Review

Luminosity determination at proton colliders

P. Grafström [a,b], W. Kozanecki [c,*]

[a] *CERN, Geneva, Switzerland*
[b] *Università di Bologna, Bologna, Italy*
[c] *DSM/IRFU/SPP, CEA-Saclay, Gif-sur-Yvette, France*

ARTICLE INFO

ABSTRACT

Luminosity is a key parameter in any particle collider, and its precise determination has proven particularly challenging at hadron colliders. After introducing the concept of luminosity in its multiple incarnations and offering a brief survey of the pp and $p\bar{p}$ colliders built to date, this article outlines the various methods that have been developed for relative-luminosity monitoring, as well as the complementary approaches considered for establishing an absolute luminosity scale. This is followed by a survey, from both a historical and a technical perspective, of luminosity determination at the ISR, the $Sp\bar{p}S$, the Tevatron, RHIC and the LHC. For each of these, we first delineate the interplay between the experimental context, the specificities of the accelerator, and the precision targets suggested by the physics program. We then detail how the different methods were applied to specific experimental environments and how successfully they meet the precision goals.

Contents

---

* Corresponding author.
  E-mail address: witold.kozanecki@cern.ch (W. Kozanecki).

1. Introduction

The word "luminosity" stems from the Latin word "lumen" (light) or "luminosus" (full of light). The term "luminosity" has been used in astronomy since long, referring to the amount of electromagnetic energy an astronomical object radiates per unit time. It was picked up by particle physicists in the late 1950's or early 1960's, in the context of the first e^+e^- collider (AdA) under construction at the Frascati Laboratory, to designate what was then called the "source" factor. This number related the electron–positron annihilation cross-section to the number of annihilation events per unit time; it was inferred from the intensities, geometry and time structure of the circulating beams. It is believed that the appellation "luminosity" was coined by B. Touschek [1], one of the "founding fathers" of the AdA project. The analogy between the accelerator and astronomical definitions is obvious and since then the proportionality factor between the event rate and the cross-section in a particle collider has been called luminosity.

By and large there are two main figures of merit for any collider. The first is the energy available in the center of mass system to produce heavier particles or probe smaller scales; the second is the luminosity, which quantifies the potential of the collider for delivering a statistically significant sample of a given class of events. These two key parameters, energy and luminosity, are to some extent coupled in their potential for discovery. Obviously a high luminosity is of utmost importance in the search for rare events and new phenomena. But in many cases the probability for a certain reaction to occur or a certain particle to be created increases with energy, and in such circumstances a somewhat lower energy can be compensated for by a higher luminosity. This can be understood in the parton picture, where the hadronic cross-section for a given process is the convolution of the parton-level cross-section with the parton distribution function (PDF). Thus the probability to find a parton with an energy above a given threshold can be increased by raising either the beam energy or the luminosity.

This article deals with the determination of the luminosity in proton-(anti)proton collisions at the five hadron colliders that have been constructed to date. These are, in chronological order, the Intersecting Storage Rings (ISR) [2] and the S$p\bar{p}$S [3,4] at CERN, the Tevatron [5] at Fermilab, the Relativistic Heavy Ion Collider (RHIC) [6] at BNL, and the Large Hadron Collider (LHC) [7] at CERN. Luminosity determination in e^+e^- and ep colliders lies beyond the scope of the present report: the methodology is very different there, basically because of the difference in the nature of the corresponding fundamental interactions. In Quantum Electrodynamics (QED), the absolute rate of simple and abundant final states such as $e^+e^- \rightarrow e^+e^-$ at LEP or $ep \rightarrow ep\gamma$ at HERA are calculable in perturbation theory to very high accuracy, providing a conceptually straightforward (if experimentally demanding) path to high-precision, absolute cross-section measurements. In contrast, Quantum Chromodynamics (QCD) calculations of even the simplest final-state topologies in hadron collisions are much less precise, and in many cases the perturbative approach is not even applicable. This fundamental difference between the electromagnetic and the strong interaction deeply influences the luminosity-measurement techniques in the different type of colliders.

This paper is structured as follows. Section 2 presents in a general way the concept of luminosity and its avatars, introduces a few requisite notions of accelerator physics together with the associated observables, and outlines the luminosity-determination methodology. In Section 3, a succinct overview of all past and operating proton colliders highlights the diversity of accelerator configurations considered in this review. The many methods devised over the years to monitor the relative luminosity and to provide absolutely-calibrated values are outlined in Sections 4 and 5 respectively. With all the concepts at hand, we then proceed, in Sections 6–10, to describe the luminosity determination at each of the five colliders considered, detailing in each case how the different methods have been applied to specific experimental environments and how successfully they meet the precision targets suggested by the corresponding physics program.

2. The concept of luminosity

The general definition of the luminosity $\mathcal{L}$ is given by the straightforward formula:

$$\mathcal{L} = R/\sigma \ [\text{cm}^{-2} \ \text{s}^{-1}] \tag{1}$$

where R is the rate of events produced per unit of time for a process with cross-section σ.

An accurate determination of the absolute luminosity is needed for any cross-section measurement, but the required precision may vary considerably from case to case (Section 2.1). Achieving the desired overall performance leads to a refined, multi-prong specification of the precision requirements, each of which is associated with a different avatar of the concept of luminosity (Section 2.2). The necessary basic concepts in accelerator physics are introduced in Section 2.3, together with the corresponding notation and formalism. Section 2.4 outlines the methodology of luminosity determination at proton colliders.

2.1. Precision requirements

A first, obvious criterion is that the experimental accuracy match the uncertainty affecting the theoretical predictions for the physical process under study. Until the turn of the century, theoretical predictions have been significantly less precise at hadron colliders than at e^+e^- colliders, both because of significant uncertainties in the parton distribution functions, and because parton-level cross-sections cannot be calculated with the same precision as QED cross-sections. This fundamental difference between hadron and electron colliders was reflected in the relatively coarser precision expected from, and typically achieved in, the corresponding luminosity measurements. A notable exception is offered by some ISR experiments (Section 6), where the demands of the early physics program [8], combined with a favorable instrumental environment, yielded the most precise luminosity determinations at a hadron machine to date. More recently, substantial progress in predicting accurately electroweak cross-sections in the LHC regime, combined with more sophisticated experimental techniques, have stimulated afresh [9] a focused effort towards high-precision luminosity determinations.

In general, the importance of precision increases when "new physics" are not directly accessible through the on-shell production of new final states, for instance because the characteristic energy scale lies beyond the kinematic reach of the collider. The purported new scale may however still manifest itself through precision measurements at a lower energy, in which case very-high accuracy luminosity measurements may be called for. The classical example of this in hadron physics is the connection to high energies via dispersion relations. Precise measurements, at a given energy, of the total cross-section σ_{tot} and of the ρ-parameter (i.e. of the ratio of the real to the imaginary part of the forward elastic-scattering amplitude) strongly constrain the behavior of σ_{tot} at higher energies. The new scale may also manifest itself through loops of virtual particles involving heavy bosons and fermions, in which case very precise cross-section (and therefore luminosity) determinations at experimentally accessible energies may become important. One can cite examples where the experimental techniques to determine signal rates are mature enough, and where the understanding of acceptances, detector biases, reconstruction efficiencies or background subtraction is at the subpercent level, so that the final precision of the physics measurement is dominated by the luminosity uncertainty. The determination of the inclusive W and Z cross-sections at the Tevatron and at the LHC falls in this category.

The achieved precision of the luminosity determination at hadron colliders typically ranges from one to fifteen percent. The 1% "precision frontier" is not linked to a single, fundamental limitation: it originates from the fact that many different sources of systematic uncertainty contribute in a complex mix. The various limiting factors will be discussed case-by-case in Sections 6–10.

2.2. The luminosity and its avatars

The major deities of the Hindu pantheon are said to manifest themselves under different incarnations, known as *avatars*, depending on the circumstances under which the deity is depicted, revered or invoked. The same applies to luminosity. The *instantaneous luminosity*, denoted here by $\mathcal{L}$ (or sometimes $d\mathcal{L}/dt$) and expressed in $\text{cm}^{-2} \ \text{s}^{-1}$, is that defined by Eq. (1). It reflects the instantaneous performance of the collider, may fluctuate on time scales from tens of nanoseconds to minutes, and typically decays with time constants of hours or even days. The *integrated luminosity*, denoted by $\int \mathcal{L} dt$, refers to the instantaneous luminosity accumulated over a certain time interval and is typically quoted in units of cm^{-2}, or equivalently in inverse barns, microbarns (μb^{-1}), picobarns (pb^{-1}), etc.

A second important distinction is that between *absolute* and *relative* luminosity. The former refers to luminosity expressed on some absolute scale that was determined, at one point in time, through a calibration procedure such as van der Meer scans or elastic pp scattering (Section 5). The second concept is invoked in the context of monitoring relative variations in instantaneous luminosity over some time interval: here long-term stability is the paramount concern, while the absolute-luminosity scale determined at calibration time is in principle irrelevant.

Both absolute and relative luminosity are important to convert a rate measurement for a given physics process to a cross-section determination. The absolute luminosity scale directly affects that of the reported cross-sections, and so it must be calibrated as accurately as possible; the stability and long-term internal consistency of the instantaneous luminosity measurements are essential to evaluate the time integral correctly. Luminosity-based accelerator diagnostics also prove of great importance: real-time instantaneous-luminosity monitoring is crucial for beam tuning and collision optimization; the absolute *specific luminosity*, defined as the luminosity per bunch and per unit bunch intensity, provides information on beam
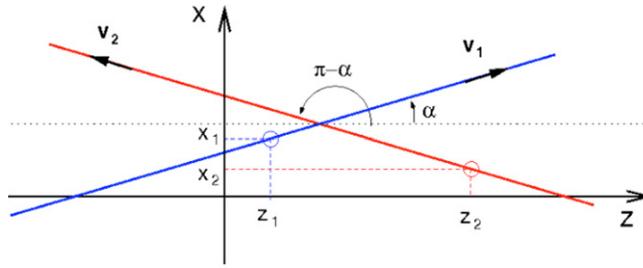
Fig. 1. Definition of coordinate axes and beam directions. The points (x_1, z_1) and (x_2, z_2) are the positions of the bunch centers at time $t = 0$. The y axis is perpendicular to the plane of the figure, such that xyz form a right-handed coordinate system.
Source: Figure reproduced from Ref. [14]$^©$ CERN.

emittances and collider optics that cannot be obtained reliably otherwise; and the long-term evolution of the integrated luminosity directly quantifies, in a global way, the overall performance of the collider.

2.3. Luminosity-related observables and beam parameters

The luminosity in circular colliders can be simply expressed in terms of the collision geometry and of the density distribution of the counter-rotating beams (Section 2.3.1). The notation used to describe the single-beam phase-space at the interaction point (IP) is defined in Section 2.3.2; it is used in Section 2.3.3 to compute both the total luminosity and the parameters of the *luminous region*, i.e. the spatial luminosity-density distribution that is mirrored by, and experimentally accessible using, the three-dimensional distribution of reconstructed collision-event vertices.

Except where specified otherwise, the following assumptions apply throughout this chapter (but will be revisited in Sections 8–10):

- the transverse and longitudinal beam densities are Gaussian;
- linear x–y coupling [10] is small enough to be neglected, both around the collider rings and (when applicable) within the solenoid of the experimental spectrometers;
- IP dispersion and angular dispersion [10] are negligible;
- beam–beam-induced effects [10] are negligible.

More extensive, but still pedagogical introductions to the accelerator issues germane to luminosity determination can be found, together with an extensive bibliography, in (for instance) Refs. [11]–[12].

2.3.1. Bunch luminosity in circular colliders

Storage-ring beam dynamics lead to operating particle colliders in *bunched* mode: each of the two beams is composed of a string of short bunches, typically numbering a few ten to a few thousand, that are focused to the smallest practical transverse dimensions and brought into collision at one or more IP's distributed around the collider ring. The one exception is that of the ISR, that operated with flat, unbunched beams: the luminosity formulas describing this unique case are documented in Ref. [13].

The *bunch luminosity* $\mathcal{L}_b$ produced by one colliding-bunch pair, with time- and position-dependent density functions $\rho_1(x, y, z, t)$ and $\rho_2(x, y, z, t)$, is given by [11,14]:

$$\mathcal{L}_b = f_r\, n_1\, n_2\, K \int \rho_1(x, y, z, t)\, \rho_2(x, y, z, t)\, dx\, dy\, dz\, dt.$$

Here, 1 and 2 refer to beam 1 and beam 2, f_r is the revolution frequency (the time integral extends over a single bunch crossing), n_1 and n_2 are the total numbers of protons in the two colliding bunches, and the particle densities are normalized such that $\int \rho_1(x, y, z, t)\, dx\, dy\, dz = \int \rho_2(x, y, z, t)\, dx\, dy\, dz = 1$ at any time t. The kinematic factor K is given by [15]

$$K = \sqrt{(\vec{v}_1 - \vec{v}_2)^2 - \frac{(\vec{v}_1 \times \vec{v}_2)^2}{c^2}}$$

under the assumption that the particles in bunch 1 (2) are all moving with the same velocity $\vec{v}_1$ ($\vec{v}_2$) in the laboratory frame (more general cases are discussed in Ref. [16]).

The two beams intersect with velocity vectors $\vec{v}_1$ and $\vec{v}_2$ (Fig. 1). The laboratory frame xyz is defined such that the x axis points in the direction of $\vec{v}_1 + \vec{v}_2$, the y axis in that of $\vec{v}_1 \times \vec{v}_2$, and the z axis is parallel to $\vec{v}_1 - \vec{v}_2$. The crossing angle lies in the xz plane, and the half angle α is defined by the two beam trajectories in that plane.

If one assumes $v_1 = v_2 = c$, the bunch luminosity simplifies to

$$\mathcal{L}_b = f_r\, n_1\, n_2\, 2c\, \cos\alpha \int \rho_1(x, y, z, t)\, \rho_2(x, y, z, t)\, dx\, dy\, dz\, dt \tag{2}$$

and the total instantaneous luminosity is given by

$$\mathcal{L} = \sum_b \mathcal{L}_b \tag{3}$$

where the sum runs over all bunch pairs colliding at the IP under consideration.

2.3.2. Single-beam IP parameters

If one assumes only that the beams are Gaussian, the normalized particle density distribution can be written in the general form

$$\rho(x, x', y, y', z, t) = \frac{1}{\sqrt{(2\pi)^5 \det \sigma_T \sigma_z}} \times \exp\left[-\frac{(z \pm ct)^2}{2\sigma_z^2}\right] \times \exp\left[-\frac{1}{2}\vec{\delta r}^T \sigma_T^{-1} \vec{\delta r}\right] \tag{4}$$

with

$$\vec{\delta r} = \begin{pmatrix} \delta x \\ \delta x' \\ \delta y \\ \delta y' \end{pmatrix} \quad \text{and} \quad \sigma_T = \begin{pmatrix} \sigma_x^2 & \sigma_{xx'} & \sigma_{xy} & \sigma_{xy'} \\ \cdots & \sigma_{x'}^2 & \sigma_{x'y} & \sigma_{x'y'} \\ \cdots & \cdots & \sigma_y^2 & \sigma_{yy'} \\ \cdots & \cdots & \cdots & \sigma_{y'}^2 \end{pmatrix}.$$

Here x, y, z, x', y' are the spatial coordinates and angles of a beam particle at time t, the vector $\vec{\delta r}$ is the deviation of the particle trajectory from the closed orbit, σ_z is the bunch length and the $\pm$ sign in Eq. (4) reflects the fact that the two beams travel in opposite directions. The beam matrix σ_T describes the transverse phase space occupied by the beam particles up to linear-correlation terms. In the absence of transverse coupling, it simplifies to two diagonal 2×2 sub-matrices:

$$\sigma_T = \begin{pmatrix} \Lambda_x & 0 \\ 0 & \Lambda_y \end{pmatrix}.$$

These are traditionally parameterized in terms of the transverse geometrical emittance ϵ and Twiss parameters α, β [10,12]:

$$\Lambda_i = \epsilon_i \begin{pmatrix} \beta_i & -\alpha_i \\ -\alpha_i & (1+\alpha_i^2)/\beta_i \end{pmatrix} \quad (i = x, y),$$

$$\alpha_i = -\frac{1}{2}\frac{\delta\beta_i}{\delta s}$$

where s is the curvilinear coordinate along the central stored-beam orbit.

The parameters determining the transverse density distributions at the IP are, for each beam B, the geometrical emittances ϵ_{iB} and IP β-functions β_{iB}^*. The individual transverse beam sizes σ_{iB} ($i = x, y$ and $B = 1, 2$) are given by:

$$\sigma_{iB}(z) = \sqrt{\epsilon_{iB}\beta_{iB}(z)}.$$

Those are usually not directly measurable at the IP, but can be inferred from beam-profile measurements elsewhere in the rings (for instance using wire scanners), provided the lattice functions are known with sufficient accuracy. They can also be determined directly, in some cases, using beam-imaging techniques, as demonstrated at PEP-II [17] and at the LHC (Section 10).

Between the optical waist location z_{iB}^w (nominally at the IP) and the first focusing element, and neglecting the experimental solenoid (if any), the betatron function obeys

$$\beta_{iB}(z) = \beta_{iB}^* + \frac{(z - z_{iB}^w)^2}{\beta_{iB}^*}.$$

Physically, this expresses the fact that in the vicinity of the IP, particle trajectories are straight lines, and that the IP angular spread σ_{iB}' induces a longitudinal dependence of the transverse beam size:

$$\sigma_{iB}^2(z) = \sigma_{iB}^{*2} + \sigma_{iB}'^2(z - z_{iB}^w)^2$$

where σ_{iB}^* is the RMS transverse beam size at the IP, and

$$\sigma_{iB}' = \sqrt{\epsilon_{iB}/\beta_{iB}^*} \tag{5}$$

is the RMS angular spread. Equivalently,

$$\sigma_{iB}^2(z) = \epsilon_{iB}\beta_{iB}^*\left(1 + \frac{(z - z_{iB}^w)^2}{\beta_{iB}^{*2}}\right). \tag{6}$$

This *hourglass effect* [10,18] is significant only when β_{iB}^* is smaller than or comparable to the bunch lengths. It had a noticeable impact on the luminosity at the S$p\bar{p}$S, the Tevatron and RHIC, and must be taken into account when calibrating the RHIC luminosity [12,19]. At the LHC, it degrades the luminosity by less than 1% under nominal conditions, and has a negligible impact on the absolute luminosity calibration [20].

2.3.3. The luminous region

Under the assumptions listed at the beginning of Section 2.3, and neglecting the hourglass effect for simplicity, the normalized particle-density distribution in a bunch becomes, after integration over the angular variables (x'_B, y'_B):

$$\rho_B(x, y, z, t) = \frac{1}{\sqrt{(2\pi)^3}\, \sigma_{xB}\, \sigma_{yB}\, \sigma_{zB}} \times \exp\left[-\frac{(x - x_B)^2}{2\sigma_{xB}^2} - \frac{(y - y_B)^2}{2\sigma_{yB}^2} - \frac{(z - ct)^2}{2\sigma_{zB}^2} \right] \tag{7}$$

where $B = 1$ and $B = 2$ are associated with beam 1 and beam 2 respectively, σ_{jB} $(j = x, y, z)$ are the transverse and longitudinal stored-beam sizes, and x_B and y_B are the transverse positions of the bunch centroids at the nominal collision point $(t = 0)$. The three-dimensional spatial luminosity density $\mathcal{L}(x, y, z)$, also known as the *luminous region* or *luminous ellipsoid*, is determined by the time-integrated product of the overlapping particle densities of the two colliding bunches.

2.3.3.1. Head-on Collisions.
For head-on collisions (no relative transverse offsets, negligible crossing angles), combining Eqs. (2) and (7) yields

$$\mathcal{L}_b = f_r\, n_1\, n_2\, 2c \int \rho_1\, \rho_2\, dx\, dy\, dz\, dt = \frac{f_r\, n_1 n_2}{2\pi\, \Sigma_x\, \Sigma_y} \tag{8}$$

where

$$\Sigma_j = \sqrt{\sigma_{j1}^2 + \sigma_{j2}^2} \quad (j = x, y, z)$$

are the *convolved beam sizes*.

If the transverse beam sizes are pair-wise equal ($\sigma_{x1} = \sigma_{x2}$, $\sigma_{y1} = \sigma_{y2}$), Eq. (8) simplifies to

$$\mathcal{L}_b = \frac{f_r\, n_1 n_2}{4\pi\, \sigma_x\, \sigma_y}. \tag{9}$$

If in addition $\epsilon_x = \epsilon_y$ and $\beta_x^* = \beta_y^*$, then the beams are round ($\sigma_x = \sigma_y$), and

$$\mathcal{L}_b = \frac{f_r\, n_1 n_2}{4\pi\, \sigma^2} = \frac{f_r\, n_1 n_2\, \gamma}{4\pi\, \epsilon_N\, \beta^*} \tag{10}$$

where $\epsilon_N = \gamma\beta\epsilon$ is called the *normalized emittance*, γ is the relativistic Lorentz factor and $\beta = \sqrt{1 - 1/\gamma^2}$ is the mean velocity of a beam particle in units of the speed of light.

In general, luminosity-weighted observables can be calculated by taking the appropriate moment of the product of bunch particle densities. For example, the horizontal *luminous centroid* is given by

$$\bar{x}_{\mathcal{L}}(z) = \frac{\int x \rho_1 \rho_2}{\int \rho_1 \rho_2}$$

where the integral runs over x, y and t. Similarly, the horizontal (or x–z) *luminous tilt* is defined as $\overline{x'}_{\mathcal{L}} = \frac{\delta \bar{x}_{\mathcal{L}}}{\delta z}$. Equivalent expressions hold for the vertical centroid and luminous tilt, and for the longitudinal centroid: these parameters define the position and orientation of the luminous ellipsoid.

The longitudinal luminosity density is given by

$$\frac{d\mathcal{L}_b}{dz} = f_r\, n_1\, n_2\, 2c \int \rho_1\, \rho_2\, dx\, dy\, dt = \frac{2 f_r\, n_1\, n_2}{\sqrt{(2\pi)^3}\, \Sigma_z\, \Sigma_x\, \Sigma_y} \exp\left(-\frac{(z - z_c)^2}{2\,(\Sigma_z/2)^2} \right) \tag{11}$$

where z_c is the longitudinal location at which the bunches collide.

The longitudinal convolved beam size Σ_z can be extracted (together with the IP β-function if the hourglass effect is strong enough) from the longitudinal luminosity density [17]. The transverse convolved beam sizes Σ_x, Σ_y can be measured by beam-separation scans, as will be detailed in Section 5.2.

A related – albeit distinct – measure of transverse phase space is supplied by the horizontal *luminous size* $\sigma_{x\mathcal{L}}$, defined by

$$(\sigma_{x\mathcal{L}})^2 = \overline{x^2}_{\mathcal{L}}(z) = \frac{\int x^2 \rho_1 \rho_2}{\int \rho_1 \rho_2},$$

which is related to the stored-beam sizes by

$$\sigma_{x\mathcal{L}} = \left(\frac{1}{\sigma_{x1}^2} + \frac{1}{\sigma_{x2}^2} \right)^{-1/2}, \tag{12}$$

with equivalent expressions for $\sigma_{y,\mathcal{L}}$. These two parameters describe the transverse shape of the luminous ellipsoid; they are directly measurable, and carry information about β-functions and emittances. In the limiting case where the transverse sizes of beam 1 and beam 2 are equal pair-wise, one recovers the familiar expressions:

$$\sigma_{i\mathcal{L}} = \sigma_{iB}/\sqrt{2} = \Sigma_i/2.$$

The *luminous length* $\sigma_{z\mathcal{L}}$ is similarly defined as

$$(\sigma_{z\mathcal{L}})^2 = \overline{z^2}_{\mathcal{L}}(z) = \frac{\int z^2 \rho_1 \rho_2}{\int \rho_1 \rho_2}.$$

It is directly related to the convolved longitudinal beam size (Eq. (11)):

$$\sigma_{z\mathcal{L}} \approx \Sigma_z/\sqrt{2}$$

up to hourglass corrections arising from the z-dependence of Σ_x and Σ_y [18].

The *specific luminosity* $\mathcal{L}_{spec}$, defined as the luminosity per bunch and per unit bunch population, simply equals (up to a known constant) the inverse product of the horizontal and vertical convolved beam sizes (Eq. (8)). It thereby constitutes a valuable, non-invasive (and sometimes the only available) monitor of the evolution of the average beam emittance during physics running.

2.3.3.2. Collisions with Non-zero Crossing Angle. For beams colliding with a half crossing-angle α in the xz plane, and with a relative transverse offset Δ_x (Δ_y) in the x (y) direction, but still neglecting the hourglass effect, Eq. (8) becomes [14]

$$\mathcal{L}_b = f_r\, n_1 n_2\, \cos\alpha\, \frac{e^{-\frac{\Delta_x^2}{2\Sigma_x^2} - \frac{\Delta_y^2}{2\Sigma_y^2}}}{2\pi\, \Sigma_x\, \Sigma_y} \tag{13}$$

where the definition of the convolved beam sizes has been generalized as follows.

Denoting by $\sigma_{\hat{x}_B}^2$, $\sigma_{\hat{y}_B}^2$, and $\sigma_{\hat{z}_B}^2$ the RMS beam sizes in the beam frames which have the $\hat{z}$ axis in the direction of motion of beam B, $\hat{y} = y$ and $\hat{x}$ such that it forms a right-handed system, the convolved beam size in the crossing plane:

$$\Sigma_x = \sqrt{(\sigma_{\hat{x}1}{}^2 + \sigma_{\hat{x}2}{}^2)\cos^2\alpha + (\sigma_{\hat{z}1}{}^2 + \sigma_{\hat{z}2}{}^2)\sin^2\alpha} \tag{14}$$

receives an additional contribution from the longitudinal spread of the beams. The other two convolved beam sizes remain unaffected:

$$\Sigma_y = \sqrt{\sigma_{\hat{y}1}{}^2 + \sigma_{\hat{y}2}{}^2}$$
$$\Sigma_z = \sqrt{\sigma_{\hat{z}1}{}^2 + \sigma_{\hat{z}2}{}^2}.$$

Eq. (13) describes the variation of the luminosity as a function of the beam separation (Δ_x, Δ_y) in the transverse plane, such as measured for instance during *van der Meer scans* (Section 5.2). The physical interpretation of the transverse convolved beam size now becomes apparent: for strictly Gaussian beams, such a luminosity-scan curve is a Gaussian function of the horizontal (vertical) beam separation Δ_x (Δ_y), and its standard deviation is given by Σ_x (Σ_y).

At zero beam separation ($\Delta_x = \Delta_y = 0$), the bunch luminosity becomes

$$\mathcal{L}_b = \frac{f_r\, n_1 n_2\, \cos\alpha}{2\pi\, \Sigma_x\, \Sigma_y}. \tag{15}$$

Compared to the zero crossing-angle case, it is reduced by the geometric factor

$$F = \cos\alpha/\sqrt{1 + \tan^2\alpha(\sigma_{\hat{z}1}{}^2 + \sigma_{\hat{z}2}{}^2)/(\sigma_{\hat{x}1}{}^2 + \sigma_{\hat{x}2}{}^2)} \tag{16}$$

that reflects the effective increase in Σ_x. This luminosity degradation can become significant (10%–40%), even for crossing angles well below a milliradian, because of the typically large longitudinal/transverse beam-size ratio.

The fully general case of bunches colliding with a relative transverse offset, in the presence of crossing angles in both planes and with non-negligible hourglass effect is treated in Refs. [11,12].

2.4. Methodology of luminosity determination at proton colliders

The determination of the relative luminosity amounts to measuring, precisely and reproducibly but on an arbitrary scale, the event rate R in Eq. (1). A wide variety of experimental techniques has been considered, that depend on the accelerator and detector context: unbunched or bunched beams, center of mass (c.m.) energy, typical collision rates, hostility of the radiation environment, etc. Even within a given particle-physics experiment, the optimal approach can be time- and beam-conditions-dependent. Redundancy – the use of several independent, mutually complementary techniques – is of the essence in order to achieve percent-level accuracy. An overview of the strategies and algorithms that have been used to monitor the relative luminosity will be presented in Section 4.

The luminosity scale itself is independent of the physics process used to measure the rate R: it depends solely on the properties of the colliding beams. This suggests resorting to those beam parameters that are measurable precisely enough to calibrate the absolute luminosity. Alternatively, one can adopt the opposite approach and determine the luminosity completely independently of the accelerator parameters by using reactions for which the cross-sections (and the corresponding detector efficiencies) are well known. A third approach relies on exploiting the optical theorem together with the measurement of the differential (anti)proton–proton elastic cross-section. These diverse absolute-calibration strategies, and their interplay, form the subject of Section 5. Their practical application, the associated instrumental choices and the determination of the resulting systematic uncertainties will be detailed in the context of the corresponding collider (Sections 6–10).

3. Overview of proton colliders

Table 1 offers, for the five proton-(anti)proton colliders built to date, a synopsis of those accelerator parameters that are germane to luminosity determination. The data is taken from the compilations by the Particle Data Group [21,22]; as that review no longer lists the ISR parameters, the latter are excerpted from Ref. [23]. RHIC and LHC operate both as pp and as ion–ion colliders; their parameters in ion–ion and p-ion mode can be found in Ref. [22].

From the startup of the ISR to the end of the first LHC run in 2013, the c.m. energy rose over a hundred-fold, and the peak instantaneous luminosity grew by a comparable factor. This spectacular improvement in total luminosity is in large part due to the increase in beam energy (Eq. (10)). It is also closely tied to two paradigm changes made possible by advances in accelerator technology and in the mastery of beam dynamics. The first was the switch from continuous, unbunched beams at the ISR to a few widely spaced, effectively isolated bunches at the S$p\bar{p}$S. The second was the transition to multibunch operation from late S$p\bar{p}$S operation onwards, that required increasingly sophisticated beam-separation schemes as well as bunch-by-bunch stabilization feedbacks, and that culminated in two-ring colliders with close to 3000 bunches in the LHC. Also crucial was, as each collider matured, the steady push towards lower β^*, somewhat tempered by the associated issues of dynamic aperture and, more recently, of collimation of high-power beams.

