ATLAS CONF Note

ATLAS-CONF-2019-021

3rd June 2019

Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC

The ATLAS Collaboration

This note describes the preliminary determination of the integrated luminosity for pp collisions at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector during Run 2 of the LHC in the years 2015–2018. The absolute luminosity scale was determined using van der Meer scans during dedicated running periods in each year, and extrapolated to the physics data-taking regime using complementary measurements from several luminosity-sensitive detectors. The total uncertainties on the integrated luminosities for each individual year of data-taking range from 2.0 to 2.4 %, and are partially correlated between years. After typical data-quality selections, the full Run 2 pp data sample corresponds to an integrated luminosity of 139 fb^{-1}, with an uncertainty of 1.7 %.

ATLAS-CONF-2019-021
03 June 2019

1 Introduction

A precise measurement of the integrated luminosity is a key component of the ATLAS physics programme at the LHC, in particular for cross-section measurements where it is often one of the leading sources of uncertainty. This note describes the preliminary luminosity calibration for pp collisions at $\sqrt{s} = 13\,\text{TeV}$ recorded by the ATLAS detector [1] during LHC Run 2. This measurement builds on the experience and techniques developed for Run 1, documented extensively in Refs. [2–4].

The luminosity measurement is based on an absolute calibration of the primary luminosity-sensitive detectors in low-luminosity runs with specially-tailored LHC conditions using the van der Meer (vdM) method [5, 6]. A calibration transfer procedure is then used to transport this calibration to the physics data-taking regime at high luminosity. The vdM calibration was performed once per year during Run 2 data-taking, and relative comparisons of the luminosities measured by different detectors were used to set limits on any possible change of the calibration through the year. Finally, the integrated luminosity and uncertainty for the whole Run 2 data-taking period was derived, taking into account correlations between the uncertainties in each of the component years. The structure of this note mirrors the calibration methodology: the various luminosity-sensitive detectors are described in Section 2, the absolute vdM calibration in Section 3, the calibration transfer procedure in Section 4, the long-term stability studies in Section 5 and the correlations and combination in Section 6. The latter also contains a summary (Table 2) of the contributing uncertainties in all individual years and the combination. Finally, conclusions are given in Section 7.

The data-taking conditions evolved significantly during Run 2, with the LHC peak instantaneous luminosity at the start of fills ($\mathcal{L}_{\text{peak}}$) increasing from 5 to $19 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$ as the number of colliding bunch pairs n_b and average bunch current were increased, together with progressively stronger focusing in the ATLAS and CMS interaction regions (characterised by the β^* parameter [7]) leading to smaller transverse beam sizes. Table 1 shows an overview of typical parameters for each year, together with the total delivered integrated luminosity. The mean number of inelastic pp interactions per crossing μ, also known as the pileup parameter, characterises the instantaneous luminosity at any given time, and this quantity averaged over all colliding bunch pairs is denoted by $\langle\mu\rangle$.[1]

The evolution of LHC conditions, together with changes in the ATLAS subdetector operation and data analysis, result in small differences in the luminosity analysis for each data-taking year. However, the overall strategy is similar for all years, and the plots shown below illustrate examples from 2016–2018 (the 2015 data comprises less than 3 % of the full sample).

2 Luminosity detectors and algorithms

The luminosity calibration relies on multiple redundant luminosity detectors and algorithms, which have complementary capabilities and different systematic uncertainties. For LHC Run 2, the primary bunch-by-bunch luminosity measurement was provided by the LUCID2 Cherenkov detector [8], upgraded from its Run 1 configuration and referred to hereafter as LUCID. This was complemented by bunch-by-bunch measurements from the ATLAS beam conditions monitor (BCM) diamond detectors, and from offline measurements of the multiplicity of reconstructed charged particles in randomly-selected bunch-crossings (track counting). The ATLAS calorimeters provided bunch-integrated measurements

[1] A reference inelastic cross-section of $\sigma_{\text{inel}} = 80\,\text{mb}$ is assumed when converting between instantaneous luminosity and μ.

