.

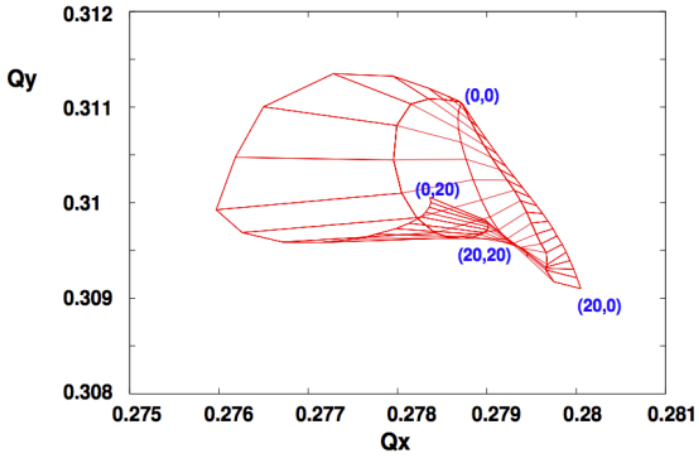


Figure 2.10: The tune footprint from long-range interactions only. Vertical separation and amplitudes between 0 and 20σ [35].

In particular, the particles within a bunch spread out in phase space and oscillate at different frequencies, forming *filaments* or *streaks*. Such tune spread is indicative of the bunch losing its coherence, acting like a collection of individual

particles, each following its own path and some may shift onto resonances. As a result, the stability of particles is not preserved anymore and can get lost from the beam and the bunch decoheres, the emittance increases, and the dynamic aperture region in which stable motion occurs is reduced. In such a scenario, the beam-beam effects set the limit on the machine's peak luminosity; thus, it is crucial to highlight the impact of the beam-separation dependence on luminosity production.

In high-energy colliders, the beam-beam effects often stand as a primary constraint on machine luminosity optimisation. This phenomenon necessitates a deep understanding of its impact on luminosity to ensure peak performance. To this end, it becomes crucial to revise the luminosity formula in Eq. (2.9), to explicitly incorporate these effects. Central to this adjustment is the beam-beam parameter ξ. This parameter emerges as a factor that offers insight into the luminosity in relation to the beam-beam effects. For equally populated, equally sized round beams, the luminosity in terms of the tune shift parameter ξ is

$$L = \frac{\gamma_r k_b f_{\text{rev}} n}{r_0 \beta^*} \xi \tag{2.27}$$

where ξ is provided in Eq. (1.48) and is measured as the largest tune shift owing to the HOI in a bunch.

Equation (2.27) shows that, for a fixed beam-beam parameter ξ, the luminosity is proportional to the number of particles per bunch n, the number of bunches, and inversely proportional to the low β-function insertion at the IP, β^*. Since LRI has been identified as the major factor limiting the dynamic aperture, which is strongly dependent on the crossing angle, β^* and beam population n, hence, lowering the beam-beam tune shifts provides the possibility of further reducing the beta functions at the IP to push the luminosity production. Nevertheless, it is essential to optimise the luminosity while ensuring that the beam-beam tune shifts remain within allowed limits. Therefore, the tolerable value of the tune shift is smaller than the typical maximum value for the HOI one.

Chapter 3

Luminosity Evolution Including Dynamic Aperture Effects

In the domain of superconducting colliders, magnet imperfections cause lattice nonlinearities, which reduce the luminosity lifetime. To quantify this, significant efforts have been devoted to developing models that shed light on the evolution of beam losses in circular colliders, based on the concept of time evolution of the dynamic aperture (DA). The dynamic aperture is a metric that gives the extent of the stable phase-space region occupied by particles in motion within the beam envelope. Referring to non-linear single-particle dynamics, models that merge the notion of dynamics aperture and that of luminosity evolution have been developed.

3.1 Beam Intensity Evolution

As detailed in the previous chapters, non-linear effects such as magnetic field errors or the beam-beam interactions, specifically the long-range beam-beam effects discussed in Section 2.3.2, lead to particle losses due to the shrinking of the region in phase space where stable motion occurs, which is the so-called DA, defined as the extent of the phase-space region where bounded motion occurs. Although analytical examination of the DA is still out of reach, a practical approach is to numerically compute the DA and establish a link between the DA and physical observables, such as beam lifetime [21].

It becomes evident when the dynamic aperture is within the phase-space

region occupied by the stable motion of the beam, the particles with higher amplitudes and lie outside the boundary of stability defined by the DA are likely to be lost after N revolution, causing a decrease in the beam intensity and an associated decay in luminosity. Interestingly, if we calculate the DA in the action space and refer to the definition of the Courant-Snyder invariant J in Section 1.3.2, a relation between the DA and beam intensity can be derived. The approach considers the collision of Gaussian bunches in the following form

$$\rho(q,p;\epsilon) = \frac{1}{2\pi\epsilon}\mathrm{e}^{-\left(\frac{q^2}{2\epsilon}+\frac{p^2}{2\epsilon}\right)} \tag{3.1}$$

and the beam distribution can be represented in the $x-y$ phase space as

$$f(x,p_x,y,p_y) = I_\mathrm{b}(1)\rho(x,p_x;\epsilon_x)\rho(y,p_y;\epsilon_y). \tag{3.2}$$

Here, (x,p_x,y,p_y) is the vector of the Courant-Snyder coordinates at a given section of the machine, and $I_\mathrm{b}(1)$ denotes the beam intensity at first turn $N=1$.

Given the relation $x^2+p_x^2=2J_x$ and $y^2+p_y^2=2J_y$ in which J_x and J_y follow the standard definition of the consistent phase-space area divided by 2π, f can be described in terms of angle-action variables (φ, J) as

$$f(J_x,J_y,\varphi_x,\varphi_y) = \frac{I_\mathrm{b}(1)}{4\pi^2\epsilon_x\epsilon_y}\mathrm{e}^{-\left(\frac{J_x}{\epsilon_x}+\frac{J_y}{\epsilon_y}\right)} \tag{3.3}$$

and

$$\hat{f}(J_x,J_y) = \int_0^{2\pi}\int_0^{2\pi} f(J_x,J_y,\varphi_x,\varphi_y)d\varphi_x d\varphi_y = \frac{I_\mathrm{b}(1)}{\epsilon_x\epsilon_y}\mathrm{e}^{-\left(\frac{J_x}{\epsilon_x}+\frac{J_y}{\epsilon_y}\right)}. \tag{3.4}$$

It is noteworthy to state that we are dealing with a distribution of particles in the beam, thus,

$$\langle J_x\rangle = \int_0^{\infty}\int_0^{\infty} J_x\hat{f}(J_x,J_y)dJ_xdJ_y = \epsilon_x \tag{3.5a}$$

$$\langle J_y\rangle = \int_0^{\infty}\int_0^{\infty} J_y\hat{f}(J_x,J_y)dJ_xdJ_y = \epsilon_y \tag{3.5b}$$

The average measure in Eqs.(3.5a) and (3.5b) define the beam envelope that will contain a specified fraction of the beam particles. Introducing the coordinates $r\in[0,\infty[$ and $\theta\in[0,\pi/2[$, we get

$$\sqrt{J_x} = \sqrt{\epsilon_x}r\cos\theta \quad \text{and} \quad \sqrt{J_y} = \sqrt{\epsilon_y}r\sin\theta \tag{3.6}$$

Since our focal point is to determine the fraction of particles contained within a specific volume given by $r \leq R$, thus, the surviving particles can be calculated by

$$I_{\mathrm{b}}(R) = \int_0^\infty \int_0^\infty \Theta_R(r,\theta)\hat{f}(J_x, J_y)dJ_x dJ_y, \quad \Theta_R(r,\theta) = \begin{cases} 1 & \text{for } r \leq R \\ 0 & \text{else.} \end{cases} \tag{3.7}$$

where r and θ are considered functions of J_x and J_y. From the definition we have

$$dJ_x dJ_y = 4\epsilon_x \epsilon_y r^3 \cos\theta \sin\theta dr d\theta\,, \tag{3.8}$$

and

$$-\frac{J_x}{\epsilon_x} - \frac{J_y}{\epsilon_y} = -r^2 \tag{3.9}$$

Inserting Eqs.(3.8), and (3.9) in Eq. (3.7) we deduce the expression for $I_{\mathrm{b}}(R)$ in a simplified way as

$$I_{\mathrm{b}}(R) = 4I_{\mathrm{b}}(1) \int_0^{\pi/2} \int_0^R \mathrm{e}^{-r^2} r^3 \cos\theta \sin\theta dr d\theta. \tag{3.10}$$

The integration over θ can directly be performed, and the integral over r can also be solved analytically, and we arrive at the following expression

$$\frac{I_{\mathrm{b}}(R)}{I_{\mathrm{b}}(1)} = 1 - (1 + R^2)\mathrm{e}^{-R^2}, \tag{3.11}$$

which represents the relative intensity contained in the region $r \leq R$. Furthermore, $R = D(N)$, where $D(N)$ is the DA value that varies with the number of turns N. As a result, the direct link between the beam intensity and dynamic aperture at some turn number N is given by

$$\frac{I_{\mathrm{b}}(N)}{I_{\mathrm{b}}(1)} = 1 - \left[1 + D^2(N)\right] \mathrm{e}^{-D^2(N)}. \tag{3.12}$$

3.1.1 Dynamic Aperture Scaling-Law

Since the late 1990s, considerable efforts have been made to study and model the evolution of the dynamic aperture as a function of the number of turns [2, 21, 34]. These studies highlighted the profound dependence between dynamic aperture, beam stability, and the resulting luminosity. Given the fact that DA is a key concept of nonlinear beam dynamics in the circular collider, the aim has been to reliably predict the DA value, as a central measure for predicting and evaluating

the overall machine performance. Therefore, it is worth mentioning that in [2] some DA models were proposed. From this array of models, the *Model 2* is particularly highlighted for its applicability in our study as an effective scaling law tied to the luminosity evolution, and it reads

$$\textbf{Model 2} \qquad \Rightarrow \qquad D(N) = \rho_* \left(\frac{\kappa}{2\mathrm{e}}\right)^{\kappa} \frac{1}{\ln^{\kappa} \frac{N}{N_0}}\,, \tag{3.13}$$

where $D(N)$ is the dynamic aperture as a function of turn number N, while ρ_*, κ and N_0 are free parameters. For simplicity, $N_0 = 1$, therefore, only two parameters are considered.

Equation (3.13) shows an inverse logarithmic decay of the dynamic aperture with the turn number N. An initial strong reduction of DA settles down when N becomes very large, the logarithmic term dominates, and the function stabilises.

3.2 DA Model for Luminosity Evolution

The effort of our study gravitates toward connecting the time evolution of luminosity with the DA model after a finite number of turns. This effort is strongly guided by the methodologies developed in [22, 23, 24], emphasising the groundbreaking connection between DA and the intensity evolution in the presence of non-linear effects, as showcased in Eq. (3.12) and by extending the DA scaling law given in Eq. (3.13), we have addressed the time evolution of beam intensity within a circular collider, thus developing a new luminosity model based on the concept of dynamic aperture.

In this work, we present the luminosity evolution as a function of the number of turns, including both effects from burn-off losses and any pseudo-diffusive effects arising from DA evolution. It is worth noting that for a direct application of the DA model, the time frame t is re-scaled into a new time variable τ, expressed as

$$\tau - 1 = f_{rev}\, t \qquad \text{giving} \qquad \frac{\mathrm{d}}{\mathrm{d}t} = f_{rev}\, \frac{\mathrm{d}}{\mathrm{d}\tau} \tag{3.14}$$

Here, τ stands as a dimensionless variable, symbolizing the number of turns, where a deviation of the origin of τ with respect to t has been introduced. In the following, the derivative with respect to τ is indicated by $\dot{}$, while $'$ indicates the derivative with respect to t.

3.2.1 Luminosity Evolution from Burn-Off Losses

In a real-world collider like the LHC, delivered luminosity is the most important figure of merit for its performance, its expression without taking into account the hourglass effect [26], formulated as

$$L = \frac{\gamma_{\mathrm{r}}\, f_{\mathrm{rev}}\, k_{\mathrm{b}}\, n_1\, n_2}{4\,\pi\epsilon^*\beta^*}\, F(\theta_c, \sigma_z, \sigma^*), \tag{3.15}$$

where the parameters in Eq. (3.15) are already introduced in Chapter 2, with γ_r is the relativistic γ-factor, f_{rev} the revolution frequency, and k_b indicates the number of colliding bunches. The variable n_i gives the particle count per bunch in each colliding beam. Furthermore, ϵ^* denotes the RMS normalized transverse emittance, while β^* represents the value of beta-function at IP. The total beam population is defined as $N_j = k_b\, n_j$ and the fact that not all bunches collide at the high-luminosity experimental points is taken into account by introducing a scale factor.

The factor F accounts for the reduction in volume overlap between the colliding bunches due to the presence of a crossing angle. F is a function of the half-crossing angle $\theta_c/2$, and the RMS dimensions σ^*andσ_z, as outlined in Section 2.3.1, thus it is expressed as

$$F(\theta_c, \sigma_z, \sigma^*) = \frac{1}{\sqrt{1 + \left(\dfrac{\theta_c}{2}\,\dfrac{\sigma_z}{\sigma^*}\right)^2}}\,. \tag{3.16}$$

Note that $\sigma^* = \sqrt{\beta^*\,\epsilon^*/(\beta_r\,\gamma_r)}$, where β_r is the relativistic factor β. Equation (3.15) is valid in the case of round beams ($\epsilon_x^* = \epsilon_y^* = \epsilon^*$) and round optics ($\beta_x^* = \beta_y^* = \beta^*$). For our scope, Eq. (3.15) will be recast in the following form:

$$L = \Xi\, N_1\, N_2, \qquad \Xi = \frac{\gamma_r f_{rev}}{4\,\pi\epsilon^*\beta^*\,k_b} F(\theta_c, \sigma_z, \sigma^*) \tag{3.17}$$

in which the dependence on the total intensity of the colliding beams is highlighted and the other quantities are included in the term Ξ.