The development of accelerators has been paralleled [24] by a major evolution in the complexity, sophistication and scale of the experimental detectors, accompanied by a major consolidation of resources into a small number of large experimental groups. Over the lifetime of the ISR, about 50 different experiments took turns in the accelerator tunnel, compared to four large ones at the LHC so far.

Although detector technologies underwent enormous progress in a global sense, luminosity instrumentation remained pretty much the same from one pp collider to the next. The forward scintillator hodoscopes with photomultiplier readout introduced at the ISR remained the technology of choice at the S$p\bar{p}$S, RHIC, the Tevatron (at least during run 1) and even the ALICE experiment at the LHC. It is only with the advent of multiple pp collisions per bunch crossing and of intense collision-induced radiation at the Tevatron and at the LHC that alternative counting technologies had to be developed.

4. Relative-luminosity monitoring methods

4.1. Methodology

The bunch luminosity of a proton collider can be expressed as

$$\mathcal{L}_b = \frac{R_{ref}}{\sigma_{ref}}$$

where R_{ref} is the rate of a reference collision process and σ_{ref} is the corresponding cross-section. The reference process can in principle be chosen arbitrarily; selecting inelastic pp collisions as a typical one, the above equation can be rewritten as

$$\mathcal{L}_b = \frac{\mu f_r}{\sigma_{inel}} \tag{17}$$

where μ is the average number of inelastic interactions per bunch crossing (BC), also called *pile-up parameter*, f_r is the bunch revolution frequency, and σ_{inel} is the inelastic pp cross-section. Thus, the instantaneous luminosity can be determined using any method that measures the ratio μ/σ_{inel}.

Table 1

Typical IP parameters of proton colliders. The LHC parameters listed here reflect the performance as of February 2013; the design parameters can be found in Table 8.

	ISR (pp)	S$p\bar{p}$S ($p\bar{p}$)	Tevatron ($p\bar{p}$)	RHIC (pp)	LHC (pp)
Physics start date	1971	1981	1987	2000	2009
Physics end date	1983	1991	2011		
Maximum beam energy (TeV)	0.031	0.315	0.980	0.255	4.0
Circumference (km)	0.943	6.9	6.3	3.8	26.7
Interaction regions	8	2	2	6	4
Geometrical emittance $\epsilon_{x,y}$ (π nm-rad)	200 (x)	p: 9	p: 3	15	0.6
	70 (y)	$\bar{p}$: 5	$\bar{p}$: 1		
β^* (m)	4.1 (x)	0.60 (x)	0.28	0.65	0.6
	0.31 (y)	0.15 (y)			
RMS transverse beam size $\sigma_{x,y}$ (μm)	900 (x)	p: 73 (x), 36 (y)	p: 28	90	18.8
	150 (y)	$\bar{p}$: 55 (x), 27 (y)	$\bar{p}$: 16		
RMS bunch length σ_z (cm)	N/A	20	p: 50	60	9.4
			$\bar{p}$: 45		
Full crossing angle 2α	14.8°	0	0	0	290 μrad
Particles/bunch (10^{10})	N/A	p: 15	p: 26	18.5	16
		$\bar{p}$: 8	$\bar{p}$: 9		
Average beam current	35000	p: 6	p: 70	257	400
per species (mA)		$\bar{p}$: 3	$\bar{p}$: 24		
Luminosity per bunch (10^{30} cm^{-2} s^{-1})	N/A	1.0	12	1.9	5.6
Pile-up μ (interactions/crossing)	N/A	1.3	15.3	1.3	37
Bunches per ring per species	unbunched	6	36	111	1380
Total luminosity (10^{30} cm^{-2} s^{-1})	140	6	431	215	7700
Time between bunch crossings (μs)	N/A	3.8	0.396	0.107	0.050
Typical luminosity decay time (h)	$\approx$100	15	6 (avg)	5.5	$\approx$6

Adopting the methodology and the notation of Refs. [25,26], techniques for luminosity determination can be classified as follows.

- *Event Counting:* here one determines the fraction of bunch crossings during which a specified detector registers an "event" satisfying a given selection requirement. For instance, a bunch crossing can be said to contain an "event" if at least one pp interaction in that crossing induces at least one observed hit in the detector being considered.
- *Hit Counting:* here one counts the number of hits (for example the number of electronic channels or energy clusters above a specified threshold) per bunch crossing in a given detector.
- *Particle Counting:* here one determines the distribution of the number of particles per beam crossing (or its mean) inferred from reconstructed quantities such as calorimeter-energy distributions, or from other observables that reflect the instantaneous particle flux traversing the detector (e.g. the total ionization current drawn by a liquid-argon calorimeter sector).

The simplest approach is that of event counting. Eq. (17) can be rewritten as:

$$\mathcal{L}_b = \frac{\mu f_r}{\sigma_{inel}} = \frac{\mu_{vis} f_r}{\varepsilon \sigma_{inel}} = \frac{\mu_{vis} f_r}{\sigma_{vis}} \tag{18}$$

where ε is the efficiency for one inelastic pp collision to satisfy the event-selection criteria, and $\mu_{vis} \equiv \varepsilon\mu$ is the average number of visible inelastic interactions per BC (i.e. the mean number of pp collisions per BC that pass that "event" selection). The visible cross-section $\sigma_{vis} \equiv \varepsilon\sigma_{inel}$ is the calibration constant that relates the measurable quantity μ_{vis} to the absolute bunch luminosity $\mathcal{L}_b$. Both ε and σ_{vis} depend on the colliding-particle species (pp or $p\bar{p}$), the c.m. energy, the pseudorapidity and transverse-momentum distributions as well as the particle composition of the collision products: they are therefore different for each luminometer and luminosity algorithm.

The measurement of the visible interaction rate μ_{vis}, which is proportional to the luminosity up to an overall scale factor, is described in the present chapter, from the low-rate limit at the ISR up to pileup parameters of several ten interactions per bunch crossing typical of LHC operation. The calibration of the absolute-luminosity scale, i.e. the precision determination of the visible cross-section σ_{vis}, is the subject of Section 5.

4.2. Interaction-rate determination

Most luminometers consist of two symmetric detector elements, placed on either side of the IP in the forward ("F") and backward ("B") arms of the 4π experimental detectors. Each side is normally further segmented into a discrete number of readout segments, typically arranged azimuthally around the beam pipe, each with a separate readout channel. For event-

and hit-counting algorithms, a threshold is applied to the analog signal output from each readout channel, and every channel with a response above this threshold is counted as containing a "hit".

4.2.1. Relative-rate monitoring

When the probability for more than one interaction to occur within a bunch crossing is low enough, the relative luminosity can be monitored simply using an arbitrary physics process with sufficiently high rate and low enough background:

$$\mathcal{L}_b = \frac{R_{vis}}{\sigma_{vis}}.$$

This condition was satisfied at the ISR[1] and in the early years at the S$p\bar{p}$S, and corresponds to the low-rate limit of the event-counting methods ($\mu_{vis} \ll 1$). In this case, the average number of visible inelastic interactions per beam crossing is given by the intuitive expression

$$\mu_{vis} \approx \frac{N}{N_{BC}}$$

where N is the (background-subtracted) number of events passing the selection criteria that are observed during a given time interval, and N_{BC} is the number of bunch crossings in that same interval.

When μ increases, the probability that two or more pp interactions occur in the same bunch crossing is no longer negligible, and μ_{vis} is no longer linearly related to the raw event count N. Instead μ_{vis} must be calculated taking into account Poisson statistics and in some cases instrumental or pile-up-related effects, as discussed below.

4.2.2. Event-counting methods

The value of μ_{vis} used to determine the bunch luminosity $\mathcal{L}_b$ in bunch pair b is obtained from the raw number of counts N and the number of bunch crossings N_{BC}, using an algorithm-dependent expression and assuming that:

- the number of pp interactions occurring in any bunch crossing obeys a Poisson distribution. This assumption drives the combinatorial formalism presented in Sections 4.2.2.1 and 4.2.2.2;
- the efficiency to detect a single inelastic pp interaction is constant, in the sense that it does not change when several interactions occur in the same bunch crossing. This is tantamount to assuming that the efficiency ε_n for detecting one event associated with n interactions occurring in the same crossing is given by

$$\varepsilon_n = 1 - (1 - \varepsilon_1)^n \tag{19}$$

where ε_1 is the detection efficiency corresponding to a single inelastic interaction in a bunch crossing (the same definition applies to the efficiencies ε^{OR}, ε^F, ε^B and ε^{AND} defined below). The validity of this assumption may be affected by subtle detector effects (Section 4.2.4); it must be validated, on a case by case basis, by evaluating the mutual consistency of luminosity measurements performed simultaneously using multiple luminosity algorithms with an intrinsically different μ-response.

The absolute bunch luminosity is then given directly and without additional assumptions[2] by

$$\mathcal{L}_b = \frac{\mu_{vis} f_r}{\sigma_{vis}}$$

using the value of σ_{vis} measured, for instance, by the van der Meer method.

4.2.2.1. Inclusive-OR algorithms. In an Event_OR algorithm, a bunch crossing is counted as containing an "event" if there is at least one hit on either the F or the B side of the detector. Since the Poisson probability for observing zero events in a given bunch crossing is $P_0(\mu_{vis}) = e^{-\mu_{vis}} = e^{-\mu \varepsilon^{OR}}$, the probability of observing at least one event is

$$P_{\text{Event\_OR}}(\mu_{vis}) = \frac{N_{OR}}{N_{BC}} = 1 - P_0(\mu_{vis}) = 1 - e^{-\mu_{vis}}. \tag{20}$$

Here the raw event count N_{OR} is the number of bunch crossings, during a given time, in which at least one pp interaction satisfies the event-selection criteria of the OR algorithm under consideration, and N_{BC} is the total number of bunch crossings during the same interval. Eq. (20) reduces to the intuitive result $P_{\text{Event\_OR}}(\mu_{vis}) \approx \mu_{vis}$ when $\mu_{vis} \ll 1$. Solving for μ_{vis} in terms of the event-counting rate yields:

$$\mu_{vis} = -\ln\left(1 - \frac{N_{OR}}{N_{BC}}\right). \tag{21}$$

---

[1] Because the ISR operated in unbunched mode, the relevant time interval was the sensitivity window of the luminometers (typically a few ten ns), rather than the bunch crossing.

[2] In contrast, providing a value for $\mu \equiv \mu_{vis}/\varepsilon = \mu_{vis}\sigma_{inel}/\sigma_{vis}$ requires the knowledge of (or an assumption on) the total inelastic cross-section.

4.2.2.2. Coincidence algorithms. For the Event_AND case, the relationship between μ_{vis} and N is more complicated. Here a bunch crossing is counted if there is at least one hit in each of the F and B detector arms. This coincidence condition can be satisfied either from a single *pp* interaction or from individual hits on either side of the detector from different *pp* interactions in the same bunch crossing. Therefore the event-counting probability no longer depends on a single efficiency: it must be written in terms of ε^F, ε^B and ε^{AND}, the efficiencies for observing an event with, respectively, at least one hit on the F-side, at least one hit on the B-side and at least one hit on both sides simultaneously. These efficiencies are related to the Event_OR efficiency by $\varepsilon^{OR} = \varepsilon^F + \varepsilon^B - \varepsilon^{AND}$.

The probability $P_{\text{Event\_AND}}(\mu)$ of there being at least one hit on both sides is one minus the probability $P_0^{\text{Zero\_OR}}$ of there being no hit on at least one side. The latter, in turn, equals the probability that there be no hit on at least side F ($P_{0F} = e^{-\mu\varepsilon^F}$), plus the probability that there be no hit on at least side B ($P_{0B} = e^{-\mu\varepsilon^B}$), minus the probability that there be no hit on either side ($P_0 = e^{-\mu\varepsilon^{OR}}$):

$$\begin{aligned} P_{\text{Event\_AND}}(\mu) &= \frac{N_{AND}}{N_{BC}} = 1 - P_0^{\text{Zero\_OR}}(\mu) \\ &= 1 - (e^{-\mu\varepsilon^F} + e^{-\mu\varepsilon^B} - e^{-\mu\varepsilon^{OR}}) \\ &= 1 - (e^{-\mu\varepsilon^F} + e^{-\mu\varepsilon^B} - e^{-\mu(\varepsilon^F + \varepsilon^B - \varepsilon^{AND})}) . \end{aligned} \tag{22}$$

This equation cannot be inverted analytically, and the best approach depends on the values of ε^F, ε^B and ε^{AND}.

If the layouts, geometries and efficiencies of the forward and backward luminometers are sufficiently similar, the above equation can be simplified under the assumption that $\epsilon_F \approx \epsilon_B$. The efficiencies ε^{AND} and ε^{OR} are *defined* as, respectively, $\varepsilon^{AND} \equiv \sigma_{vis}^{AND}/\sigma_{inel}$ and $\varepsilon^{OR} \equiv \sigma_{vis}^{OR}/\sigma_{inel}$; the average number of visible inelastic interactions per BC is computed as $\mu_{vis} \equiv \varepsilon^{AND}\mu$. Eq. (22) then becomes

$$\frac{N_{AND}}{N_{BC}} = 1 - 2e^{-\mu(\varepsilon^{AND} + \varepsilon^{OR})/2} + e^{-\mu\varepsilon^{OR}} = 1 - 2e^{-(1 + \sigma_{vis}^{OR}/\sigma_{vis}^{AND})\mu_{vis}/2} + e^{-(\sigma_{vis}^{OR}/\sigma_{vis}^{AND})\mu_{vis}} . \tag{23}$$

The value of μ_{vis} is then obtained by solving Eq. (23) numerically using the values of σ_{vis}^{OR} and σ_{vis}^{AND} extracted (for instance) from van der Meer scans.

If the efficiency is high and $\varepsilon^{AND} \approx \varepsilon^F \approx \varepsilon^B$, as is typically the case for large-acceptance scintillator hodoscopes, Eq. (22) can be approximated by

$$\mu_{vis} \approx -\ln\left(1 - \frac{N_{AND}}{N_{BC}}\right) .$$

The μ-dependence of the probability function $P_{\text{Event\_AND}}$ is controlled by the relative magnitudes of ε^F, ε^B and ε^{AND} (or of the corresponding measured visible cross-sections). This is in contrast to the Event_OR case, where the efficiency ε_{OR} factors out of Eq. (21).

4.2.3. Hit-counting methods

When $\mu_{vis} \gg 1$, event-counting algorithms lose sensitivity as fewer and fewer bunch crossings in a given time interval report zero observed interactions: this is known as *saturation* or *zero starvation*. In the limit where $N/N_{BC} = 1$, it is no longer possible to use event counting to determine the interaction rate μ_{vis}, and more sophisticated techniques must be brought to bear. One example is *hit counting*, where the number of hits in a given detector is counted rather than just the total number of events. This provides more information about the interaction rate per event, and pushes up the luminosity at which the algorithm saturates.

Under the assumption that the number of hits in one *pp* interaction follows a binomial distribution and that the number of interactions per bunch crossing follows a Poisson distribution, one can calculate the average probability to have a hit per bunch crossing in one of the detector channels as

$$P_{\text{HIT}}(\mu_{vis}^{\text{HIT}}) = \frac{N_{\text{HIT}}}{N_{BC}N_{CH}} = 1 - e^{-\mu_{vis}^{\text{HIT}}}, \tag{24}$$

where N_{HIT} and N_{BC} are the total numbers of hits and bunch crossings during a time interval, and N_{CH} is the number of detector channels [26]. The expression above allows μ_{vis}^{HIT} to be calculated from the number of hits as

$$\mu_{vis}^{\text{HIT}} = -\ln\left(1 - \frac{N_{\text{HIT}}}{N_{BC}N_{CH}}\right) \tag{25}$$

from which the bunch luminosity can be inferred in the same way as for event-counting algorithms:

$$\mathcal{L}_b = \frac{\mu_{vis}^{\text{HIT}}f_r}{\sigma_{vis}^{\text{HIT}}} .$$

Hit-counting algorithms are typically more sensitive than event-counting methods to instrumental imperfections such as threshold effects, instrumental noise, channel-to-channel efficiency variations, long-term gain drifts or cross-talk. In particular, the binomial assumption used to derive Eq. (25) is only true if the probability to observe a hit in a single channel is independent of the number of hits observed in the other channels. Even so, such algorithms (e.g. pixel-cluster counting in the main tracker) have been used successfully in the hostile experimental environment of the LHC, where very high values of the pile-up parameter μ render event counting impractical for large-acceptance luminometers.

4.2.4. Pile-up-related instrumental issues

The intrinsic non-linearity of the counting methods, which is a direct consequence of Poisson statistics, has fundamental implications for luminometers meant to operate in a regime where the probability of detecting more than one inelastic *pp* interaction per bunch crossing cannot be neglected.

Firstly, and fundamentally because of the lack of synchrotron-radiation damping, bunch-to-bunch intensity and emittance fluctuations of 10%–20% are not uncommon in proton colliders. The resulting fluctuations in bunch luminosity $\mathcal{L}_b$, coupled with the non-linearities apparent in Eqs. (21), (22) and (25), are often violent enough that bunch-averaged luminosity measurements become impractical. For large enough values of μ_{vis}, a naive procedure that averages the event probabilities ($P_{\mathrm{Event\_OR}}$, $P_{\mathrm{Event\_AND}}$, P_{HIT} ...) over all colliding bunch pairs and only then applies the Poisson combinatorial formalism above, can bias the measurement by tens of percent compared to the correct procedure of computing first the bunch luminosity $\mathcal{L}_b$ for each bunch separately, and then summing over all colliding bunch pairs. It is therefore essential, at least in an LHC-like environment, that the front-end electronics and dedicated data-acquisition (DAQ) systems associated with luminometers be designed from the start to cope with bunch-by-bunch luminosity determination.

Next, the applicability of the Poisson formalism depends critically on the validity of the assumption, summarized by Eq. (19), that the efficiency for detecting an inelastic *pp* interaction is independent of the number of interactions that occur in each crossing. This requires in particular that the threshold for registering a hit in an individual readout channel be low enough compared to the average single-particle response. When the threshold is too high, a particle from a single *pp* interaction will occasionally fail to be detected, while two such particles from different *pp* interactions in the same bunch crossing traversing the same detector channel, may produce large enough an analog output to register a hit. This effect, called *migration* [25], intrinsically affects hit counting more severely than event counting. Depending on the detailed response of the luminometer considered (e.g. the hit threshold) it can cause the luminosity to be overestimated at high pile-up parameters. It then manifests itself by a μ-dependence of the luminosity ratios between different luminometers, as reported e.g. in Refs. [27,28].

4.2.5. Tracker-based algorithms

The advent of large-acceptance, high-precision silicon trackers at the LHC has enabled the development of additional algorithms that are based on reconstructing charged-particle tracks and from these, *pp* interaction vertices. In all cases, the full readout of the tracking detectors is initiated by a low-level trigger that either randomly selects beam crossings of filled bunch pairs where collisions are possible, or applies a loose selection requirement based on scintillator hodoscopes.

4.2.5.1. Track algorithms. *Track-based event counting* consists in measuring the fraction of bunch crossings with a minimum number of tracks (typically 1 or 2) reconstructed in the silicon trackers. This inclusive event-counting algorithm is the baseline method adopted by the LHCb experiment [29], which operates at low enough pile-up parameters to remain insensitive to potential μ-dependent tracking-efficiency variations. This method was also used to compare the luminosity delivered in 2010 to the ATLAS and CMS experiments (again at moderate pile-up parameters), using track-selection criteria that were accessible to both detectors [25,30].

Track counting is conceptually similar to hit counting and is in some respects more robust, e.g. for what regards background subtraction. It amounts to counting the number of well-reconstructed tracks per bunch crossing, with the track quality ensured by tight selection criteria. At high pile-up parameters, this method appears better behaved than the vertex-counting algorithms outlined below.

4.2.5.2. Vertex algorithms. *Vertex-based event counting* methods [25,26,31] keep a tally of the rate of events with at least one reconstructed vertex, formed using typically a minimum of two to five tracks above a moderate transverse-momentum threshold and satisfying track-quality criteria optimized for stability of response. It is fundamentally an inclusive event-counting algorithm, and the conversion from the observed event rate to μ_{vis} follows Eq. (21).

In *vertex-counting* algorithms, the average number of visible interactions per bunch crossing is determined by counting the number of reconstructed vertices found in each bunch crossing. In its principle, the method is intrinsically linear and is applicable to much higher μ regimes than vertex-based event counting. But it suffers from nonlinear behavior with increasing pile-up parameter, primarily due to two effects: vertex masking and fake vertices [26]. Depending on the detailed performance of the tracking systems, the associated corrections can become large enough, at pile-up levels typical of recent LHC operation, to render percent-level luminosity monitoring very delicate.

4.2.6. Particle-flux methods

A related strategy is that of *particle-counting algorithms*, where some observable is directly proportional to the rate of particles interacting in the detector. These should be the most linear of all of the algorithm types, in that the visible interaction rate is directly proportional to the particle flux.

Transverse-energy distributions in calorimeters have been considered [29,32], but proved less reliable than event-counting methods. A more global, and surprisingly precise approach [26] is to measure the particle flux as reflected by the total ionization current flowing through a well-chosen set of liquid-argon calorimeter cells, or the current drawn by the photomultipliers (PMT's) of an iron-scintillator hadronic calorimeter. Although such measurements do not, strictly speaking, count individual particles, the recorded currents depend linearly on the luminosity to sub-percent accuracy. The weakness of the technique is that it does not give access to bunch-by-bunch luminosity information, but only to the total (i.e. bunch-averaged) luminosity. Because the method is intrinsically linear in μ, this averaging does not introduce any bias; but it somewhat dilutes the diagnostic power of the measurement. Nevertheless, calorimeter-current techniques provide highly valuable relative-luminosity monitoring at the LHC, that is totally independent from, and affected by very different systematics than the bunch-capable luminometers.

5. Absolute-luminosity calibration methods

In order to use the measured interaction rate μ_{vis} as an *absolute* luminosity monitor, each luminometer and algorithm must be calibrated by determining its visible cross-section σ_{vis}.

The simplest approach is to compare the visible interaction rate to the absolute luminosity computed, at the same point in time, from measured beam parameters. A first estimate can be obtained by measuring, at a well-chosen location away from the IP, the parameters of the two beams separately, and extrapolating them to the collision point (Section 5.1). Much better precision can be obtained by determining the beam-overlap area directly at the IP using the well-established technique of beam-separation scans (Section 5.2). More recently, the advent of vertexing-based beam-imaging techniques opened up interesting perspectives (Section 5.3).

An alternative strategy is to measure, in dedicated runs, the small-angle elastic *pp* cross-section and exploit the optical theorem to extract simultaneously the total *pp* cross-section and the luminosity (Section 5.4), thereby fixing the absolute scale of the visible interaction rate reported by the luminometer(s).

A third possibility is to rely on reference physics processes for which not only the cross-section, but also the acceptance and efficiency of the relevant subdetectors, are sufficiently well known (Section 5.5).

5.1. Single-beam parameters from accelerator instrumentation

The bunch luminosity (Eqs. (8)–(9)) can in principle be inferred from the bunch populations n_B and the transverse IP beam sizes σ_{iB} ($i = x, y$; $B = 1, 2$). While beam-current monitoring now achieves percent-level (or better) precision, single-beam profiles are notoriously difficult to measure accurately at the IP because of instrumental resolution and space limitations. One normally resorts to beam-profile monitors, typically based on wire scanners [33–35] or synchrotron-light telescopes [35,36], that are installed in diagnostic accelerator sections far from the collision point to measure emittances. These instruments usually report the projected horizontal and vertical RMS beam sizes, which can be extrapolated to the IP using an optical model of the collider lattice. While conceptually straightforward, this technique suffers from several limitations.

- Although *relative* beam-profile measurements are relatively straightforward, instrumental systematics (wire scanning speed, beam-induced heating, distortions of optical mirrors, resolution effects, . . .) often make the precise determination of the *absolute* transverse beam size at the monitor rather challenging.
- Extrapolating the projected single-beam sizes to the IP requires the knowledge of (at the minimum) the β functions and betatron phases both at the monitor and at the IP. These parameters must be measured in separate, dedicated machine-development sessions; their combination typically contributes a significant uncertainty to each of the four IP single-beam sizes σ_{iB}.
- The extrapolation procedure can become rather delicate in cases where transverse coupling, dispersion [37] or dynamic-β effects [17] play a significant role.
- Because of instrumental limitations (primarily resolution and dynamic range), the technique effectively assumes that each beam can be modeled by a Gaussian that is factorizable in *x* and *y*, in some cases a demonstrably incorrect hypothesis (see Sections 10.4–10.5).

5.2. van der Meer scans

At the ISR, RHIC and the LHC, the primary technique to determine the absolute luminosity scale is based on dedicated beam-separation scans, whereby the absolute luminosity is inferred, at one point in time, from the measurable parameters of the colliding bunches. By comparing the known luminosity delivered at the peak of the *van der Meer* (*vdM*) scan to the visible interaction rate μ_{vis}, the visible cross-section of inelastic *pp* collisions can be determined from Eq. (18). To achieve the desired accuracy on the absolute luminosity, these scans are sometimes performed not during normal physics operations, but under carefully controlled conditions and with beam parameters optimized for the purpose.
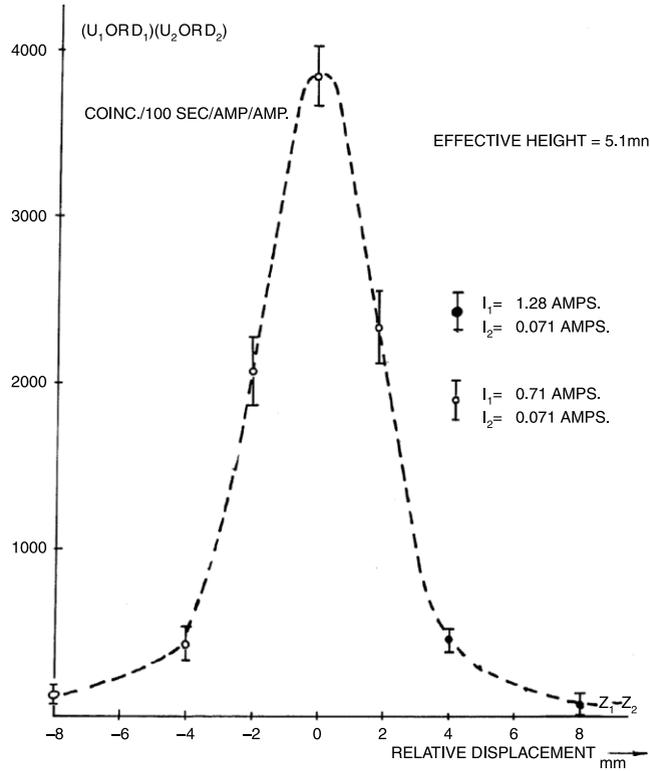
Fig. 2. van der Meer method at the ISR: monitor rate $R_y(\delta_y)$ as a function of the relative vertical displacement δ_y of the two beams. *Source:* Figure reproduced from Ref. [40]© CERN.

5.2.1. Absolute luminosity from beam parameters

In terms of colliding-beam parameters and for beams colliding with zero crossing angle, the bunch luminosity $\mathcal{L}_b$ is given (after integrating Eq. (2) over time and over the longitudinal coordinate z) by[3]

$$\mathcal{L}_b = f_r n_1 n_2 \int \rho_1(x, y)\, \rho_2(x, y)\, dx\, dy \tag{26}$$

where f_r is the machine revolution frequency, $n_1 n_2$ is the bunch-population product, and $\rho_{1(2)}(x, y)$ is the normalized particle density in the transverse $(x$–$y)$ plane of beam 1 (2) at the IP. Under the assumption that the particle densities can be factorized into independent horizontal and vertical components ($\rho(x, y) = \hat{\rho}_x(x)\hat{\rho}_y(y)$), Eq. (26) can be rewritten as

$$\mathcal{L}_b = f_r n_1 n_2\, \Omega_x(\hat{\rho}_{x1}, \hat{\rho}_{x2})\, \Omega_y(\hat{\rho}_{y1}, \hat{\rho}_{y2}) \tag{27}$$

where

$$\Omega_y(\hat{\rho}_{y1}, \hat{\rho}_{y2}) = \int \hat{\rho}_{y1}(y)\, \hat{\rho}_{y2}(y)\, dy \tag{28}$$

is the beam-overlap integral in the y direction (with an analogous definition in the x direction). In the method proposed by van der Meer [38] for continuous ribbon beams and generalized by Rubbia to elliptical bunched beams [39], the overlap integral (for example in the y direction) can be calculated as

$$\Omega_y(\hat{\rho}_{y1}, \hat{\rho}_{y2}) = \frac{R_y(0)}{\int R_y(\delta_y)\, d\delta_y}.$$

Here $R_y(\delta_y)$ is the luminosity (or equivalently μ_{vis}) – at this stage in arbitrary units – measured during a vertical scan (Fig. 2) at the time the two beams are separated by the vertical distance δ_y, and $\delta_y = 0$ represents the case of zero beam separation.

Defining the parameter Σ_y as

$$\Sigma_y = \frac{1}{\sqrt{2\pi}} \frac{\int R_y(\delta_y)\, d\delta_y}{R_y(0)}, \tag{29}$$

---

[3] This exposition largely mirrors that published by the ATLAS Collaboration [25,26].

and similarly for Σ_x, the bunch luminosity in Eq. (27) can be rewritten as

$$\mathcal{L}_b = \frac{f_r n_1 n_2}{2\pi \, \Sigma_x \Sigma_y},$$
(30)

which enables the luminosity to be extracted from machine parameters by performing a pair of beam-separation scans. Eq. (30) is quite general: Σ_x and Σ_y, as defined in Eq. (29), depend only upon the area under the luminosity curve, and make no assumption as to the shape of that curve. In the special case where the luminosity curve $R_y(\delta_y)$ is Gaussian, Σ_y coincides with the standard deviation of that distribution (Section 2.3.3).