Parameter	2015	2016	2017	2018
Maximum number of colliding bunch pairs (n_b)	2232	2208	2544/1909	2544
Bunch spacing (ns)	25	25	25/8b4e	25
Typical bunch population (10^{11} protons)	1.1	1.1	1.1/1.2	1.1
β^* (m)	0.8	0.4	0.3	0.3–0.25
Peak luminosity $\mathcal{L}_{\text{peak}}$ ($10^{33}\,\text{cm}^{-2}\text{s}^{-1}$)	5	13	16	19
Peak number of inelastic interactions/crossing ($\langle\mu\rangle$)	~ 16	~ 41	$\sim 45/60$	~ 55
Luminosity-weighted mean inelastic interactions/crossing	13	25	38	36
Total delivered integrated luminosity (fb^{-1})	4.0	38.5	50.2	63.4

Table 1: Selected LHC parameters for pp collisions at $\sqrt{s} = 13\,\text{TeV}$ in 2015–2018. The values shown are representative of the best accelerator performance during normal physics operation. In 2017, the LHC was run in two modes: standard 25 ns bunch train operation with long trains, and '8b4e', denoting a pattern of eight bunches separated by 25 ns followed by a four bunch-slot gap. Values are given for both configurations. The instantaneous luminosity was levelled by beam separation to about $\mathcal{L}_{\text{peak}} = 16 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$ for part of the 8b4e period. The $0.1\,\text{fb}^{-1}$ of physics data delivered during 2015 with 50 ns bunch spacing is not included.

based on quantities proportional to instantaneous luminosity: liquid-argon (LAr) gap currents in the case of the endcap electromagnetic (EMEC) and forward (FCal) calorimeters, and photomultiplier currents from the scintillating-tile hadronic calorimeter (TILE). Additional measurements were provided by cluster counting in a network of TimePix readout sensors (TPX) distributed throughout the ATLAS cavern. The LUCID and track-counting measurements are discussed further below, and more details on the other detectors can be found in Refs. [4, 9].

The LUCID detector contains 16 photomultiplier tubes (PMTs) in each forward arm of the ATLAS detector (side A and side C), placed at approximately $z = \pm 17\,\text{m}$ from the interaction point.[2] Cherenkov light is produced in the quartz windows of the PMTs, which are coated with ^{207}Bi radioactive sources that provide a calibration signal. The LUCID detector was read out with dedicated electronics which provided luminosity counts for each of the 3564 nominal LHC bunch slots where a colliding bunch pair could be present. These counts were integrated over well-defined time periods (called luminosity blocks) with a typical length of 60 seconds. Several luminosity 'algorithms' were used to convert the raw signals from the PMTs to a luminosity measurement, combining the information from several PMTs in various ways. In 2015, the baseline luminosity estimates were derived from BiEventORA, an algorithm requiring a hit in any of four bismuth-calibrated PMTs on the A-side of the detector (analogous algorithms were available for the C-side and the combination of both sides, but the C-side suffered from significant timing drifts as the PMT high-voltage settings were adjusted). In 2016 and 2017, the BiHitOR algorithm was used, counting the average number of hits per bunch crossing summed over four bismuth-calibrated PMTs on the A-side and four more on the C-side. In 2018, a significant number of PMTs stopped working during the course of the data-taking year, and a single PMT on the C-side (C12) was used for the baseline luminosity estimate, because it showed good stability throughout the year and gave similar results to a more complicated HitOR-type combination of the remaining seven working PMTs.

In all cases, the raw hit counts were converted into a visible interaction rate per bunch crossing μ_{vis}, which is proportional to the instantaneous luminosity. For example, for the HitOR algorithm combining

[2] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector. The x axis direction points towards the centre of the LHC ring, the y axis vertically upwards, and the z axis lies along the beam line. Pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan\theta/2$, and transverse momentum and energy are defined relative to the beamline as $p_{\text{T}} = p\sin\theta$ and $E_{\text{T}} = E\sin\theta$.