In the context of an ideal collider, excluding any levelling gymnastics or dynamic-beta effects, only the emittances and the bunch intensities can change over time, luminosity burn-off stands out as the main mechanism that drives beam losses during collisions. Under these conditions which are well fulfilled for the parameters of the (HL)-LHC luminosity L at a collision point, Eq. (3.15) is more

correctly interpreted as the peak luminosity at the beginning of the fill. Furthermore, given the definition of the instantaneous luminosity as the number of events per second normalised to the interaction cross section, its evolution over time is closely related to its peak value and the beam intensity, which turns out to be derived from the following equation

$$N'(t) = -\sigma_{int}\, n_c\, L(t) = -\sigma_{int}\, n_c\, \Xi\, N^2(t) \tag{3.18}$$

where σ_{int} represents the inelastic $p-p$ cross section for the interaction of charged particles. Assume that there are two colliding beams with equal intensities, where n_c represents the number of collision points. The solution of Eq. (3.18) provides the time-dependent bunch population, namely

$$N(t) = \frac{N_i}{1 + \sigma_{int}\, n_c\, \Xi\, N_i\, t}, \tag{3.19}$$

Here, N_i stands for the initial intensity. From Eq. (3.19), we deduce that the luminosity gradually decreases as

$$L(t) = \frac{\Xi\, N_i^2}{\left(1 + \sigma_{int}\, n_c\, \Xi\, N_i\, t\right)^2}\,. \tag{3.20}$$

In the most general case, where the beams can have different intensities, the intensity evolution is described by the following equations

$$\begin{cases} N_1'(t) &= -\sigma_{int}\, n_c\, \Xi\, N_1(t)\, N_2(t) \\ N_2'(t) &= -\sigma_{int}\, n_c\, \Xi\, N_1(t)\, N_2(t) \end{cases} \tag{3.21}$$

The solution of Eq. (3.21), indicated as $N_{1,2}^{bo}(\tau)$ to highlight that it only includes the burn off contribution, can be obtained by re-writing:

$$\begin{cases} \dot{N}_1^{bo}(\tau) + \dot{N}_2^{bo}(\tau) &= -2\,\varepsilon\, N_1^{bo}(\tau)\, N_2^{bo}(\tau) \\ \dot{N}_1^{bo}(\tau) - \dot{N}_2^{bo}(\tau) &= 0 \end{cases} \tag{3.22}$$

with

$$\varepsilon = \frac{\sigma_{int}\, n_c\, \Xi}{f_{rev}} \tag{3.23}$$

and from which one finds

$$\begin{cases} N_1^{bo}(\tau) &= N_2^{bo}(\tau) + \xi \\ \dot{N}_2^{bo}(\tau) &= -\,\varepsilon\, N_2^{bo}(\tau)\, \left[N_2^{bo}(\tau) + \xi\right] \end{cases} \tag{3.24}$$

Equation (3.24) has two solutions depending on the value of ξ. If $\xi = 0$ then

$$\begin{cases} N_1^{bo}(\tau) &= \dfrac{N_i}{1 + \varepsilon\, N_i\,(\tau - 1)} \\ N_2^{bo}(\tau) &= N_1^{bo}(\tau), \end{cases} \tag{3.25}$$

where $N_i = N_{i,1} = N_{i,2}$ stands for the initial beam intensity.

Otherwise, if $\xi \neq 0$ then

$$\begin{cases} N_1^{bo}(\tau) &= \xi\, \dfrac{1}{1 - N_r\, \mathrm{e}^{-\varepsilon\, \xi\, (\tau - 1)}} \\ \\ N_2^{bo}(\tau) &= \xi\, \dfrac{N_r\, \mathrm{e}^{-\varepsilon\, \xi\, (\tau - 1)}}{1 - N_r\, \mathrm{e}^{-\varepsilon\, \xi\, (\tau - 1)}}\,, \end{cases} \tag{3.26}$$

where $\xi = N_{i,1} - N_{i,2}$ and $N_r = \frac{N_{i,2}}{N_{i,1}}$. It is worth noting that the outcomes are invariant to the sign of ξ, although, from a computational point of view a negative exponent is preferred, as τ can grow to very large values. Note also that Eq. (3.25) can be easily recovered from Eq. (3.26) by expanding the exponential and setting $N_{i,1} = N_{i,2}$.

For colliding beams with equal intensities, and considering the beam losses are exclusively due to the burn-off effect, the luminosity can be modeled as

$$L_{bo}(\tau) = \frac{L_i}{\left[1 + \varepsilon\, N_i\,(\tau - 1)\right]^2} \tag{3.27}$$

where L_i, is the peak luminosity at the start of the filling scheme. The observation in Eq. (3.27) demonstrates that luminosity gradually decreases with the number of turns τ and the intensity of the beam N_i.

The drawback of the luminosity burn-off model lies in the fact that a wide range of additional processes contribute to beam degradation. Effects such as intra-beam scattering (IBS) and synchrotron radiation are just the tip of the iceberg. Notably, the beam-beam interactions and the machine non-linearities add another layer of variability to this scenario. However, even this list is not all-inclusive. The interplay of these factors impacts the intensity of the beam at different stages during the beam cycle within the collider.

3.2.2 Luminosity Evolution Including Pseudo-Diffusive Effects

Achieving a definitive formulation for luminosity is not trivial, as its performance is limited by a series of effects that come into play in its degradation over time. In fact, particles within the dynamic aperture may still be subjected to losses because of diffusive mechanisms. The diffusive processes cause the particle amplitudes to increase arbitrarily, which in turn can lead to particles hitting the physical aperture of the machine and eventually getting lost. As a result, these diffusive effects can further degrade the beams and affect the luminosity model grounded only on the burn-off processes. Hence, the model suggested in Eq. (3.27) is not valid and a more refined model based on the concept of DA is proposed that is not entirely phenomenological.

In a diffusive framework, the luminosity can be described mathematically as follows

$$L(\tau) = L_{bo}(\tau) \cdot [1 + \zeta(\tau) \cdot f(\mathrm{DA}(\tau))] \tag{3.28}$$

where $L(\tau)$ represents the luminosity evolution with the number of turns τ, L_{bo} denotes the burn-off component of luminosity. The term ζ is a coefficient capturing the pseudo-diffusive effects on the luminosity, and $f(\mathrm{DA}(\tau))$ is a function that maps the evolution of the DA with the number of turns, while the product $\zeta(\tau) \cdot f(\mathrm{DA}(\tau))$, provides a holistic view of how DA impacts luminosity over time.

This approach assumes that besides the inevitable beam losses due to the particle burn-off and emittance change during the collision at IPs, it is plausible to model all possible pseudo-diffusive effects within the dynamical system of the beam. To this end, a more comprehensive model of luminosity is developed at the end of this section (see Eq. (3.33)).

At the LHC, the proton burn-off predominantly targets the core of the beam distribution, an area densely populated with particles. Meanwhile, diffusive processes exert their influence primarily on the tails of the beam distribution. Essentially, proton burn-off and diffusive actions take place in distinct regions of the beam distributions and at different intervals during the beam cycle, ensuring that they act independently without significant overlap.

Then, under these assumptions, it is justified to describe the intensity evolution as

$$\begin{cases} N_1'(t) &= -\sigma_{int}\, n_c\, \Xi\, N_1(t)\, N_2(t) - \hat{\mathcal{D}}_1(t) \\ N_2'(t) &= -\sigma_{int}\, n_c\, \Xi\, N_1(t)\, N_2(t) - \hat{\mathcal{D}}_2(t)\,. \end{cases} \tag{3.29}$$

The terms $\hat{\mathcal{D}}_i$ represent the intensity-independent pseudo-diffusive effects. The

reason for the origin change of τ with respect to t is now clear: it is needed so that $t = 0$ corresponds to $\tau = 1$ and then $D(1) = +\infty$ according to Eq. (3.13).

Equation (3.21) becomes

$$\begin{cases} \dot{N}_1(\tau) &= -\varepsilon\, N_1(\tau)\, N_2(\tau) - \mathcal{D}_1(\tau) \\ \dot{N}_2(\tau) &= -\varepsilon\, N_1(\tau)\, N_2(\tau) - \mathcal{D}_2(\tau)\,, \end{cases} \tag{3.30}$$

with $\mathcal{D}_i = \hat{\mathcal{D}}_i / f_{rev}$.

The explicit expression for $\mathcal{D}_i(\tau)$ can be found by noting that these functions are solutions of

$$\begin{cases} \dot{N}_1(\tau) &= -\mathcal{D}_1(\tau) \\ \dot{N}_2(\tau) &= -\mathcal{D}_2(\tau) \end{cases} \tag{3.31}$$

and that the explicit solution has been assumed to be of the form (3.12). Therefore, one obtains

$$\mathcal{D}_j(\tau) = -2N_{i,j}\, D_j(\tau)\, \dot{D}_j(\tau)\, \mathrm{e}^{-D_j^2(\tau)} \left(2 + D_j^2(\tau)\right) \qquad j = 1, 2\,. \tag{3.32}$$

The detail of the derivation can be found in [22], but under the assumptions that the initial beam intensities are the same as well as the terms $\mathcal{D}_j$, then an explicit expression at the lowest order in ε (see Eq.(3.23)) can be given for luminosity, namely

$$\frac{L(\tau)}{L_i} = \frac{1}{\left[1 + \varepsilon N_i(\tau - 1)\right]^2} - \left[1 + D^2(\tau)\right] \mathrm{e}^{-D^2(\tau)} \left\{ 2 - \left[1 + D^2(\tau)\right] \mathrm{e}^{-\frac{D^2(\tau)}{2}} \right\} \tag{3.33}$$

where L_i is the initial luminosity value, given by $L_i = \Xi\, N_i^2$. In the forthcoming analysis and chapters, we will refer to this mathematical model in Eq. (3.33) as DA Model−2.

The experimental verification and validation of the DA Model−2 will be discussed in the next chapter.4, using the experimental data from LHC Run−2 as a benchmark to assess the effectiveness of our model.

Chapter 4

Dynamic Aperture Model Qualification

The proposed scaling law for the time evolution of the luminosity in Chapter 3 has been verified numerically by applying the model in Eq. (3.33) to the LHC Run 2 data collected from the ATLAS detector [7]. In this chapter, we first provide a brief overview of the extracted luminosity data. Afterward, we stress out the results obtained by implementing the luminosity model fitting procedure.

4.1 LHC Run 2 Data

The luminosity data used in our study were extracted from the ATLAS detector during the LHC Run 2 physics run, i.e. 2016, 2017, and 2018. The data considered exclusively targets proton-proton collision events, and each recorded physics data during a specific time frame run block is catalogued in a file as a luminosity dataset [9], termed as physics fill. For a detailed insight into the nature of these data files, one can delve into the comprehensive descriptions provided in [6, 19].

As already mentioned in the previous chapter, the peak luminosity is recorded at the start of the fill, and given that the time evolution of luminosity heavily depends on its initial value and the decay of the beam intensity, this results in a gradual decrease of the luminosity over the fill duration.

In particular, fills that start with a higher peak luminosity tend to decay

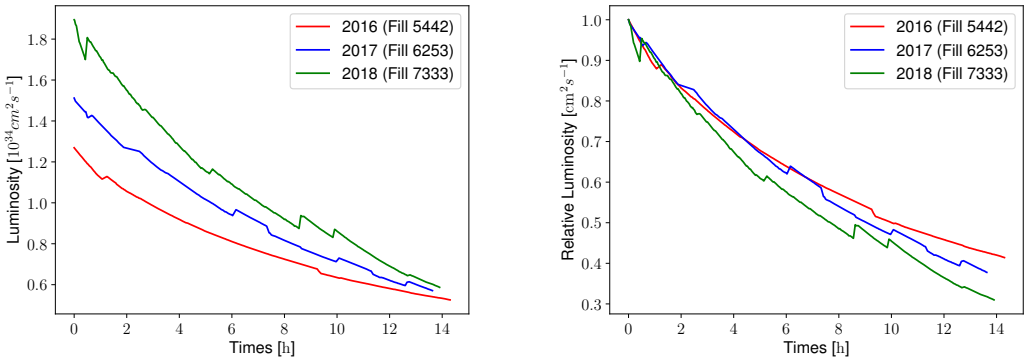


Figure 4.1: Measured luminosity evolution at the ATLAS experiment over typical fills of 2016, 2017, and 2018 LHC operation. (a) Absolute luminosity (b) Relative luminosity.

faster as the fill progresses, as illustrated in Fig.4.1 for an example of LHC fills in 2016, 2017, and 2018.

4.1.1 Selection Criteria

For the purposes of our study, which focuses strictly on physics fills, we introduce some selection criteria that have been taken into account to evaluate the entire dataset collected during LHC Run 2.

Following the selection procedure in [19], we additionally further our analysis by defining cut-off values on the initial luminosity and fill length; thus, only fills with values beyond this limit are considered. Fill selection based on peak luminosity and fill length is illustrated, respectively, in Figs. 4.2a and 4.2b, while the values below the cutoff (dashed lines) are removed.

Furthermore, given the detailed annual report of the full Run 2 dataset in [8], we extend our investigation by checking all physics fills and removing the fills that are preceded by any interruptions due to special events, such as technical breakdowns.

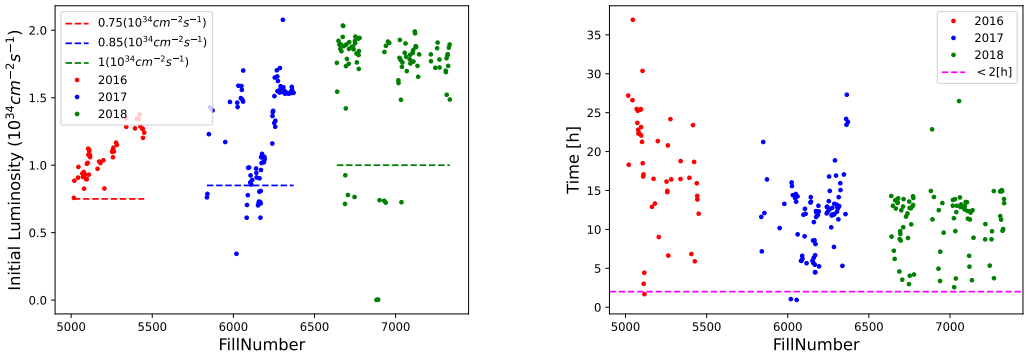


Figure 4.2: Fills selection during LHC Run 2. The selection is based on: (a) Initial luminosity, (b) Fill length.