In the more general case where the particle densities (or more precisely the dependence of the luminosity on the beam separation (δ_x, δ_y)) cannot be factorized into a product of uncorrelated x and y components, the formalism can be extended to yield [39]

$$(\Sigma_x \Sigma_y) = \frac{1}{2\pi} \frac{\int R_{x,y}(\delta_x, \delta_y) \, d\delta_x \, d\delta_y}{R_{x,y}(0, 0)}$$
(31)

with Eq. (30) remaining formally unaffected.

5.2.2. Luminosity calibration by beam-separation scans

To calibrate a given luminosity algorithm, one can equate the absolute luminosity computed using Eq. (30) to the luminosity measured by a particular algorithm at the peak of the scan curve using Eq. (18) to get

$$\sigma_{\text{vis}} = \mu_{\text{vis}}^{\text{MAX}} \frac{2\pi \, \Sigma_x \Sigma_y}{n_1 n_2},$$
(32)

where $\mu_{\text{vis}}^{\text{MAX}}$ is the visible interaction rate per bunch crossing observed at the peak of the scan curve as measured by that particular algorithm. Eq. (32) provides a direct calibration of the visible cross-section σ_{vis} for each algorithm in terms of the peak visible interaction rate $\mu_{\text{vis}}^{\text{MAX}}$, the product of the convolved beam widths $\Sigma_x \Sigma_y$, and the bunch population product $n_1 n_2$.

In the presence of a crossing angle in one of the scanning planes only (say the horizontal), the conclusions remain unaltered and Eqs. (29)–(32) remain valid [11,12]. The non-zero crossing angle widens the luminosity-scan curve by a factor that depends on the bunch length, the transverse beam size and the crossing angle, but reduces the peak luminosity by the same factor. The corresponding increase in the measured value of Σ_x is exactly cancelled by the decrease in $\mu_{\text{vis}}^{\text{MAX}}$, so that no correction for the crossing angle is needed in the determination of σ_{vis}. The vdM-scan method actually remains applicable in the generalized case of arbitrary crossing angles and beam-scanning planes [41].

A useful quantity that can be extracted from the vdM scan data for each luminosity method and that depends only on the convolved beam sizes, is the specific luminosity $\mathcal{L}_{\text{spec}}$:

$$\mathcal{L}_{\text{spec}} = \mathcal{L}_b/(n_1 n_2) = \frac{f_r}{2\pi \, \Sigma_x \Sigma_y}.$$
(33)

Comparing the specific luminosity values (i.e. the inverse product of the convolved beam sizes) measured in the same scan by different luminometers and algorithms provides a direct check on the mutual consistency of the absolute luminosity scale provided by these methods.

5.3. Beam imaging

5.3.1. Beam-gas imaging

An alternative way of measuring directly the absolute luminosity is to reconstruct the spatial distribution of beam-gas interaction vertices, in order to measure the transverse profiles of the individual beams and determine their overlap integral (Eq. (2)); the technique also provides a direct measurement of the relative beam separation and crossing angle. This absolute luminosity is then combined with the simultaneously measured rate of any relative-luminosity monitor to calibrate its visible cross-section σ_{vis}. This method was first tried at the ISR (see Section 6.2), but remained marginal until the advent of precise enough microvertex detectors. The LHCb experiment pioneered this technique at the LHC with remarkable success, as originally proposed in Ref. [42] and reported in Section 10.5.

5.3.2. Luminous-region imaging

Under the assumptions that the beams are Gaussian, that they are perfectly centered on each other, and that the horizontal and vertical sizes of beams 1 and 2 are equal pairwise, the bunch luminosity can be expressed in terms of the transverse luminous sizes (Eqs. (12)–(13)):

$$\mathcal{L}_b = \frac{f_r n_1 n_2}{8\pi \, \sigma_{x\mathcal{L}} \sigma_{y\mathcal{L}}}.$$
(34)

This suggests to infer the absolute luminosity from the luminous (or "beamspot") widths ($\sigma_{x\mathcal{L}}$, $\sigma_{y\mathcal{L}}$), provided the transverse distribution of *pp* collision vertices can be reconstructed to sufficient accuracy. In practice however, the assumption of equal beam sizes is rarely valid, so that Eq. (34) by itself provides only an upper limit on the actual luminosity. This method has been used at the ISR, in combination with beam-gas imaging, as a consistency check.

5.3.3. Beam–beam imaging

This method harnesses the combined power of high-precision vertexing and of beam-separation scans. Here, one measures the evolution, during a *vdM* scan, of the transverse distribution of reconstructed *pp* collision vertices, effectively using each beam to map the transverse particle density of the other. The technique can be pictured by first considering the case where beam 1 (say) is infinitely narrower than beam 2. Scanning beam 1 transversely across beam 2 and recording at each step, and with perfect instrumental resolution, the spatial distribution of collision vertices would provide a measurement of the transverse profile of beam 2. In practice however, the two beams are of comparable widths, with each one serving as a probe of the other. In addition, the beam-profile information is partially diluted by the vertexing resolution, which must be deconvoluted from the measured distributions. From the reconstructed normalized beam profiles one can again determine their overlap integral and from there the absolute luminosity.

Beam–beam imaging uses only the beam-separation dependence of the normalized shape of the vertex distributions, while the traditional *vdM* method extracts the overlap integrals from the evolution of the relative luminosity during the scan. The two methods are affected by very different systematics, and can therefore be considered complementary. The most recent and most complete proposal is detailed in Ref. [41], and its first application at the LHC published in Ref. [29].

5.4. Elastic pp scattering at small angles

The connection between elastic scattering on the one hand, and the luminosity and total cross-section on the other, was already used in the early days of the ISR. The optical theorem, which is a general law of scattering theory and follows from the conservation of probability in quantum mechanics, states that the total cross-section σ_{tot} is proportional to the imaginary part of the forward elastic scattering amplitude f_{el}:

$$\sigma_{tot} = 4\pi \, \mathrm{Im}\,[f_{el}\,(t=0)].$$

For small values of the scattering angle θ, the momentum transfer squared t can be approximated by $t = -(p\theta)^2$, where p is the beam momentum. The optical theorem implies that a measurement of elastic scattering in the forward direction will always provide information on the luminosity, either when combined with a measurement of the total *pp* interaction rate (Section 5.4.1), or by exploiting the interference between the strong and the electromagnetic elastic-scattering amplitudes (Section 5.4.2).

5.4.1. Measuring elastic scattering and the total interaction rate

The differential elastic cross-section is related to the strong-interaction scattering amplitude by

$$\frac{d\sigma_{el}}{dt} = \pi \, |f_{el}\,(\theta)|^2 \quad \text{with } f_{el}\,(\theta) = \mathrm{Re}\,[f_{el}\,(\theta)] + \mathrm{i}\,\mathrm{Im}\,[f_{el}\,(\theta)]. \tag{35}$$

Applying the optical theorem yields

$$\sigma_{tot}^2 = \frac{16\pi}{1+\rho^2} \left(\frac{d\sigma_{el}}{dt}\right)_{t=0} \tag{36}$$

where ρ is defined as

$$\rho = \mathrm{Re}\,[f_{el}\,(t=0)]/\mathrm{Im}\,[f_{el}\,(t=0)].$$

Noting that $\sigma_{tot} = R_{tot}/\mathcal{L}$ and $d\sigma_{el}/dt = (1/\mathcal{L})\,dR_{el}/dt$, and exploiting the fact that the left-hand side (resp. the right-hand side) of Eq. (36) depends quadratically (resp. linearly) on the cross-section, one can express both the total cross-section and the luminosity in terms of the total interaction rate R_{tot} and of the elastic event rate dR_{el}/dt extrapolated to zero momentum transfer:

$$\sigma_{tot} = \frac{16\pi}{1+\rho^2} \frac{(dR_{el}/dt)_{t=0}}{R_{tot}} \tag{37}$$

$$\mathcal{L} = \frac{1+\rho^2}{16\pi} \frac{R_{tot}^2}{(dR_{el}/dt)_{t=0}}. \tag{38}$$

Because the ρ parameter remains small in the energy range of all the colliders considered here, and because it enters the equations above as a quadratic correction only, its experimental uncertainty barely affects the measurement: for a typical value of $\rho \approx 0.1 \pm 0.02$, the uncertainty induced on the luminosity and on the total cross-section is less than 0.5%.
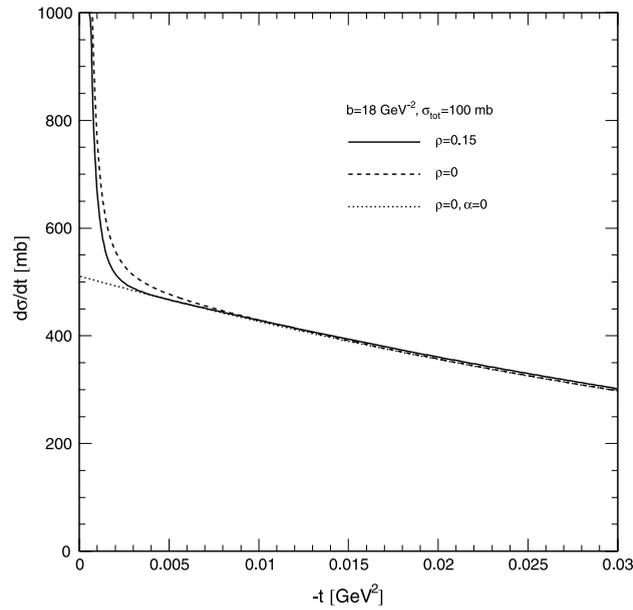
Fig. 3. Coulomb-nuclear interference method at the LHC: t-dependence of the elastic cross-section for a possible parameter set. The differential cross-section is plotted for $\rho = 0$ and $\rho = 0.15$ to highlight the interference region, as well as for $\alpha = 0$ to isolate the nuclear contribution.
Source: Figure reproduced from Ref. [43]© CERN.

To minimize the extrapolation error to $t = 0$, it is important to measure the differential elastic cross-section at very small angles, which is possible only if the intrinsic beam angular spread at the IP is small compared to the typical scattering angles that fall within the experimental acceptance. This, in turn, leads to the requirement of performing the measurement with low-emittance beams and in a special optical configuration with high β^* (see Eq. (5)). Also essential is a very good rapidity coverage by the subdetectors used to measure the total interaction rate, which in practice can prove challenging in the forward direction. Extrapolating the inelastic rate into uncovered rapidity intervals then requires a good knowledge of the diffractive and double-diffractive differential cross-sections. The model-dependent uncertainties associated with this extrapolation can, in some cases, dominate the accuracy of the luminosity determination.

5.4.2. Measuring elastic scattering in the Coulomb-interference region

The problem of estimating accurately the inelastic rate in the very forward direction can be circumvented by measuring the elastic cross-section down to angles small enough for the t-dependence to become sensitive to the Coulomb amplitude. This electromagnetic amplitude is calculable to high precision, thereby providing an absolute normalization. Using Eq. (35) and adding the Coulomb contribution, the elastic-scattering rate at small t can be written as

$$\frac{dR_{el}}{dt} = \pi \mathcal{L} \, |f_C + f_{el}|^2 = \pi \mathcal{L} \left| -\frac{2\alpha}{|t|} + \frac{\sigma_{tot}}{4\pi} \, (i + \rho) \, e^{-b|t|/2} \right|^2 \tag{39}$$

where the first term corresponds to the Coulomb amplitude and the second to the strong-interaction contribution (observe that the latter is written in such a way that the optical theorem is taken into account in an indirect fashion). By fitting this expression to the measured differential elastic rate, both the luminosity and the total cross-section can be determined without having to measure the inelastic rate. The fit also provides a measurement of the ρ parameter and of the slope b of the nuclear amplitude.

For the sake of clarity, Eq. (39) excludes the proton form factor and the Coulomb phase. It also assumes, for simplicity, that the imaginary part of the nuclear amplitude depends on t as a single exponential, that the real and imaginary parts of the nuclear amplitude exhibit the same t-dependence, and that spin effects can be neglected (see e.g. [44]). The t ranges dominated by Coulomb scattering, Coulomb-nuclear interference and nuclear scattering are illustrated in Fig. 3. The interference region, where some sensitivity to the Coulomb amplitude starts to develop, lies around $-t \leq 0.001 - 0.002$ GeV^2; this onset varies slightly from one collider to the next because of the energy dependence of the total cross-section. As $t = -(p\theta)^2$, the sensitivity threshold expressed in terms of the scattering angle θ is inversely proportional to the beam energy: at the ISR the Coulomb interference region could be reached with scattering angles of a couple of milliradians, while at the LHC angles as small as 3.5 μrad are needed. This illustrates well the difficulty of this method as the center of mass energy increases.

5.5. Use of reference physics processes

5.5.1. Inelastic hadronic interaction rate

Measuring the rate of events from a process with a well-known cross-section is a straightforward way of determining the luminosity. The problem is that when a new collider opens up an unexplored energy regime, hadronic cross-sections are not known *a priori*, and calculating e.g. the total inelastic cross-section with adequate precision is not possible due to the non-perturbative character of soft hadronic interactions. A second difficulty is that the phase-space coverage of any luminometer is in practice limited, thus requiring a precise estimate of its acceptance for the physics process of interest.

The inelastic cross-section was used both at the S$p\bar{p}$S (Section 7.4) and at the Tevatron (Section 8.4), where technical limitations made *vdM* calibrations impractical. In both cases the inelastic cross-section was determined by measuring simultaneously the inelastic rate and the differential elastic cross-section (Section 5.4.1), thereby providing a "luminosity-independent" value of the inelastic cross-section.

5.5.2. Semileptonic W and/or Z decays

Using *W* or *Z* production for luminosity measurements has been considered both for the Tevatron and for the LHC. At $\mathcal{L} = 10^{33}$ cm^{-2} s^{-1} and $\sqrt{s} = 14$ TeV the *W* production rate is about 10 Hz and the *Z* rate roughly ten times smaller; the leptonic decay modes are experimentally clean and the background processes controllable. The integrated luminosity is given by

$$\int \mathcal{L}dt = \frac{N - B}{\sigma \, A \, C}$$

where *N* is the number of observed events, *B* the estimated number of background events, σ the product of the theoretical cross-section by the leptonic branching fraction, *A* the geometric acceptance that relates the cross-section measured in the fiducial volume to that extrapolated to the full kinematic range and *C* the event-reconstruction efficiency.

W and *Z* counting have mainly been considered for relative-luminosity monitoring, in which case the precise values of σ and *A* become irrelevant; instrumental effects may however induce a time dependence in the efficiency *C*, that must be carefully corrected for. But absolute luminosity measurements have also been envisaged, in which case the values of *C*, *A* and σ must be accurately known on an absolute scale. A combination of Monte-Carlo simulations and data-driven methods can provide a precision as good as 1%–1.5% on the efficiency *C*; extrapolating to the whole phase space brings an additional contribution of 1.5%–2% from *A* [45]. The real issue is rather the precision of the theoretically-predicted cross-section σ.

At the energy scale of the Tevatron or the LHC, the parton constituents of the proton can be considered as quasi-free, and the *W*/*Z* production cross-sections written as the convolution of the parton–parton cross-section with the parton distribution functions (PDF). The parton–parton cross-section has been calculated in next-to-next-to-leading order with an accuracy better than 1% (see e.g. [46] and references therein). The main uncertainty in the *W* and *Z* production cross-section is associated with the knowledge of the PDF's. In the pre-LHC era, PDF's were extracted from measurements of deep inelastic scattering in fixed-target experiments and at the HERA $e^{\pm}p$ collider, and of Drell–Yan lepton-pair production at $p\,\overset{(-)}{p}$ colliders. The uncertainty has diminished considerably over the last decade, especially thanks to more precise HERA data. Today one considers that PDF-related uncertainties on the *W* and *Z* production cross-section are at the 5% level [47], though the exact number is still being debated. Further improvements are expected once all the available LHC data will have been systematically exploited.

A different approach, suggested in Ref. [48], would be to consider a parton–parton luminosity instead of the conventional proton–proton luminosity. The principle would be to measure simultaneously the pseudorapidity distribution of *W* and *Z* leptonic decay products and the event rate of weak boson production. Such measurements would constrain the *x* distribution of quarks and antiquarks and allow percent-level predictions of other $q\bar{q}$-related scattering processes without requiring the knowledge of the proton–proton luminosity.

One can use *W*/*Z* counting either to measure absolute weak-boson production cross-sections (with a *pp* luminosity calibrated e.g. by *vdM* scans) and thereby constrain theoretical predictions, or to provide an integrated luminosity to which other cross-sections can be normalized; but one cannot do both simultaneously. The consensus at the Tevatron and the LHC has been to adopt the first approach, both because that is conceptually cleaner and because the achieved experimental precision on the absolute luminosity scale has outpaced the uncertainties on theoretical predictions. Relative-luminosity monitoring using $Z \to \mu\mu$ decays, however, is free of theoretical inputs and offers promising prospects as an independent check on the long-term consistency of luminosity measurements at the LHC.

5.5.3. Two-photon production of lepton pairs

Before LHC turn-on, exclusive production of lepton pairs by two-photon fusion has been repeatedly discussed as a promising candidate for absolute-luminosity determination (see e.g. [49–51]). This pure-QED process is well understood theoretically, the leptons are unaffected by strong interactions, proton–proton rescattering effects can be controlled by proper kinematic cuts, and the cross-section is known to better than 1% [50,51].

In principle, centrally-produced muon pairs provide a clean signal. Using the ATLAS detector as an example, the minimum transverse momentum for a muon to reach the muon chambers and provide reasonable trigger efficiency is $p_T > 6$ GeV/c. However, it is important that the total transverse momentum of the pair be very small ($p_T^{\mu\mu} < 50$ MeV/c) in order to control rescattering effects and suppress background: this means that the muons must be essentially back to back in the transverse plane. The main backgrounds originate from the scattering debris of the interacting protons, from Drell–Yan dimuon production and from π/K decays. The latter two sources are suppressed by vertex quality cuts and by requiring that the event contain no additional charged track.

But there are several difficulties. In order to make the $p_T^{\mu\mu}$ cut efficient, the measurement resolution on the individual muon momenta would have to be quite good. In addition, the high pile-up environment of the LHC makes it extremely challenging to isolate such an exclusive process. More importantly however, the usable event rate appears very low: the effective cross-section after kinematic cuts is of the order of 1 pb, more than three orders of magnitude lower than W or Z production. Muon pair-production via two-photon fusion is a theoretically precise method but it is both experimentally challenging and statistically limited, and it has not been employed at the LHC so far.

One way to circumvent the statistical limitation might be to resort to e^+e^- rather than $\mu^+\mu^-$ pairs, as analyzed in detail in [52] and references therein. Electrons are in principle detectable down to much lower transverse momenta than muons, and reducing the p_T cut from 6 to 0.5 GeV/c increases the rate by three orders of magnitude. However the e^+e^- pairs would have to be identified in a ferocious environment with hadronic rates still many order of magnitude larger. In addition, the present layout of the LHC detectors would not lend itself, without major modifications, to the detection of low-p_T electrons. At this point no practical proposal has been put forward by any LHC experiment.

6. Luminosity determination at the ISR

6.1. The experimental context

6.1.1. Collider overview

The ISR [2,23] ran for physics from 1971 to 1983, at beam energies from 15 to 31 GeV. What was learnt there came to be the basis for all future hadron colliders. The facility consisted of two independent rings intersecting at eight points, with experiments at seven of these and one utility insertion. The first pp collider ever built displayed unique features (Table 1): unbunched coasting beams crossing with a large horizontal angle of 14.8° (this implied that only the vertical beam size, but not its horizontal width, influenced the luminosity); a momentum spread as high as 3.7% (that required a generous horizontal aperture); and impressive luminosities (4×10^{30} cm^{-2} s^{-1} design, 1.4×10^{32} cm^{-2} s^{-1} achieved) made possible by the very large beam currents (20 A design, 57 A achieved) and by the installation, during the mature phase of the program, of low-β insertions in some of the interaction regions.

6.1.2. Experimental physics program

The luminosity determination at the ISR is interesting for many reasons. The machine was the first of its kind: by no means was it evident how to measure the luminosity. Many of the tools that would be used at later colliders were developed here. Early discussions [53] included contributions from P. Darriulat, C. Rubbia, W. Schnell and J. Steinberger (among others); these pilot proposals were later superseded by a simple observation made by S. van der Meer [38]. The list of prominent people interested in how to correctly measure the luminosity indicates the importance attached to the issue at the time. This is of course related to the physics context prevailing in the early seventies: the principal question was the high-energy behavior of the total cross-section and at what point "asymptopia" would be reached (it was common belief that at high enough an energy, all cross-sections would become energy-independent). Essentially all the precise luminosity measurements at the ISR were driven by total cross-section measurements in pp or $p\bar{p}$ collisions.

It is also worth pointing out that many of the best luminosity measurements were made during the first part of the ISR life span. Towards the end of the seventies, the paradigm changed: high-p_T physics gradually took over and the interest in accurate luminosity measurements diminished very significantly. The focus had become jets and other new high-p_T phenomena, for which the absolute normalization was less important. This tendency somehow persisted at the S$p\bar{p}$S and at RHIC (except again for pp total cross-section measurements). With the start of the LHC, precision luminosity measurements returned to the fore, as the accuracy of perturbative QCD calculations had reached a level where absolute cross-sections measurements in specific channels can provide sensitive tests of the Standard Model.

A total of about 50 different experiments were carried out at the ISR. From the viewpoint of luminosity measurements, they fall into three distinct categories: the first generation of elastic scattering and total cross-section measurements, the second generation of elastic scattering and total cross-section measurements, and experiments focused on high-p_T phenomena. We will discuss each of these in turn.

6.2. van der Meer scans

6.2.1. Coasting-beam formalism

As the ISR operated in unbunched mode, the bunched-beam formalism of Section 5.2 needs to be translated for coasting, ribbon-like beams. In terms of accelerator parameters, the luminosity becomes [8,13]

Table 2
Overview of luminosity determinations by the van der Meer method for total cross-section measurements at the ISR.

Experiment	$\sqrt{s}$ (GeV)	Year	Point-to-point error on σ_{tot}		Absolute luminosity scale	
			Luminosity contribution	Total	Uncertainty on vertical scale	Vertical scale determination
Aachen–CERN–Harvard–Genova–Torino [54]	30	1971	5%	8%		
CERN–Rome [55]	23–53	1973	2%	2.5%	2%	Pick-up electrodes
Pisa–Stony Brook [56]	23–53	1973	2%	2.0%	2%	Spark chambers
Aachen–CERN–Heidelberg–MPI Munich [57]	24–63	1975	1.5%	1.6%–2.0%	0.6%	High-precision scrapers
CERN–Rome–Pisa–Stony Brook [58,59]	24–63	1976	0.5%	0.6%–0.7%	0.5%	High-precision scrapers
Aachen–CERN–Harvard–Genova–Munich–Riverside [60]	24–63	1978	0.6%	0.8%	<1%	High-precision scrapers
CERN–Napoli–Pisa–Stony Brook [61]	31–63	1984		0.7%	0.4%	High-precision scrapers
Louvain-Northwestern [62]	31–63	1985	0.4%	0.5%	0.4%	High-precision scrapers

$$\mathcal{L} = \frac{I_1 I_2}{e^2 c \tan \alpha} \frac{1}{h_{eff}} \qquad (40)$$

where I_1 and I_2 are the total currents in the two coasting beams, e and c denote respectively the proton charge and the speed of light, and α is the horizontal half crossing-angle. The geometrical overlap of the two beams is quantified by what was then called the effective height h_{eff}:

$$\frac{1}{h_{eff}} = \int \hat{\rho}_{y1}(y)\hat{\rho}_{y2}(y)dy \qquad (41)$$

where $\hat{\rho}_{y1}$ and $\hat{\rho}_{y2}$ are the normalized vertical density distribution of the two beams (see Eq. (28)). This expression does not depend on the horizontal beam size, because the beams are unbunched, collide at an angle and all the particles in one beam traverse the entire cross-section of the other beam.

Early proposals to measure the effective height discussed different ways of determining separately the two vertical distributions. Those ideas were superseded by van der Meer's simple but ingenious proposal to scan one beam against the other in small and precisely known vertical steps while observing the *pp* collision rate in an appropriate monitor [38]. As the effective height is simply the inverse of the vertical overlap integral (Section 5.2.1), it is proportional to the vertical convolved beam size Σ_y, and is given by

$$\frac{1}{h_{eff}} = \frac{R(0)}{\int R(\delta_y)d\delta_y}.$$

Here δ_y is the relative vertical displacement of the two beams, $R(0)$ is the rate at the peak of the displacement curve and the integral equals the area under that curve.

This method was first used in dedicated machine studies; one of the first documented scans [40] is shown in Fig. 2. Soon thereafter, the Aachen–CERN–Harvard–Genova–Torino collaboration applied the technique to their 1971 measurement of the elastic cross-section at $\sqrt{s} = 30$ GeV [54]. The authors made the simplifying assumption that the detection efficiency and the beam shapes do not depend on the beam separation. If the two individual beam distributions are strictly Gaussian, then the luminous region (called at the time the *beam source distribution*) is Gaussian as well, and its height is independent of δ_y. The beams were indeed found to be largely Gaussian by detecting beam particles elastically scattered off the residual gas in the vacuum pipe. It was also argued that neither the emittance nor the local β-function changed with δ_y, thus conserving the single-beam shapes during the scan. The beams were displaced symmetrically by equal amounts $\pm\delta_y/2$, thus both the shape and the average position of the luminous region were independent of δ_y. An "ad hoc" estimate of the systematic errors yielded a luminosity uncertainty of 5%.

Luminosity calibration by the *vdM* method was used by most total cross-section experiments at the ISR. The luminosity entered either directly in combination with the total interaction rate ($\sigma_{tot} = R_{tot}/\mathcal{L}$), or indirectly via the optical theorem. In the latter case, the differential elastic cross-section was extrapolated to $t = 0$ and the total cross-section had a weaker dependence on the luminosity ($\sigma_{tot} \sim 1/\sqrt{\mathcal{L}}$). The results are summarized in Table 2 and discussed in more detail below.

6.2.2. First-generation total cross-section experiments

In 1973 the CERN-Rome group [55] and the Pisa-Stony Brook collaboration [56] published at the same time their ground-breaking results on the total cross-section. Both collaborations observed an increase of the total cross-section of about 10% over the 23–53 GeV c.m. energy range. The two groups based their cross-section normalization on *vdM* scans; in both cases, the dominant cross-section uncertainty was associated with the luminosity measurement. As it was the energy dependence of the cross-section that was at stake, it was important to separate the point-to-point (i.e. the $\sqrt{s}$-dependent) error on the luminosity from the energy-independent uncertainty on the absolute luminosity scale. The point-to-point error was typically estimated on the basis of reproducibility and internal consistency considerations. The luminosity-scale uncertainty, which dominated the absolute-scale uncertainty on the total cross-section, was almost entirely due to the uncertainty on the

absolute vertical beam separation during the scans, the determination of which required considerable effort and multiple novel techniques.

The CERN-Rome group [55] used three independent monitor systems, each separately calibrated by *vdM* scans, that were based on scintillation counters of various sizes placed about 5 m from the IP and with a high acceptance for inelastic collisions. The calibration constant σ_{vis} of each monitor was determined for a variety of beam conditions; the most stable monitor never varied by more than $\pm 2\%$ over the period of the experiment. The overall point-to-point error on the total cross-section was estimated at about 2.5%, with slight differences from energy to energy, and dominated by the $\pm 2\%$ from the luminosity calibration of the best monitor at each setting. The scale error, common to all beam energies, was estimated at $\pm 3\%$, of which $\pm 2\%$ were associated with the calibration of the magnets used to displace the beams vertically. The beam displacements during the scans were measured using pick-up electrodes and compared to the "ISR scale" calculated from the magnets and the optics. Note that the beam currents I_1 and I_2 (Eq. (40)) were known to an absolute accuracy better than a per mil, so that the corresponding uncertainty was neglected in all luminosity measurements at the ISR [63].[4]

The Pisa-Stony Brook collaboration [56] used a very similar approach to luminosity determination and likewise employed three separate sets of scintillator monitors to detect inelastic-collision products in different regions of phase space. They assigned a comparable 2% point-to-point error to their *vdM*-based calibration. The overlap integral (Eq. (41)) was also determined in two additional, scan-independent ways. First, titanium atoms were evaporated from a source close to the IP, and a spark-chamber telescope was used to select quasi-elastic recoil protons from *p–Ti* scattering and to reconstruct the density distributions of the two beams separately. The second method used the same spark chambers to directly reconstruct the luminous region. Both approaches gave results consistent with those of the *vdM* method, albeit with larger errors. In the end it was decided to use only *vdM* results and to treat the alternative methods as consistency checks. The vertical-displacement scale was checked by using the spark chamber telescope to visualize the interaction region and compare the measured beam displacements to the "ISR scale"; the corresponding scale error was estimated at 2%.

6.2.3. Second-generation total cross-section experiments

The two collaborations joined forces in a second-generation experiment [58] aimed at measuring more accurately the rise of the total cross-section. They profited from the fact that the operation of the ISR had improved considerably since the first years, and a number of refinements were adopted in the *vdM* procedure. For instance, the scan protocol was modified so as to minimize hysteresis effects, and scans were performed only one intersection at a time in order to eliminate cross-talk between IP's. A large number of independent monitor telescopes covered different ranges to record inelastic collisions, with visible cross-sections varying from 0.3 to 25 mb for different monitors. Fig. 4 shows the general layout of the experiment and the different hodoscopes which were combined to form the separate monitors.