N_{PMT} PMTs, the average probability P_{HIT} to have a hit in any given PMT during the N_{BC} colliding bunch crossings of one luminosity block is given from the total number of hits summed over all PMTs N_{HIT} by

$$P_{\text{HIT}} = \frac{N_{\text{HIT}}}{N_{\text{BC}} N_{\text{PMT}}} = 1 - e^{-\mu_{\text{vis}}} ,$$

leading to $\mu_{\text{vis}} = -\ln(1 - P_{\text{HIT}})$ [3]. The resulting μ_{vis} is related to the per-bunch instantaneous luminosity $\mathcal{L}_b$ via

$$\mathcal{L}_b = \frac{\mu_{\text{vis}} f_r}{\sigma_{\text{vis}}} , \tag{1}$$

where f_r is the LHC revolution frequency (11246 Hz for protons). The total instantaneous luminosity is the sum of $\mathcal{L}_b$ over all colliding bunch pairs. The algorithm-specific visible cross-section σ_{vis} in Eq. (1) is a calibration constant which represents the absolute luminosity calibration of the given algorithm, and is determined experimentally via the vdM calibration discussed in Section 3. An analogous formalism was used for the bunch-by-bunch measurements from the four diamond BCM sensors on each side of the interaction point.

The track-counting method takes the per-bunch visible interaction rate μ_{vis} from the mean number of reconstructed tracks per bunch crossing averaged over a luminosity block. The measurement was derived from randomly-sampled colliding-bunch crossings, where only the data from the SCT and pixel detectors (including the Inner B-Layer (IBL) [10, 11] installed before Run 2) were read out, typically at 200 Hz during normal physics data-taking and at much higher rates during vdM scans. These events were saved in a dedicated event stream which was then reconstructed offline. The track selection evolved over the course of Run 2, in order to improve the robustness against time-dependent detector inefficiencies, and reduce the efficiency loss and fake-track contamination at high pileup. For 2015 data, tracks were required to have transverse momentum $p_T > 0.9$ GeV and to satisfy the 'TightPrimary' selection of Ref. [12]. In 2016, a requirement $|d_0/\sigma_{d_0}| < 7$ was added, where d_0 is the transverse impact parameter of the reconstructed track with respect to the beamline, and σ_{d_0} is the uncertainty on the measured d_0. This requirement rejects tracks that are inconsistent with originating directly from a pp collision. For 2017 and 2018, tracks were restricted to the barrel region of the ATLAS inner detector, $|\eta| < 1.0$, and up to one pixel 'hole' was allowed, i.e. a track with a pixel hit missing where one was expected. The average number of reconstructed tracks per $\sqrt{s} = 13$ TeV inelastic pp collision was about 3.7 for the 2015–16 selections, and 1.7 for the 2017–18 selection.

3 Absolute luminosity calibration

The absolute luminosity calibration of LUCID and BCM, corresponding to the determination of the visible cross-section σ_{vis} for each of the LUCID and BCM algorithms, was derived using dedicated vdM scan sessions during special LHC fills in each data-taking year. The calibration methodology and main sources of uncertainty are the same as those for $\sqrt{s} = 8$ TeV as described in Ref. [4], and only a brief overview is given here.

The instantaneous luminosity for a single colliding bunch pair, $\mathcal{L}_b$, is given in terms of LHC beam parameters by

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

where f_r is the LHC revolution frequency, n_1 and n_2 are the numbers of protons in the beam-1 and beam-2 colliding bunches, and Σ_x and Σ_y are the convolved beam sizes in the horizontal and vertical planes [6]. The quantity Σ_x can be determined using the instantaneous luminosity $R(\Delta x)$ measured as a function of the separation Δx between the two beams in the horizontal plane:

$$\Sigma_x = \frac{1}{\sqrt{2\pi}} \frac{\int R(\Delta x)\, \mathrm{d}\Delta x}{R(0)}\,, \tag{3}$$

i.e. the ratio of the integral over the horizontal scan to the peak instantaneous luminosity $R(0)$ (and similarly for Σ_y with a vertical scan). If the form of $R(\Delta x)$ is Gaussian, Σ_x is equal to the standard deviation of the distribution, but the method is valid for any functional form of $R(\Delta x)$. However, the formulation does assume that the particle densities in each bunch can be factorised into independent functions of x and y, and the effect of violations of this assumption (i.e. non-factorisation) is quantified below. Since the normalisation of $R(\Delta x)$ cancels out in Eq. (3), any quantity proportional to luminosity can be used to determine the scan curve. The calibration of a given algorithm (i.e. its σ_{vis} value) can then be determined by combining Eqs. (1) and (2) to give