4.1.2 DA Model 2 Fitting Results

The DA Model 2, as described in Eq. (3.33), has been successfully validated through experimental data. Its direct application to the luminosity measurements from the LHC Run 2 is clearly seen to be in good agreement, as shown in Fig. 4.3 for a typical LHC fill, respectively, for 2016, 2017 and 2018.

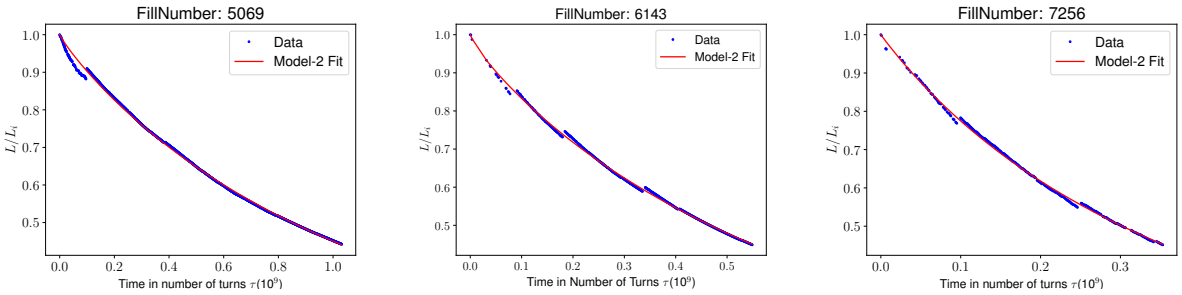


Figure 4.3: The DA Model−2 fit to normalised luminosity L/L_i as a function of the number of turns for typical LHC fills: 5069, 6143 and 7256, respectively, for years 2016, 2017 and 2018.

All other selected fills have been studied and are more or less similar in fit quality. Therefore, for a global evaluation, in Fig. 4.4 we present, the peak luminosity value (Fig. 4.4a), the fit quality R^2_{adj} (Fig. 4.4e), and χ^2 in logarithmic scale (Fig. 4.4f). The values of the fit parameters: κ (Fig. 4.4c), $\rho*$ (Fig. 4.4d), and $\varepsilon\, N_i$

(Fig. 4.4b) are also shown.

The statistical measure used in this study to assess the effectiveness of our model is the so-called adjusted coefficient of determination, R^2_{adj} , mathematically expressed as

$$R^2_{\text{adj}} = 1 - \frac{n-1}{n-p-1}\frac{\Sigma^2}{\sigma^2} \tag{4.1}$$

where n stands for the total number of observations, and p represents the count of adjustable parameters within the fit, while Σ^2 and σ^2 are, respectively, the squared sum of residuals and the total sum of squares which is proportional to the data variance, namely

$$\Sigma^2 = \sum_{i=1}^{n} (y_i - f_i)^2 \tag{4.2}$$

$$\sigma^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{4.3}$$

with $\bar{y}$ is the average overall measured data y_i, while f_i is the fit model.

If the value of R^2_{adj} is significantly less than one, or even negative, this means that the model is of poor quality and that the mean of the data gives a better fit to the given model. A desirable fit is considered optimal if the R^2_{adj} closely approaches one, i.e. $R^2_{\text{adj}} \to 1$. In our analysis, R^2_{adj} serves as an essential metric for evaluating the goodness of our fitting model to luminosity evolution at LHC.

Therefore, one can see in Fig. 4.4e, all fits are of good quality with $R^2_{\text{adj}} > 98\%$. Another important observation is the growth of the fit parameter εN_{i} on a yearly basis, reflecting the increasing trend of the peak luminosity.

As previously reported in Chapter.4, the εN_{i} parameter is fundamentally related to the core process burn-off phenomena, and the initial beam intensities N_{i}. Indeed, it is worth noting that two major aspects arise from this:

- The strong correlation between peak luminosity and initial beam intensities is mirrored in the mutual variation of εN_{i} with peak luminosity.
- The εN_{i} parameter is not static and tends to change as the beam intensity decays. Its time dependence therefore requires an in-depth evaluation to determine the most suitable approach for data analysis.

In line with the study presented in Ref.[23], the time dependence of εN_{i} can be safely neglected in our model. Such a decision stems from a thorough comparison

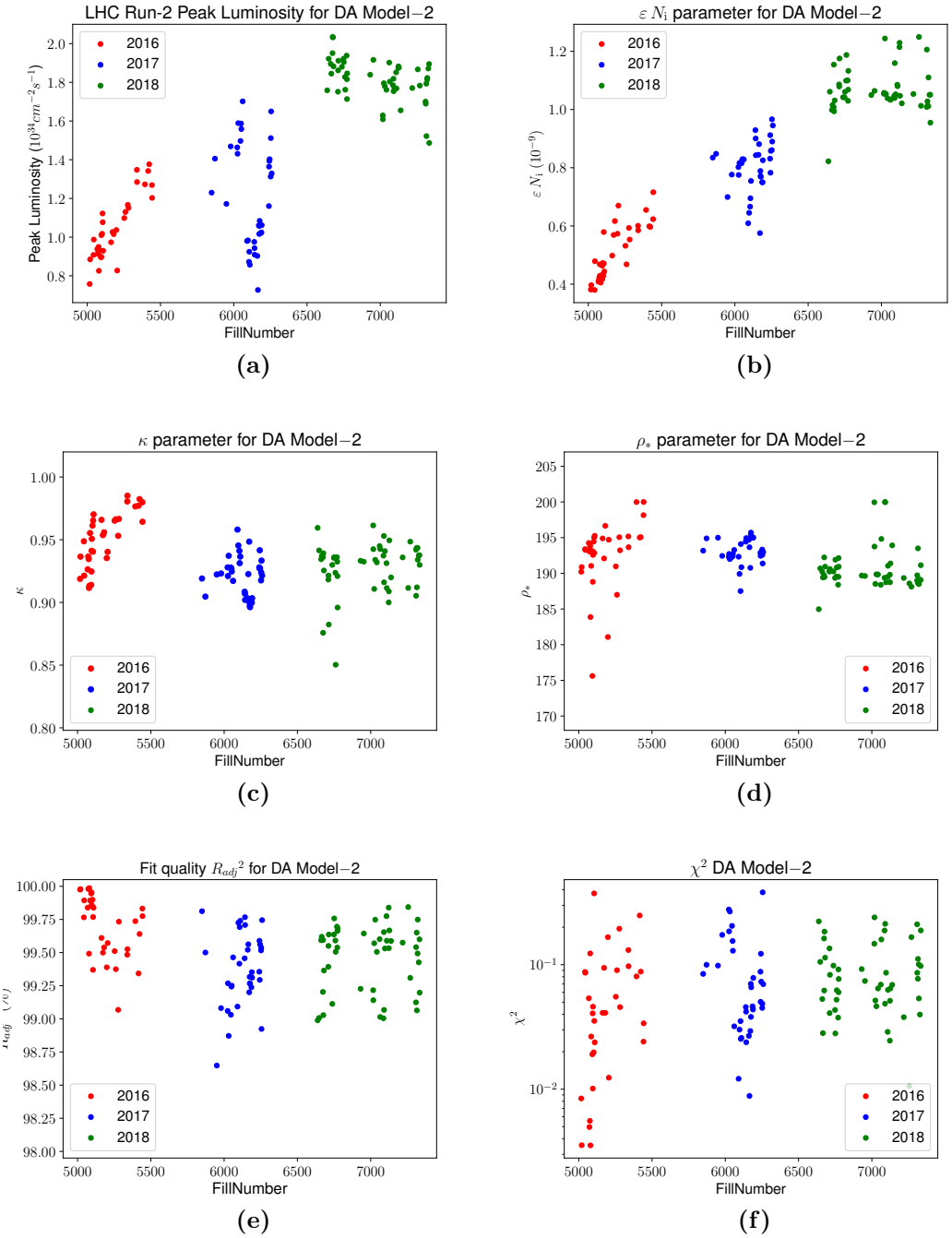


Figure 4.4: Peak luminosity, fit quality R^2_{adj}, *Reduced*-χ^2 and fit parameters of DA Model 2 (κ, $\rho*$, and $\varepsilon\, N_i$) for LHC during Run 2.

between models neglecting the temporal variability of $\varepsilon N_{\mathrm{i}}$ cross-checked against models in which $\varepsilon N_{\mathrm{i}}$ is assumed to be time-dependent. The resultant differences were found to be minimal, thereby validating the choice to overlook the time-sensitive aspect of $\varepsilon N_{\mathrm{i}}$ in our analysis.

4.2 DA Model 2 parameters evolution with time

Determining whether or not the DA Model 2 free parameters are time independent is crucial for the model's qualification. In an effort to address this, a study on the parameter evolution over time has been carried out by taking a small amount of data and then gradually increasing the number of data points until the end of the physics fill.

Fig. 4.5 shows the time evolution of the fit quality R^2_{adj} and DA Model 2 free parameters κ, ρ_*, and $\varepsilon\, N_i$, combining all the fills in the same plot for a given year of LHC Run 2 (2016, 2017, and 2018). The plots in Fig. 4.5 show that the parameters tend to remain stable over time, which is not the case when applying a double exponential model (see Fig. 4.6b and Fig. 4.7b).

In our endeavour to validate the robustness of the proposed model, a comparative analysis was conducted between the DA Model 2 and the Double Exponential model. he Double Exponential model for luminosity evolution description, given as follow

$$L(t) = a\exp(-bt) + c\exp(-dt) \tag{4.4}$$

where a, b, c, d are the fit parameters. $a + c$ represents the peak luminosity, while b and d are the time constants of the luminosity decay.

To investigate how the fit parameters of the considered models behave in a typical physics fill, we made a summary plot that displays the fit quality R^2_{adj} , the normalised fit parameters, as well as the normalised luminosity and the fitted model over time. The outcome of this analysis for the physics fills 5069 and 6161 are presented, respectively, in Fig. (4.6) and Fig. (4.7).

Examination of the Model 2 fit parameters evolution as a function of the fill length as shown in Fig. 4.6a and Fig. 4.7a, shows that the parameters remain within a stable region, with slight changes at some points due to changes in data blocks. In contrast, for the Double Exponential model, the fit parameters are not stable through the fill length. As depicted in Fig. 4.6b and Fig. 4.6b, the parameters jump abruptly from one wide range of values to another.

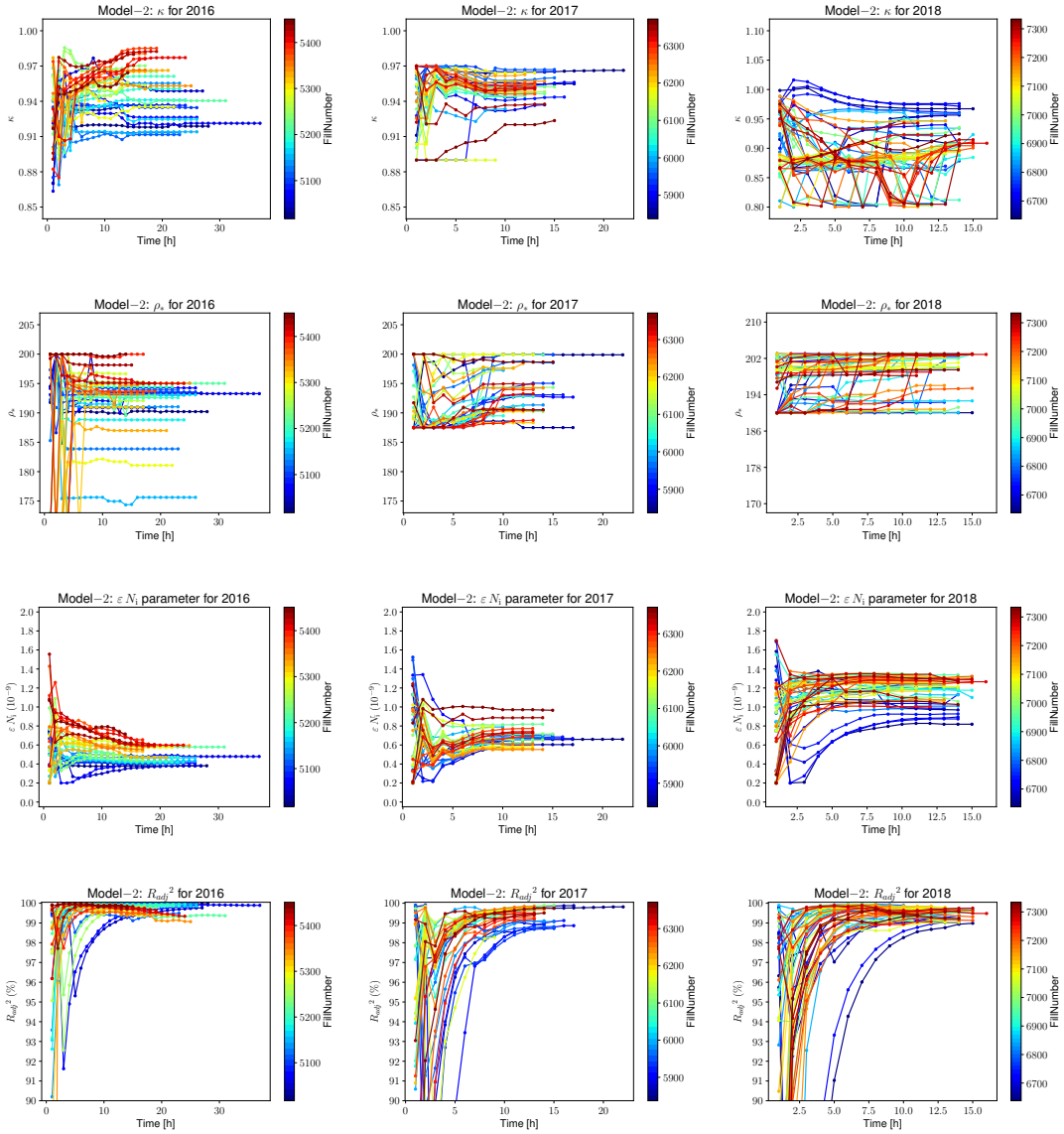
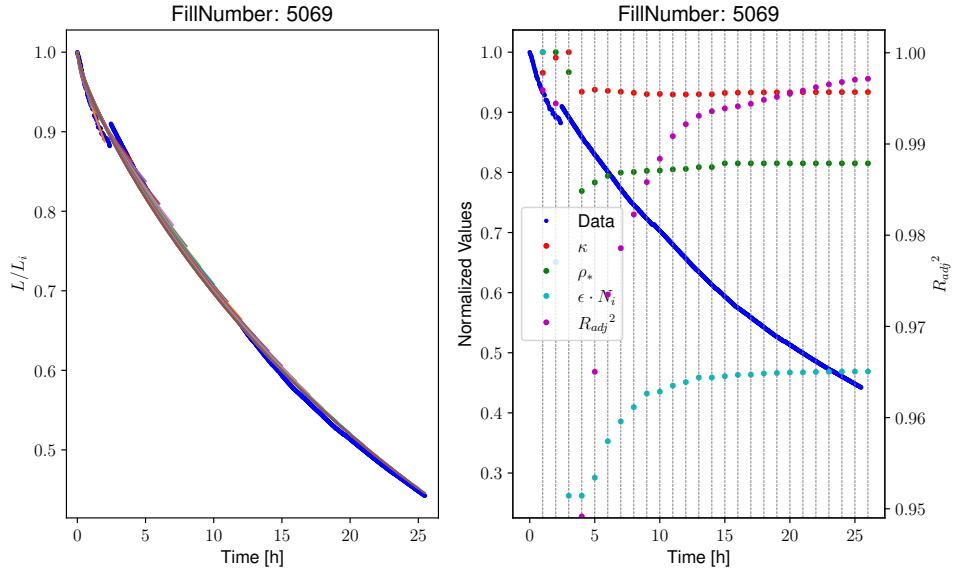
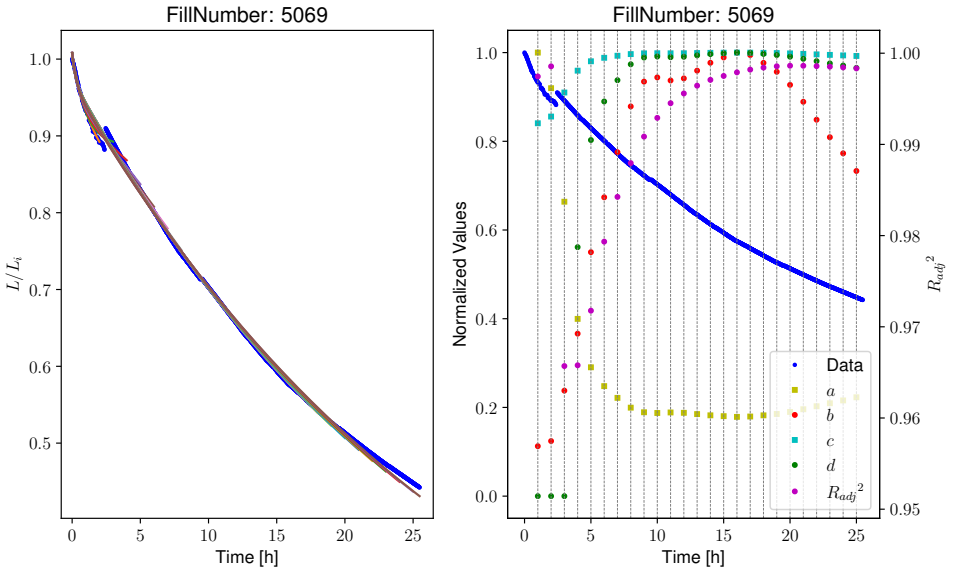