From the internal consistency of the results, it was concluded that the point-to-point error could be reduced to 0.5%, compared to the previous estimate of 2%. There was also progress on the dominant contribution to the (energy independent) error on the absolute luminosity scale, which arose from the uncertainty on the vertical separation scale. High-precision beam scrapers, consisting of precision-machined tantalum blades, with edges straight to within $10\,\mu m$ over the 20 cm blade length and driven by a precision screw, were developed, that could move in above and below the beam with a positioning precision of $5\,\mu m$ [64]. The beam center was found by recording the current lost at each step, and the beam-positioning error was estimated at $50\,\mu m$ over 1 cm. In this manner the "ISR scale" could be calibrated to within a 0.5% uncertainty, compared to the previous estimate of 2%–3%. That the 0.5% estimate was basically correct was verified in a manner independent of *vdM* scans (see Section 6.4).

With the advent of antiproton–proton collisions towards the end of the ISR lifetime, the total $p\bar{p}$ cross-section could also be determined. These experiments also re-measured the total pp cross-section very precisely [61,62]. Considerable effort had been invested [65,66] to push down the scale error, using a dedicated set of the above-mentioned scrapers installed within a few meters of the IP where the cross-section was being measured. The scale error was ultimately reduced from 0.5% to 0.4%.

On the whole, all the second-generation total cross-section experiments at the ISR claimed an error below 1% on their luminosity measurement.

6.2.4. High transverse-momentum experiments

The paradigm changed in the later years of the ISR: hard processes with high-p_T leptons or hadrons in the final state became more and more important, while the interest in soft processes faded away. A precise cross-section determination was less important for those exploratory measurements of new phenomena. Combing through all the publications from that period, it is hard to find more than a couple of lines relating to luminosity measurements. The standard procedure was to use a couple of scintillator telescopes on both sides of the IP to monitor the collision rate. The calibration was always based on *vdM* scans and the quoted error varied from 3% to 10% depending on the experiment. The size of this error probably reflects how much attention was paid to the measurement and how carefully the calibration was carried out. In many cases a luminosity measurement was not even needed as the data were presented as ratios of cross-sections.

---

[4] In contrast, the scale errors at different LHC energies (or even at different β^* settings) are uncorrelated, and the bunch-population product contributes a non-negligible uncertainty (see Section 10).
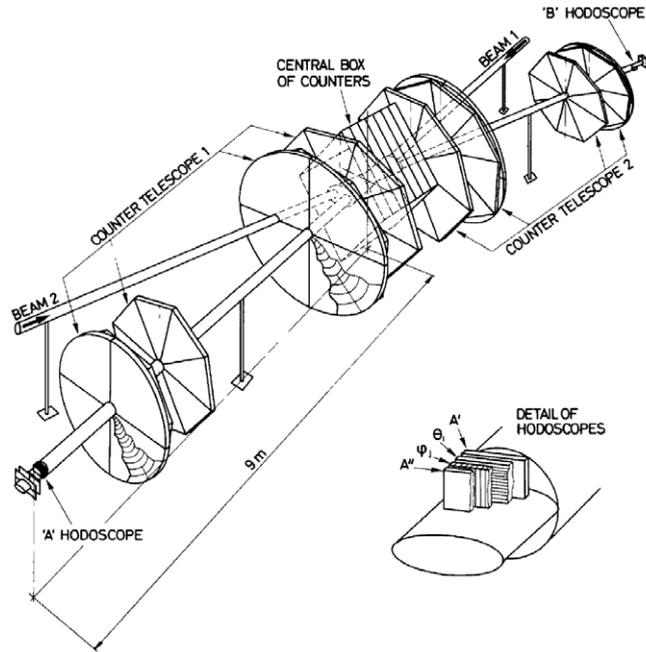
Fig. 4. Layout of the CERN–Rome–Pisa–Stony Brook experiment.
Source: Figure reproduced from Ref. [58], with permission from Elsevier.

6.3. Elastic pp scattering in the Coulomb-interference region

The earliest measurement of the energy-dependence of the total *pp* cross-section in the ISR energy range was published in 1973 [44] by the CERN-Rome collaboration, using the Coulomb method (Section 5.4.2) for the first time. At a c.m. energy of 30 GeV, the Coulomb and nuclear cross-sections are equal at an angle of about 3 mrad, which corresponds to a displacement from the beam axis of 3 cm at the end of the ISR 10 m long straight section. To be able to come so close to the beam, scintillation-counter hodoscopes were placed in movable, thin-wall sections of the ISR vacuum chamber. Those inserts were later to be called "Roman Pots", named after the Rome group in the collaboration. Different versions of this type of insertable devices have since been used at every *pp* collider.

Fig. 5 (top) illustrates the experimental layout and the "Roman Pot" concept. The experiment covered an angular range of 2–8 mrad and took data at 23 and 31 GeV c.m. energy. Typical angular distributions are shown in Fig. 5 (bottom). The total cross-section, the ρ parameter and an overall normalization constant (proportional to the luminosity) were left free in the fit to the differential elastic cross-section; the slope parameter *b* was taken from earlier ISR measurements. The total cross-section extracted from the fit had a precision of about 2%. The value and the error on the luminosity that could be extracted from the fitted normalization constant are not quoted in the paper; but in this kind of analysis the luminosity error typically comes out roughly twice as large as the error on the total cross-section, because the elastic rate depends quadratically on σ_{tot} (see Eq. (39)). Thus this method gave the luminosity with a precision of roughly 4%. The authors noted an increase of the total cross-section with energy but, given the size of the errors, remained very prudent in their conclusion. One year later they confirmed the increase by halving the error of the luminosity using the *vdM* method (see Section 6.2) and by extending the beam energy range.

6.4. Elastic pp scattering, optical theorem and total interaction rate

The so-called "luminosity-independent" method, based on the simultaneous measurement, at the same IP, of the total interaction rate and of the differential elastic collision rate (Section 5.4.1) was put into practice by the joint CERN–Rome–Pisa–Stony Brook collaboration [59] to measure simultaneously the total cross-section (Eq. (37)) and the corresponding luminosity (Eq. (38)). These results, which do not rely on any external luminosity determination, were compared to separate total cross-section measurements normalized to luminosities calibrated by the *vdM* method. In the ideal case, the ratio $\lambda = \sigma_{tot}(\mathcal{L}_{indep})/\sigma_{tot}(\mathcal{L}_{vdM})$ equals 1. If, however, a scale error biased significantly the *vdM*-based luminosity measurements, then λ would deviate from unity in the same direction at all energies.

No such systematic deviation can be found in the data (Fig. 6). The authors interpreted the energy-averaged ratio $\langle \lambda \rangle = 0.9944 \pm 0.0093$ as a confirmation of the reliability of the *vdM* method for determining the absolute luminosity at about the 0.9% level. Since the absolute *vdM* displacement scale had been independently verified, they also alternatively
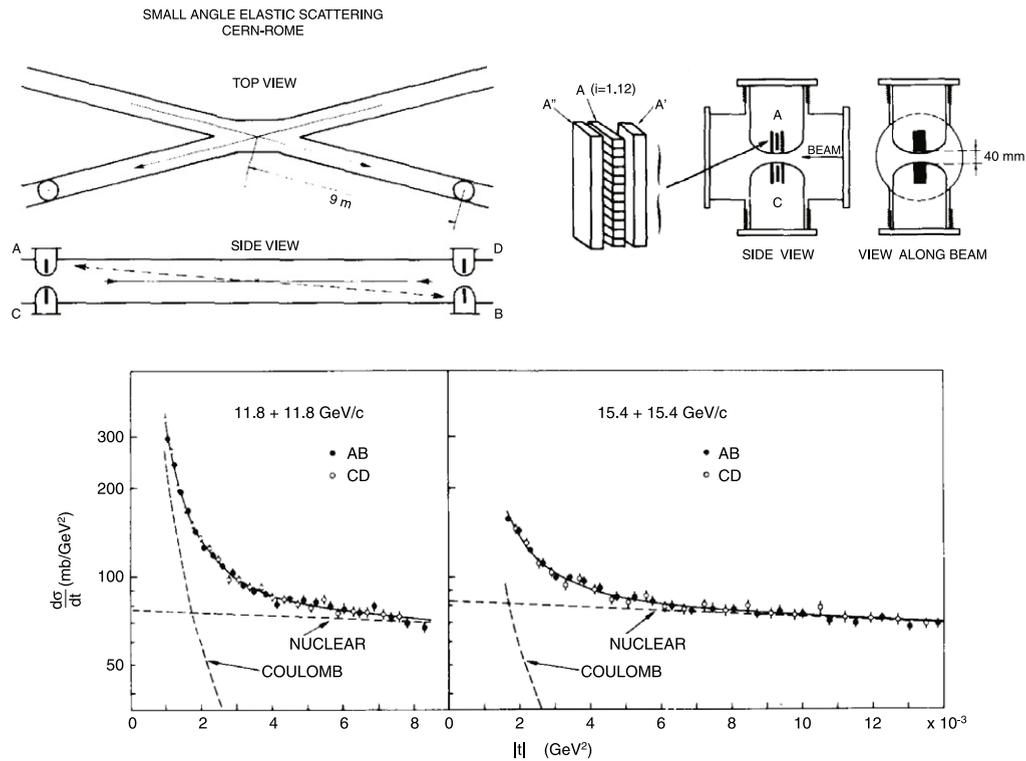
Fig. 5. Top: general layout of the CERN-Rome experiment (left) and sketch of the Roman Pots (right). Bottom: typical angular distributions measured in the Coulomb region at two different energies; the separate contributions from Coulomb and nuclear scattering are indicated.
Source: Figures reproduced from Ref. [44], with permission from Elsevier.
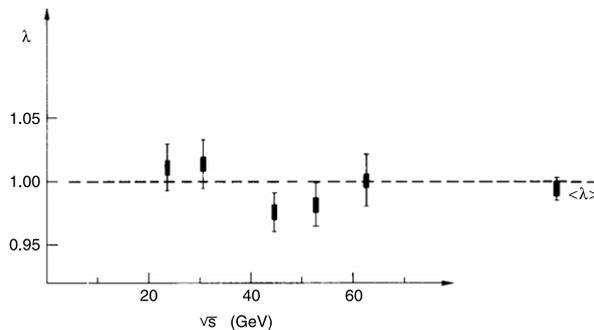

Fig. 6. Results obtained for the ratio λ used to monitor a possible energy-independent scale error on the total cross-section.
Source: Figure reproduced from Ref. [59], with permission from Elsevier.

interpreted the result as a proof that the optical theorem holds true to $\pm 0.5\%$ for pp collisions at ISR energies. The factor of 2 in the precision (compared to the 0.9% above) arises from the fact that the total cross-section enters Eq. (36) quadratically.

The luminosity-independent method was also used by the Aachen–CERN–Harvard–Genova–Munich–Riverside collaboration [60], with results completely consistent with measurements based on *vdM* calibrations. However, because of the smaller uncertainties achieved by the *vdM* method, the *vdM*-based cross-sections were chosen for the final result.

7. Luminosity determination at the CERN $p\bar{p}$ collider

7.1. The experimental context

7.1.1. Collider overview

In the summer of 1978, the CERN Council approved the construction of an antiproton storage facility at CERN, as well as the SPS modifications necessary for protons and antiprotons to be accelerated and brought into collision. The first $p\bar{p}$ collisions were observed in July 1981 and the first W candidates detected in autumn 1982. This was an unprecedented

Table 3
Evolution of luminosity uncertainties at the S$p\overline{p}$S using accelerator instrumentation.

	1982 (UA1) [68,69]	1983 (UA1) [70]	1995 (UA4/2) [71]
Beam profile measurement		7%	3%
Extrapolation of transverse beam size to IP		2%	–
Bunch-population product		4%	4%
Total uncertainty	25%–30%	8%	5%

success story. The reasons for this amazing achievement were many. The impact of the knowledge acquired during the construction and operation of the ISR was of course enormous, and the smooth transition of the SPS from a single-beam accelerator to a collider was fundamentally due to the "gold-plated" technology with which the SPS had been built. Most important however was the invention of stochastic cooling and its concrete realization in the Antiproton Accumulator ring: without cooling, antiprotons could never have been stored in the SPS. Already during the early years, bunch intensities of about 10^{10} $\overline{p}$ were achieved, growing over time to a maximum of 8×10^{10} $\overline{p}$/bunch.

The design luminosity of 10^{30} cm^{-2} s^{-1} was surpassed late in the eighties, and had reached values as high as 6×10^{30} cm^{-2} s^{-1} by the time the S$p\overline{p}$S shut down in 1991. The luminosity gains were mainly due to three reasons: the upgrade of the Accumulator complex, the increase from 3 to 6 colliding-bunch pairs in the SPS, and the reduction of the IP beam size. Over the years, the IP β-functions were gradually reduced from the initial values of $\beta_x^* = 7$ m, $\beta_y^* = 3.5$ m to $\beta_x^* = 0.6$ m, $\beta_y^* = 0.15$ m. In this regime, the 20 cm bunch length was comparable to β^*, so that the hourglass-induced luminosity degradation (Section 2.3.2) became significant.

To start with, the maximum possible energy was 270 GeV/beam. Extra water pumps were installed in the tunnel in 1984 to improve the magnet cooling, allowing an increase to a maximum of 315 GeV/beam. The SPS pulsing scheme (see Section 7.5) pushed the peak energy to 450 GeV/beam.

7.1.2. Experimental physics program

By the time the S$p\overline{p}$S turned on, the context was radically different from that at ISR startup. The physics landscape had evolved considerably: the main goal had become the hunt for the intermediate vector bosons that were expected to be produced at the energy of the new collider. Altogether, high-p_T physics dominated, while issues like elastic scattering and total cross-section receded into the background. Discovery physics does not in general require very accurate luminosity measurements; the scant attention given to the topic can clearly be seen in most publications of the time. The more precise luminosity determinations at the S$p\overline{p}$S were, like at the ISR, essentially driven by the total cross-section measurements performed by the UA1 and UA4 experiments.

The experimental environment had also changed compared to the ISR. The epoch of a large number of experiments replacing each other frequently with different detectors moving in and out on short notice had come to an end. The two multi-purpose detectors UA1 and UA2 operated over a timescale of a decade rather than a year. A couple of smaller experiments ran for shorter periods, such as UA5 (exploratory characterization of $p\overline{p}$ collisions using streamer chambers) and UA4 (elastic scattering and total cross-section measurements), but the main thrust was definitely on large detectors with seamless coverage over a wide angular range.

7.2. Single-beam parameters from accelerator instrumentation

Using directly measured single-beam parameters, without resorting to beam-separation scans, is the most obvious approach to luminosity determination. This method was never tried in earnest at the ISR, as it turned out that *vdM* scans provided the most accurate results from the very beginning. But transposing the van der Meer method to the S$p\overline{p}$S was far from straightforward, because p's and $\overline{p}$'s shared the same beam pipe and were therefore acted upon by the same magnetic elements. Counter-rotating beams stored in the same ring cannot be magnetically separated.

As one beam could not be scanned against the other by steering correctors, an alternative proposal using electrostatic separators was put forward by C. Rubbia as early as 1977 [39], but the scheme was never employed for physics-related measurements. One difficulty was that the electrostatic separators only acted in the horizontal plane and were grounded in a way such that the beams could be scanned in one direction only. Another limitation was their strength. The *vdM* method requires separating the beams by up to several times their transverse size; but in the initial period of S$p\overline{p}$S operation, β^* was quite large, and no electrostatic separators were available with a strength matching the large beam sizes and the considerable separations required [67]. No attempt to use *vdM* scans is reported in any physics publication over the lifespan of the $p\overline{p}$ collider.

Instead, the first determination of the luminosity at the S$p\overline{p}$S, which was reported by the UA1 experiment, was based on calculating the luminosity directly from single-beam parameters. In the first publications the total error was quite large: it included contributions from both bunch-intensity and single-beam profile measurements and was estimated to lie in the 25%–30% range [68,69]. This method, used by all of UA1, UA4 and UA4/2, was gradually refined, ultimately leading to an absolute luminosity uncertainty well below 10% (Table 3).
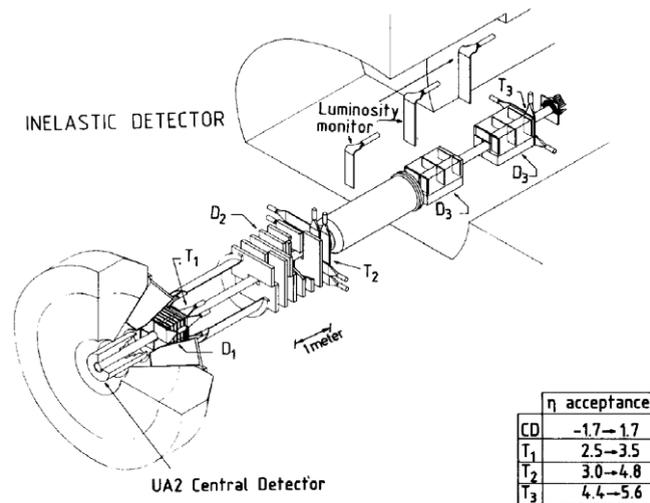
	η acceptance
CD	−1.7 → 1.7
T₁	2.5 → 3.5
T₂	3.0 → 4.8
T₃	4.4 → 5.6

Fig. 7. Layout of one arm of the UA4 experiment, showing the forward telescopes (T1–T3) and the central detector of UA2 (CD).
Source: Figure reproduced from Ref. [74], with permission from Elsevier.

The transverse single-beam profiles were measured by the wire-scan technique. A fine wire moved through the beam at high speed (4 m/s) and the profile was measured by detecting the secondaries produced. The scattering products had a strong directionality and were detected by scintillators equipped with PMT's placed close to the beam pipe on both the p and the $\bar{p}$ side. The first version of this instrument used at the S$p\bar{p}$S contained a Be wire [72]. But at higher intensities the heating of the wire became prohibitive, and in a later version a 25 μm carbon filament was employed instead [33,34]. In their first total cross-section measurement, UA1 quoted an error of 7% [70] on the effective width and height of the beams measured by the wire-scan technique. A decade later, UA4/2, referring to the same method, reported uncertainties in the 3% range [71], the exact value depending on the actual beam size.

The extrapolation of the four single-beam profiles from the wire-scanner location to the IP was estimated to contribute a couple of percent to the UA1 luminosity uncertainty [70], a rather challenging achievement in view of the currently known limitations of the method (Section 5.1).[5] This source of error was eliminated by UA4/2 by installing a wire scanner at the IP itself.

The bunch-intensity measurements were based upon directional couplers, that make it possible to separate the signals from the two beams. The couplers were specially adapted to correctly measure the low-intensity antiprotons (0.5×10^9 $\bar{p}$/bunch) in the presence of the intense proton beam ($1-2 \times 10^{11}$ p/bunch), and were also equipped with specially developed synchronous RF-receivers [73]. For redundancy and more precise measurements several of those couplers were placed around the ring, eventually achieving a precision of 3% per bunch [70,71], i.e. 4% on the bunch-population product.

Adding the three contributions above, UA1 reported a total error of 8% on their luminosity measurement [70]. The UA4/2 experiment did not quote a distinct number, but it can be inferred from their publication on the total cross-section [71] that they estimated the error on the total luminosity to be around 5%. This uncertainty is smaller than that quoted by UA1 primarily because of the higher precision of beam-profile measurement towards the end of the S$p\bar{p}$S operation.

7.3. Elastic $p\bar{p}$ scattering, optical theorem and total interaction rate

In addition to the direct luminosity-dependent method, the UA4 experiment used the simultaneous measurement of the total interaction rate and of the differential elastic rate (Section 5.4.1) to normalize its total cross-section measurement [74]. To reconstruct small-angle elastic scatters, they perpetuated the ISR tradition by installing movable vacuum-chamber sections connected to the standard beam pipe by bellows. These Roman Pots, located about 40 m from the crossing point symmetrically on both sides of the IP, were movable in the vertical plane only. Each pot was instrumented with a wire chamber and a scintillator hodoscope. Elastic scattering was measured down to $t = -0.05$ GeV^2. The total interaction rate was monitored by the combination of forward telescopes (Fig. 7) and of the UA2 central detector, that covered the large polar-angle region.

The authors quote an overall error of 2.4% on their total cross-section measurement (Eq. (37)). Translating this uncertainty into a luminosity error using Eq. (38) together with the individual sources of uncertainty yields a luminosity precision in the 3%–4% range. This is significantly better than the 10% luminosity determination extracted from directly-measured beam parameters for the same data set [74], and comparable to the ultimate precision achieved later by the direct method (Table 3).

---

[5] For instance, state-of-the-art lattice measurements of high-luminosity optics at the LHC would typically contribute a 5 to 10% uncertainty to each of the four extrapolated single-beam sizes.

Table 4
Evolution of luminosity uncertainties at the S$p\bar{p}$S using the inelastic interaction rate.

	1983 (UA2) [77,78]	1985 (UA2) [79]	1990 (UA2) [80]
Total uncertainty	20%	8%	5%–6%

7.4. Use of reference physics processes: the inelastic interaction rate

UA2 took a different approach to luminosity determination. They monitored the rate of non-diffractive events in small-angle hodoscopes surrounding the vacuum pipe at ± 10 m on either side of the IP [75]. Each array covered one unit of pseudorapidity around $\eta = \pm 4.7$. The integrated luminosity was evaluated from the coincidence rate of the two arms. To interpret this rate in terms of an absolute luminosity required knowing the acceptance of the hodoscope for non-diffractive events, as well as the non-diffractive cross-section itself. The acceptance was estimated at about 60% using early UA5 measurements of the rapidity dependence of the cross-section [76]. The non-diffractive cross-section ($\sim$38 mb) was first calculated as the difference between the total cross-section and the sum of the elastic and diffractive cross-sections, all of which were extrapolated from lower-energy data.

Luminosity measurements were initially assigned a 20% uncertainty, estimated from the observed fluctuations during different running conditions and from the overall uncertainty on the cross-section visible to the small-angle hodoscope [77,78]. This was later reduced (Table 4) by replacing the extrapolated cross-sections by direct UA4 measurements: the uncertainty on the visible cross-section was estimated at 4% and the fluctuations at 6%, yielding an overall uncertainty estimate of 8% [79].

The luminosity coincidence above was also included in the overall trigger menu as a "minimum bias" (MB) trigger to eliminate backgrounds from sources other than $p\bar{p}$ collisions. This was both a weakness and a strength of the approach. The advantage with this arrangement was that any inefficiency of the luminosity hodoscopes would cancel in cross-section determinations. The downside was that any correlation between the MB selection defined using the luminosity counters and the physics process under study would, to some extent, bias the event selection and thus the cross-section measurement [75,77]. At a later stage the generation of an MB trigger signal was separated from the luminosity measurement by introducing special time-of-flight counters. The luminosity uncertainty quoted in these later papers was 5 to 6%, dominated by the uncertainty on the inelastic cross-section (4.7%) and that on the acceptance of the luminosity telescopes (2.3%) [80]. By that time the pile-up parameter μ was large enough that the probability for multiple $p\bar{p}$ interactions to occur in a given bunch crossing was by no means negligible, and had to be taken into account in the luminosity algorithm (Section 4.2.2.2).

It is interesting to note that the luminosity evaluation of UA2 depended on the total cross-section measurement of UA4, which in turn depended on the UA4 measurement of ratio ρ of the real to the imaginary part of the forward scattering amplitude. It later turned out that the value of ρ first reported by UA4 was too high, implying the need for a 4% correction to the UA2 cross-sections [80,81].

7.5. Pulsed operation and luminosity ratios

Starting in 1985, the S$p\bar{p}$S operated in pulsed mode, with the goal to characterize $p\bar{p}$ collisions at the highest possible c.m. energy. In DC mode, the Ohmic heating in the coils of the main magnets limited the energy to 315 GeV per beam. This could be circumvented by a pulsing scheme where the maximum SPS energy of 450 GeV could be sustained during a flat top of about 4 s, combined with a flat bottom at a lower energy (chosen to be 100 GeV). This allowed to measure quite accurately ratios of cross-sections at $\sqrt{s} = 200$ and 900 GeV.

It was demonstrated that within 1%, the luminosity was directly proportional to the energy. The key point was that the normalized emittance and the bunch populations could be kept practically constant [82] during an energy cycle (an impressive achievement from the accelerator viewpoint). In this case, and provided the IP β-functions remain unchanged, the luminosity depends linearly on the relativistic factor γ (Eq. (10)). The ratio of the luminosities at flat top and flat bottom was thus directly given by the ratio of the corresponding c.m. energies, which was known to about 0.3%. It was shown that the β-functions were indeed equal at the two energies within ± 1% in both the x and y planes, yielding an uncertainty on the luminosity ratio of the order of ± 1%. This operating mode of the S$p\bar{p}$S was used by both UA5 and UA1 to measure cross-section ratios to very high accuracy [76,83].

7.6. Comparison of luminosity-measurement strategies at the S$p\bar{p}$S

The sub-percent precision achieved at the ISR could not be replicated here, as van der Meer scans proved impractical in a single-ring collider. Table 5 offers a summary of the performance achieved by the strategies implemented at the S$p\bar{p}$S. The best results all lie in the 3%–6% range. Linking elastic scattering and total interaction rate through the optical theorem fared slightly better (3%–4%) than the other methods. The UA2 approach of normalizing to the non-diffractive cross-section

Table 5
Best performance of absolute-luminosity measurements at the CERN $p\bar{p}$ collider.

Method	Uncertainty	Experiment
Elastic scattering and total interaction rate	3%–4%	UA4
Normalization to non-diffractive cross-section	5%–6%	UA2
Single-beam parameters	5%	UA1, UA4/2

Table 6
Tevatron luminosity-performance summary [84]. The listed luminosity represents the average initial luminosity achieved over the last several months of the run.

	Period	Luminosity ($cm^{-2}\,s^{-1}$)	$\sqrt{s}$ (TeV)	Colliding bunches	Beam intensities ($p/\bar{p}$ per bunch)	β^* (m)
Initial years	1987–1989	1.6×10^{30}	1.8	6	$7/3 \times 10^{10}$	1.0–0.55
RUN I	1992–1996	16×10^{30}	1.8	6	$23/6 \times 10^{10}$	0.50
RUN II	2001–2011	340×10^{30}	1.96	36	$29/8 \times 10^{10}$	0.28

implied a precision a couple of percent worse (5%–6%). The claimed accuracy of the single-beam parameter approach bottomed out at about 5%. In the pulsed mode of operation, luminosity ratios at different energies could be measured to higher precision (1%) than the corresponding absolute luminosities.

8. Luminosity determination at the Tevatron

8.1. The experimental context

8.1.1. Collider overview

The Tevatron, the first superconducting collider ever built, was the highest-energy accelerator in the world until the LHC turned on. Protons and antiprotons were accelerated up to $\sqrt{s} = 1.96$ TeV in a common beam pipe. Physics started in 1987 and ended on 29 September 2011. There were three major stages in the operation of the Fermilab collider (Table 6), during which many of the accelerator parameters changed drastically. The initial years, in the late eighties, were followed by what has been called Run I (1992–1996) and Run II (2001–2011).[6] Over this 25-year period, the average instantaneous luminosity increased by a factor of over 200. Runs I and II were separated by a five-year shutdown, allowing significant upgrades to both the experiments and the accelerator. The c.m. energy was raised by 10%. The $\bar{p}$ source and the injector complex underwent massive upgrades [84], the number of colliding-bunch pairs increased from 6 to 36, paving the way for major luminosity improvements, and the $\bar{p}$ intensity progressively grew from 3×10^{10} to almost 10^{11} $\bar{p}$/bunch. Additional luminosity gains were achieved by reducing the transverse coupling and by lowering the IP β-function by almost a factor of two, as well as by numerous incremental improvements in the instantaneous performance and the operational efficiency of the accelerator complex [37,84].

8.1.2. Experimental physics program

When the Tevatron started up, the $Sp\bar{p}S$ was still operating and the physics panorama had not changed much yet. But over the years and with the results from LEP, the Standard Model became better and better established. Over the nineties and the first decade of the 21st century, the physics program of the Tevatron thus evolved around two main themes: direct searches for new phenomena, and increasingly precise measurements of Standard Model parameters as part of the quest for new physics.

As at the $Sp\bar{p}S$, the Tevatron experimental program was dominated by two multipurpose, mammoth detectors at separate interactions points, CDF (Collider Detector at Fermilab) and D0 (named for its location on the accelerator ring). Both exhibited the same general structure: an inner layer for reconstructing charged-particle trajectories in a magnetic field, a calorimeter layer for measuring electromagnetic and hadronic energy depositions, and a muon spectrometer. Although the overall architecture was similar, the technologies chosen for each layer were very different in the two experiments.

Initially, and during the whole of Run I, both CDF and D0 relied on scintillator hodoscopes to measure the luminosity: this had been the technology used until then by every hadron collider experiment. For Run II however, with a design luminosity of 2×10^{32} $cm^{-2}\,s^{-1}$, pile-up and radiation hardness both became significant issues. CDF then switched to Cherenkov technology. D0 continued with scintillators, but with finer-granularity hodoscopes for improved background rejection and vertex-position measurements by timing.

---

[6] For historical reasons, Run II is often further split into Run IIa (2001–2005) and Run IIb (2006–2011).

Two smaller experiments installed at the E0 interaction point, E710 [85] and its successor E811 [86], played a central role in luminosity determination at the Tevatron (Section 8.3).

8.1.3. Specific challenges in luminosity determination

Combining the achieved transverse emittances with the bunch populations and β^* values listed in Table 6 results in bunch luminosities such that the mean number of interactions per bunch crossing lies well above one. At the design Run-II luminosity of 2×10^{32} cm^{-2} s^{-1}, there were on the average 6 inelastic interactions per crossing. Such large pile-up parameters never had to be faced before. Luminosity projections for Run-II drove an upgrade of the luminometers during the 1996–2001 shutdown and, more generally, strongly influenced the luminosity-measurement strategy of the experiments.[7]

Luminosity calibration by *vdM* scans was as impractical at the Tevatron as at the CERN $p\bar{p}$ Collider, for similar reasons. The determination of the luminosity from accelerator parameters also proved very challenging (Section 8.2). These limitations naturally led to a strategy akin to that of UA2 at the S$p\bar{p}$S: use the total inelastic $p\bar{p}$ cross-section as measured by the luminosity-independent method (Section 8.3), and infer the luminosity from the inelastic interaction rate (Section 8.4).