$$\sigma_{\mathrm{vis}} = \mu_{\mathrm{vis}}^{\mathrm{max}} \frac{2\pi\Sigma_x\Sigma_y}{n_1 n_2}, \tag{4}$$

where $\mu_{\mathrm{vis}}^{\mathrm{max}}$ is the visible interaction rate per bunch crossing at the peak of the scan curve. A single pair of x–y vdM scans thus suffices to measure σ_{vis} for each algorithm active during the scan. Since the quantities entering Eq. (4) are different for each colliding bunch pair, it is essential to perform a bunch-by-bunch analysis to determine σ_{vis}, in practice limiting the vdM absolute calibration in ATLAS to the LUCID and BCM luminosity algorithms (track-counting cannot be used since only a limited rate of bunch crossings can be read out for offline analysis).

In order to minimise uncertainties, vdM scans were not performed in standard physics running, but rather in dedicated low-luminosity vdM scan fills during each year of Run 2. These were performed with special LHC optics settings with zero crossing angle and the parameter β^* set to 19.2 m rather than the 0.25–0.8 m used during normal physics running, and larger beam emittances of 3–4 μm-rad. These parameters were chosen to increase the transverse single-beam sizes to about 90 μm, thus reducing the uncertainties in evaluating the non-factorisation corrections discussed below. Filling schemes with 30–140 isolated bunches were used, to avoid the effects of long-range encounters between incoming and outgoing bunches away from the interaction point that occur in normal bunch train running. Furthermore, moderate bunch currents of $\sim 0.8 \times 10^{11}$ protons/bunch were employed to minimise beam-beam effects. Special care was taken in the LHC injector chain to produce beams with Gaussian-like transverse profiles in x and y to minimise non-factorisation effects in the scans. These configurations typically resulted in $\langle\mu\rangle \approx 0.5$ at the peak of the scan curves, whilst maintaining measurable rates in the tails of the scans at up to 6σ nominal separation. A typical scan pair took 2×20 minutes for x and y scans, and a vdM fill lasted up to one day, with alternating groups of scans in ATLAS and CMS. The final dataset for each year consists of several pairs of on-axis x–y scans, spaced through one or more fills in order to study reproducibility, together with one or two pairs of off-axis scans, where the beam was scanned in x with an offset in y of e.g. 300 μm (and *vice versa*) to study the beam profiles in the tails in order to constrain non-factorisation effects.

A typical x scan curve from 2017 using the LUCID BiHitOR algorithm is shown in Figure 1. The scan profile was fitted using an analytic function (e.g. a Gaussian multiplied by a polynomial), after subtracting backgrounds. These are dominated by the bismuth calibration source signal, estimated from the counting rate in the preceding empty bunch slot, followed by beam-gas interactions, estimated from 'unpaired' bunch

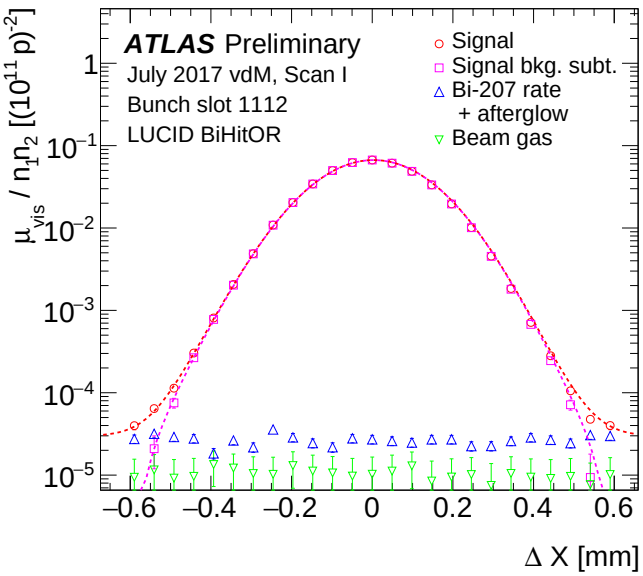
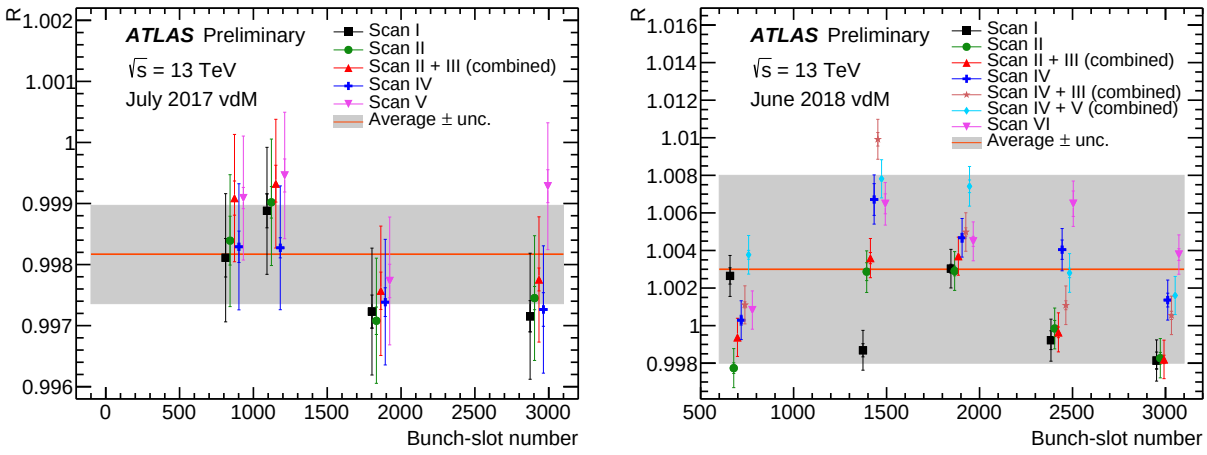