Figure 4.5: Fit quality R^2_{adj} and fit parameters κ, ρ_*, and $\varepsilon\, N_i$ of Model 2, for years 2016 (left), 2017 (middle) and 2018 (right).

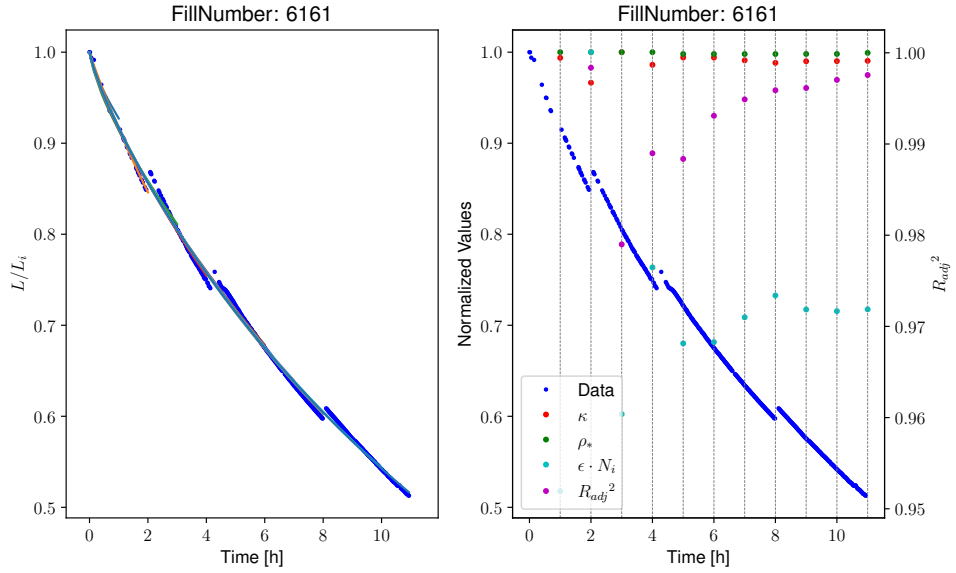


(a)

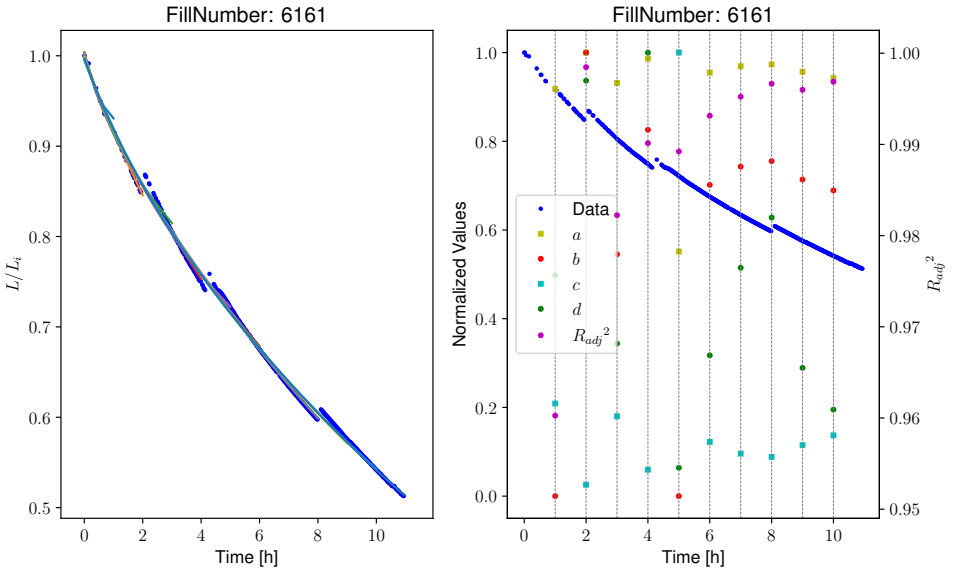


(b)

Figure 4.6: Physics fill 5069: Time evolution of normalised luminosity L/L_i, fit quality R^2_{adj} and normalised fit parameters for: (a) Model 2, and (b) Double Exponential Model.

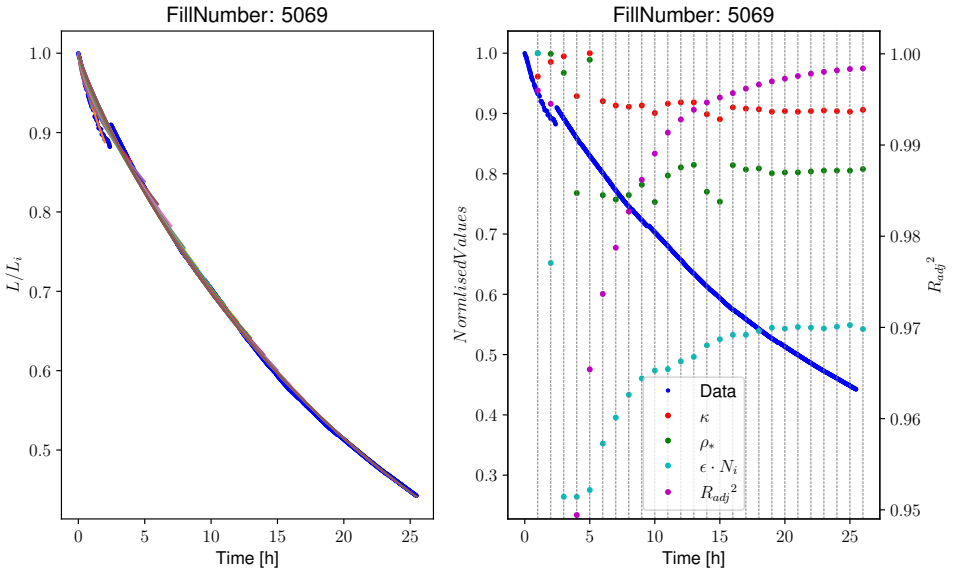


(a)

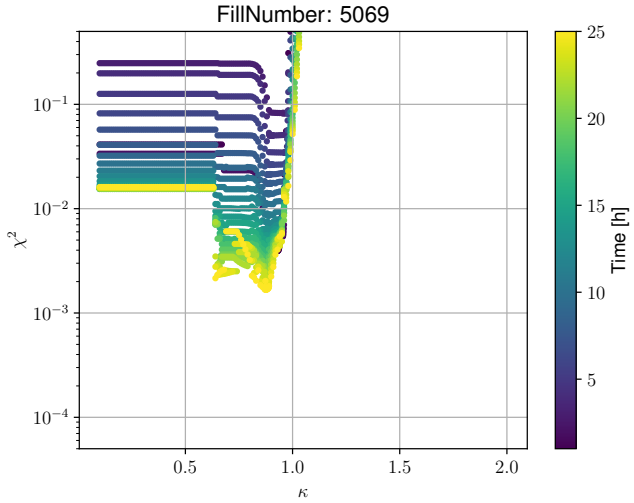


(b)

Figure 4.7: Physics fill 6161: Time evolution of normalised luminosity L/L_i, fit quality R^2_{adj} and normalised fit parameters for: (a) Model 2, and (b) Double Exponential Model.

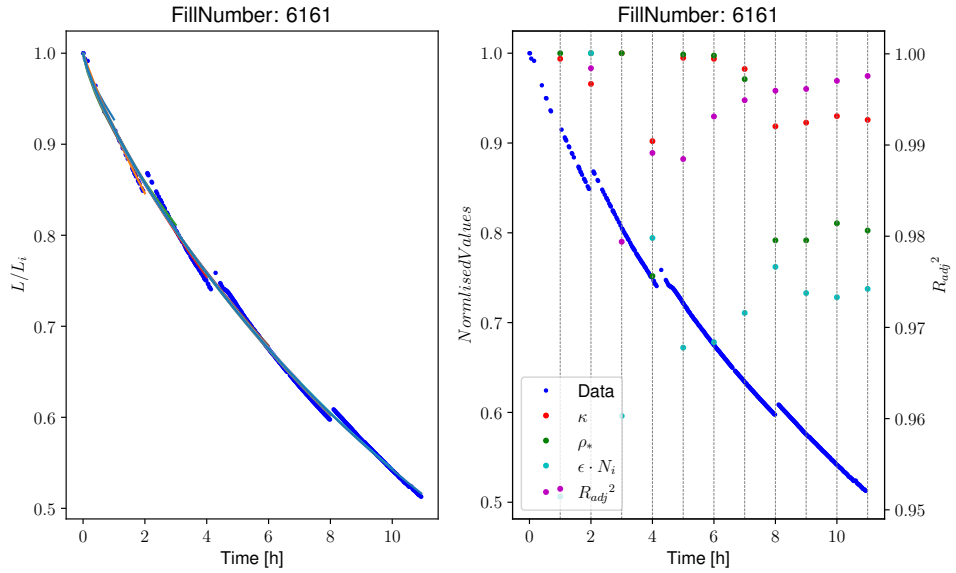


(a)

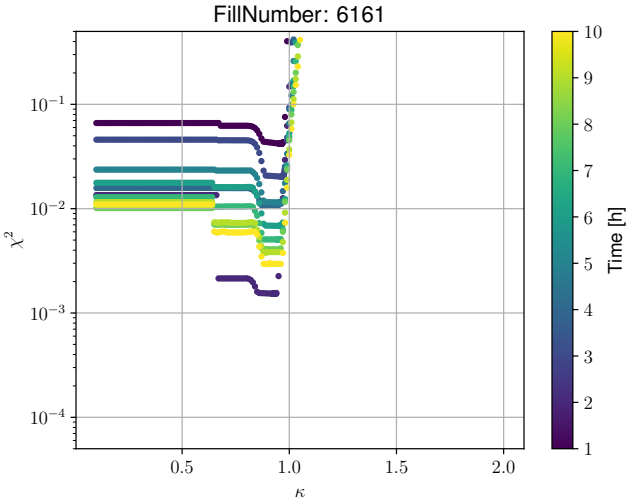


(b)

Figure 4.8: Physics fill 5069: (a) Time evolution of normalised luminosity L/L_i, fit quality R^2_{adj} for Model 2 with a scan over κ and fit parameters (ρ_* and εN_{i}) in normlised scale. (b) The χ^2 in logarithmic scale as function of scanned κ.



(a)



(b)

Figure 4.9: Physics fill 6161: (a) Time evolution of normalised luminosity L/L_i, fit quality R^2_{adj} for Model 2 with a scan over κ and fit parameters (ρ_* and εN_{i}) in normlised scale. (b) The χ^2 in logarithmic scale as function of scanned κ.

Furthermore, in order to improve the fit quality of Model 2, we did consider scanning over one parameter κ and fitting only two parameters: ρ_* and εN_i. The approach is to find the global χ^2 minimum and take the corresponding κ values as input for Model 2, to check the evolution of the fit parameters over time, see Fig. 4.8a and Fig. 4.9a. Additionally, the plots of χ^2 as a function of κ on a logarithmic scale are presented in Fig. 4.8b and Fig. 4.9b, respectively for the same physics fills as above, namely 5069 and 6161. Interestingly, the global χ^2 minimum is clearly visible in both fills.