8.2. Single-beam parameters from accelerator instrumentation

The method based on the direct measurement of single-beam parameters never gave very good accuracy at the Tevatron. In the first $p\bar{p}$ cross-section measurement at the new collider, the E710 collaboration quoted [87] a precision of $\pm 15\%$ on the luminosity (unfortunately without documenting it in detail); this was by far the dominating error on this early measurement of the total cross-section at $\sqrt{s} = 1.8$ TeV. Subsequent measurements of the luminosity using accelerator parameters remained in the same range.

The luminosity determination [37] was based on separate measurements of optical lattice parameters, beam emittances and bunch intensities. The optical functions were measured by resonant excitation using the Orbit Response Matrix formalism [88], yielding in the best case a precision of about 5% on the β functions. The emittances were extracted from transverse beam profiles measured using several complementary techniques (wire scanners, synchrotron-light and ionization profile monitors), but never reached a precision better than 15%. The presence of non-negligible residual dispersion at the IP's, and of significant nominal dispersion at the profile monitors, complicated the extrapolation of the beam sizes from the profile monitors to the collision points and implied that in addition to the β functions, both the dispersion functions and the momentum spread of each beam had to be measured accurately.

The beam intensities were measured by wall-current monitors (WCM), that detect the image charge induced by the beam in the vacuum chamber. The WCM signals were processed by a Fast Bunch Integrator [89], allowing measurements of individual bunch populations. The absolute intensity scale was ultimately anchored to that of a direct-current (DC) current transformer (DCCT), that measured the total $(p + \bar{p})$ beam current with an absolute accuracy of about 1%–2%.

An additional limitation arose from the 50 cm bunch length, first comparable to, and later larger than the β-function at the IP (Table 1), that forced a significant hourglass correction (Section 2.3.2). Combining all the above uncertainties resulted in an overall precision on the luminosity from single-beam parameters in the 15%–20% range.

8.3. Elastic $p\bar{p}$ scattering, optical theorem and total interaction rate

Determination of the total cross-section via simultaneous measurements of small-angle elastic scattering and of the total inelastic rate (Section 5.4.1) was pursued both by the CDF collaboration [90] and by the E710/E811 tandem [85,86], two consecutive experiments dedicated to this purpose.

For the elastic part of the experiment, both CDF and E710/E811 used Roman Pots to approach the beam within 2–4 mm of its axis. The CDF pots were equipped with silicon detectors and drift chambers for tracking, while E811 used high-resolution scintillating fibers. The inelastic detectors of E710/811 were annular scintillation counters that surrounded the beam pipe over the pseudorapidity range $3.9 < |\eta| < 6.5$. The CDF detectors covered a similar kinematic range $(3.2 < |\eta| < 6.7)$ with scintillators; in addition, a system of small time-projection chambers around the beam pipe provided tracking capability over part of the η range.

Even though the goal of these three experiments was to measure the total and inelastic $p\bar{p}$ cross-sections rather than calibrate the luminosity scale, their results strongly impacted all absolute luminosity determinations at the Tevatron, and thereby the precision of many CDF and D0 cross-section measurements. E710 measured a total $p\bar{p}$ cross-section of 72.1 ± 3.3 mb, later superseded by the more precise but almost identical result of E811 (71.7 ± 2.0 mb). CDF on the other hand reported a value of 80.0 ± 2.2 mb, resulting in a 2.8 σ discrepancy (Fig. 8) that was never resolved. It is interesting to note that CDF also performed a measurement at $\sqrt{s} = 546$ GeV, the same c.m. energy as for the UA4 experiment at the S$p\bar{p}$S, and obtained good agreement with the UA4 results.

---

[7] It is worth noting, for instance, that because of pile-up-related non-linearities (Section 4.2.4) and of bunch-dependent luminosity backgrounds, it became important to compute the total instantaneous luminosity as the sum of individual bunch-by-bunch luminosity measurements, rather than extract it from a naive sum of the counting rate over all bunches.
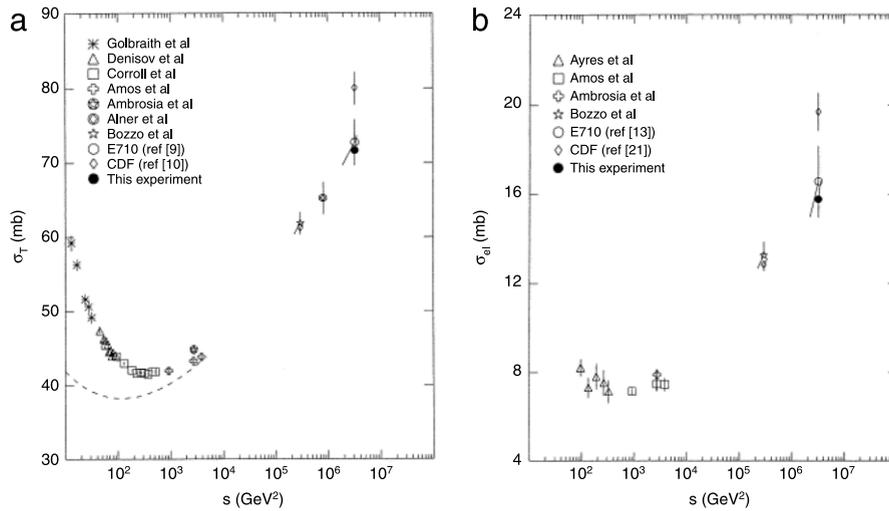
Fig. 8. Compilation of the energy dependence of (a) total and (b) elastic $p\bar{p}$ cross-section measurements. "This experiment" refers to E811.
Source: Figure reproduced from Ref. [86], with permission from Elsevier.

8.4. Use of reference physics processes: the inelastic interaction rate

Converting rates to luminosity, with the absolute scale anchored to the measured inelastic cross-section and the luminometer acceptance inferred from Monte Carlo simulations, was the main method employed at the Tevatron. Both the central value of this normalization factor and its uncertainty were always affected by the discrepancy discussed above. It would have been natural for CDF to use their own measurement of the inelastic cross-section. But to facilitate the comparison between CDF and D0 results, the two Collaborations agreed to use a common value, derived using a somewhat involved averaging procedure [91] of the CDF and E811 measurements at $\sqrt{s} = 1.8$ TeV, and extrapolated to a c.m. energy of 1.96 TeV to provide a luminosity calibration applicable to Run II data. The resulting inelastic cross-section of 60.7 ± 2.4 mb was thus used for all luminosity determinations in Run II; its 4% error, which takes into account [91] the discrepancy between the CDF and E811 results, dominates the uncertainty of all Tevatron luminosity measurements.

Though CDF and D0 based their luminosity normalization on the same inelastic cross-section, in Run II they resorted to very different techniques to measure the inelastic event rate. CDF was the first collider experiment to abandon the traditional scintillator hodoscopes, turning instead to gas Cherenkov counters operating near atmospheric pressure [92–94]. Each counter consisted of a 1–2 m long aluminum tube, constructed out of $100\,\mu$m aluminized Mylar and filled with isobutane. There were 48 such tubes on each side of the IP, arranged in three concentric layers of 16 tubes each around the beam pipe and covering the pseudorapidity range $3.7 < |\eta| < 4.7$ (Fig. 9a). The Cherenkov light was collected at the far end of the tube by fast, 2.5 cm diameter PMT's with thin quartz windows. The single-particle response, the time resolution[8] and the readout architecture were optimized for strong background rejection and for luminosity determination by particle counting.

The proponents of this approach put forward several advantages compared to traditional scintillator counters, that come in addition to the obvious radiation-hardness argument. First, the geometry of these long tubes is such that primary particles from the $p\bar{p}$ interaction travel almost parallel to the tube axis, while secondaries produced in the beam pipe cross the counters at different angles, with correspondingly shorter path lengths and smaller signals. A second advantage is related to the Cherenkov process itself, that makes the detector less sensitive to shower debris (most remain below threshold) and to background from incoming-beam halo (thanks to the directionality of the light-production mechanism).[9] A third argument is related to pile-up. At high bunch luminosity, two particles from two distinct $p\bar{p}$ collisions can traverse simultaneously the same tube. If one only considers the number of counters hit, the raw hit count is not linear with luminosity (as expected from Eq. (24)). In contrast, and because there are no Landau fluctuations in Cherenkov (as opposed to scintillation) light, the peaks corresponding to multiple traversing particles are more easily separated in a Cherenkov counter. This makes it possible to count particles rather than hits, thereby mitigating the uncertainties associated with the modeling of the non-linearities (Sections 4.2.3–4.2.4). This argument is illustrated by Fig. 9b: although not negligible, the intrinsic non-linearity of the particle-counting algorithm is much less pronounced than that associated with naïve hit counting [94].

---

[8] The combination of long bunches and of an excellent time resolution (< 100 ps) opened the possibility of measuring μ directly by counting the number of time clusters [92].

[9] The design of the ATLAS LUCID luminometer [25,28] at the LHC was driven by the same two considerations.
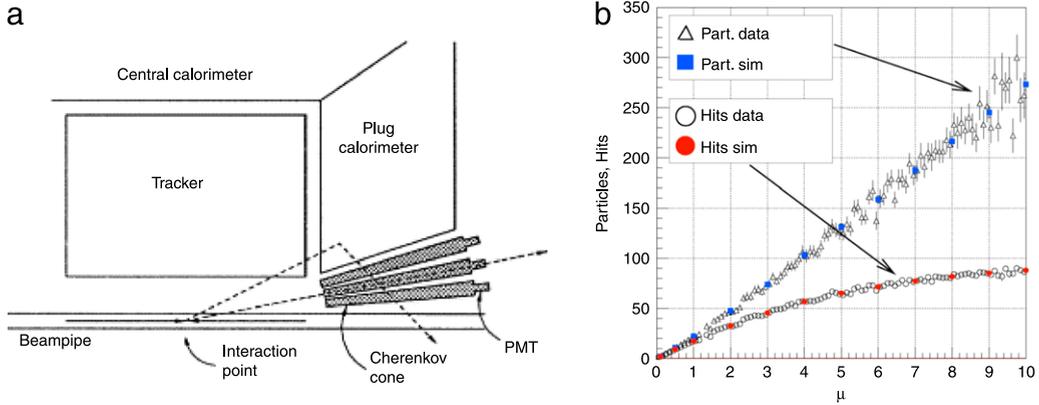
Fig. 9. (a) Schematic view of the luminosity monitor inside a CDF quadrant. Figure reproduced from Ref. [92], with permission from Elsevier. (b) μ dependence of the average number of hits (data: open circles; simulation: red filled circles) and of detected particles (data: open triangles; simulation: blue filled squares) reported by the CDF Cherenkov luminosity counter. Figure reproduced from Refs. [37,94], with permission from Elsevier.
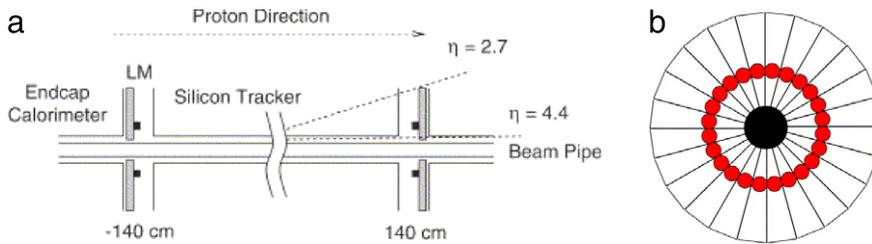

Fig. 10. (a) Schematic view showing the location of the D0 luminosity monitors (LM) in the r–z plane. (b) Geometry of one of the luminosity-monitor arrays in the plane perpendicular to the collision axis. The location of the PMT's is indicated by red filled circles.
Source: Figures reproduced from Ref. [96], with permission from Elsevier.

An alternative strategy to measure the luminosity in the presence of multiple interactions per bunch crossing is that of event- (or zero-) counting algorithms. As detailed in Section 4.2.2, the principle is to measure the fraction of bunch crossings with no detected $p\bar{p}$ interaction, from which one can infer, under the assumption of Poisson statistics, the mean number μ_{vis} of visible inelastic collisions per crossing. The method has the advantage that one need not distinguish between one, two or more interactions; the downside is that it becomes unusable at very high pile-up because of zero starvation. In practice, coincidence-based event counting (Section 4.2.2.2) was chosen as the baseline by both CDF and D0. Hit- and particle-counting ended up being used for cross-checks only, presumably because they are intrinsically more sensitive to instrumental drifts and to threshold effects.

The D0 experiment also had to adapt its luminosity-measurement strategy to cope with the harsher environment of Run II: shorter bunch spacing, higher pile-up, larger radiation-dose rates, axial magnetic field, etc. The upgraded luminosity monitor [95,96] consisted of two arrays of 24 scintillators each, one on each side of the IP (Fig. 10a) and covering the pseudorapidity range $2.7 < |\eta| < 4.4$. Background rejection and vertex-position reconstruction by timing drove the design. The wedge-shaped scintillators (Fig. 10b) were read out by fine-mesh PMT's designed to operate in high magnetic fields, and positioned so as to optimize timing resolution and minimize tube aging.

Table 7 summarizes the uncertainties affecting the absolute integrated luminosity reported by CDF and D0. For Run IIb,[10] the overall uncertainty of the CDF luminosity measurement was estimated at 5.8% [37], almost evenly shared between the 4% uncertainty on the inelastic cross-section and the combination of several smaller sources. The most prominent of these were the assumed relative contributions of diffractive and non-diffractive processes (2%) and the description of the detector material (3%) in the simulation, both of which impacted the calculation of the luminometer efficiency.

The final error reported by D0 is 6.1% for Run IIa, and 4.3% for Run IIb; a detailed breakdown of the uncertainties is documented in Refs. [98,99] respectively. For Run IIb, the D0 error is dominated by the 4% uncertainty on the inelastic cross-section, with an additional 0.9% associated with the fraction of the diffractive cross-section attributed to the single-diffractive process. The time resolution, bunch-by-bunch capabilities and granularity of the luminometer system proved crucial in controlling the sum of all detector-related uncertainties down to 1%: effects such as material-budget uncertainties, geometrical imperfections and determination of the luminometer acceptance and efficiency contribute 0.7%, while radiation damage and instrumental stability add about 0.8% to the overall D0 luminosity uncertainty.

---

[10] The CDF Run-I luminosity analysis can be found in Ref. [97].

Table 7
Luminosity uncertainties for Run IIb of the Tevatron, using the inelastic interaction rate.

Source of uncertainty	CDF	D0
Reference inelastic $p\bar{p}$ cross-section	4.0%	4.0%
Diffractive fraction	2.0%	0.9%
Simulation of detector material (CDF)	3.0%	
Other small contributions (CDF)	2.0%	
Geometry, material budget, luminometer acceptance (D0)		0.7%
Radiation damage & instrumental stability (D0)		0.8%
Total	5.8%	4.3%

Table 8
Beam parameters for the 2013 *pp* run at RHIC [22] *vs.* LHC design parameters [7].

	RHIC (2013)	LHC (design)
Center-of-mass energy (GeV)	510	14000
Number of colliding-bunch pairs	111	2808
Bunch intensity (10^{11} p/bunch)	1.85	1.15
Normalized emittance (μm-radians)	4.1	3.75
β^* (cm)	65	55
RMS bunch length σ_z (cm)	60	7.5
σ_z/β^*	0.92	0.14
Nominal full crossing angle 2α (μrad)	0	285

8.5. Comparison of luminosity-measurement strategies at the Tevatron

Luminosity determination based on the direct measurement of single-beam parameters achieved an accuracy of 15% at best. Monitoring the inelastic interaction rate normalized to luminosity-independent measurements of the inelastic cross-section fared significantly better (4%–6%), with the associated uncertainty dominated by the discrepancy between CDF and E811 on the reference cross-section.

9. Luminosity determination at RHIC

9.1. The experimental context

9.1.1. Collider overview

The genesis of the Relativistic Heavy-Ion Collider [6] dates back to 1971, when design studies started for the ISABELLE project, a superconducting $200 + 200$ GeV *pp* collider to be sited at Brookhaven National Laboratory (BNL). Following major difficulties with the superconducting magnets, taking into account the strong competition from the highly successful S$p\bar{p}$S and the upcoming Tevatron, and even though the tunnel had already been excavated, ISABELLE was cancelled in 1983 in favor of building the ill-fated Superconducting Super Collider (SSC). When the ISABELLE project was discontinued, the nuclear-physics community seized the opportunity to use the existing tunnel and BNL accelerator infrastructure for a heavy-ion collider. The project was approved in 1991 and has been in operation since 2000, providing ion–ion (both equal and different species) and *pp* collisions.

RHIC consists of two independent rings, with the beams crossing at six IP's, four of which can house an experiment. Parameters representative of recent *pp* physics running are compared to nominal LHC values in Table 8. The normalized emittances and IP β-functions are similar; but the RHIC bunches are much longer, which complicates the luminosity determination (Section 9.2).

9.1.2. Experimental physics program

The two general-purpose detectors for heavy-ion physics, STAR and PHENIX, concentrate on hadronic final states and electromagnetic probes respectively [100]. Two smaller experiments, now completed, pursued more specialized physics goals: BRAHMS and PHOBOS. There was also a dedicated proton–proton experiment called PP2PP, now associated with STAR and aimed at measuring the total and the differential elastic *pp* cross-section, including polarization effects.

A unique feature of the RHIC proton program is the possibility to accelerate and collide polarized protons. The spin programs of PHENIX and STAR (and of BRAHMS until 2006) center around the measurement of single and double spin-asymmetries (Section 9.3), both longitudinal and transverse, in elastic scattering and inclusive hadron production. Some of the results are of high relevance for understanding the gluon contribution to the spin of the proton.

Relative-luminosity monitoring at RHIC relies on scintillator hodoscopes (dubbed BBC's, for beam–beam counters) and on zero-degree calorimeters (ZDC's). Both detectors are operated in coincidence-counting mode (Section 4.2.2.2) to suppress

single-beam backgrounds. The BBC's are installed at about ± 3.5 m (± 1.4 m) on either side of the STAR (PHENIX) IP, with large enough an acceptance for inelastic *pp* collisions to provide an adequate counting rate at moderate luminosity. The ZDC's are in both cases located approximately ± 18 m from the IP, beyond the transition from a common to two distinct beam pipes. Their primary purpose is the detection of the forward neutrons that are abundantly produced in ion–ion collisions; in *pp* mode, the ZDC visible cross-section is one to two orders of magnitude smaller than that of the BBC's.

Because the RHIC bunches are very long ($\sigma_z \sim 0.6 - 1$ m), the PHENIX BBC's are located too close to the IP to provide full coverage for *pp* collisions, and their acceptance is sensitive to small changes in the length or the longitudinal position of the luminous region. The BBC's also showed symptoms of detector and/or electronics saturation effects at high luminosity [101]. The PHENIX and STAR ZDC's are now used by the accelerator team as luminometers not only for ion–ion runs, but also during *pp* operation when the luminosity is high enough for them to provide an adequate counting rate; the BBC's, however, continue to provide the primary luminosity information for many physics analyses.

While the baseline strategy for absolute-luminosity calibration, discussed below, relies on *vdM* scans, the PP2PP Collaboration proposed to exploit the Coulomb-nuclear interference method (Section 5.4.2) to achieve an absolute precision of 1% on the absolute luminosity [102].

9.2. van der Meer scans

Absolute luminosity calibration by van der Meer scans is a natural choice at RHIC, both because the two-ring configuration makes magnetic beam separation relatively straightforward, and because the *vdM* method is equally applicable to *pp* and ion–ion collisions. Compared to the ISR, the achieved precision has been limited by a combination of intrinsic challenges and of RHIC-specific operational constraints.

9.2.1. Bunch-intensity and collision-rate measurements

Precise beam-intensity measurements are more demanding for bunched than for continuous beams because of the higher frequencies the electronics have to deal with. The chosen strategy, detailed in [103] and references therein, is to base the absolute beam-intensity scale on highly precise total-current measurements performed by DCCT's, as at the LHC (Section 10.3). The single-bunch population is measured by wall-current monitors (WCM's), that can distinguish individual bunches but suffer from larger scale uncertainties. The total intensity reported by the WCM is cross-calibrated to that of the corresponding DCCT at the end of the acceleration ramp before partial debunching can occur, and the difference between the DCCT- and WCM-based beam intensities, as it evolves during the fill, provides an estimate of the debunched-beam component in each ring. This procedure results in a 1%–3% systematic uncertainty being assigned to the average bunch population in each beam.

Absolute-luminosity calibrations at RHIC are based on the collision rate summed over all bunches. As the current varies from one bunch to the next and as not all bunches collide at all IP's, a pattern-dependent correction of up to 10% [12,103] must be applied to the summed bunch-population product to account for the actual intensity of those bunch pairs that collide at a given IP. This procedure relies on the assumption that bunch-to-bunch emittance and pile-up variations can be neglected, a hypothesis that is not tenable at the LHC (Section 10.1.3.1).

9.2.2. Scan protocol

Operating with bunched beams instead of the continuous "ribbon" beams of the ISR, forces to scan in two orthogonal directions rather than just vertically. Not only does this double the number of closed-orbit bumps and steering correctors (each of which is a potential source of calibration and hysteresis errors): it also potentially introduces cross-talk between horizontal and vertical scans via residual *x–y* coupling in the ring optics. Fringe-field-induced orbit distortions during the scans, either of the scanned beam in the direction orthogonal to the scan, or of the unscanned beam, also proved a concern.[11]

Diagnosing the relevant issues primarily relies on the beam-position monitor (BPM) system, as does the evaluation of the associated corrections and systematic uncertainties. As experienced at other colliders, *vdM* scans at RHIC put stringent demands on the BPM system, in terms both of reproducibility and of absolute beam-position accuracy. The interplay of BPM performance limitations with optics-reproducibility and orbit-stability issues remain a major contributor to the absolute-luminosity uncertainty at RHIC [19,101,103,104].

9.2.3. Beam-separation scale

The beam displacements during the scans are carried out using closed orbit bumps that involve four steering correctors per IP, per plane and per beam. The absolute beam separation can in principle be computed by combining the measured magnetic response of the correctors, the calibration of the magnet power supplies and the nominal properties of the optical lattice. Calibrating the nominal beam displacements against those of the luminous centroid was attempted in the early

---

[11] These issues were expected to similarly affect *vdM* scans at the LHC, but most of them have proven negligible so far. The LHC benefits from significantly smaller alignment and optical imperfections and from outstanding orbit stability. It also profited from the *vdM*-scan experience accumulated on previous CERN machines and at RHIC.
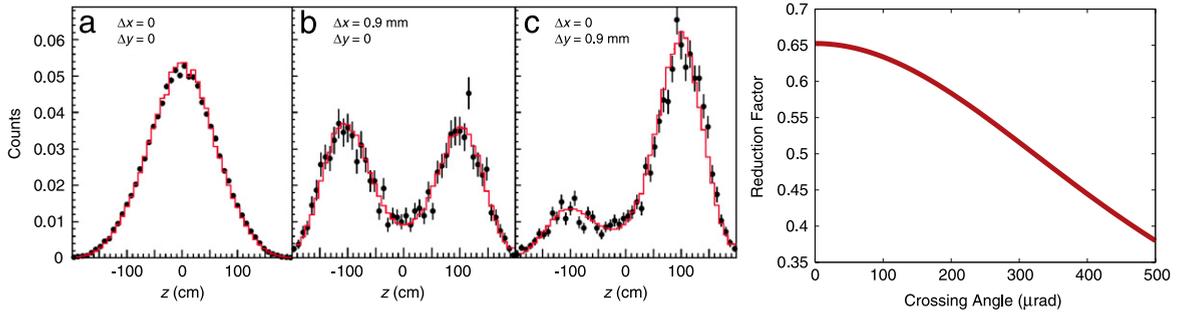
Fig. 11. Left: *z*-distribution of collision vertices measured by the PHENIX ZDC's during a *vdM* scan at $\sqrt{s} = 200$ GeV (black points), and *z*-dependence of the luminosity (red histograms) calculated by convolution of the longitudinal bunch-intensity profiles including the hourglass effect and assuming $\beta^* = 1$ m, $\sigma_z = 1$ m, $\sigma_{x,y} = 0.3$ mm, for (a) head-on beams; (b) one beam displaced horizontally by 0.9 mm relative to the other (illustrates the hourglass effect); (c) one beam displaced vertically by 0.9 mm (the asymmetry is caused by the 0.15 mrad vertical crossing angle). Figure reproduced from Ref. [106] with permission,© (2009) by the American Physical Society. Right: luminosity-reduction factor due to the combined impact of the crossing angle and of the hourglass effect, for the parameters of the RHIC 2009 *pp* run [12]. Figure reproduced from Ref. [12] © CERN.

years using vertices reconstructed by the STAR TPC [105]. The absolute separation scale is now systematically cross-checked against the beam displacements measured by BPM's located 8.3 m on either side of each IP, inboard of the beam-separation dipoles (DX). The consistency of the nominal beam displacements (as input to the RHIC control system) with those inferred from the BPM readings, and the reproducibility of such comparisons, provide measures of the systematic uncertainty on the absolute beam-separation scale. The agreement differs from IP to IP, varies from one scan session to the next, and is typically at the level of a few percent in each of the horizontal and the vertical plane [12].

9.2.4. Crossing angle and hourglass corrections

Three additional, RHIC-specific effects impact the accuracy of the absolute luminosity calibration by the *vdM* method: the presence of hard-to-measure residual crossing-angles in both the horizontal and the vertical plane; a substantial hourglass-induced luminosity reduction caused by the large σ_z/β^* ratio; and the interplay of the hourglass- and of the crossing-angle-related corrections to the convolved beam sizes measured during the scans.[12]

If the beam-crossing plane coincides with either scan plane, i.e. in the presence of a crossing angle in one scan plane only, the geometrical luminosity-reduction factor is fully accounted for by the increase in the measured convolved beam width (Eqs. (14)–(16)), so that no explicit knowledge of the crossing angle is necessary. But if the beams cross with a non-zero relative angle in both the horizontal and the vertical plane, the reference luminosity (Eq. (30)), or equivalently the visible cross-section (Eq. (32)) must be corrected by a factor that depends on the bunch lengths, the transverse beam sizes and the two projected crossing angles [11]. This correction typically amounts to 0.5%–1.5% for RHIC parameters; the associated systematic error is of comparable magnitude because of the BPM-related uncertainties on the measured crossing angles [12].

As the bunch length σ_z is comparable to β^* (Table 8), the transverse beam sizes increase with the distance from the waist (Eq. (6)) rapidly enough to cause a substantial luminosity reduction (Fig. 11). While this is automatically accounted for by normalizing the counting rate of the luminometer to the reference luminosity computed from beam parameters, the hourglass effect also slightly biases the measured convolved beam sizes that are used to compute this reference luminosity [12,18], even when assuming zero crossing angle [101]. The associated corrections to the visible cross-section, which depend on β^*, the bunch lengths and the crossing angles, are of the order of 3%–5%. Because of the uncertainty on the measured crossing angles, these effects contribute an additional luminosity-calibration uncertainty that ranges from 0.2% to 1.3% for the 2009 *pp* run at 250 GeV [12,19].

9.2.5. Beam dynamics

RHIC is the first hadron collider where two tell-tale signatures of beam-dynamical effects were first observed [12,19], that are now highly relevant (Section 10.4.3.2) to luminosity determination at the LHC.

Luminosity-scan curves, which since the ISR days (Fig. 2) had always been described as single Gaussians, were found at RHIC to sometimes contain a significant non-Gaussian component. The physical origin of these tails does not matter as long as the overlap integral (Eq. (29)) can be evaluated to an acceptable precision. This is handled at RHIC by fitting the scan curves with the sum of two Gaussians (when necessary). At the LHC, similar observations seeded extensive studies that revealed additional systematic effects, including in particular the importance of controlling non-factorization biases.

The mutual electromagnetic deflection of two bunches that collide with a non-zero transverse offset was first observed in e^+e^- collisions at the SLC, and used at LEP as an optimization tool. Even though the beam–beam interaction is much weaker in hadron colliders, the associated orbit distortions became visible at RHIC [19], and are now sizeable enough to be corrected for [107].

---

[12] Shorter bunches, better orbit stability and specially-tailored scan procedures make these effects negligible at the LHC.

Table 9
Luminosity-calibration uncertainties $\sigma_{\mathscr{L}}/\mathscr{L}$ for the 2009 *pp* run at RHIC (250 GeV/beam), using the PHENIX (P) and STAR (S) ZDC's as luminometers.
Source: Table adapted from Refs. [12,19].

Uncertainty source	$\sigma_{\mathscr{L}}/\mathscr{L}$	Comments
Bunch-population product	3%	Correlated from scan to scan
Scan protocol & beam-separation scale	2%–10%	Both dominated by BPM systematics
Crossing angle & hourglass corrections	0.2%–1.3%	and scan-to-scan reproducibility
Hysteresis in closed-bump magnets	1%	Correlated from scan to scan
Beam–beam deflections	1%	Partially correlated scan-to-scan
Luminometer statistics	3%	Uncorrelated scan-to-scan
Total uncertainty, per scan session	4.3 to 11.6%	Smallest & largest reported uncertainty
Overall calibration uncertainty	P: 6.5%, S: 6.9%	Weighted average over 7 scan sessions

Table 10
Luminosity-calibration uncertainties $\sigma_{\mathscr{L}}/\mathscr{L}$ for recent PP2PP [104] & PHENIX/STAR [107] runs. Blank entries refer to contributions that are not explicitly discussed in the corresponding reference.