Figure 1: Visible interaction rate μ_{vis} per unit bunch population product $n_1 n_2$ *vs.* beam separation Δx in the horizontal plane, as measured by the LUCID BiHitOR algorithm for bunch slot 1112 in scan I of the July 2017 vdM session. The total rate is shown by the red circles, the background-subtracted rate by the magenta squares, and the fit functions by the dashed lines. The estimated background contributions from the bismuth calibration source and beam-gas interactions are shown by the blue and green triangles. The error bars are statistical, and in most cases smaller than the marker sizes.

crossings, where only one beam has a bunch present. The fitted scan curve was used to determine the peak interaction rate $\mu_{\text{vis}}^{\text{max}}$ and fit integral, from which Σ_x was calculated using Eq. (3). Systematic uncertainties arise from the choice of fit model, estimated by using different functional forms, and from uncertainties in the background subtraction procedure, explored e.g. by modelling the background as a constant term included in the fit function (but not the $\Sigma_{x,y}$ integral) rather than using an explicit background correction to the data before fitting.

The determination of the bunch populations n_1 and n_2 in Eq. (4) was based on the LHC beam instrumentation, supplemented by measurements of beam-gas event rates by the LHCb experiment [13]. The total intensity in each beam was measured by the LHC DC current transformers (DCCT), which have an absolute precision of better than 0.1 % but lack the ability to resolve individual bunches. The per-bunch intensity was determined from the fast beam-current transformers (FBCT) which can resolve the current in each of the 3564 nominal 25 ns bunch slots in each beam, normalised to the DCCT measurement at each vdM scan step. Systematic uncertainties were assessed via comparisons with the ATLAS beam-pickup timing system, which can also resolve the relative populations in each bunch. Further corrections were made using LHCb beam-gas measurements of ghost charge, i.e. small numbers of protons circulating in nominally-empty bunch slots which contribute to the DCCT current, but are below the threshold of the FBCT, and using the LHC longitudinal density monitor to measure satellites, i.e. protons captured in the 2.5 ns wide RF buckets in the same bunch slot as the nominally filled buckets. The ghost and satellite corrections were typically $O(0.1)$ % with systematic uncertainties below half the size of the correction.