The analysis conducted on the LHC Run 2 instantaneous luminosity data using different models, namely DA Model 2 and Double Exponential model, combined with the observation of their free parameters evolution during a physics fill, provide an initial understanding of our model's performance. In the following section, the extrapolation approach serves as a first step towards predicting the evolution of luminosity in a circular collider.

4.3 Luminosity Extrapolation

For the sake of the accuracy of our study, starting with a small amount of data, using various models for extrapolation up to the end of the physics fill, we assess the validity of our techniques. Fig. 4.10 shows that the relative difference in integrated luminosity $\mathcal{L}(t)$ between the LHC data and the chosen model tends to decrease over time with increasing data points. While for instantaneous luminosity $L(t)$ the relative difference is greater, as shown in Fig. 4.11. More importantly, the results of this analysis can be summarised as follows:

- DA Model 2 relative difference with reference to LHC data continuously is steadily decreasing, maintaining its inherent trend. Essentially, this means that as we integrate more data points, our predictive values get closer to the actual luminosity values, reflecting the model's ability to account for any diffusive effect within the machine.
- Double Exponential model relative difference in comparison to the LHC data decreases with the number of data points that have been taken into account. However, notable fluctuations are observed throughout the filling time, analogously to the time-dependent evolution of its fitting parameter. This suggests that the model is sensitive to specific machine dynamics, which compromises its ability to effectively predict the evolution of luminosity, particularly given the non-linear challenges in circular devices.

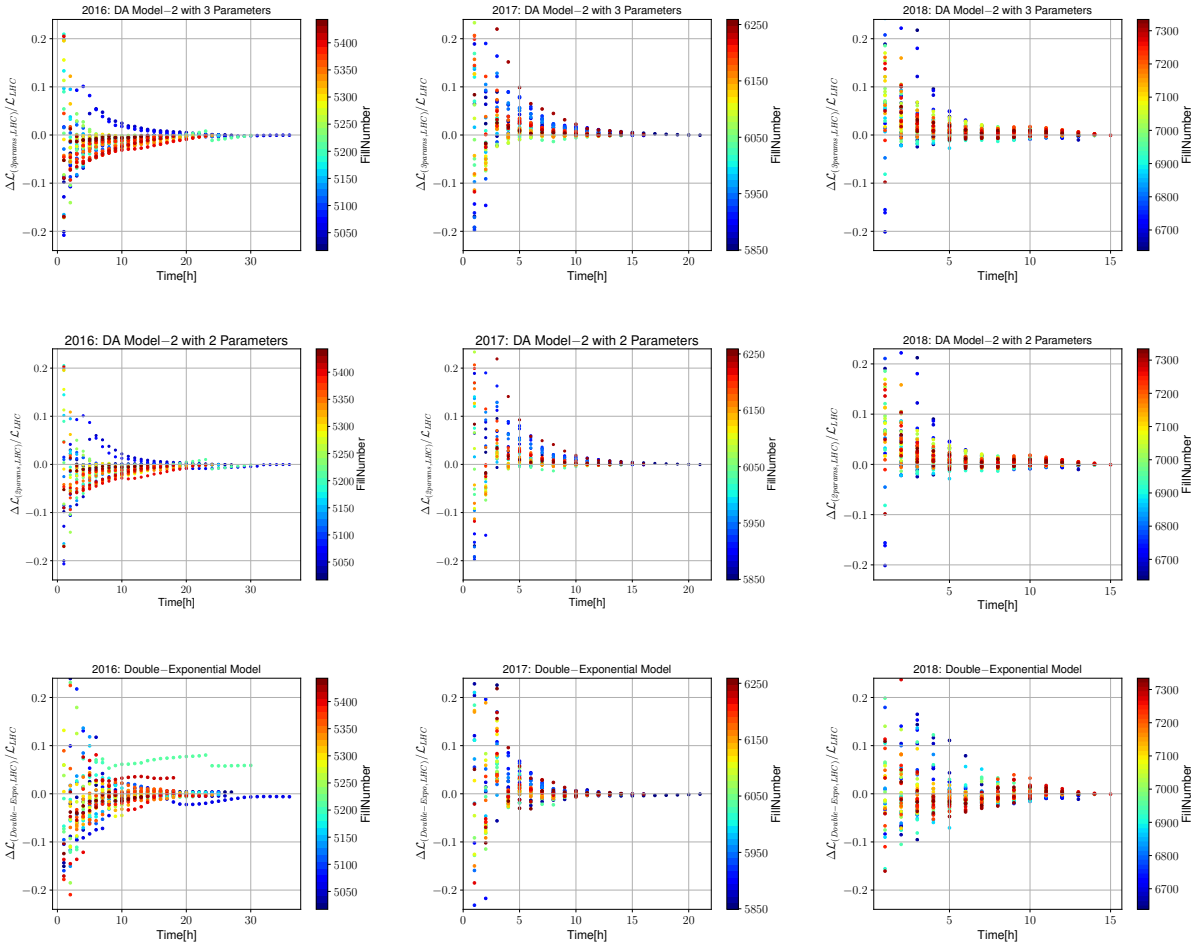


Figure 4.10: The relative difference in terms of integrated luminosity $\mathcal{L}(t)$ over time between the LHC data and the extrapolated data for 2016 (left), 2017 (middle), 2018 (right). Extrapolated data using: Model 2 with three free parameters (κ, ρ_*, and εN_i) (top), Model 2 with two free parameters (ρ_*, and εN_i) and a scan of κ (centre), and double exponential model with four parameters (bottom).

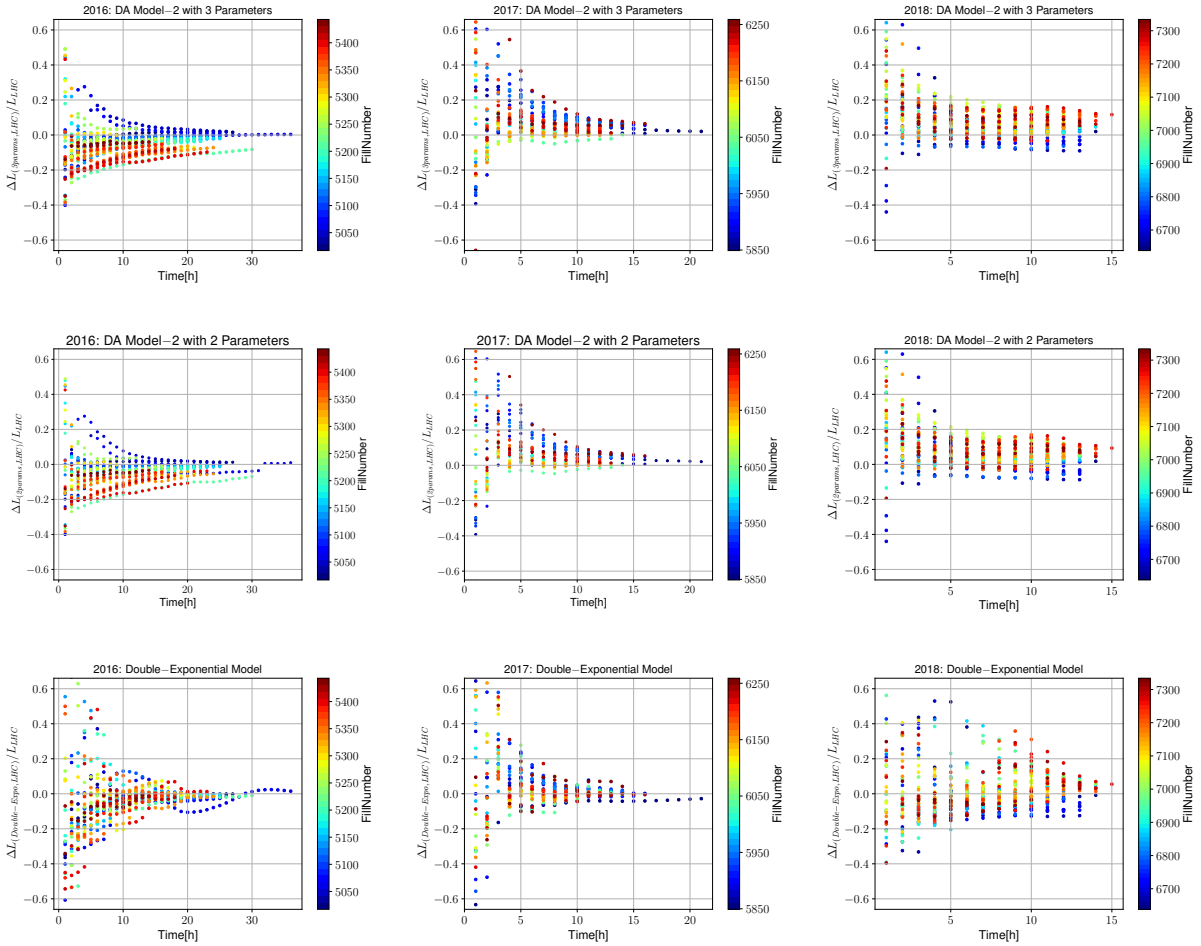


Figure 4.11: The relative difference in terms of instantaneous luminosity L_{inst} over time between the LHC data and the extrapolated data for 2016 (left), 2017 (middle), and 2018 (right). Extrapolated data using: Model 2 with three free parameters (κ, ρ_*, and εN_i) (top), Model 2 with two free parameters (ρ_*, and εN_i) and a scan over κ (centre), and double exponential model with four parameters (bottom).

In summary, for luminosity modelling, any functional form fitted to the measured luminosity evolution can be used, providing a good fit quality. Although a Double Exponential model initially provides a good empirical description of the evolution of luminosity over time, in practice, complexities arise. Factors such as particle losses due to diffusive phenomena can affect the beams, leading to radical changes in the model parameters, and making the model inconsistent. We therefore advocate a model based on the concept of DA, which includes all pseudo-diffusive effects on the model. Given these insights, reinforce our conviction in the DA Model 2 as a potential candidate to predict the evolution of luminosity in circular colliders.

Chapter 5

Optimisation Strategies

This chapter presents the main computational techniques that could be applied to the LHC to enhance its overall integrated luminosity. First, we implemented a numerical optimiser that predicts luminosity evolution, making optimal decisions based on the beforehand accessible LHC Run-2 data and DA Model 2, i.e. Off-line Strategy. Then, for real-time luminosity predictions and to determine optimal measures, we propose an optimisation strategy that operates during the LHC Run filling scheme, i.e. On-line Strategy. In both approaches, the optimisation is performed with the constraint that the sum of the optimised fill lengths equals that of the actual fill lengths.

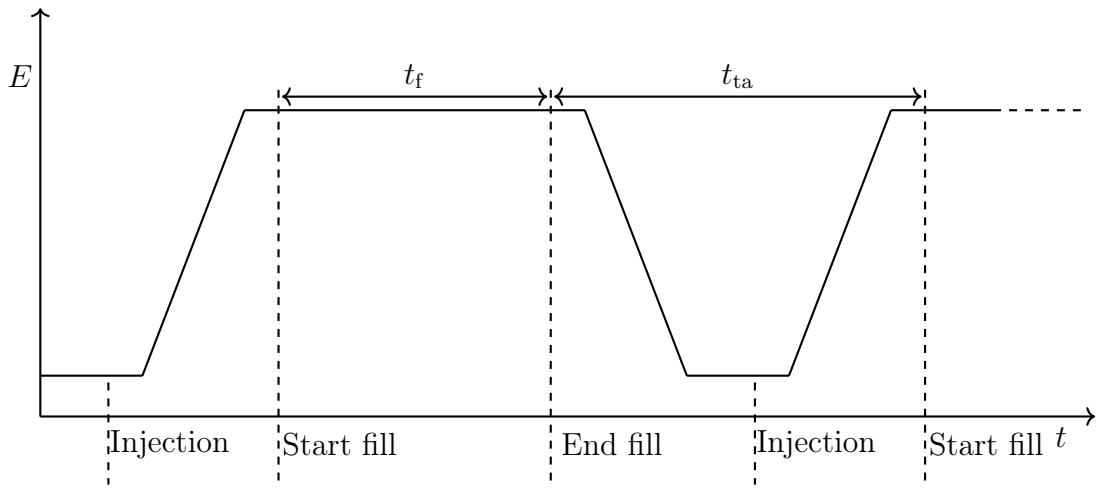


Figure 5.1: Sketch of the luminosity production process in a circular collider.

The process of luminosity production in a circular collider is schematically illustrated in Fig.5.1, highlighting the magnetic cycle and introducing process-related main components.

From an analytical point of view, the luminosity measurement $L(t)$ throughout a fill can be described by adopting any mathematical model that aims to provide a customised approach for collecting luminosity in a circular collider.

Since the goal is to maximise the integrated luminosity, in deterministic case, achieving this aim, means maximising the following

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

where $\mathcal{L}_{\mathrm{tot}}(t_{\mathrm{f}})$ is the total integrated luminosity in the machine, N_{f} refers to the number of fills considered in our study, while T is the total time allocated for physics, t_{ta} refers to the so-called turnaround, i.e. the time between the end of a physics fill and the start of the subsequent one, t_{f} is the length of an individual physics fill. Special attention must be given to the second component of Eq. (5.1) to ensure accuracy and meticulous handling. Ideally, the expression should be presented as

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

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

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

Henceforth, we consider the possibility of optimising the performance of a circular collider, focusing on maximising the annual integrated luminosity at LHC in the dynamic aperture framework. The computational foundation of our study relies on the adaptation of the proposed strategies discussed in Ref.[19], applying the DA Model 2 scaling law to derive the analytical expressions for the optimized integrated luminosities for the optimum fill lengths.

5.1 Off-line Optimisation

The numerical optimiser in this study uses the Python's minimize routine. Specifically, we employ the Sequential Least Squares Programming (SLSQP) Algorithm, a gradient-based optimization technique that addresses constrained optimisation problems through a sequential quadratic programming approach[19].

Given the intriguing challenge that no single model can accurately describe all fills, it becomes imperative to compute the total integrated luminosity by summing the individual integrated luminosities, $\mathcal{L}(t_{\mathrm{f}})$ obtained from the integral of their respective fill-fitting model, which, for our study, is predominantly based on the DA Model 2.

Utilizing insights from the DA framework, we have scaled and refined the DA Model 2 to optimally project integrated luminosities for the optimum fill lengths at the LHC.