Uncertainty source	*pp*, 100GeV/beam [104] $\beta^* = 22$ m (2009)	*pp*, 255GeV/beam [107] $\beta^* = 0.65$ m (2013)
Bunch-population product	2.0%	Loss of bunched beam: 2.0% Fill-pattern variations: 1.0%
Scan protocol & beam-separation scale	1.5%	2.0%
x–y orbit cross-talk during scan	1.0%	–
Ring-to-ring cross-talk during scan	1.0%	–
Crossing angle & hourglass corrections	negligible	2.0%
Beam–beam deflections	–	corrected for
Correction for accidental coincidences	1.0%	4.0%
Luminometer statistics	0.8% (BBC's)	2.0%–3.5% (ZDC's)
Overall calibration uncertainty	3.1%	5.7%–6.4%

9.2.6. Precision of the absolute luminosity determination at RHIC

One of the most extensively documented analyses [12,19] of absolute luminosity calibration in *pp* collisions at RHIC is summarized in Table 9. The total (statistical + systematic) uncertainty on the absolute luminosity scale is approximately 7%; comparable errors were reported in earlier – albeit terser – accelerator publications [101,103,105].

This precision applies to absolute calibrations performed in a controlled accelerator environment. The uncertainty reported by RHIC experiments on the integrated luminosity used in absolute cross-section measurements is often larger, in part because of trigger- and other detector-related uncertainties. For instance, and for the same running period at $\sqrt{s} = 500$ GeV, STAR quotes a 13% luminosity uncertainty, dominated by possible non-Gaussian components in the beam profiles (10%), potential gain drifts in the barrel calorimeter (5%) and bunch-current uncertainties (4%) [108]. Similarly, an early PHENIX publication at $\sqrt{s} = 200$ GeV quotes a precision of 9.6%, broken down in a 3.2% "absolute" error and a 6.4% scan-to-scan error added linearly [109]. At $\sqrt{s} = 62$ GeV, PHENIX reports an 11% uncertainty [106]: the combination of a long luminous region, of the hourglass effect and of the *z*-dependent BBC efficiency dominate this error (10%), with the bunch-intensity and the beam-separation scale contributing an additional 4%, and the hourglass corrections to the convolved beam sizes 2%.

More precise absolute calibrations (3%–4%) have been reported by the RHIC accelerator group for a special *pp* run dedicated to total cross-section measurements [104]. This result is compared in Table 10 to the precision achieved in the most recent high-luminosity *pp* run. The large β^* value required by the PP2PP physics program all but eliminates the systematic uncertainties associated with crossing-angle measurements and hourglass corrections; the corresponding luminosity per bunch is much lower, significantly reducing the pile-up related uncertainty from 4% to 1%, and the specialized fill pattern and lower beam currents contribute to a somewhat reduced error on the bunch-population product.

The luminosity-calibration methodology for ion collisions closely parallels that in the *pp* case, and the precision achieved on the absolute luminosity scale is comparable, as illustrated for instance in Ref. [110].

9.3. Relative luminosities for different helicities

As each beam is polarized, both single and double spin-asymmetries can be measured. The total polarization is 80% at the source; it is preserved through the acceleration phase thanks to "Siberian Snakes" and, after years of effort, reached over 50% in collision at 255 GeV/beam. The stable spin axis is vertical, but longitudinal polarization can be provided by spin rotators

close to the IP. The spin orientations are distributed along the bunch trains such that all four possible helicity combinations of colliding protons occur in regular succession.

Measuring single spin asymmetries requires the determination of the relative-luminosity ratio between all the spin-up and all the spin-down bunches of one beam while averaging over the spin states of the other beam. In the case of double spin asymmetries, the four different combinations of possible relative luminosities reduces to one relative measurement if parity-violating effects are neglected [111]: the relevant ratio is then that of the relative luminosities of protons colliding with the same, and with the opposite, helicity.

Spin asymmetries have been measured for a number of final states, both for transversely and for longitudinally polarized beams. In general, single (resp. double) spin asymmetries are measured for transversely (resp. longitudinally) polarized protons. The systematic uncertainty on the relative luminosity always lies in the range of 10^{-3}–10^{-4} for both single and double asymmetries. Such small errors are essentially due to the fact that the luminometers can resolve the bunch structure and that each bunch crossing is uniquely related to a spin orientation. The systematic error is estimated by comparing the response of different luminometers, that select different physics processes and are sensitive to different kinematic ranges [106]. To reduce potential uncertainties associated with correlations of the measurement within the bunch structure, the pattern of spin combinations is cycled between four different configurations [112]. In some cases the systematic uncertainty on the double spin asymmetry is in addition checked by measuring a single spin asymmetry that is expected to be zero [113]. The error on the relative luminosity sometimes dominates the total error on the asymmetry, while in other cases it can be neglected.

10. Luminosity determination at the LHC

Compared to what was accomplished at previous *pp* colliders, luminosity determination at the LHC has already achieved significant jumps in the sophistication and the actual precision of the measurement and analysis techniques. In order to highlight the interesting recent developments and the key challenges in this lively and rapidly-moving field, the discussion below will delve into technical issues in more depth than some of the previous chapters.

This section is structured as follows. The harsh experimental environment, combined with renewed physics demands on high-precision measurements, raised unprecedented challenges in luminosity determination (Section 10.1), that prompted the development of multiprong strategies for luminosity monitoring and calibration (Section 10.2). The high-precision determination of bunch populations (Section 10.3) is a central piece of the two otherwise independent methods that have been brought to bear to calibrate the absolute-luminosity scale: van der Meer scans (Section 10.4) and beam imaging (Section 10.5). Absolute calibration using elastic *pp* scattering is also being pursued (Section 10.6). Finally, the precision of a one-time absolute calibration is only part of the story: the long-term stability and the internal consistency of relative-luminosity measurements must be taken into account when evaluating the total uncertainty that affects the integrated luminosity associated with a physics data set (Section 10.7).

10.1. The experimental context

10.1.1. Collider overview

The LHC [7], at today's energy frontier, represents in many respects a major leap relative to the Tevatron and to RHIC. The physics program, centered on elucidating electroweak symmetry breaking and the search for new physics, set ambitious goals: a factor of seven increase in c.m. energy, and four (resp. two) orders of magnitude improvement in instantaneous luminosity compared to the Tevatron design (resp. achieved) performance. These targets posed enormous challenges in terms of superconducting magnet technology, industrialization and quality control; of large-scale cryogenic systems operating with superfluid He in a high-radiation environment; of safe handling of unprecedented stored-beam energy by multi-stage, high-power collimation systems; of impedance control and mitigation; of production and stabilization of high-brightness beams; and of low-β optics and beam–beam effects—to mention but a few. The experimental detectors proved equally demanding.

During its first run (December 2009–February 2013), the LHC delivered *pp* collisions for physics at c.m. energies of 0.9, 2.76, 7.0 and 8.0 TeV, as well as *PbPb* and *pPb* collisions at c.m. energies per nucleon of 2.76 and 5.02 TeV respectively. The luminosity measured during physics running varied from about 2×10^{25} (in lead–lead collisions) to 7.7×10^{33} cm^{-2} s^{-1} in high-energy *pp* collisions. From the viewpoint of luminosity determination, the LHC combines and in most cases magnifies the challenges encountered at the Tevatron and at RHIC: higher-radiation environment, more-than-doubled pile-up parameters, over a thousand colliding bunches, significant nominal crossing angles, beam-dynamics-related biases. The saving grace is that the combination of much shorter bunches, greatly improved beam instrumentation, lower single-beam backgrounds and outstanding reproducibility of accelerator conditions mitigated some of the difficulties encountered at previous hadron colliders—at least so far.

10.1.2. Experimental physics program

The LHC rings include eight straight sections, four of which are utility insertions and four house experiments. Two general-purpose detectors optimized for high-luminosity discovery physics, ATLAS and CMS, occupy the diametrically
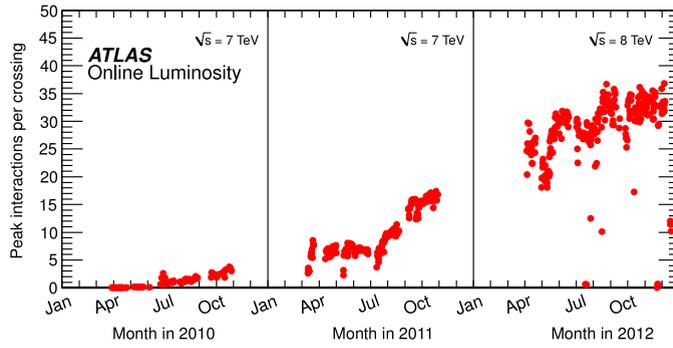
Fig. 12. Run-1 history of the maximum pile-up parameter in *pp* collisions at IP1, averaged over all bunch crossings and restricted to stable-beam periods (courtesy ATLAS Collaboration).

opposed IP's 1 and 5. The ALICE and LHCb experiments, devoted respectively to heavy-ion physics and to *B* physics and CP violation, are installed at interaction points 2 and 8, and operate at luminosities one to three orders of magnitude lower than ATLAS and CMS.

The smaller TOTEM experiment shares IP5 with CMS. It is dedicated to diffractive physics and to elastic and total cross-section measurements. The LHCf detector, that consists of two calorimeters installed 140 m on either side of IP1 beyond the beam-separation dipoles, aims at measuring forward production of neutral particles over a range of c.m. energies relevant to modeling cosmic-rays showers.

10.1.3. Specific challenges in luminosity determination

The demands of luminosity determination at the LHC represent a major step compared to what had been achieved at earlier hadron colliders, because of the combination of collision parameters, radiation environment and precision goals.

10.1.3.1. Collision parameters. First and foremost, the pile-up parameter μ at the ATLAS and CMS IP's has been rising from around 1 (typical of late S$p\bar{p}$S operation) at LHC startup, to beyond 35 (about twice the Tevatron record) by late 2012 (Fig. 12). At the same time, specialized measurements such as that of the elastic *pp* cross-section require beam conditions that result in $\mu \ll 1$. The linearity of the luminometers over this unprecedented range of pile-up parameters thus becomes a central consideration, be it in terms of instrumental design requirements, detector stability, calibration strategy or systematic uncertainties.

Next, the significant bunch-to-bunch intensity and emittance variations typical of hadron accelerators, combined with the intrinsic non-linearities induced by pile-up, impose that the luminosity infrastructure and the associated absolute-calibration protocols support bunch-by-bunch luminosity determination (Section 4.2.4). LHC luminometers must therefore be able to accommodate bunch-to-bunch spacings as short as 25 ns but with occasionally much longer gaps between bunches. This requirement, combined with the unprecedented LHC collision rates and particle densities, carries significant challenges in terms of detector technology tradeoffs, electronic signal shaping, baseline restoration, linearity of response both within and between bunch trains, data-acquisition architecture, etc.

The higher the efficiency of a luminosity algorithm, the lower the pile-up parameter at which this algorithm saturates (Section 4.2): high-μ operation therefore disfavors wide-acceptance, low-granularity luminometers such as the scintillator arrays used at previous colliders. But narrow-acceptance devices such as the ATLAS BCM (see Table 11), which remain reasonably well behaved up to the highest pile-up parameters reached so far, become statistically limited in the low-luminosity regime of elastic-scattering measurements, or in some cases even in the tails of van der Meer scans. Similar considerations apply to luminosity monitors whose rate capabilities are constrained by trigger-bandwidth considerations. This leads to the concept of *preferred luminosity algorithm*: luminosity determination no longer is the purview of a single device or method, but is best shared among multiple, complementary approaches. These are cross-checked against each other to evaluate systematic uncertainties, and the "best-suited" one is chosen to provide the reference absolute luminosity for a specific running period or set of beam conditions.

10.1.3.2. Radiation environment. In contrast to previous experience, the particle flux through the LHC particle detectors is dominated [114] by *pp* collision products, and dwarfs the single-beam halo created by aperture losses or beam-gas interactions. The resulting integrated radiation dose effectively precludes the use of radiation-soft technologies such as scintillator arrays for luminosity determination at the high-luminosity IP's.[13] It also induces significant aging effects (sometimes accompanied by partial recovery during beam-off periods) in some subdetectors, that need to be carefully

---

[13] The ALICE and LHCb detectors do rely in part on scintillator arrays, but they operate at luminosities one to three orders of magnitude lower than the ATLAS and CMS experiments.

Table 11

Luminometer technologies and luminosity algorithms at the LHC. The luminosity algorithms are defined in Section 4.2; the other abbreviations are explained in the text.

| Luminosity from	Luminometer			Luminosity Algorithm	Ref.
	Detector type	Data flow	Acronym		
ALICE	Scintillator array	bbb, DTI	V0	Event_AND	[118]
ATLAS	pCVD-diamond pads	bbb, DTI	BCM	Event_OR	[26]
				Event_AND	[26]
	Gas/quartz Cherenkov tubes	bbb, DTI	LUCID	Event_OR	[25]
				Event_AND	[25]
				Hit counting	[26]
	Si strip +pixel tracker: # vertices	bbb, PD	"Vtx"	Vtx counting	[26]
	Si strip +pixel tracker: # tracks	bbb, PD	"Trks"	Trk counting	[26]
	Fwd LAr calorimeter: gap currents	ba, DCS	FCal	Particle flux	[26]
	TILE calorimeter: PMT currents	ba, DCS	TILE	Particle flux	[26]
CMS	Pixel tracker: # clusters	bbb, SD	"Pixel"	Hit counting	[115]
	Si strip +pixel tracker: # vertices	bbb, SD	"Vtx"	Vtx counting	[115]
	Forward Fe/quartz calorimeter	bbb, DTI	HF	Hit counting	[27]
LHCb	VELO vertex tracker: ≥ 2 tracks	bbb, ST	"Tracks"	Event_OR	[31]
	VELO vertex tracker: ≥ 1 vertex	bbb, ST	"Vertex"	Event_OR	[31]
	Si pile-up tracker: ≥ 2 hits	bbb, ST	PU	Event_OR	[29]
LHC diagnostics	Ar-N_2 ionization chamber	bbb, DTI	BRAN	Event_OR	[119]
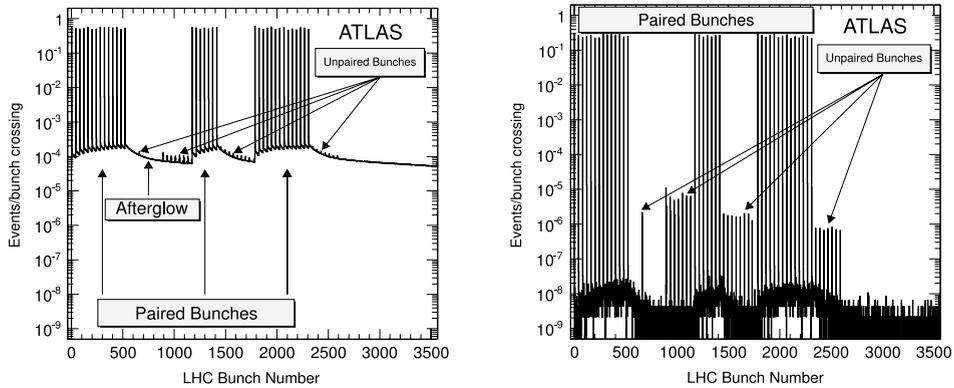
Fig. 13. Bunch-by-bunch event rate per bunch crossing in LHC fill 1309, as recorded by an ATLAS algorithm that requires at least one hit in either detector arm (Event_OR algorithm, left), or at least one hit in both arms (Event_AND algorithm, right) within the same BCID.
Source: Figure reproduced from Ref. [25], with permission from the European Physical Journal.

monitored and precisely corrected for. Even the dose rate on time scales of minutes to hours can have a noticeable impact, for instance via charge-pumping effects in diamond sensors [26].

The high luminosity and tight bunch spacing induces a peculiar type of *afterglow* background, that affects luminosity measurements at levels of up to a few percent during physics running, depending on the detector and algorithm considered [25,26,115,116]. Fig. 13 displays the event rate per bunch crossing, as measured as a function of the bunch-crossing identifier[14] (BCID) and time-averaged over an early, low-luminosity run that lasted about 15 h. For this run, 35 bunch pairs collided at the ATLAS IP: these are called "colliding" (or "paired") BCIDs. Bunches that do not collide at the considered IP are labeled "unpaired". The structures observed in this figure are visible in the bunch-by-bunch luminosity distributions of all LHC detectors, albeit with relative magnitudes and time constants that depend on the instrumental characteristics and on the local material distribution.

Requiring at least one hit in either detector arm reveals a complex time structure (Fig. 13, left). The colliding bunches are clearly distinguished, with a rate of about four orders of magnitude above background. BCIDs from unpaired bunches appear as small spikes above the afterglow background. These spikes are the result of beam-gas and beam-halo interactions; they may also contain a very small fraction of *pp* collisions between an unpaired bunch in one beam and a satellite-, ghost- or debunched-proton component in the opposing beam.[15]

---

[14] The LHC beam is subdivided into 35 640 RF-buckets of which nominally every tenth can contain a bunch. Subtracting abort and injection gaps, up to 2808 of these 3564 "bunch slots", which are 25 ns long, can be filled with beam. Each of these possible crossings is labeled by an integer beam-crossing identifier.

[15] In hadron rings, a small fraction of the stored bunch may slowly diffuse out of its RF-bucket, coalescing into very low-intensity "satellite" (resp. "ghost") bunches that are separated from a nominal bunch by a few (resp. a few ten) buckets, and/or generating a barely detectable unbunched beam component.

The colliding bunches are followed by a long tail[16] where the rate builds up when the paired BCIDs follow each other in close succession, but decays slowly when no collisions occur for a sufficiently long time. This "afterglow" vanishes when beams are out of collision. Its level is proportional to the instantaneous luminosity, but depends on the bunch pattern because of the long-decaying tail: it remains around 0.02% for the bunch pattern shown in Fig. 13, but exceeds 0.5% (for this same detector and algorithm) during physics operation when every other BCID is filled. Requiring a coincidence between the two arms of the luminometer suppresses the signal by several orders of magnitude (Fig. 13, right), indicating that the hits are randomly distributed. These observations suggest that the afterglow is due to photons from nuclear de-excitation, which in turn is induced by the hadronic cascades initiated by *pp* collision products. This interpretation is supported by FLUKA simulations of very similar observations in the CMS beam-conditions monitor [117].

10.1.3.3. Precision goals. The physics program of the general-purpose detectors is focused on elucidating electroweak symmetry breaking (for which the discovery of the Higgs was but a first step), and on the search for new physics (supersymmetry, new gauge bosons, extra dimensions, etc.). As such measurements are statistics- and/or background-dominated, they only require a moderately precise luminosity determination. However, QCD predictions for some Standard Model processes (e.g. W, Z and $t\bar{t}$ production) have become crisp enough to warrant high-accuracy measurements of the corresponding cross-sections.

Ref. [9] offers "a pedagogical introduction to the physics implications of a precise knowledge of the LHC luminosity, defining the goals and some benchmark accuracy targets". One of the main drivers for precision is the determination of the W and Z cross-sections. As discussed in Section 5.5.2, the dominant uncertainty on the corresponding theoretical predictions is associated with the knowledge of the PDF's; it currently hovers around 5%. Luminosity set aside, the next precision barrier arises from the determination of the experimental acceptance and detection efficiency, which together contribute a 2%–2.5% uncertainty. An absolute luminosity determination significantly better than 5% – say around 2% – can therefore provide valuable constraints on the PDF's, a precise knowledge of which is of crucial importance for the experimental determination of a number of other fundamental parameters, such as the W mass, the top mass or the electroweak coupling $\sin^2 \theta_W$ [9].

Other topics that benefit from a precise absolute luminosity determination are the total inelastic cross-section (where the luminosity uncertainty typically dominates), the dependence of QCD and electroweak cross-sections on the *pp* c.m. energy, and the comparison of absolute cross-sections in *pp*, *pPb* and *PbPb* collisions at the same c.m. energy per nucleon.

10.2. Luminosity monitoring and calibration strategies

10.2.1. Luminometers and luminosity algorithms

Table 11 summarizes the main technologies and algorithms used for luminosity monitoring at the LHC in *pp* collisions at $\sqrt{s} = 7$ and 8 TeV. For each of the four large LHC detectors, the "preferred algorithm" that provides the integrated luminosity for most *pp* physics analyses, is listed first. A different luminometer and/or algorithm may be used under special conditions such as elastic *pp* scattering measurements or heavy-ion runs: these are documented in the references provided.

Most luminometers deliver bunch-by-bunch ("bbb") luminosity measurements. This capability is essential not only because of the non-linearities intrinsic to high-pile-up operation (Section 10.1.3), but also because precision calorimetric measurements of e/γ energies, of jet transverse momenta and of missing transverse energy (which are central to many aspects of the LHC physics program) require delicate bunch-by-bunch, μ-dependent corrections of the calorimeter response. Bunch-averaged ("ba") luminosity measurements, in turn, provide crucial consistency checks, based for instance on the electrical current flowing through the liquid argon (LAr) gaps of the ATLAS forward calorimeter, or supplied to the PMT's of the hadronic calorimeter.

The "Data flow" column of Table 11 also specifies whether the luminosity information flows through:

- the standard trigger ("ST") path. In this case the luminosity measurement is integrated in the general trigger system, with a fraction of the bandwidth attributed to *luminosity triggers* that initiate the readout of the trigger information. The visible interaction rate μ_{vis} is proportional to the rate of this luminosity trigger. The strategy is subject to rate constraints (typically a few ten to a few hundred Hz) and to dead-time corrections, that under certain circumstances limit the statistical (and even sometimes the systematic) accuracy of the luminosity measurement.
- a dedicated, trigger-independent ("DTI") path. Here the luminometer readout is activated on every bunch crossing independently of the state of the main trigger and data-acquisition (TDAQ) chain. The raw luminosity information is entirely contained in the bunch-by-bunch event- or hit-counting rates averaged over a few seconds to a couple of minutes; its volume is modest enough for the data to be continuously archived in an online database. This approach makes full use of the potentially available statistics, avoids at least some of the pitfalls of dead-time corrections, and by using multiple, redundant algorithms lends itself to sophisticated consistency checks (Section 10.7).

---

[16] A much shorter tail of purely instrumental origin (signal reflections on cables, electronic pulse shaping, ion feedback in PMT's) is also present in some cases [115,116].

- the standard ("SD") or a partial ("PD") data stream. The SD approach is adopted when the luminosity determination requires access to the full event information. Here, similarly to the ST case, a full-detector readout is typically initiated by a luminosity trigger: the ST and SD strategies therefore share similar limitations. A powerful extension of this method is provided by partial data (PD) streams: rather than reading out and storing the full detector information on each luminosity trigger, only the data from those subdetector sections used by the luminosity algorithm (e.g. the silicon trackers) are archived to disk in a separate data stream. This reduces the I/O overhead by a large factor, allowing to store the full tracking information at rates exceeding 10 kHz.
- the detector control system ("DCS"). This part of the data-taking infrastructure monitors and archives detector parameters such as temperatures, voltages or currents, on time scales of seconds to minutes; it is the natural home of calorimeter-based, bunch-averaged luminosity measurements.

10.2.2. Absolute calibration strategies

In the first few months of LHC operation, the inelastic cross-section had not yet been measured in this new energy regime, and neither had the detector simulations been validated yet. Early attempts at determining the luminosity by combining the measured interaction rate with Monte Carlo-based visible cross-sections were systematically limited at the 20% level, primarily because of inconsistencies between the various physics models of inelastic pp collisions [25]. The van der Meer technique offered more promising prospects [120]; it has since become the method of choice for ALICE, ATLAS and CMS (Section 10.4). With its unique vertexing capabilities motivated by B-physics, LHCb pioneered the beam-gas imaging (BGI) of single-bunch profiles, that have been cross-checked against, and proved at least as precise as, the vdM method (Section 10.5). Both strategies require dedicated luminosity-calibration sessions, described below, as well as highly accurate measurements of bunch populations (Section 10.3).

Originally conceived as the ultimate luminosity-calibration strategy, elastic scattering methods are operationally very challenging, They have the advantage of being fully independent from both the vdM and the BGI method, and have provided interesting cross-checks already (Section 10.6). Whatever their ultimate accuracy, however, extrapolating an absolute luminosity calibration over six orders of magnitude, from the extremely low luminosity regime of elastic scattering to nominal, high pile-up LHC conditions, may prove a daunting task.

The beam conditions for vdM and BGI calibrations are different from those in normal physics fills, with fewer bunches colliding, lower bunch intensities and in some cases a special optical configuration at the IP; they are optimized to reduce various systematic uncertainties in the calibration procedure.

By the end of 2012 an LHC physics fill typically contained 1380 bunches separated by 50 ns and densely packed in a string of 11 bunch trains. In contrast, luminosity calibrations are normally performed with at most 50 bunches separated by a minimum of 1 μs. Using individual, widely spaced bunches rather than trains is dictated by the need to eliminate the parasitic crossings between incoming and outgoing bunches in the shared beam pipe within ±150 m of the IP: such encounters have been shown to cause intensity- and emittance-dependent orbit kicks and tune shifts induced by the long-range beam–beam interaction [121,122], thereby unacceptably distorting the luminosity-scan curves. Operating with few bunches also carries several advantages:

- improved control over transverse and longitudinal phase-space quality in the LHC injector chain (reduced non-Gaussian tails, weaker satellites bunches, less ghost charge);
- improved flexibility with respect to the IP orbits, in particular with respect to the crossing angle;
- reduction of the afterglow level (Section 10.1.3) by about two orders of magnitude.

The typical bunch intensity is also lowered to about half of that during physics running, in order to:

- minimize the satellite- and ghost-charge fraction that arise in part from space-charge effects at the low-energy end of the injector chain,
- minimize the systematic uncertainties affecting the total beam-intensity measurements [123],
- minimize the scan-curve distortions induced by single-bunch beam–beam effects (dynamic β, beam–beam deflections) at the IP [12,26],

while still providing a usable counting rate in the tails of vdM scans performed at large β^*.

The magnetic configuration near the IP can also be adjusted to optimize two other aspects of the luminosity calibration:

- the use of widely spaced bunches makes it possible to collide with zero nominal crossing angle, which in turn allows to estimate the satellite fraction by reconstructing longitudinally displaced vertices;
- luminous-region and beam-gas imaging measurements (Sections 10.4.3.2 and 10.5) all benefit from larger IP beam sizes, as the latter minimize the impact of the finite vertex resolution. A consensus developed to perform luminosity calibrations with injection optics ($\beta^* = 10 - 11$ m), rather than at the physics setting ($\beta^* = 0.6 - 3$ m). This configuration corresponds, for typical injected emittances, to single-beam (resp. luminous) widths of approximately 90 (resp. 60) μm, to be compared to a typical vertexing resolution of 20–50 μm.

10.3. Bunch-population determination

The LHC bunch currents are determined in a multi-step process due to the different capabilities of the available instrumentation. First, the total intensity of each beam is monitored by two identical and redundant DC current transformers (DCCT) [123] which are high-accuracy devices but have no ability to distinguish individual bunch populations. Each beam is also monitored by two fast beam-current transformers (FBCT) [124] which measure relative bunch currents individually for each of the 3564 nominal 25 ns slots in each beam; these fractional bunch populations are converted into absolute bunch currents using the overall current scale provided by the DCCT. Finally, corrections are applied to account for out-of-time charge present in a given BCID but not colliding at the interaction point.

The dominant uncertainty in early LHC running was associated with the absolute scale of the DCCT. A precision current source with a relative accuracy of 0.1% was used to calibrate the system at regular intervals, but the peak-to-peak variation of the measurements made in 2010 set an uncertainty on the bunch-current product of $\pm 2.7\%$ [125]. For later data, a much more detailed evaluation of the various sources of systematic uncertainty and a dedicated measurement and instrumental-improvement campaign to constrain these sources resulted in a DCCT-related uncertainty of 0.2%–0.3% RMS on the absolute-luminosity scale [123].

Because of the much more demanding bandwidth specifications dictated by bunch-to-bunch current measurements, the FBCT response is potentially sensitive to the frequency spectrum radiated by the circulating bunches, timing adjustments with respect to the RF phase, bunch-to-bunch intensity variations, etc. Dedicated laboratory measurements and beam experiments, comparisons with the response of other bunch-aware beam instrumentation systems, as well as imposing constraints on the bunch-to-bunch consistency of the measured visible cross-sections, resulted in a 0.2%–0.5% systematic luminosity-calibration uncertainty arising from the relative-intensity measurements [124].

Additional corrections to the bunch-by-bunch population are made to correct for "ghost charge" and "satellite bunches". Ghost charge refers to protons that are present in nominally empty bunch slots at a level below the FBCT threshold (and hence invisible), but still contribute to the current measured by the more accurate DCCT. Impressively precise measurements of these tiny currents (normally at most a few per mil of the total intensity) have been achieved by comparing the number of beam-gas vertices reconstructed by LHCb in nominally empty bunch slots, to that in non-colliding bunches whose current is easily measurable [31,126,127]. This by-product of the beam-gas imaging technique (Section 10.5) is a crucial input to the absolute-luminosity calibration of all LHC experiments.

Satellite bunches describe out-of-time particles present in collision BCIDs that are measured by the FBCT, but that remain captured in an RF-bucket at least one period (2.5 ns) away from the nominally filled LHC bucket, and as such experience at most longitudinally-displaced encounters with the nominally-filled bunches in the other beam. Their measurement relies primarily on the Longitudinal Density Monitor (LDM) [128,129].

10.4. van der Meer scans

This exposition adopts the notation of, and largely mirrors the methodology published by, the ATLAS Collaboration [25,26,130]. Scan protocols are spelled out in Section 10.4.1, the determination of the visible cross-section is outlined in Section 10.4.2 and the associated systematic uncertainties are discussed in Section 10.4.3. Luminosity determination by the *vdM* method at the ALICE IP is detailed in Refs. [116,118,131], that of CMS in [27,32,115,132–134], and that of LHCb in [29,127].