A number of additional effects need to be taken into account compared to the simplified description above, leading to corrections and additional uncertainties as discussed in detail e.g. in Ref. [4]. Orbit drifts of



(a) (b)

Figure 2: Non-factorisation correction factors R for several colliding bunch pairs and vdM scan sets in the (a) 2017 and (b) 2018 vdM sessions. The results are extracted from combined fits to the luminosity scan curves and reconstructed primary vertex data (see text), using either on-axis scans or combined fits to on- and off-axis scans. The red lines show the error-weighted mean corrections, and the grey bands the uncertainties assigned from the spread of the measured R values. The results from different scans for each colliding bunch pair are slightly offset on the horizontal axis for clarity.

up to $O(10)\,\mu$m in the position of one or both beams have been observed during the course of a single vdM scan, which change the actual beam separation from the intended one induced by the LHC corrector magnets. These were monitored using two independent beam position monitor (BPM) systems, one located in the LHC arcs, and one at the entrance to the final-triplet quadrupole at $z = \pm 22$ m from the interaction point. Residual uncertainties were estimated by comparing the results of various arc– and triplet-based BPM correction strategies. The arc BPMs were also used to set an upper limit on beam position jitter, i.e. modulation of the actual beam separation on time scales shorter than the typical 30 s of a single vdM-scan step. The electromagnetic interaction between the two beams produces two effects that distort the overlap integral of the colliding bunches: a mutual deflection that induces a non-linear distortion of the intended beam separation, and the defocusing of one beam by the other that modulates the optical demagnification from the LHC arcs to the interaction point. The associated corrections depend on the beam separation, the beam energy, the transverse beam sizes, the bunch currents, the actual β^* values and the fractional LHC tunes. They were calculated using MADX simulations [14], giving total upward corrections of 1.3–1.7 % on σ_{vis}. The associated uncertainties were calculated using 20 % variations in the assumed β^* and ± 0.01 in the tune values, as in Ref. [4], and appear as 'Beam-beam effects' in Table 2. A further uncertainty arises from potential variations of the horizontal and/or vertical beam emittances over the course of an x–y scan pair; as discussed in Ref. [4]; this only leads to a bias if the horizontal and vertical emittances grow at different rates. No study of this effect with the Run 2 vdM scans has been carried out yet, so a common uncertainty of 0.2 % was assigned to all years, based on the largest effect seen in any of the November 2012 scans at $\sqrt{s} = 8$ TeV, excluding scan set X for the reasons discussed in Ref. [4].

The vdM formalism described by Eqs. (2–4) assumes that the particle densities in each bunch can be factorised into independent horizontal and vertical components, such that the term $1/2\pi\Sigma_x\Sigma_y$ in Eq. (2) fully describes the overlap integral of the two beams. Evidence for non-factorisation was clearly seen

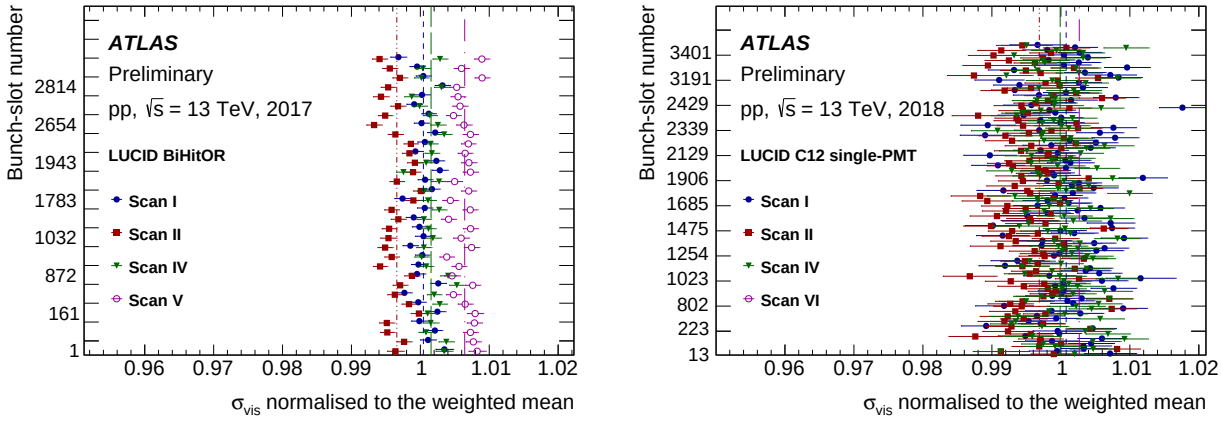
during Run 1 [4], especially when dedicated beam-tailoring in the LHC injectors was not used. Following the Run 1 analysis, non-factorisation corrections were evaluated in each Run 2 vdM scan session from a combined fit to the beam-separation dependence of the LUCID luminosity and of the position, orientation and shape of the luminous region, characterised by the three-dimensional spatial distribution of the primary collision vertices formed from tracks reconstructed in the ATLAS inner detector. Since this procedure requires a large number of reconstructed vertices per scan step, it was carried out for only a handful of colliding bunch pairs, whose tracking information was read out at an enhanced rate. As discussed in Ref. [4], non-factorisable single-beam profiles were fitted to these data, and used in simulated vdM scans to determine the ratio R between the apparent luminosity scale (measured assuming factorisation) and the true luminosity scale determined from the actual beam overlap integral. These ratios were then used to correct the visible cross-section from the standard factorisable analysis: $\sigma_{\text{vis}}^{\text{corr}} = \sigma_{\text{vis}}/R$. Examples of the per-bunch and per-scan corrections are shown for the 2017 and 2018 vdM scan sessions in Figure 2, from fits to on-axis scans alone, or combined fits to an on-axis scan and a close-in-time off-axis scan. The final corrections applied to the vdM scans for each year 2015–2018 are $R = 1.00 \pm 0.01$, 1.006 ± 0.004, 0.998 ± 0.002 and 1.003 ± 0.005 respectively, where the errors were evaluated from the spread of corrections from different bunches and scan sets.