Given that the storage of luminosity data is presented in terms of time rather than the number of turns, the most efficient computational approach to derive the pseudo-diffusive contribution to luminosity prediction data involves applying the DA Model 2 model in such a way the integrated luminosity per fill reads

$$\mathcal{L}(t_i) = \frac{L_i}{f_{rev}} \int_0^{t_i} \mathrm{d}t \, L_{norm}(t) \tag{5.4}$$

where $\mathcal{L}(t_i)$ is the maximum integrated luminosity per fill, computed from $t = 0$ to the optimal fill time $t = t_i$ defined by the optimiser. $L_{norm}(t_i)$ is the fill normalised luminosity evolution form provided by the DA Model 2 in Eq.(3.33). Hence, to remove the normalization and go from the dimensionless turn variable to the time variable, a fill-related scaling factor is introduced and given by L_i/f_{rev}, with L_i is the initial luminosity value and f_{rev} is the revolution frequency of the machine.

It is straightforward to express the total integrated luminosity $\mathcal{L}_{\mathrm{tot}}$ to be optimised as

$$\mathcal{L}_{\mathrm{tot}} = \sum_{i}^{N_{\mathrm{f}}} \mathcal{L}(t_i) \tag{5.5}$$

with imposing the constraint which assumes the total optimum fill lengths determined by the optimisation algorithm equal to the total actual machine fill lengths,

namely

$$\sum_{i}^{N_{\rm f}} t_i - \sum_{i}^{N_{\rm f}} t_{ri} = 0. \tag{5.6}$$

where $N_{\rm f}$ as already mentioned is the number of fills selected by the optimiser, and t_{ri} is the fill time that took place at LHC.

5.2 On-line Optimisation

Now, let's examine the potential of real-time optimization and prediction of luminosity production during the filling process. Before diving deep into solving the Eq.(5.3) for the purpose of enhancing the machine luminosity, for the purpose of enhancing the machine luminosity, we aim first to get a firm grasp on our primary issue by tackling the statistical nature of both the turnaround time and the physics fill length. Focusing on a scenario where we have n exponentially-distributed values t_i representing n realisations of the turnaround time, the goal is to maximise

$$\mathcal{L}_{\rm tot}(\hat{t}) = n \int_0^{\hat{t}} {\rm d}t \, L(t) \,, \tag{5.7}$$

Interestingly, the underlying premise here is that all fills should be of equal length, even though the turnaround times might vary, subjected to the constraint

$$\sum_{i=1}^{n} t_i + n \, \hat{t} = T \,. \tag{5.8}$$

One can simplifying the sum of the exponentially-distributed values with a term $n\,\tau$, resulting in

$$n = \frac{T}{\tau + \hat{t}} \tag{5.9}$$

Hence, inserting Eq. (5.9) in Eq. (5.7), the optimisation of Eq. (5.7) becomes of the same type as the problem (5.1). It is worth noting that the implicit assumption here, is that τ is the average value of the turnaround distributions, $\tau = \langle t_i \rangle$, for further invisi.

It can be checked *a posteriori* that the assumption of optimising the integrated

luminosity by using equal fills lengths is the correct one. In case one would consider

$$\begin{aligned}\mathcal{L}_{\rm tot}(\hat{t}_1,\hat{t}_2) &= (n-1)\int_0^{\hat{t}_1}\mathrm{d}t\,L(t)+\int_0^{\hat{t}_2}\mathrm{d}t\,L(t)\\ &= \frac{T-(\hat{t}_2+\tau)}{\tau+\hat{t}_1}\int_0^{\hat{t}_1}\mathrm{d}t\,L(t)+\int_0^{\hat{t}_2}\mathrm{d}t\,L(t)\,,\end{aligned} \tag{5.10}$$

where the constraint (5.8) has been adapted to this new case. The maximisation of $\mathcal{L}_{\rm tot}(\hat{t}_1,\hat{t}_2)$ is obtained by solving

$$\begin{cases}\dfrac{\partial\mathcal{L}_{\rm tot}}{\partial\hat{t}_1} &= 0\\[2ex] \dfrac{\partial\mathcal{L}_{\rm tot}}{\partial\hat{t}_2} &= 0\,,\end{cases} \tag{5.11}$$

and by direct inspection, one finds that $\hat{t}_1=\hat{t}_2$.

Since the fill length is a key measure for evaluating the performance of a circular collider, an in-depth analysis of how to optimise this quantity to reach the maximum integrated luminosity has been presented in Refs.[3, 19].

Clearly, in the realistic case, the n values of the turnaround are not known *a priori*, which modifies the scheme previously described. Let us assume that t_j, $1\le j\le i$ and $\hat{t}_j$, $1\le j\le i-1$ are the turnaround 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

$$\mathcal{L}_{\rm tot}(\hat{t}_i,\hat{t}) = \sum_{j=1}^{i-1}\mathcal{L}(\hat{t}_j)+\int_0^{\hat{t}_i}\mathrm{d}t\,L(t)+\frac{T-\sum_{j=1}^{i}\left(t_j+\hat{t}_j\right)}{\frac{1}{i}\sum_{j=1}^{i}t_j+\hat{t}}\int_0^{\hat{t}}\mathrm{d}t\,\mathcal{L}_{\rm mp}(t)\,, \tag{5.12}$$

where and additional optimisation parameter has been introduced, namely $\hat{t}$, which represents the optimum fill time of future fills. Here, $\mathcal{L}$ stands for the integrated luminosity in a single fill. The third term in Eq. (5.12) introduces a relationship between $\hat{t}_i$ and $\hat{t}$. We remark that $\mathcal{L}_{\rm mp}(t)$ 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

$$\frac{1}{i}\sum_{j=1}^{i}t_j \tag{5.13}$$

is intended to provide an estimate of the average turnaround time.

The optimisation is performed as usual, namely

$$\begin{cases} \dfrac{\partial \mathcal{L}_{\text{tot}}}{\partial \hat{t}_i} & = L(\hat{t}_i) - \dfrac{1}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \displaystyle\int_0^{\hat{t}} \mathrm{d}t\, \mathcal{L}_{\text{mp}}(t) = 0 \\ \\ \dfrac{\partial \mathcal{L}_{\text{tot}}}{\partial \hat{t}} & = L(\hat{t}) - \dfrac{1}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \displaystyle\int_0^{\hat{t}} \mathrm{d}t\, \mathcal{L}_{\text{mp}}(t) = 0\,, \end{cases} \tag{5.14}$$

which gives

$$\begin{cases} \hat{t}_i & = \hat{t} \\ \\ L(\hat{t}) & = \dfrac{1}{\frac{1}{i}\sum_{j=1}^{i} t_j + \hat{t}} \displaystyle\int_0^{\hat{t}} \mathrm{d}t\, \mathcal{L}_{\text{mp}}(t)\,. \end{cases} \tag{5.15}$$

The proposed strategies have been evaluated by optimizing the fill length using a numerical optimizer while retaining the same number of fills as for the actual run. Afterward, we tested the online strategy designed to optimize the integrated luminosity for a single year of run. The following chapter.6 will provide a detailed verification and showcase the outcomes of these strategies in enhancing the integrated luminosity.

Chapter 6

Optimisation Results

In the previous chapter, we discussed several optimisation strategies that could be applied to the LHC to improve its luminosity performance. In this chapter, we have undertaken the integration of these techniques in such a way that the sum of the optimised fill lengths is equivalent to the sum of the actual fill lengths. The main objective of this initiative is to predict and explore the possibility of enhancing the integrated luminosity using the proposed DA Model 2 compared with the combined integrated luminosity collected over the three years of LHC Run 2. To achieve this, we implemented a numerical optimiser, which will be further elaborated in the following section.

6.1 Off-line Results

It is noteworthy that for operational accelerators, offline optimisation plays a crucial role, particularly for facilities that require frequent recalibrations or changes in configuration, such as the LHC with its multiple experiments and varied operational modes. Although many accelerator can, in theory, access the optimisation resources, practical constraints, especially computational ones, often prevent in-depth studies, i.e. LHC is expensive in terms of time. Therefore, an optimisation procedure, taking into account these boundary conditions becomes essential.

Once the operational characteristics of the machine are well understood, numerical optimisation using the established procedures, as was done for the early store, can indeed improve the machine's performance and safety.

As an initial step, we employed a numerical optimiser following the procedure outlined in Refs.[19, 31]. However, instead of using the Double Exponential model, we adopted the DA Model 2 proposed by our study. Hence, the numerical optimisation presented herein is rooted in the offline approach, as detailed in Sec. 5.1. The purpose of this approach is to optimise the fill length, ensuring maximum integrated luminosity from each run. This optimisation is achieved while maintaining the constraint condition that the total optimised fill lengths are equivalent to the total considered run fill lengths, as shown in Fig. 6.2. Aditionally, the optimised individual fill time compared to the actual fill time is showcased in Fig. 6.1.

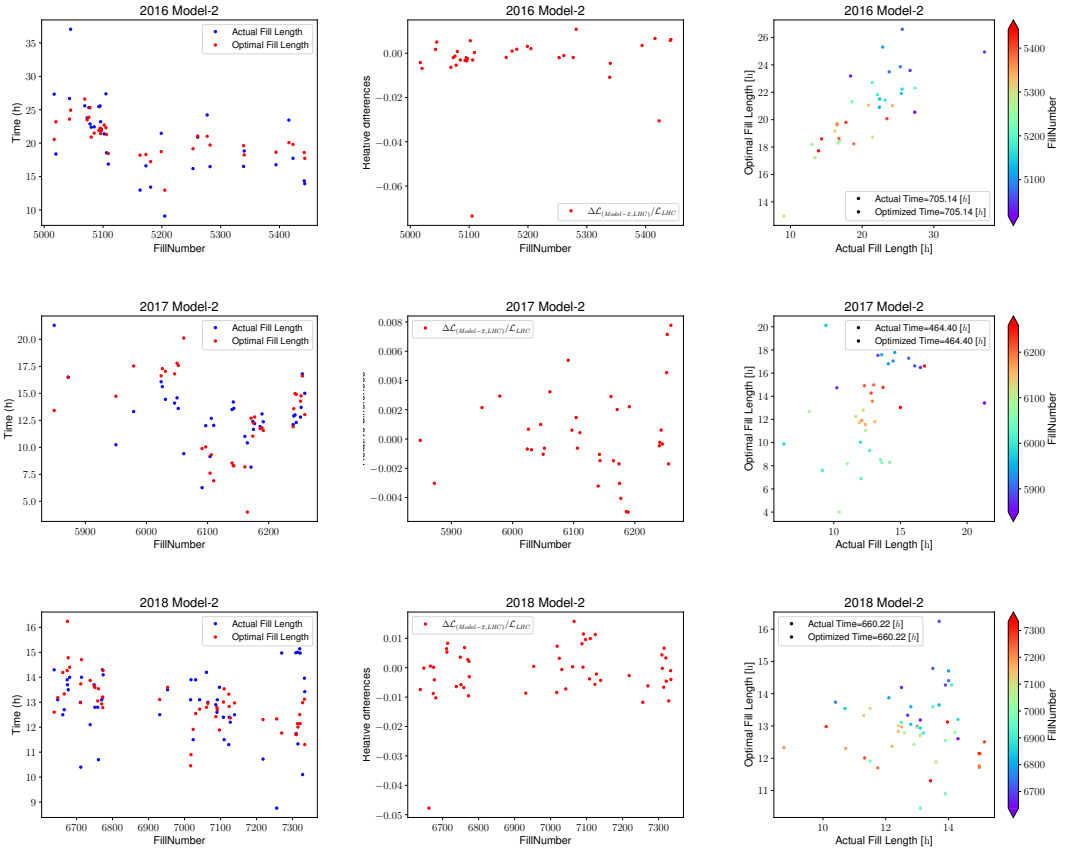


Figure 6.1: (left) Comparisons between the Actual fill times that took place at LHC during Run 2 and the optimal fill times chosen by the DA Model 2. (middle) The relative difference on integrated luminosity $\mathcal{L}(t)$ by the optimiser with reference to the measured one at LHC during Run 2 (right) Optimal fill times vs Actual fill times that took place at LHC during Run 2.

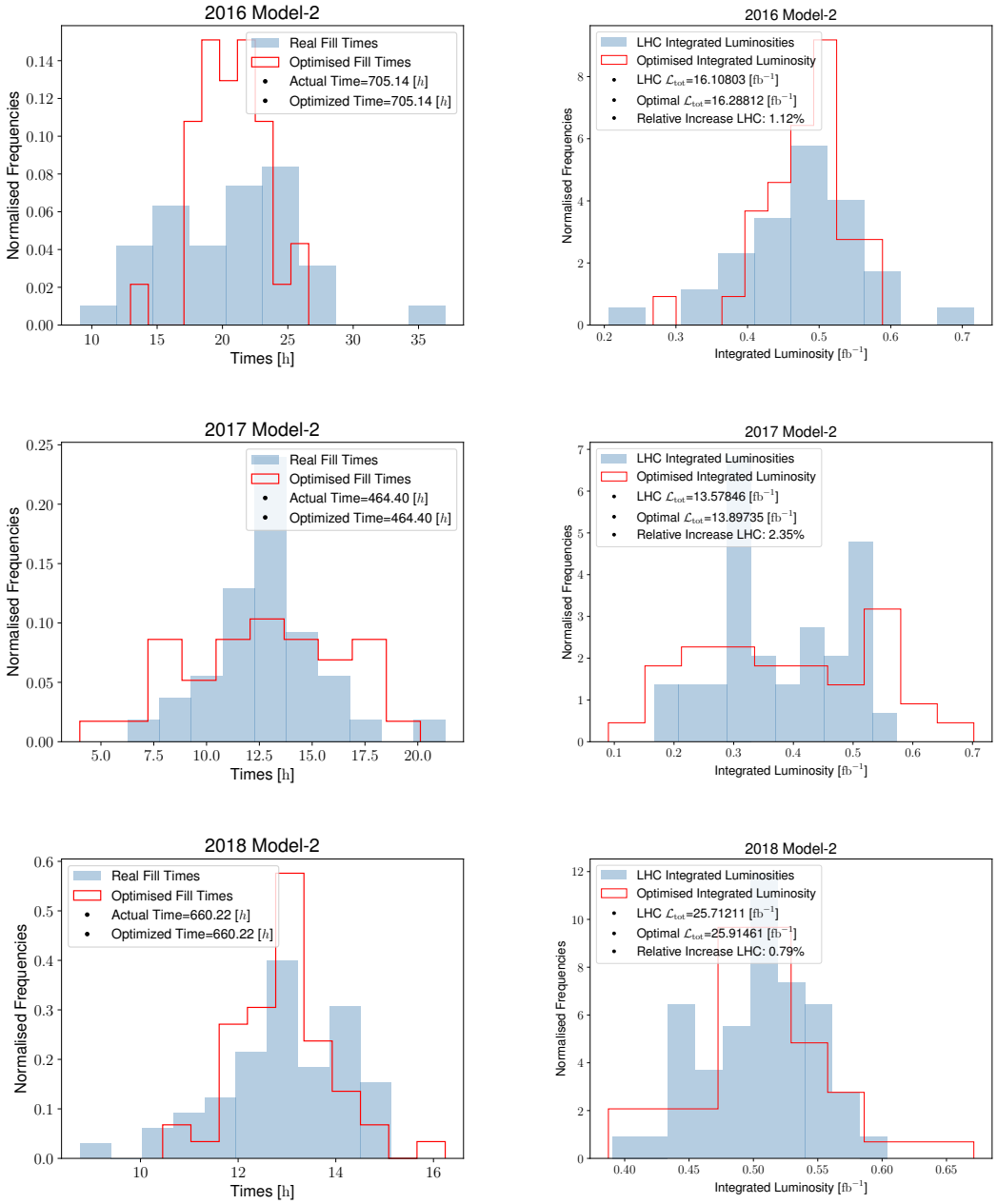


Figure 6.2: (left) Comparisons between the fill times that took place at LHC during Run 2 and the optimal fill times chosen by DA Model 2; (right) Comparisons between the measured integrated luminosities at LHC during Run 2 and the optimal ones obtained by Model 2.