10.4.1. Scan protocols

10.4.1.1. Beam-separation scans for absolute luminosity calibration. The determination of the visible cross-section (Eq. (32)) requires the simultaneous measurement, for each colliding-bunch pair, of the interaction rate with the beams fully centered on each other ($\mu_{\text{vis}}^{\text{MAX}}$), of the convolved transverse beam sizes (Σ_x, Σ_y) and of the bunch intensities. It therefore necessitates a minimum of two separate beam scans, one where the beams are separated horizontally by up to $\pm 6\sigma_{xB}$ while being kept centered in y, and a second where the beams are separated vertically while remaining centered in x. The beams are moved in typically 25 scan steps, and data recorded for 20–30 s at each step to obtain a statistically significant measurement in each luminometer. A typical scan curve is presented in Fig. 14a. To help assess experimental systematic uncertainties, at least two such *vdM* scan pairs are usually performed in short succession to provide independent calibrations under similar beam conditions.

10.4.1.2. Beam-displacement scans for absolute length-scale calibration. Another key input to the *vdM* scan technique is the knowledge of the absolute beam separation at each scan point. The ability to measure $\Sigma_{x,y}$ depends upon knowing the precise distance by which the beams are separated during the scan, which is controlled by a set of closed orbit bumps applied locally near the IP using steering correctors.

To determine this beam-separation scale, dedicated length-scale calibration measurements are performed close in time to each *vdM* scan set using the same collision-optics configuration at the interaction point. Beams are displaced transversely, by the same amount and in the same direction, in typically five steps over a range of up to $\pm 3\sigma_{iB}$. Because the beams remain in collision during these scans, the actual position of the luminous region can be reconstructed with high accuracy, at each
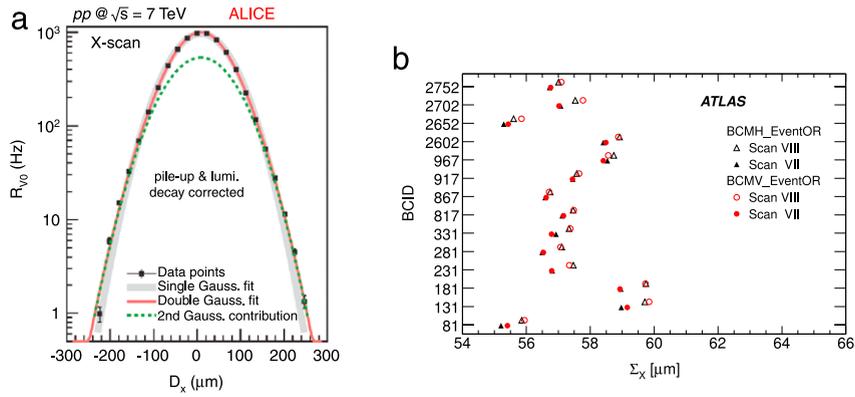
Fig. 14. (a) Measured visible interaction rate versus nominal beam separation during a horizontal scan at IP2, with single- and double-Gaussian fits overlaid. The dashed line shows the wider component of the double-Gaussian fit. Figure reproduced from Ref. [135], with permission. (b) Fitted horizontal convolved beam size Σ_x per colliding-bunch pair, as measured by two different Event_OR algorithms in consecutive scan pairs during a scan at IP1. The statistical uncertainty on each measurement is approximately the size of the marker.
Source: Figure reproduced from Ref. [26], with permission from the European Physical Journal.

step, using the primary vertex position reconstructed by the tracking detectors. The nominal beam displacement entered into the accelerator control system is thereby calibrated against the measured displacement of the luminous centroid. Since each of the four bump amplitudes (two beams in two transverse directions) depends on different magnet and lattice functions, distance-scale calibration scans are performed so that each of these four calibration constants can be extracted independently at each LHC IP. The resulting xy length-scale product is found to differ from unity by up to 2%–3%.

10.4.2. Visible cross-section determination

For each bunch pair, the specific visible interaction rate $\mu_{\mathrm{vis}}/(n_1 n_2)$ is measured as a function of the actual beam separation, equal to the product of the "nominal" separation (as specified by the LHC control system for each scan step) by the length-scale calibration factor discussed above. The value of μ_{vis} is determined from the raw rate using the prescriptions of Section 4.2. Each scan curve is fit to a characteristic function that, depending on beam conditions, is a single Gaussian plus a constant, a double Gaussian plus a constant, a Gaussian multiplied by a polynomial, or other variations. The peak value of that function determines $\mu_{\mathrm{vis}}^{\mathrm{MAX}}$, and its integral the convolved beam widths Σ_x, Σ_y (Eq. (29)).

The measured interaction rate must first be corrected for instrumental noise and single-beam backgrounds (e.g. halo from beam-gas interactions or tail scraping in the incoming-beam lines). In the simplest approach, these are accounted for by assuming that any constant term fit to the observed scan curve originates from luminosity-independent sources, and therefore contributes neither to the peak interaction rate nor to the overlap integrals. In some cases, a more detailed correction procedure is justified, whereby the instrumental noise, the luminosity-induced afterglow and the single-beam backgrounds are first measured directly using nominally empty bunch slots and non-colliding bunches. Each of these contributions (scaled if appropriate by the measured ratio of bunch intensities) is then subtracted separately from the colliding-bunch signal. Depending on the physical location and instrumental characteristics of the luminometer (timing resolution, response to low-energy collision debris, dark current, etc...) and on the luminosity algorithm considered, the total subtraction ranges from a few 10^{-4} to over 20% of the luminosity signal proper.

Transverse drifts of the individual beam orbits at the IP during a scan can distort the luminosity-scan curve and, if large enough, bias the determination of the overlap integrals. Such effects can be monitored by extrapolating to the IP beam-orbit segments measured using beam-position monitors located outboard of the interaction region [136]. This procedure is applied to each beam separately and provides a measurement of the relative drift of the two beams during the scan, that can be used to correct the beam separation at each scan step.

The beam separation and the measured interaction rate must also be corrected for, respectively, electromagnetic beam–beam deflections and the dynamic-β effect: this will be discussed in Section 10.4.3.2.

The bunch populations and the convolved beam sizes can all vary by several percent from one bunch to the next (Fig. 14b). Therefore the luminosity is different for each colliding-bunch pair, and the determination of $\mu_{\mathrm{vis}}^{\mathrm{MAX}}$, Σ_x and Σ_y must be performed independently for each BCID. But each BCID should measure the same σ_{vis} value. The average over all BCIDs is taken as the σ_{vis} measurement for that scan pair, and the agreement among the σ_{vis} values extracted from different bunches provides a consistency check on the calibration procedure.

10.4.3. Calibration uncertainties

The uncertainties affecting absolute luminosity calibrations by the *vdM* method are summarized in Table 12. The precision on the luminosity scale was originally limited by the accuracy of bunch-population measurements [25,29,32,116]. By now, it is dominated by uncertainties associated with the modeling and the reproducibility of beam conditions, and to a lesser extent by instrumental effects.

Table 12
Example breakdowns of the fractional systematic uncertainties affecting the determination of the visible pp cross-section σ_{vis} by the vdM method at the LHC. Blank entries correspond to cases where the uncertainty is either not applicable to that particular experiment or scan session, is considered negligible by the authors, or is not mentioned in the listed reference. In some cases, uncertainties quoted separately in the original publication have been regrouped to fit in a common classification.

Experiment	ALICE	ATLAS	CMS	LHCb
Reference	[118]	[26]	[133]	[127]
pp running period	2011	2011	2012	2012
$\sqrt{s}$ (TeV)	2.76	7.0	8.0	8.0
Total beam intensity	0.34%	0.23%	0.3%	0.23%
Bunch-to-bunch fraction	0.08%	0.20%	–	0.10%
Ghost charge and satellite bunches	0.45%	0.44%	0.2%	0.23%
Subtotal, bunch-population product	0.57%	0.54%	0.4%	0.34%
Orbit drift & beam-position jitter	–	0.32%	0.1%	0.32%
Bunch-to-bunch σ_{vis} consistency	–	0.55%	–	–
Emittance growth & scan-to-scan reproducibility	0.64%	0.67%	0.2%	0.80%
Dynamic β & beam–beam deflections	0.40%	0.50%	0.7%	0.28%
vdM fit model	–	0.28%	2.0%	0.54%
Non-factorization effects	0.60%	0.50%	in fit model	0.80%
Subtraction of luminosity backgrounds	0.30%	0.31%	–	0.14%
Subtotal, beam conditions	1.01%	1.24%	2.2%	1.33%
Difference of reference $\mathcal{L}_{spec}$ across luminometers	–	0.29%	–	–
μ-dependent non-linearities during vdM scans	–	0.50%	–	–
Other instrumental effects	0.20%	–	–	0.09%
Statistical uncertainty	–	0.04%	0.5%	0.04%
Subtotal, instrumental effects	0.20%	0.58%	0.5%	0.10%
Absolute beam-separation scale	1.41%	0.42%	0.5%	0.50%
Total systematic uncertainty on σ_{vis}	1.84%	1.53%	2.3%	1.47%

10.4.3.1. Bunch-population product. The bunch current measurement techniques and the associated uncertainties have been outlined in Section 10.3. The total systematic uncertainty on the bunch-population product now typically lies around 0.3%–0.5%, with occasionally larger error bars associated with unusual satellite or ghost-charge contributions. The determination of individual error contributions is detailed in Refs. [31,123,124,128,129].

10.4.3.2. Stability and modeling of beam conditions. Firstly, considerable effort has been devoted to identify and mitigate accelerator sources of calibration non-reproducibility [26,31,130].

- Slow orbit drifts can distort individual scan curves, as well as cause originally well-centered beams to move slightly out-of-collision between consecutive scans, thereby biasing the measured interaction rate at the peak of the scan. In addition, at each step of a scan, the actual beam separation may be affected by random jitter of the beam positions from their nominal setting, which in turn induces fluctuations in the luminosity measured at each scan point.
- While proton population and emittance differ significantly from bunch to bunch, the visible cross-section should be independent of the specific bunch(es) used to measure it. Bunch-to-bunch variations of σ_{vis} of non-statistical origin have been repeatedly observed, as illustrated in Fig. 15a. The RMS spread across bunch-by-bunch σ_{vis} measurements within one scan pair is 0.6%. Comparing the first scan pair (circles, labeled scan VII) to the second (squares, scan VIII), it is clear that some of the fluctuations are not statistical in nature, but correlated by BCID; they will therefore contribute to the systematic calibration uncertainty.
- The impact of emittance growth within and between scans has been studied in some detail [26,133]. The variation in bunch-averaged visible cross-section between scan pairs reflects the reproducibility of the calibration procedure during a single fill. It amounts to 0.7% in the example of Fig. 15a; but significantly larger scan-to-scan variations have been observed in other scan sessions and at other IP's. The non-reproducibility of the visible cross-section between consecutive scan pairs remains poorly understood, and sometimes constitutes the dominant source of systematic uncertainty. It has been traced to multiple causes, including operational subtleties in executing the scan sequence, orbit drifts, as well as non-linear x–y correlations in the phase space of each beam which invalidate the factorization assumption that underpins the standard vdM method.

Taken together, the bunch-to-bunch and scan-to-scan variations of σ_{vis} measurements quantify the stability of the calibration technique. The non-statistical component of the RMS spread of bunch-by-bunch σ_{vis} measurements within a given scan pair is taken, at least by some authors, as a systematic uncertainty in the calibration technique, as is the scan-
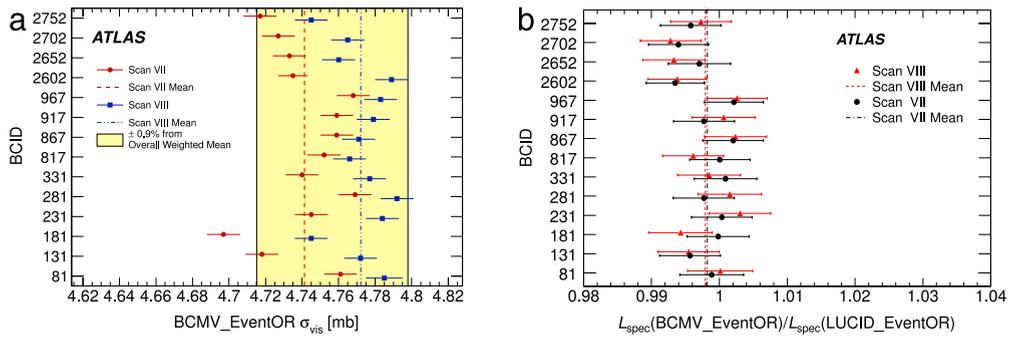
Fig. 15. (a) Measured σ_{vis} values for the BCMV_EventOR algorithm, per colliding-bunch pair for two consecutive scans pairs at IP1. The shaded band indicates a $\pm 0.9\%$ variation from the average, which is the systematic uncertainty evaluated from the per-bunch and per-scan σ_{vis} consistency. (b) Bunch-by-bunch ratio of the specific luminosity determined at the ATLAS IP by independent luminometers (BCM, LUCID) in the same scan pairs. In both figures, the error bars represent statistical uncertainties only, and the vertical lines indicate the weighted average over colliding bunches for the two scan pairs separately.
Source: Figures reproduced from Ref. [26], with permission from the European Physical Journal.

to-scan reproducibility of bunch-averaged visible cross-sections. The band in Fig. 15a, which represents a range of $\pm 0.9\%$, shows the quadrature sum of these two systematic uncertainties for this particular IP and scan session.

Next, the challenging precision goals outlined in Section 10.1.3 have motivated a careful investigation of beam-dynamics effects that might induce significant distortions of the scan curves.

- Hourglass corrections, which are significant at RHIC, remain negligible at LHC because $\sigma_z \ll \beta^*$.
- When charged-particle bunches collide, the electromagnetic field generated by a bunch distorts the individual particle trajectories in the corresponding bunch of the other beam. This so-called *beam–beam interaction* manifests itself in several ways, two of which are relevant here.

  First, when the bunches are not exactly centered on each other in the *x–y* plane, their electromagnetic repulsion induces a mutual angular kick that distorts the closed orbits and modulates the actual transverse separation at the IP in a manner that depends on the separation itself. If left unaccounted for, these *beam–beam deflections* [136] would bias the measurement of σ_{vis} by up to 4%–5%, depending on the bunch parameters.

  The second phenomenon, called *dynamic β* [137], arises from the mutual (de)focusing of the two colliding bunches: this effect is tantamount to inserting a small quadrupole at the collision point. The resulting fractional change in β^*, or equivalently the optical demagnification between the LHC arcs and the IP, varies with the transverse beam separation, slightly modifying the collision rate at each scan step. If left uncorrected, the resulting distortion of the *vdM* scan curve would bias the measurement of σ_{vis} [26,130] by as much as -1%.

  The amplitude and the beam-separation dependence of both effects depend similarly on the beam energy, the ring tunes, the unperturbed IP β-functions, as well as the bunch charge and transverse beam sizes. These beam–beam distortions are modeled and corrected for by a combination of analytical formulas and of beam–beam simulations [26]; the associated systematic uncertainty on σ_{vis} typically amounts to 0.5%–1%.

- Variations in the functional form used to model the luminosity-scan curves and to extract the overlap integrals contribute diversely to the calibration uncertainty. It should be stressed, however, that this baseline *vdM* procedure, as detailed in Sections 5.2.1, 10.4.2 and 10.4.1.1, assumes that the particle densities in each bunch can be factorized into independent horizontal and vertical components.

- If this factorization assumption is violated, the horizontal (resp. vertical) convolved beam width Σ_x (Σ_y) is no longer independent of the vertical (resp. horizontal) beam separation δ_y (δ_x); similarly, the transverse luminous size in one plane ($\sigma_{x\mathcal{L}}$ or $\sigma_{y\mathcal{L}}$), as extracted from the spatial distribution of reconstructed collision vertices, depends on the separation in the other plane. Such non-factorization effects have been unambiguously observed in multiple scan sessions, either from the δ_x- (δ_y-) dependence of $\sigma_{y\mathcal{L}}$ ($\sigma_{x\mathcal{L}}$) during a standard horizontal (vertical) scan, or from significant differences in Σ_x (Σ_y) between a standard scan, and an off-axis scan during which the beams are separated by up to $4\sigma_{yB}$ ($4\sigma_{xB}$) in the non-scanning plane [130]. The strength of the effect varies widely across *vdM* scan sessions, and has been seen to differ from one bunch to the next as well as to evolve with time within one LHC fill. In addition, some of the scan-to-scan irreproducibility discussed above has now been demonstrated to correlate with large non-factorization effects [127,130]. Overall, the body of available observations can be explained neither by residual linear *x–y* coupling in the LHC optics, nor by crossing-angle or beam–beam effects; instead, it points to non-linear transverse correlations in the phase space of the individual bunches. This phenomenon was never considered at previous colliders, and its study is only beginning. A recent breakthrough is the direct and quantitative demonstration, by the LHCb beam-gas imaging analysis (Section 10.5), of the non-factorizability of individual bunch density distributions.

  While the generalized *vdM* formalism (Eq. (31)) can in principle handle arbitrary two-dimensional luminosity distributions, determining the latter by performing an *x–y* "grid" scan (rather than two one-dimensional *x*- and *y*- scans) would be prohibitively expensive in terms of beam time, as well as limited by emittance-growth systematics. The present

strategy, therefore, is to retain the standard *vdM* technique (that assumes factorization) as the baseline calibration method, and to constrain possible non-factorization biases using the data themselves. Short of input from beam-gas imaging measurements,[17] most powerful so far has been the modeling of the transverse-density distributions of the two beams by fitting the evolution, during *vdM* scans, not only of the luminosity itself but also of the position, orientation and shape of its spatial distribution, as reflected by that of reconstructed *pp*-collision vertices [26,130]. Luminosity profiles are then generated for simulated *vdM* scans using these fitted beam parameters, and analyzed in the same fashion as real *vdM* scan data. The impact of non-factorization in the single-beam distributions is determined from the difference between the 'true' luminosity from the simulated overlap integral at zero beam separation, and the 'measured' luminosity from the fits to the one-dimensional simulated luminosity profiles (assuming factorization). This technique is tantamount to beam–beam imaging (Section 5.3.3), but with the notable difference that it is much less sensitive to resolution effects because it is used only to estimate a small fractional correction to the overlap integral, rather than its full value. The resulting systematic uncertainty ranges from 0.1 to 3%, depending on the beam conditions during the scan session. Also promising (but still highly empirical) is the tuning of the LHC injector chain to produce more Gaussian, less correlated beam profiles [138].

Finally, the subtraction of the single-beam backgrounds from the luminosity signal may, depending on the detector technology and the luminosity algorithm, contribute noticeably to the systematic uncertainty.

10.4.3.3. Instrumental uncertainties. Since the specific luminosity (Eq. (33)) depends only on the convolved beam sizes, results for a given scan should be consistent, to high precision, across all luminometers and luminosity algorithms. Fig. 15b shows, for each scan pair, the bunch-by-bunch ratio of the $\mathcal{L}_{spec}$ values reported by two different ATLAS luminometers. For each bunch, the two algorithms appear statistically consistent; in both scans however, the bunch-averaged $\mathcal{L}_{spec}$ ratio is consistently low by a fraction of percent, suggesting a detector-dependent bias that must be covered by a systematic uncertainty on the reference absolute luminosity.

A complementary check is provided by the consistency, during the scan and as a function of the pile-up parameter μ, of the luminosity values reported by different methods. Relative non-linearities in the luminosity response can be indicative of (for instance) inconsistent noise subtraction (potentially important in the tails of the scan), unaccounted-for combinatorics effects in the calibration of coincidence algorithms (Section 4.2.2.2) or migration-related biases (Section 4.2.4) near the peak of the scan. Each of those effects has been observed at some level, and the residual μ-dependence of the calibrated-luminosity ratios provides a global estimate of their potential impact on the absolute luminosity scale.

The statistical uncertainty affecting the visible cross-section remains in most cases below 0.1%, except for some pixel-tracker-based algorithms because of dataflow-related rate limitations (Section 10.2.1).

10.4.3.4. Length-scale calibration. The determination of the absolute beam-separation scale has been described in Section 10.4.1.2. The associated systematic uncertainty on the absolute luminosity scale varies between ± 0.4 and $\pm 2\%$, depending on the details of the scan procedure and on the reproducibility of beam conditions.

10.4.3.5. Total calibration uncertainty. The discussion above summarizes the detailed analyses published by the ATLAS Collaboration [25,26,130]; similar conclusions are reported by ALICE [118,139], CMS [115] and LHCb [127]. The systematic uncertainty affecting the determination of visible *pp* cross-sections by the *vdM* method spans a range from ± 1.5 to $\pm 4.4\%$, depending on the beam conditions, the data set considered, as well as the instrumental and analysis techniques that were brought to bear. It should be noted, however, that the methodology (and in particular the quantitative criteria) for setting systematic uncertainties remain to be unified across the LHC Collaborations.

Comparable or better precision has been achieved by the beam-imaging technique described in Section 10.5. But the calibration uncertainty is only one of the components of the total absolute-luminosity uncertainty associated with a running period, as will be demonstrated in Section 10.7.

10.5. Beam imaging

The principle of the beam-gas imaging technique is outlined in Section 5.3.1. Its first application by LHCb is reported in Ref. [140]. Soon thereafter, beam-gas imaging provided an absolute luminosity accuracy comparable to that of the *vdM* method [29]. It has since been considerably refined, providing direct experimental evidence, for the first time ever, for non-factorization effects in single-beam density distributions, as well as the most precise absolute luminosity determination to date [31,127] at a bunched hadron collider.

The BGI method is made possible by the high resolution of the LHCb "vertex locator" (VELO) and its extremely close proximity to the beams [141]. The technique is based on reconstructing beam-gas interaction vertices to measure the positions, angles and transverse shapes of individual bunches in the two beams separately (Fig. 16a), as well as the

---

[17] The LHCb *vdM* results summarized in Table 12 are based on a non-factorizable single-beam model that is constrained by the beam-imaging results performed in the same fill and described in Section 10.5. This approach is unique to LHCb.
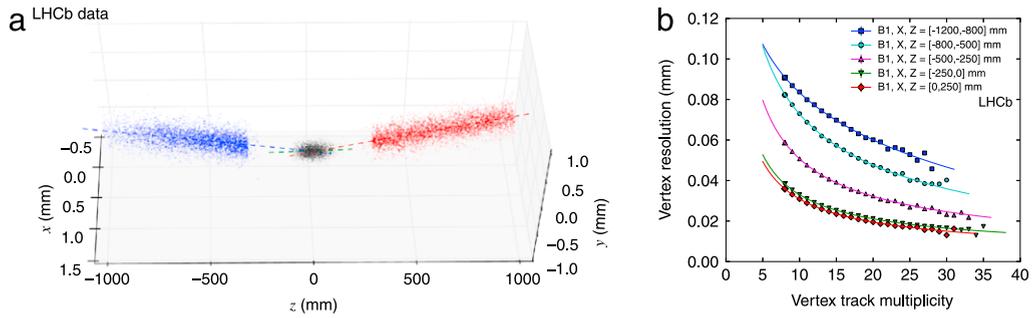
Fig. 16. (a) 3-d view of vertices reconstructed in LHCb: beam-gas interactions satisfying beam-1 selection criteria (left cluster, blue) or beam-2 criteria (right cluster, red) are clearly distinguished from the narrower distribution of *pp* collision vertices (central cluster, black). Beam-gas candidates are excluded from the central region ($|z| < 250$ mm) as they cannot be unambiguously separated from beam–beam events. The two dashed lines are straight-line fits to the two beam-gas vertex distributions. The crossing angle between the beams is clearly visible. Figure reproduced from Ref. [31] © CERN. (b) Parameterization of the beam-gas vertex resolution for beam 1, for different bins in longitudinal position *z*. Single markers indicate a resolution measurement for a given track multiplicity at the vertex. The true single-beam width is approximately 0.09 mm.
Source: Figure reproduced from Ref. [127],© SISSA Medialab Srl. CC BY-NC-SA.

three-dimensional distribution of *pp* collision vertices. These measurements fully determine the spatial density distributions of the two colliding bunches, thereby allowing the determination of their overlap integral and of the corresponding bunch luminosity (Eq. (2)). This absolute luminosity is then combined with the simultaneously measured rate of any relative-luminosity monitor to calibrate its visible cross-section σ_{vis}.

In order to accumulate adequate beam-gas statistics in a time short enough for the beam emittances and orbits not to vary significantly, a local pressure bump must be created in the vicinity of the collision point. Neon at a pressure of a few 10^{-7} mbar is injected under tightly controlled conditions for the duration of the luminosity-calibration session [31]; special care is taken to ensure that the gas density distribution is uniform enough not to bias the spatial distribution of reconstructed vertices.

As the precision of the luminosity calibration depends on measuring, to sub-percent accuracy, the transverse size of bunches a few ten microns across, it relies on an exquisite control of the vertexing resolution. The BGI method therefore requires a dedicated beam-optics setup ($\beta^* \geq 10$ m), to ensure that the beam sizes are sufficiently large compared to the resolution for the latter not to dominate the raw vertex distributions. Even so, considerable effort had to be invested [29,31, 126,127] to map and validate, using the data themselves, the position- and track-multiplicity-dependence of the vertexing resolution separately for beam-1 gas, beam-2 gas and collision events. Shown in Fig. 16b is the beam-gas resolution for beam 1; the resolution for beam-2 vertices is somewhat worse, because the LHCb detector is highly asymmetric with respect to the IP. The resolution for *pp* collisions is significantly better, because of the larger track multiplicity and their higher average momentum.

The bunch-density distributions, described by the analytical functions discussed below, are convolved with the resolution functions above to predict the spatial distribution of beam-gas and *pp* vertices. The parameters controlling these predicted distributions, i.e. the position, direction, shape and size of each bunch, are extracted from a simultaneous fit to the measured 3-d distributions of beam-gas and collision vertices. The resulting set of bunch parameters is then used to compute the overlap integral, which combined with the measured bunch currents provides the absolute bunch luminosity, and finally the visible cross-sections for the luminosity algorithms of interest.

The density distribution of each bunch is modeled by the superposition of three-dimensional double-Gaussians, with both a component that is factorizable in (x, y, z) and a non-factorizable component; the relative strength of these two is controlled by a "factorizability" parameter that is extracted from the fit. The analysis of the reconstructed density distributions [31,127] demonstrates unambiguously the presence, during some of the 2012 and 2013 BGI calibration sessions, of a large non-factorizable component in the transverse beam-density distributions. The measured non-factorizability varies from one LHC fill to another, differs from bunch to bunch, and decays over the duration of the fill, as independently inferred from the beam-separation dependence of luminous-region parameters during *vdM* scans (Section 10.4.3.2). The impact of such non-factorization effects is illustrated in Fig. 17. A two-dimensional fit to the vertex distributions that accounts for non-factorization, yields visible cross-section that are highly consistent bunch-to-bunch and fill-to-fill (Fig. 17a). In contrast, determining transverse shapes by one-dimensional fits to horizontal and vertical projections of the vertex distributions, i.e. ignoring possible *x*–*y* correlations in the transverse density distributions of the colliding bunches, results in ∼3% inconsistencies in the absolute luminosity scale (Fig. 17b).

The systematic uncertainties affecting the beam-imaging method are summarized in Table 13. With the exception of the bunch-population product, the VELO transverse scale and the background subtraction, they are either totally or largely uncorrelated with those affecting *vdM* scans; in particular, the effects associated with beam dynamics either are directly measured, or become irrelevant because no beam displacement is needed. The consistency of visible cross-sections measured simultaneously by the very different *vdM* and BGI methods provides a powerful check [127] on the robustness of the most precise luminosity determination at a proton collider since the ISR.
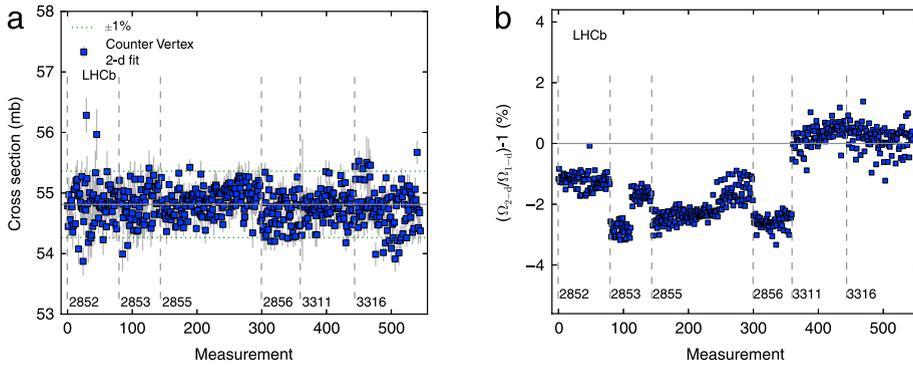
Fig. 17. (a) Visible cross-section for the LHCb vertex-based event-counting luminosity algorithm, extracted from 2-d fits to the transverse distribution of beam-gas and collision event vertices. Each point is an independent measurement using data from a single colliding-bunch pair integrated over 20 minutes. The error bars reflect the statistical uncertainty on the overlap integral. The measurements are sorted by time and BCID. The first four LHC fills (2852–2856) took place in July 2012, the last two (3311, 3316) in November 2012. The dotted horizontal lines indicate a $\pm 1\%$ deviation from the central value. (b) Ratio of the overlap integrals (i.e. inverse ratio of the visible cross-sections), for 2-d fits (in which the degree of factorizability of each bunch is left floating in the fit) and 1-d fits (which assume that the density distribution of each bunch is factorizable in x and y).
Source: Figures reproduced from Ref. [127],© SISSA Medialab Srl. CC BY-NC-SA.

Table 13
Systematic uncertainties affecting the LHCb absolute luminosity calibration by the BGI method at $\sqrt{s} = 8$ TeV [31,127].