The $\Sigma_{x,y}$ measurements also require knowledge of the length scale, i.e. the actual beam separation produced by a particular setting of the steering magnets. As in Run 1, this was measured by displacing both beams transversely by five steps over a range of up to ± 3 times the nominal beam sizes, at each step keeping the beams well centred on each other in the scanning plane, and by measuring the actual displacement of the luminous region in the ATLAS inner detector using the positions of reconstructed primary collision vertices. The uncertainties in this procedure are dominated by orbit drifts during the calibration scans, estimated using triplet and arc BPMs as for the vdM scans, and are typically $\sim 0.3\,\%$. A further uncertainty of $0.1\,\%$ arises from the absolute length scale of the inner detector, determined from simulation studies of various realistic misalignment scenarios, including the impact of the precise measurements from the new IBL pixel layer.

The ability to determine σ_{vis} separately from each colliding bunch pair and x–y scan set allows the stability and reproducibility of the calibration to be checked. Figure 3 shows the σ_{vis} values measured for each bunch-pair and scan in 2017 and 2018 data, after all corrections have been applied, and normalised to the weighted mean of all bunch pairs and scan sets. The relative spread of σ_{vis} values within one scan was used to set the per-year 'bunch-by-bunch σ_{vis} consistency' uncertainty shown in Table 2, after correcting for the spread expected due to statistical fluctuations. The largest bunch-averaged difference seen between any pairs of scans, using a variety of luminosity algorithms, was symmetrised and used to set the 'scan-to-scan reproducibility' uncertainty in each year. This gave the leading uncertainty on the absolute vdM calibration in the 2017 analysis ($1.2\,\%$), but only half that in the other years, similar to other uncertainties on the vdM calibration. Results from different LUCID and BCM algorithms were also considered. These are not expected to give comparable σ_{vis} values (as the acceptance of each algorithm is different), but should give consistent values of $\Sigma_{x,y}$ for each bunch, as quantified by the specific luminosity

$$\mathcal{L}_{\text{spec}} = \frac{\mathcal{L}_b}{n_1 n_2} = \frac{f_r}{2\pi \Sigma_x \Sigma_y}\,, \tag{5}$$

obtained from the per-bunch luminosity in Eq. (2). As an example, Figure 4 shows the distribution of per-bunch $\mathcal{L}_{\text{spec}}$ ratios for the C5 and baseline C12 LUCID single-PMT algorithms for all scan sets in 2018, C5 being chosen as it showed the largest discrepancies with respect to C12. The ratios of $\mathcal{L}_{\text{spec}}$ values measured by different algorithms and averaged over all colliding-bunch pairs were used to define the 'reference specific luminosity' uncertainties shown in Table 2.