The optimisation results are presented by the histograms in Fig. 6.2. Indeed, we can observe an improvement in the integrated luminosity by the optmiser using DA Model 2 relative to the integrated luminosity accumulated over the three years of the LHC Run 2. This implies that with the same operational time and fill lengths, and using the DA Model 2 to predict the luminosity evolution within the machine, the LHC could increase the collision rates, thus offering a gateway to more experimental outcomes.

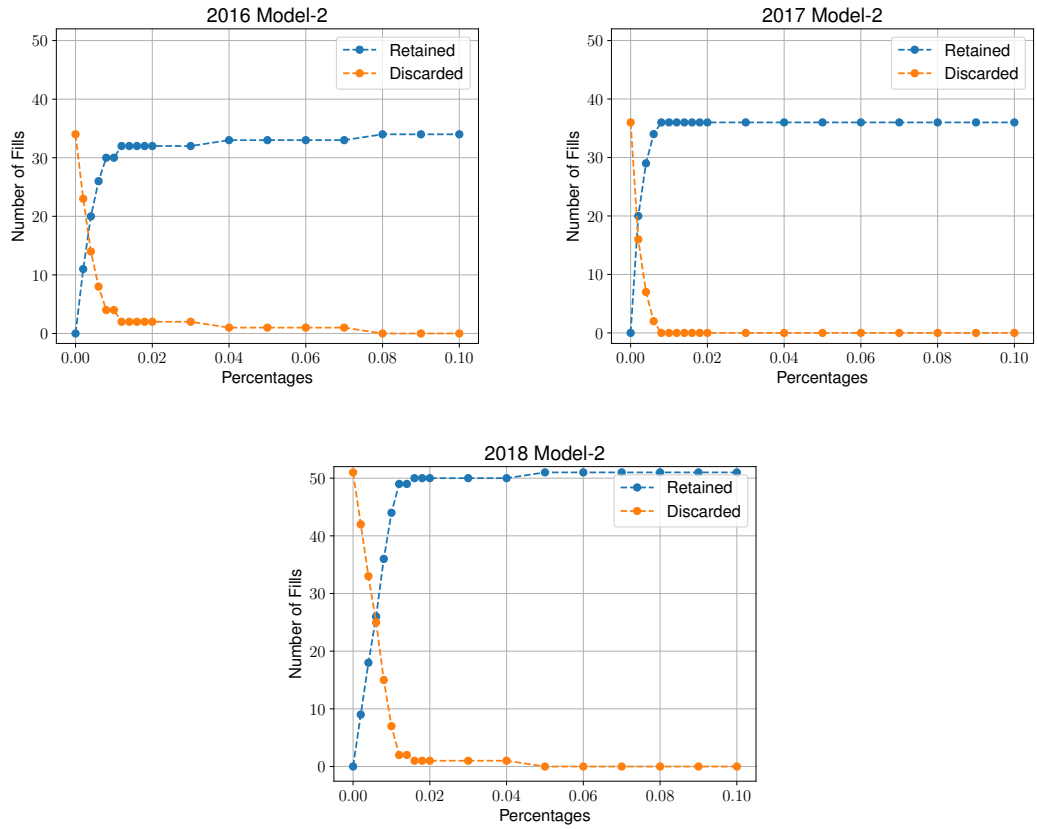


Figure 6.3: Plots of the convergence study for the outliers removal for the three years of Run 2.

Moreover, despite the strict fill selection procedure we carried out in Sec. 4.1.1, there may still exist outliers in our dataset. These outliers, overlooked by our methodology, might inadvertently amplify or attenuate the expected results. Therefore, to ensure the accuracy of our study, we made an internal check to identify any potential outliers in the luminosity data. In the Fig. 6.1 we represent the relative difference in terms of the integrated luminosity $\mathcal{L}_{\text{num}}(t)$ from the numerical optmiser with reference to the actual luminosity $\mathcal{L}_{\text{LHC}}$ recorded at LHC, i.e. the LHC

	2016	**2017**	**2018**
Relative Increase Before Cut	1.12%	2.35%	0.79%
Relative Increase After Cut	1.04%	2.35%	0.83%

Table 6.1: Integrated luminosity Relative increase before and after cutting procedure.

integrated luminosity for each fill stored in the ATLAS data directory summary files [9].

The plots of the relative difference in terms of the integrated luminosity $\mathcal{L}(t)$, as illustrated in Fig. 6.1, show that there were two visible outliers in 2016, and one in 2018. Subsequent to this analysis, our focus was solely on fills that met the criteria

$$\frac{|\mathcal{L}_{\mathrm{mod}_2} - \mathcal{L}_{\mathrm{LHC}}|}{\mathcal{L}_{\mathrm{LHC}}} \leq 0.02 \tag{6.1}$$

The premise here in Eq. (6.1) is we attempt to remove all outliers by applying a cut that discards values above 0.02. The underlying logic behind choosing a threshold of 0.02 is shown in shown in Fig. 6.3.

Therefore, the optimisation results after performing the cut-off in Eq. (6.1) are provided in Fig. 6.4. In order to evaluate the effectivness of the numerical optimiser both before and after implementing the cut-off in Eq. (6.1), we have outlined in Tab.6.1 the relative increase in integrated luminosity from the optmiser with respect to the actual integrated luminosity at LHC. This comparison reveals variations in gains both across different years and depending on whether the cut-off has been applied. For instance, in 2016, there was a slight reduction in the gain post-cut, which can be attributed to the removal of some fills. In contrast for 2018, after removal of the outliers, the gain increased slightly. The interpretation here is that outliers can sometimes distort or skew the apparent performance, leading to misleading conclusions. Thus, the deeper our knowledge of the data, the greater the potential gains in integrated luminosity.

To have a general view of how the offline optimiser predicts luminosity, we have illustrated the modelled luminosity as a function of the optimal time for consecutive optimised fills in Fig. 6.5, shown in red. At the same time, this figure displays the nominal luminosity evolution at the LHC during the actual run and its fitted curve using DA Model 2, shown in blue and green, respectively. Furthermore,

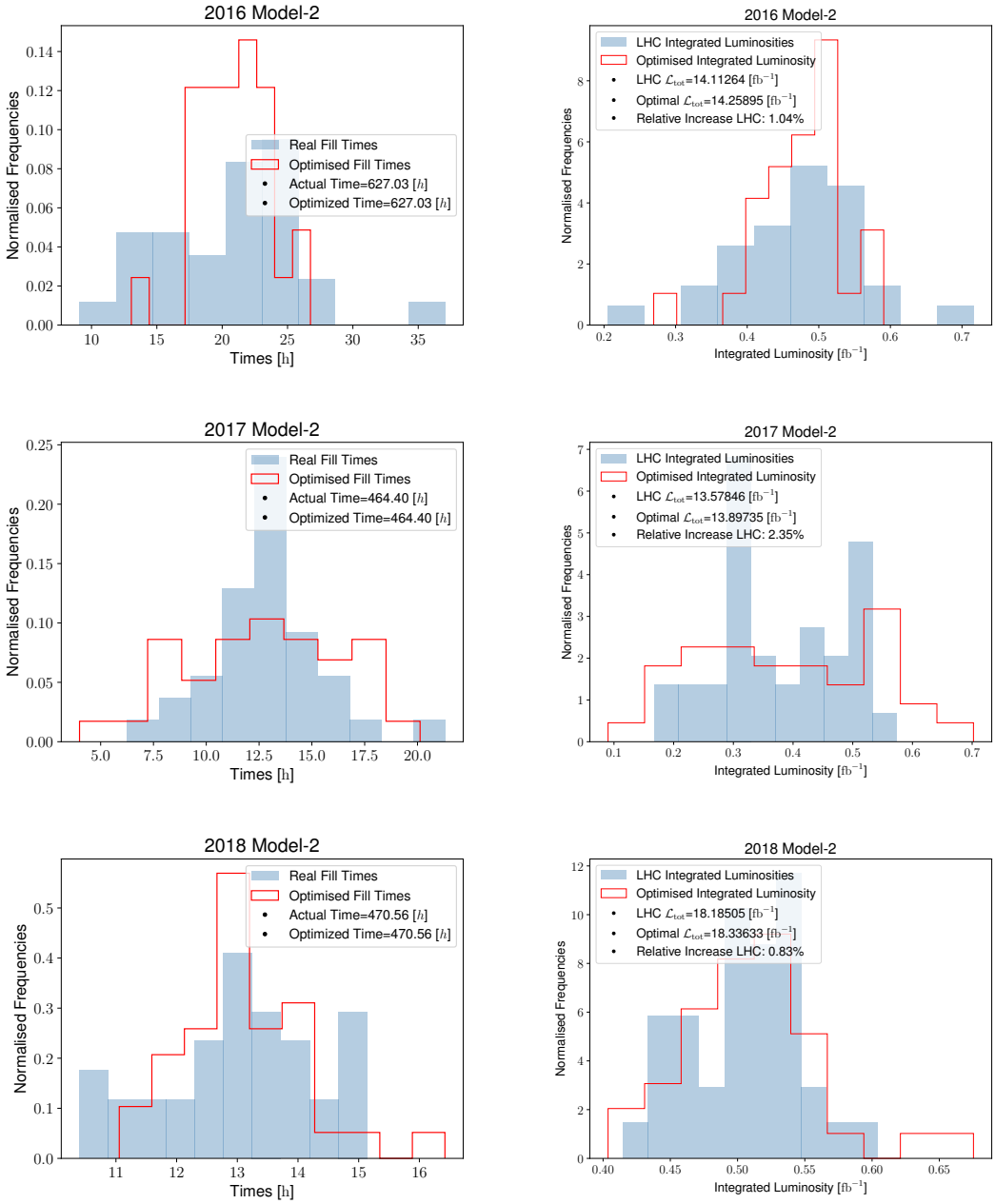


Figure 6.4: (left) Comparisons between the fill times that took place at LHC during Run 2 and the optimal fill times chosen by DA Model 2; (right) Comparisons between the measured integrated luminosities at LHC during Run 2 and the optimal ones obtained by Model 2.

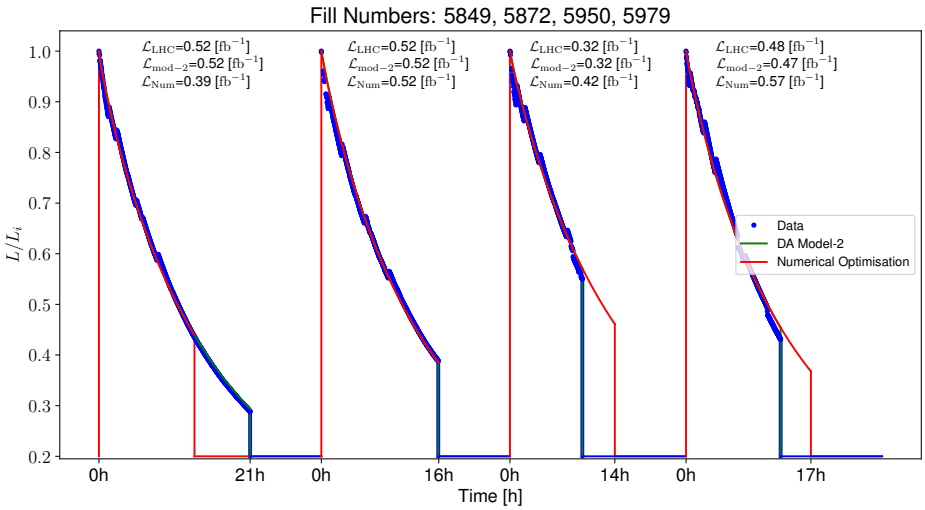


Figure 6.5: Luminosity evolution as a function of time, at LHC (in blue) and from the fit DA Model 2 (in green), and the predicted from the numerical optimiser (in red) for successive fill numbers 5849, 5872, 5950, 6024.

the integrated luminosity value at the end of the fill, obtained from numerical optimisation, DA Model 2, and the actual run, are also presented.

It can be concluded from this schematic view that the optmiser decides if the fill is optimal or not, i.e. whether to continue the fill to collect more luminosity or to interrupt it.

6.2 On-line Results

At the start of each LHC fill, the beams at the IPs are adjusted to ensure the beam head-on collisions. This is done to achieve the maximum number of collisions, thereby maximising the luminosity production.

Throughout LHC Run 2, the configuration of the LHC machine configuration configuration was significantly modified, which had a significant impact on its luminosity performance. An overview of the evolution of the main parameters of the machine during LHC Run 2 is provided in Refs.[37]. Indeed, not only were different beam types used, but also several new operational features known as control parameters were introduced, i.e. β^*, ϵ^*, aperture [4]. In 2018, after careful

optimisation, the machine moved into an advanced operational mode, reaching a remarkable peak luminosity around $2 \cdot 10^{34}$ cm^{-2} s^{-1}.