Source of uncertainty	Uncertainty (%)	Correlated with vdM
Bunch-population product	0.23	Yes
Vertexing resolution: beam–beam events	0.93	No
Vertexing resolution: beam-gas events	0.55	No
Detector alignment & crossing angle	0.45	No
VELO transverse scale	0.05	Yes
Bunch-shape model	0.50	Yes
Longitudinal reconstruction efficiency	0.04	Yes
Pressure gradient	0.03	No
Convolved bunch length	0.05	No
Background subtraction ("Vertex" algorithm)	0.20	Yes
Bunch-to-bunch & fill-to-fill σ_{vis} consistency	0.54	No
Calibration transfer to "Tracks" algorithm	0.20	No
Statistical uncertainty	0.01	No
Total systematic uncertainty on σ_{vis}	1.43	

10.6. Elastic pp scattering and the optical theorem

The TOTEM experiment [142] has been optimized for diffractive physics studies, and for the determination of the total cross-section using the simultaneous measurement of elastic *pp* scattering and of the total inelastic collision rate (Section 5.4.1).

The Roman pot (RP) stations used in the elastic cross-section measurements are located at 215–220 m symmetrically on either side of the CMS IP (Fig. 18a). Each station is composed of two units separated by about 5 m; a unit consists of 3 RPs, two approaching the outgoing beam vertically and one horizontally. The pots are instrumented with stacks of Si strip detectors specifically designed to reduce to only a few ten microns the insensitive area at their beam-facing edge, such that scattered protons can be detected a couple of mm from the beam center. A dedicated optics setup ($\beta^* = 90$ m) made it possible to measure the differential elastic cross-section in the $0.005 < |t| < 0.2$ GeV^2 range at $\sqrt{s} = 7$ TeV. This $|t|$ range, the lower edge of which corresponds to a scattering angle of only 20 µrad, covers about 90% of the nuclear elastic rate, allowing a quite precise extrapolation to $|t| = 0$.

The total inelastic rate is measured by two telescopes integrated in the CMS detector (Fig. 18b): T1, located 9 m from the IP and equipped with Cathode Strip Chambers, and T2 at 13.5 m, that uses Gas Electron Multipliers. Together these cover the $3.1 < |\eta| < 6.5$ range, i.e. about 95% of the acceptance for inelastic-collision products. A low-mass, single-diffractive component at larger $|\eta|$ escapes detection; its contribution to the inelastic rate is evaluated by Monte-Carlo simulations.

At $\sqrt{s} = 7$ TeV and using the luminosity-independent method (Eq. (37), (38)), TOTEM reports [143] an accuracy of 2.6% on the total cross-section σ_{tot} and a 3.8% error bar on the corresponding integrated luminosity.[18] The latter uncertainty is

---

[18] TOTEM also published a σ_{tot} measurement at 8 TeV, but does not quote the corresponding luminosity uncertainty.
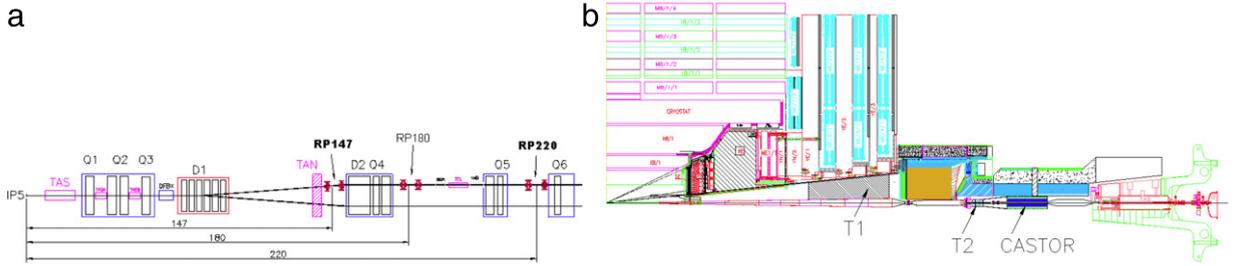
Fig. 18. (a) The LHC beam line on one side of IP5, showing the existing TOTEM Roman pots at about 147m (RP147) and 220m (RP220). (b) The TOTEM forward telescopes T1 and T2, embedded in the CMS detector.
Source: Figures reproduced from Ref. [142],© SISSA Medialab Srl. CC BY-NC-SA.

dominated by that on the total inelastic rate, which in turn reflects the precision of the extrapolation of the undetected diffractive rate beyond the T2 acceptance. The extrapolation of the elastic cross-section to $t = 0$ contributes a comparable uncertainty; the limited knowledge of the ρ parameter has little impact. The TOTEM luminosity above agrees within 1% with the simultaneous, *vdM*-based CMS measurement that carries a 4% uncertainty.

ATLAS is similarly equipped with vertical RP's, instrumented with scintillating fibers and located symmetrically at ± 240 m from IP1 to measure very-small angle elastic scattering [43]. By comparison to TOTEM, however, the ATLAS $|\eta|$ coverage is somewhat limited in the forward direction: the tracking coverage stops at $|\eta| = 2.5$ and the calorimeter coverage at $|\eta| = 5$. This would make a precise determination of the total inelastic rate significantly more delicate. Tuning the Monte Carlo physics model using diffractive data from TOTEM may help improve the extrapolation of the inelastic rate into the blind region.

The baseline ATLAS strategy, however, is to exploit the optical theorem in conjunction with elastic-scattering measurements in the Coulomb-interference region (Section 5.4.2), an approach equally accessible to TOTEM. The challenge here is to achieve small enough a beam angular divergence at the IP to measure precisely scattering angles of a couple of microradians. As the angular divergence scales like $1/\sqrt{\beta^*}$, a dedicated optical lattice with $\beta^* \sim 1$–2 km is required. A magnetic configuration with $\beta^* = 1000$ m has been tested in 2012, with encouraging results.

10.7. Total luminosity uncertainty

Luminosity calibrations are performed only once or twice per year, typically with pile-up parameters much smaller ($\mu \sim 0.5$–2) than those in routine physics running. In addition to the absolute-calibration issues discussed in Sections 10.2–10.5, three sources of uncertainty affect the precision of the integrated luminosity: the linearity of the calibrated luminosity with respect to the pile-up parameter μ (Section 10.7.1), the relative long-term stability of the luminosity reported by each algorithm (Section 10.7.2), and other corrections or instrumental effects that depend on the total luminosity (Section 10.7.3).

10.7.1. Interaction-rate dependence

A fundamental ingredient of the strategy to assess and control the systematic uncertainties affecting the absolute integrated luminosity over a running period is to compare the measurements, at a given IP, of several luminometers, some of which use more than one counting technique. These multiple detectors and algorithms are characterized by significantly different acceptance, response to the pile-up parameter μ, sensitivity to instrumental effects and to beam-induced backgrounds. To evaluate the integrated luminosity used in physics analyses, a single algorithm is chosen to provide the central value for a certain range of time; the remaining calibrated algorithms provide independent measurements that can be used to assess the consistency and stability of the results.

A first cross-check is the level of agreement between the calibrated luminosity algorithms as a function of the pile-up parameter μ: instrumental biases typically increase with pile-up, and it is important to verify that the various algorithms still provide an accurate and linear measurement of the luminosity up to the highest values of μ observed in the physics data. Several tests of the μ-dependence are discussed in Refs. [26,27,130]. One of them is illustrated in Fig. 19a, which shows the luminosity ratio between algorithms as a function of $\langle\mu\rangle$ for a single LHC fill. The shapes of the curves are directly sensitive to variations in the linearity as a function of $\langle\mu\rangle$, while the overall shifts of each algorithm up or down result from the long-term calibration drifts discussed in Section 10.7.2. In this example, a systematic uncertainty of $\pm 0.5\%$ has been applied by the authors to account for any possible μ-dependence in the extrapolation from the low-μ *vdM* calibration to the higher-μ regime of physics data-taking.

10.7.2. Long-term stability

Another uncertainty is associated with the assumption that the σ_{vis} calibration determined in a set of *vdM* scans is stable across the entire running period, typically one-year long. Several effects could degrade the long-term stability of a given
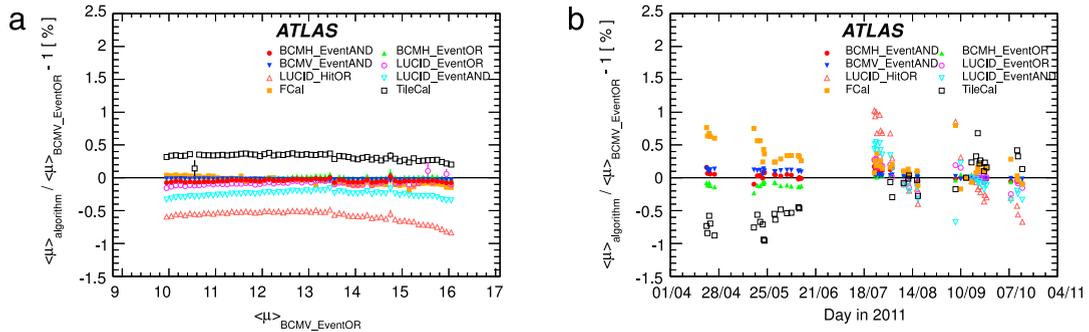
Fig. 19. (a) Fractional deviation of the bunch-averaged mean interaction rate per crossing $\langle\mu\rangle$ obtained using various algorithms from that reported by the reference algorithm BCMV_Event_OR, as a function of $\langle\mu\rangle$ during a single LHC fill. (b) History of the fractional deviation, with respect to the reference algorithm, of the value of $\langle\mu\rangle$ reported by various algorithms. Each point shows the deviation of the bunch-averaged interaction-rate ratio in a single LHC fill from that in a reference fill in September 2011. Statistical uncertainties are typically smaller than the size of the markers.
Source: Figures reproduced from Ref. [26], with permission from the European Physical Journal.

luminometer, including slow drifts in the instrumental response and sensitivity to varying LHC beam conditions, particularly the total number of colliding bunches.

Fig. 19b displays the history, across the 2011 running period, of the ratio of the bunch-averaged interaction rate $\langle\mu\rangle$ (which is proportional to the mean bunch luminosity) reported by various ATLAS luminometers, to that measured by the reference luminosity algorithm. Shown is the relative variation of this ratio over time, compared to a single fill that is used to provide a anchor point and that comes approximately four months after the *vdM* scan that was performed in May 2011. Based on the observed time variation between the various algorithms, a systematic uncertainty on long-term stability, which includes any effects related to the dependence on the number of colliding bunches or other operational conditions during the 2011 running period, is set at $\pm 0.7\%$ [26] in this particular example.

10.7.3. Luminosity-dependent corrections and instrumental effects

A variety of detector-specific corrections or potential biases may complicate the quantitative interpretation, in terms of absolute luminosity, of uncalibrated relative-luminosity readings recorded during routine physics running. A few examples are listed below.

While afterglow (Section 10.1.3.2) is mostly negligible during luminosity-calibration sessions because of the special beam conditions, it can reach several percent during physics running at the ATLAS and CMS IP's, depending on the bunch pattern and the luminosity algorithm considered. The associated uncertainty on the absolute luminosity is typically a few per mil. Afterglow has been less of an issue at the ALICE and LHCb IP's because of the much lower bunch luminosities.

In CMS, very high charged-track multiplicities or level-1 trigger rates can saturate the readout buffers of the pixel detector, inducing luminosity-dependent inefficiencies of the cluster-counting luminosity algorithm. These can be controlled at the sub-percent level [133].

At the LHC, *vdM* scans are always performed at low total luminosity, and sometimes after an extended no-beam period. A known feature of diamond radiation sensors (such as the ATLAS BCM) is a tendency for the gain to increase, under moderate irradiation levels, up to a stable asymptotic value at high dose rates. This so-called "pumping" is generally ascribed to the filling of charge traps in the diamond sensors with continued irradiation until enough charge has been sent through the device to fill essentially all the traps. Depending on the detailed dose-rate history at the IP, enough self-annealing of the diamond sensors may occur prior to, during, or shortly after the scan sessions for the efficiency of the diamonds to evolve between the time the *vdM* calibration was performed and the restart of high-luminosity physics running, resulting in an additional luminosity-calibration uncertainty of the order of 0.2% [26]. In the more recent, higher-luminosity 2012 LHC run, percent-level gain variations of either sign have been observed in these diamond sensors, between fills with comparable bunch luminosity but widely different numbers of bunches [130]; such poorly understood efficiency jumps require painstaking corrections and degrade the precision of the luminosity scale.

10.7.4. Total systematic uncertainty on the integrated luminosity

Table 14 illustrates the contributions to the total systematic uncertainty affecting the absolute luminosity during an LHC running period. This table is not meant to offer an exhaustive compilation of all the available results, but a snapshot of some of the best performance to date.[19] The most precise integrated-luminosity determinations lie well below 2%, offering promising prospects for precision Standard Model cross-section measurements and better constraints on the PDF's. While no single uncertainty source is totally dominant, overall the precision of the *vdM* method (Table 12) appears limited by the

---

[19] The ALICE calibration uncertainty quoted here is larger than that in Table 12, because it is based on an earlier data set for which the luminosity-calibration uncertainty was dominated by that on the bunch-population product.

Table 14
Examples of total fractional uncertainty on the delivered integrated luminosity.

Experiment	ALICE	ATLAS	CMS	LHCb		
pp running period	2010	2011	2012	2012		
$\sqrt{s}$ (TeV)	7.0	7.0	8.0	8.0		
Reference	[118,144]	[26]	[133]	[127]		
Absolute-calibration method	*vdM*	*vdM*	*vdM*	*vdM*	Combined	BGI
Calibration uncertainty $\Delta\sigma_{vis}/\sigma_{vis}$ (%)	3.5	1.53	2.3	1.47	1.12	1.43
μ-dependence (%)	–	0.50	<0.1		0.17	
Long-term stability (%)	1.5	0.70	1.0		0.22	
Subtraction of luminosity backgrounds (%)	3.0	0.20	0.5		0.13	
Other luminosity-dependent effects (%)	1.5	0.25	0.5		–	
Total luminosity uncertainty (%)	5.0	1.8	2.6	1.5	1.2	1.5

reproducibility and the modeling of beam profiles (and in particular by the non-factorization effects), and that of the BGI technique (Table 13) by the vertexing resolution. In addition, long-term instrumental stability is a significant issue at the high-luminosity IP's [27,130], presumably because of the very-high rate, intense-radiation environment.

11. Conclusions

The van der Meer technique was invented at the ISR, where it achieved the most precise luminosity determinations ever at any *pp* collider, with small-angle elastic scattering a close second. In the two single-ring colliders that followed, the S$p\bar{p}$S and the Tevatron, beam-separation scans proved impractical; there the absolute luminosity scale had to rely first on single-beam density measurements, and later on elastic scattering and the inelastic rate. The advent of two-ring colliders (RHIC and LHC) brought the *vdM* method back to the fore, with beam-gas imaging a major contender—but so demanding instrumentally that it became feasible only as a by-product of the highly specialized physics goals of one experiment. While LHC experiments are unlikely to break the "1% barrier" overcome at the ISR, they sometimes come close, thanks to several key developments: a large investment in the bunch-by-bunch capabilities of beam-current and luminosity instrumentation, the realization that beam-dynamics effects (non-factorization, beam–beam) have a significant impact on the calibration techniques, the maturation of beam imaging methods made possible by state-of-the-art microvertex detectors, and (at least in some cases) the emphasis on the use of several independent, redundant luminometers and algorithms to ambush biases and reduce systematic uncertainties.

Acknowledgments

References

[1] L. Bonolis, Nuovo Cimento 28 (2005) 1–60.
[2] K. Johnsen, Nucl. Instrum. Methods 108 (1973) 205–223.
[3] C. Rubbia, P. MacIntyre, D. Cline, Proc. Int. Neutrino Conf. Aachen (1976) 683.
[4] J. Gareyte, Proc. CERN Acc. School on Antiprotons for Colliding Beam Facilities (1983) 291–318.
[5] H.T. Edwards, Annu. Rev. Nucl. Sci. 35 (1985) 605–660.
[6] M. Harrison, et al., Nucl. Instr. Meth. A 499 (2003) 235–244.
[7] L. Evans, P. Bryant (Eds.), JINST 3 (2008), S08001.
[8] G. Giacomelli, M. Jacob, Phys. Rep. 55 (1979) 1–132.
[9] M.L. Mangano, Motivations and precision targets for an accurate luminosity determination at the LHC, in: CERN-Proceedings-2011-011, https://cds.cern.ch/record/1347440.
[10] A. Chao, M. Tigner, Handbook of Accelerator Physics & Engineering, World Scientific, 1999.
[11] W. Herr, B. Muratori, CERN Yellow Report 2006-002.
[12] S.M. White, CERN-THESIS-2010-139.
[13] K. Potter, CERN Yellow Report 94-01, 1994, pp. 117–130.
[14] M. Ferro-Luzzi, LHCb-PUB-2012-016, 2012.
[15] C. Moller, K. Danske, Vidensk. Selk. Mat.-Fys. Medd. 23 (1945) 1.

[16] O. Napoly, Part. Acc. 40 (1993) 180.
[17] W. Kozanecki, et al., Nucl. Instr. Meth. A 607 (2009) 293–321.
[18] M. Venturini, W. Kozanecki, RPPH141, in: Proc. Part. Acc. Conference (PAC'01), Chicago, USA, 2001.
[19] A. Drees, S.M. White, MOPEC013, in: Proc. Int. Part. Acc. Conference (IPAC'10), Kyoto, Japan, 2010.
[20] S.M. White, Luminosity scans at the LHC, in: CERN-Proceedings-2011-011, https://cds.cern.ch/record/1347440.
[21] K. Hagiwara, et al., (Particle Data Group), Phys. Rev. D66 (2002) 010001.
[22] J. Beringer, et al., (Particle Data Group), Phys. Rev. D86 (2012) 010001, and http://pdg.lbl.gov/2013/reviews/rpp2013-rev-hep-collider-params.pdf.
[23] K. Hübner, Eur. Phys. J. H (2012) 509–522, and private communication.
[24] P. Grannis, P. Jenni, Phys. Today 66 (6) (2013) 38.
[25] ATLAS Collaboration, Eur. Phys. J. C71 (2011) 1630.
[26] ATLAS Collaboration, Eur. Phys. J. C73 (2013) 2518.
[27] CMS Collaboration, CMS-PAS-EWK-11-001, 2011.
[28] J. Groth-Jensen, CERN-THESIS-2010-200.
[29] LHCb Collaboration, JINST 7 (2012) P01010.
[30] B. Heinemann, Comparison of the ATLAS and CMS Luminosity Measurements during *pp* Collisions in 2010 at $\sqrt{(s)} = 7$ TeV, in: CERN-Proceedings-2011-011. https://cds.cern.ch/record/1347440.
[31] C. Barschel, CERN-THESIS-2013-301.
[32] CMS Collaboration, CMS-PAS-EWK-10-004, 2010.
[33] J. Bosser, et al., Nucl. Instr. Meth. A 235 (1985) 475.
[34] J. Bosser, et al., CERN-SPS-87-13-ABM, 1987.
[35] J. Bosser, LHC Project Report 370, 2000.
[36] T. Lefevre, et al., MOPE057, in: Proc. Int. Part. Acc. Conference (IPAC'10), Kyoto, Japan, 2010.
[37] V. Papadimitriou, FERMILAB-CONF-11-199-AD.
[38] S. van der Meer, CERN-ISR-PO-68-31, 1968.
[39] C. Rubbia, CERN-p$\overline{p}$-Note-38, 1977.
[40] K.M. Potter and J.C. Sens (Eds.), CERN-NP-71-6, 1971.
[41] V. Balagura, Nucl. Instr. Meth. A 654 (2011) 634–638.
[42] M. Ferro-Luzzi, Nucl. Instr. Meth. A 553 (2005) 388–399.
[43] ATLAS Collaboration, CERN/LHCC/2008-004, 2008.
[44] U. Amaldi, et al., Phys. Lett. 43B (1973) 231.
[45] ATLAS Collaboration, Phys. Rev. D85 (2012) 072004.
[46] C. Anastasiou, et al., Phys. Rev. D69 (2004) 094008.
[47] G. Watt, arXiv:1106.5788v2, 2011.
[48] M. Dittmar, et al., Phys. Rev. D56 (1997) 7284–7290.
[49] V.M. Budnev, et al., Nucl. Phys. B 63 (1973) 519.
[50] V.A. Khoze, et al., Eur. Phys. J. C19 (2001) 313.
[51] A.G. Shamov, V.I. Telnov, Nucl. Instr. Meth. A 494 (2002) 51.
[52] M.W. Krasny, J. Chwastowski, K. Slowikowski, Nucl. Instr. Meth. A 584 (2008) 42.
[53] U. Amaldi, et al., CERN Yellow Report CERN-2012-004, 2012.
[54] M. Holder, et al., Phys. Lett. 35B (1971) 361.
[55] U. Amaldi, et al., Phys. Lett. 44B (1973) 112.
[56] S.R. Amendolia, et al., Phys. Lett. 44B (1973) 119.
[57] K. Eggert, et al., Nucl. Phys. B 98 (1975) 93.
[58] U. Amaldi, et al., Phys. Lett. 62B (1976) 460.
[59] U. Amaldi, et al., Nucl. Phys. B 145 (1978) 367.
[60] L. Baksay, et al., Nucl. Phys. B 141 (1978) 1.
[61] G. Carboni, et al., Nucl. Phys. B 254 (1984) 697.
[62] N. Amos, et al., Nucl. Phys. B 262 (1985) 689.
[63] K. Unser, IEEE Trans. Nucl. Sci. NS28 (1981) 3.
[64] K. Potter, S. Turner, IEEE Trans. Nucl. Sci. NS22 (1975) 1589.
[65] J.-Y. Hemery, F. Lemeilleur, R. Olsen, K. Potter, CERN-ISR-OP-CO-ES/81-12 and IEEE Trans. Nucl. Sci. NS28 (1981) 3034.
[66] P. Bryant, K. Potter, CERN-ISR-ES-BOM/82-15, 1982.
[67] W. Herr, R. Schmidt, Experience at CERN with Luminosity Monitoring and Calibration, in: CERN-Proceedings-2011-011. https://cds.cern.ch/record/1347440, and private communication.
[68] G. Arnison, et al., Phys. Lett. 121B (1982) 77.
[69] G. Arnison, et al., Phys. Lett 118B (1982) 167.
[70] G. Arnison, et al., Phys. Lett 128B (1983) 336.
[71] C. Augier, et al., Phys. Lett. B344 (1995) 451.
[72] A. Barisy, et al., IEEE Trans. Nucl. Sci. NS28 (1981) 2180.
[73] R. Bossart, et al., IEEE Trans. Nucl. Sci. NS32 (1985) 1899.
[74] M. Bozzo, et al., Phys. Lett. 147B (1984) 392.
[75] M. Banner, et al., Phys. Lett. 115B (1982) 59.
[76] G.J. Alner, et al., Z. Phys. C 32 (1986) 153.
[77] M. Banner, et al., Phys. Lett. 122B (1983) 322.
[78] P. Bagnaia, et al., Z. Phys. C 20 (1983) 117.
[79] J.A. Appel, et al., Phys. Lett. 160B (1985) 349.
[80] J. Alitti, et al., Z. Phys. C 47 (1990) 11.
[81] C. Augier, et al., Phys. Lett. B316 (1993) 448.
[82] R. Lauckner, IEEE Trans. Nucl. Sci. NS32 (1985) 1653.
[83] C. Albajar, et al., Nucl. Phys. B 335 (1990) 261.
[84] S.D. Holmes, Remembering the Tevatron: The Machine(s), in: Proc. DPF 2011 Conference, Providence, RI, Aug. 8–13, 2011.
[85] N.A. Amos, et al., Phys. Lett. B243 (1990) 158.
[86] C. Avila, et al., Phys. Lett. B445 (1999) 419.
[87] N.A. Amos, et al., Phys. Rev. Lett. 63 (1989) 2784.
[88] J. Safranek, Nucl. Instr. Meth. A 388 (1997) 27.
[89] T. Meyer, et al., FERMILAB-PUB-11-320-AD.
[90] F. Abe, et al., Phys. Rev. D50 (1994) 5550.
[91] S. Klimenko, J. Konigsberg, T.M. Liss, FERMILAB-FN-0741, 2003.
[92] J. Elias, et al., Nucl. Instr. Meth. A 441 (2000) 366.
[93] D. Acosta, et al., Nucl. Instr. Meth. A 461 (2001) 540.

[94] D. Acosta, et al., Nucl. Instr. Meth. A 494 (2002) 57.
[95] C.-C. Miao, FERMILAB-CONF-98/395, 1998.
[96] V.M. Abazov, et al., Nucl. Instr. Meth. A 565 (2006) 463.
[97] D. Cronin-Hennessy, et al., Nucl. Instr. Meth. A 443 (2000) 37.
[98] T. Andeen, et al., arXiv:1204.0461, FERMILAB-TM-2365, 2007.
[99] B. Casey, et al., Nucl. Instr. Meth. A 698 (2013) 208.
[100] T. Ludlam, Nucl. Instr. Meth. A 499 (2003) 428–432. The same volume offers several in-depth articles dedicated to the RHIC program.
[101] A. Drees, T. D'Ottavio, WE6PFP003, in: Proc. Part. Acc. Conference (PAC'09), Vancouver BC, Canada, 2009.
[102] S. Bültmann, et al., http://www.rhic.bnl.gov/pp2pp/documents/prop03.pdf, 2003.
[103] A. Drees, et al., TPPB032, Proc. Part. Acc. Conference (PAC'03), Portland, USA, 2003.
[104] A. Drees, BNL-96818-2011-IR, C-A/AP/441, 2011.
[105] A. Drees, Z. Xu, RPAH116, Proc. Part. Acc. Conference (PAC'01), Chicago, USA, 2001.
[106] A. Adare, et al., Phys. Rev. D 79 (2009) 012003. http://link.aps.org/doi/10.1103/PhysRevD.79.012003.
[107] A. Drees, BNL-102438-2013-IR, C-A/AP/488, 2013.
[108] L. Adamczyk, et al., Phys. Rev. D85 (2012) 092010.
[109] S. Adler, et al., Phys. Rev. Lett. 91 (2003) 241803.
[110] A. Drees, BNL-103466-2013-IR, C-A/AP/495, 2013.
[111] S. Adler, et al., Phys. Rev. Lett. 93 (2004) 202002.
[112] A. Adare, et al., Phys. Rev. Lett. 103 (2009) 012003.
[113] A. Adare, et al., Phys. Rev. D84 (2011) 012006.
[114] ATLAS Collaboration, JINST 8 (2013) P07004.
[115] CMS Collaboration, CMS-PAS-LUM-12-001, 2012.
[116] K. Oyama, arXiv:1305.7044v1 [nucl-ex], 2013.
[117] S. Mueller, Doctoral thesis, IEKP-KA/2011-01, Universitat Karlsruhe (TH) Karlsruhe, 2011.
[118] ALICE Collaboration, Eur. Phys. J. C73 (2013) 2456.
[119] R. Miyamoto, et al., WEOCN3, Proc. 2011 Part. Acc. Conference, New York, USA, 2011.
[120] H. Burkhardt, P. Grafstrom, LHC-PROJECT-Report-1019, 2007.
[121] W. Herr, LHC Project Report 628, 2003.
[122] W. Herr, et al., WEODA01, in: Proc. Int. Part. Acc. Conference (IPAC'11), San Sebastian, Spain, 2011.
[123] C. Barschel, et al., CERN-ATS-Note-2012-026 PERF, https://cdsweb.cern.ch/record/1425904/.
[124] G. Anders, et al., CERN-ATS-Note-2012-028 PERF, https://cdsweb.cern.ch/record/1427726/.
[125] G. Anders, et al., CERN-ATS-Note-2011-016 PERF, https://cdsweb.cern.ch/record/1333997/.
[126] P. Hopchev, CERN-THESIS-2011-210.
[127] LHCb Collaboration, JINST 9 (2014) P12005. http://dx.doi.org/10.1088/1748-0221/9/12/P12005.
[128] A. Alici, et al., CERN-ATS-Note-2012-029 PERF, https://cdsweb.cern.ch/record/1427728/.
[129] A. Boccardi, et al., CERN-ATS-Note-2013-034 TECH, https://cds.cern.ch/record/1556087.
[130] ATLAS Collaboration, Luminosity Determination in pp Collisions at $\sqrt{s} = 8$ TeV using the ATLAS Detector at the LHC, in preparation, to be submitted to Eur. Phys. J. C.
[131] ALICE Collaboration, arXiv:1402.4476 [nucl-ex], 2014.
[132] CMS Collaboration, CMS-PAS-SMP12-008, 2012.
[133] CMS Collaboration, CMS-PAS-LUM-13-001, 2013.
[134] CMS Collaboration, CMS-PAS-LUM-13-002, 2014.
[135] ALICE Collaboration, Nucl. Part. Phys. 38 (2011) 124131. ©CERN 2011. Published under licence by IOP Publishing Ltd.
[136] W. Kozanecki, T. Pieloni, J. Wenninger, CERN-ACC-NOTE-2013-0006, 2013.
[137] W. Herr, Beam–Beam effects and Dynamic β^*, in: Proc. Lumi Days 2012, CERN, 2012. http://indico.cern.ch/event/162948/session/4/contribution/27/material/paper/.
[138] H. Bartosik, G. Rumolo, CERN-ACC-NOTE-2013-0008, 2013.
[139] ALICE Collaboration, JINST 9 (2014) P11003.
[140] LHCb Collaboration, Phys. Lett. 693B (2010) 69.
[141] LHCb Collaboration, JINST 3 (2008) S08005.
[142] TOTEM Collaboration, JINST 3 (2008) S08007. http://dx.doi.org/10.1088/1748-0221/3/08/S08007.
[143] TOTEM Collaboration, Europhys. Lett. 101 (2013) 21004.
[144] ALICE Collaboration, arXiv:1403.3648 [nucl-ex], 2014.