Given the running conditions, the control parameters have a direct correlation with the machine's peak luminosity performance. These parameters with the target luminosity ensure that the instantaneous luminosity with the help of dedicated detectors aligns with the predefined target luminosity throughout the fill's duration or until the instantaneous luminosity reaches the natural luminosity decay. This results in thus, the luminosity production depending on the control parameter that defines the design luminosity. Our task in this section is to attack the main challenge in the LHC control room, namely the online optimisation and control of the delivered luminosity to the machine. In fact, our goal is to implement an online optimisation to improve luminosity production following the strategy proposed in Section 5.2.

The approach relies on scanning all physics fills from the previous year, as described in Eq. (5.13), and then determining the most probable value of the luminosity distribution for every specific fill time defined during the scan. It then determines the most probable value of the luminosity distribution for each specific fill time defined during the scan. The evaluated mode is updated whenever a new fill occurs in the ongoing year. Subsequently, the DA Model 2 is employed to derive the analytical representation required for luminosity. Before initiating a new fill, the algorithm checks the imposed constraints to ensure that the combined sum of turnaround times and fill times executed during the run does not exceed the total time allocated for physics. This is followed by finding the optimal times, $\hat{t}$ that solve Eq. (5.15).

Note that the luminosity provided by the DA Model 2 in Eq. (3.33) is normalised to its initial value. In offline techniques, it is simple to obtain the initial luminosity value; it is extracted from the stored data for each fill. In the case of online optimisation, the predictions are done in real-time, thus here in our study, we simulate the luminosity evolution at LHC by conidering two distinct perspectives regarding the initial luminosity:

- Simulation with the assumption that initial luminosity L_i as the one provided by the control parameters to model the target luminosity, thus, $L_i = \Xi\, N_i^2$ as given in Eq. (3.17).

- The simulation, assuming that the initial luminosity for each fill is consistent with what was observed during the run.

Online simulations are performed on LHC Run 2 data for 2018, with complete knowledge of the fitting parameters. The results of our analysis are presented by violin plots as illustrated in Fig. 6.6, in which we can observe the distributions of the total integrated luminosity per day for the three distinct studies, ie, LHC, online optimisation, and numerical optimisation.

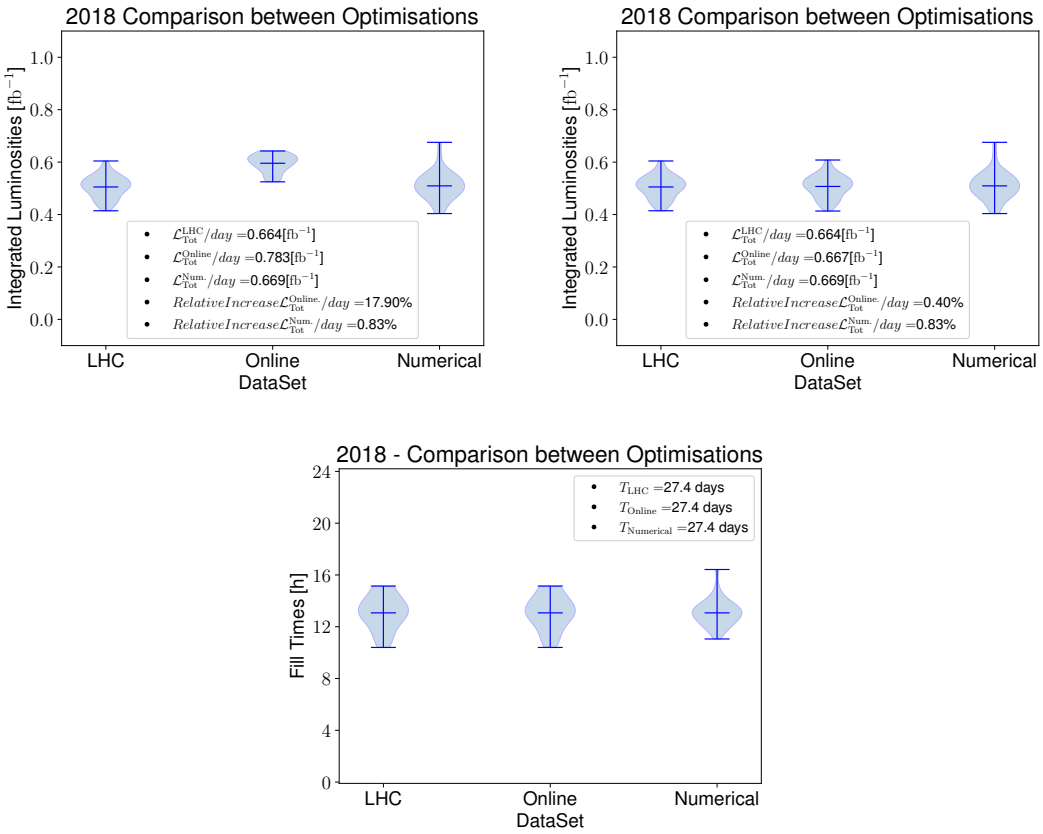


Figure 6.6: On-line optimisation results with complete knowledge of the fit model for 2018 (left). The Initial luminosity was estimated from the LHC Run 2 control parameters provided during the run for luminosity prediction and modelling (right). The initial luminosity is the one recorded at LHC during Run 2 in 2018. The bottom plot represents the fill time during the run and the optimised fills.

The two top figures in Fig. 6.6 show comparative results in terms of integrated daily luminosity. The difference in the gains between the two figures stems from the chosen initial luminosity. In the top left Fig. 6.6 the initial luminosity provided by the control parameters, one can see a significantly more improvement in daily integrated luminosity with a gain in integrated luminosity/day of 17.9%.

This relative increase underscores the impact of the control parameters on luminosity predictions and how crucial it is to adjust and align them with real-time requirements.

However, the top-right figure gives a more conservative estimate by considering the already provided initial luminosity. Although the gain here seems relatively modest, it underlines the importance of the online optimisation strategy, even when starting with a predetermined initial luminosity. Even a marginal increase of 0.40% in integrated luminosity per day can have significant implications in high-energy physics experiments where every collision event counts.

The bottom figure focusses on the overall fill duration, indicating that the sum of optimised fill times by adopting the offline or on-line strategy can achieve the total actual fill lengths during LHC Run 2, without demanding extended operational hours. This efficiency is indicative of the success of the implemented optimisation techniques.

Conclusions

Throughout this work, we have proposed a candidate model based on the concept of dynamic aperture for luminosity optimisation within the Large Hadron Collider (LHC). The mathematical framework of the dynamic aperture offers a comprehensive model for the evolution of luminosity within a colliding machine containing diffusive effects (see Chapter .3 and 4). This approach has an interesting potential for optimising integrated luminosity, which consequently enhances the performance of circular colliders.

The encouraging results obtained from the offline optimisation, followed by the promising results from the online optimisation (see Chapter .5 and 6), not only validate the reliability of the proposed model, but also pave the way for future efforts that could potentially incorporate more sophisticated algorithms to further refine the predictions.

Bibliography

[1] G. Apollinari et al. *High-Luminosity Large Hadron Collider (HL–LHC).* Vol. 4. CERN Yellow Rep. Monogr. Geneva: CERN, 2017. DOI: 10.23731/CYRM-2017-004.

[2] A. Bazzani et al. 'Advances on the modeling of the time evolution of dynamic aperture of hadron circular accelerators'. In: *Phys. Rev. Accel. Beams* 22 (10 Oct. 2019), p. 104003. DOI: 10.1103/PhysRevAccelBeams.22.104003. arXiv: 1909.09516 [physics.acc-ph]. URL: https://link.aps.org/doi/10.1103/PhysRevAccelBeams.22.104003.

[3] M. Benedikt, D. Schulte and F. Zimmermann. 'Optimizing integrated luminosity of future hadron colliders'. In: *Phys. Rev. ST Accel. Beams* 18 (2015), p. 101002.

[4] R. Bruce et al. 'Review of LHC Run 2 Machine Configurations'. In: *9th LHC Operations Evian Workshop*. CERN. Evian Les Bains, France, 2019, pp. 187–197.

[5] O. S. Bruning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1.

[6] J.W. Carter and ATLAS Collaboration. 'Luminosity measurement with the ATLAS experiment at the LHC'. In: *40th International Conference on High Energy physics (ICHEP2020) - Parallel: Operation, Performance and Upgrade of Present Detectors*. Vol. 390. Pre-published on: January 18, 2021. 2021. URL: https://doi.org/10.22323/1.390.0726.

[7] CERN. *ATLAS Experiment*. URL: https://atlas.cern/.

[8] CERN. *Fill Table*. URL: https://lpc.web.cern.ch/annotatedFillTable.html.

[9] CERN. *Massi File Specifications*. URL: https://lpc.web.cern.ch/doc/MassiFileDefinition_v2.html.

[10] The ALICE Collaboration. 'The ALICE experiment at the CERN LHC'. In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[11] The ATLAS Collaboration. 'The ATLAS Experiment at the CERN Large Hadron Collider'. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[12] The CMS Collaboration. 'The CMS experiment at the CERN LHC'. In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[13] The LHCb Collaboration. 'The LHCb Detector at the LHC'. In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[14] M. P. Crouch. 'Luminosity Performance Limitations Due to the Beam-Beam Interaction in the Large Hadron Collider'. PhD thesis. Manchester University, 2017.

[15] H. Damerau et al. *LHC Injectors Upgrade, Technical Design Report*. CERN, Dec. 2014. DOI: 10.17181/CERN.7NHR.6HGC. URL: https://cds.cern.ch/record/1976692.

[16] G Dugan. *USPAS Jan 2002 Accelerator School*. 2002. URL: https://www.classe.cornell.edu/~dugan/USPAS/.

[17] G. Erskine and M. Bassetti. *Closed expression for the electrical field of a two-dimensional Gaussian charge*. Technical Report CERN-ISR-TH-80-06, ISR-TH-80-06. Geneva: CERN, 1980.

[18] Hamish Maclean Ewen. 'Modelling and correction of the non-linear transverse dynamics of the LHC from beam-based measurements'. PhD Thesis. Hertford College, University of Oxford, 2014, p. 268.

[19] G. Faletti. 'Optimisation of LHC Integrated Luminosity'. MA thesis. University of Bologna, 2021.

[20] M. A. Furman. *The hourglass reduction factor for asymmetric colliders*. Tech. rep. SLAC-ABC-81. 1991.

[21] M. Giovannozzi. 'A proposed scaling law for intensity evolution in hadron storage rings based on dynamic aperture variation with time'. In: *Phys. Rev. ST Accel. Beams* 15 (2012), p. 024001. DOI: 10.1103/PhysRevSTAB.15.024001.

[22] M. Giovannozzi and F.F. Van der Veken. 'Description of the luminosity evolution for the CERN LHC including dynamic aperture effects. Part I: the model'. In: *Nucl. Instrum. and Methods in Physics Research Vol. A.* 905 (2018), pp. 171–179. DOI: 10.1016/j.nima.2019.01.072. arXiv: 1806.03058.

[23] M. Giovannozzi and F.F. Van der Veken. 'Description of the luminosity evolution for the CERN LHC including dynamic aperture effects. Part II: application to Run 1 data'. In: *Nucl. Instrum. and Methods in Physics Research Vol. A* 908 (2018), pp. 1–9. DOI: 10.1016/j.nima.2018.08.019. arXiv: 1806.03059 [physics.acc-ph].

[24] M. Giovannozzi and C. Yu. 'Proposal of an Inverse Logarithm Scaling Law for the Luminosity Evolution'. In: *IPAC12 Proceedings*. TUPPC078. 2012, p. 1353.

[25] W. Herr. *Proceedings of the CAS-CERN Accelerator School on Intensity Limitations in Particle Beams*. Vol. 3. CAS-CERN, 2017.

[26] W. Herr and B. Muratori. 'Concept of Luminosity'. In: *CAS - CERN Accelerator School: Intermediate Accelerator Physics*. CERN. 2006, pp. 361–378. DOI: 10.5170/CERN-2006-002.361.

[27] M. Hostettler. 'LHC Luminosity Performance'. PhD thesis. Bern University, 2018.

[28] Ewa Lopienska. *The CERN accelerator complex, layout in 2022*. CERN-GRAPHICS-2022-001. 2022.

[29] Carlo Emilio Montanari. 'Diffusive models and chaos indicators for non-linear betatron motion in circular hadron accelerators'. Dottorato di ricerca in Fisica, 35 Ciclo. PhD thesis. Alma Mater Studiorum Università di Bologna, 2023. URL: https://doi.org/10.48676/unibo/amsdottorato/10811.

[30] B. D. Muratori. *Luminosity and luminous region calculations for the LHC*. Technical report Note 301. LHC Project, 2002.

[31] *Numerical Optimization*. https://github.com/GiuliaFaletti/NumericalOptimization.git. GitHub repository. 2021.

[32] J.L. Sirvent Blasco. 'Beam secondary shower acquisition design for the CERN high accuracy wire scanner'. PhD thesis. Barcelona University, Sept. 2018.

[33] S. Tanaji. 'Luminosity and beam-beam tune shifts with crossing angle and hourglass effects in $e^+ - e^-$ colliders'. In: *arXiv:2208.08615 [physics.acc-ph]* (2022). DOI: 10.48550/arXiv.2208.08615. URL: https://arxiv.org/abs/2208.08615.

[34] E. Todesco and M. Giovannozzi. 'Dynamic aperture estimates and phase-space distortions in nonlinear betatron motion'. In: *Phys. Rev. E* 53 (1996), p. 4067.

[35] R. Wanzenberg. *Nonlinear Motion of a Point Charge in the 3D Space Charge Field of a Gaussian Bunch*. Deutsches Elektronen-Synchrotron DESY, 2010.

[36] J. Wenninger. *Operation and Configuration of the LHC in Run 2.* CERN-ACC-NOTE 2019-0007. CERN, Mar. 2019.

[37] Jorg Wenninger. *Operation and Configuration of the LHC in Run 2.* CERN-ACC-NOTE 2019-0007. Subject category: Accelerators and Storage Rings; Accelerator/Facility, Experiment: CERN LHC. CERN. Geneva. ATS Department, 2019.