(a) (b)

Figure 3: Ratios of bunch-by-bunch visible cross-sections to the weighted mean of σ_{vis} for all colliding bunch pairs and on-axis scans in (a) the 2017 vdM scan session, using the LUCID BiHitOR algorithm, and (b) the 2018 vdM scan session, using the LUCID C12 single-PMT algorithm. The vertical lines show the mean ratio for each individual scan set. The uncertainties are statistical, and are larger in 2018 due to the use of a single PMT rather than the combination of eight PMTs in the BiHitOR algorithm.

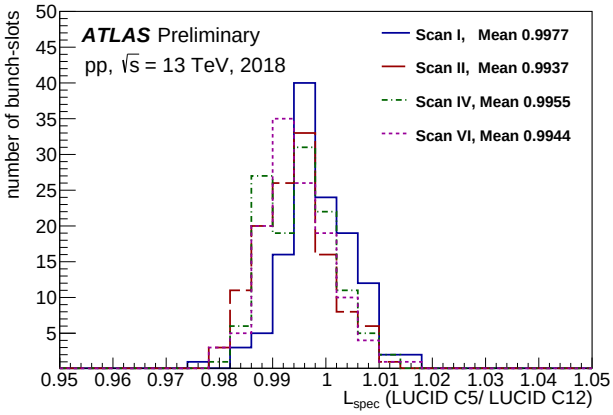
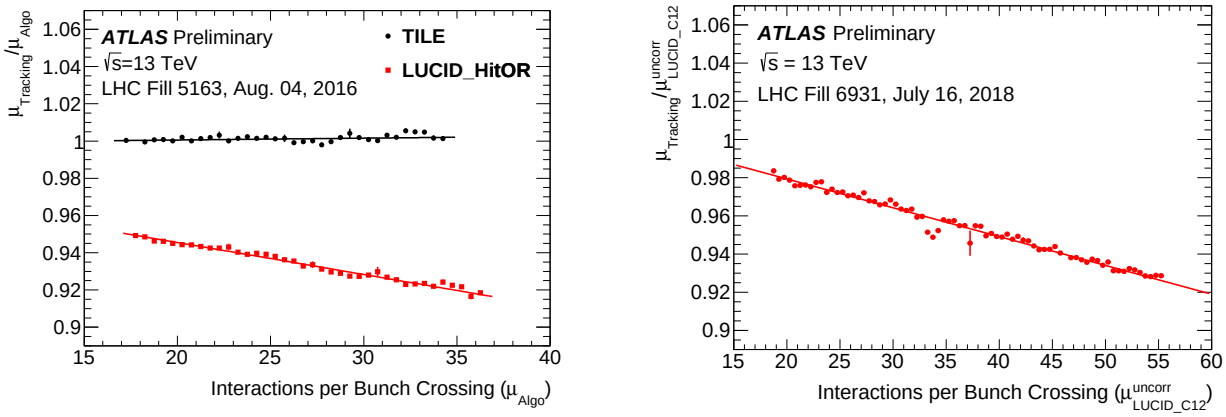


Figure 4: Ratios of specific luminosities measured using the LUCID C5 and C12 single-PMT algorithms for different colliding-bunch pairs and scan sets in the 2018 vdM scan session. The mean ratios over all bunch pairs for each scan set are also shown.

The total uncertainties on the absolute vdM calibration are listed in the 'Subtotal vdM calibration' row of Table 2, and range from 1.1% to 1.5 % across the different years of the Run 2 data-taking.

4 Calibration transfer to physics data-taking conditions

The procedures discussed above provide the absolute calibration of the LUCID (and BCM) luminosity measurements for low-μ data-taking with a limited number of isolated bunches, where the instantaneous luminosities are at least three orders of magnitude smaller than those typical of normal physics running with



(a) (b)

Figure 5: (a) Ratios of instantaneous luminosities measured by track-counting and LUCID BiHitOR (red points), and track-counting and TILE D6 cells (black points), as a function of the number of interactions per bunch-crossing μ measured by LUCID or TILE, in LHC physics fill 5163 with 2064 colliding bunches in 2016. The TILE integrated luminosity is normalised to that of track-counting during this fill. The lines show linear fits to the points. (b) Analogous ratio of track-counting to LUCID C12 single-PMT luminosity in LHC physics fill 6931 with 2544 colliding bunch-pairs, recorded in 2018.

$\mu \approx 50$ and 25 ns bunch trains (see Table 1). The LUCID detector suffers from significant non-linearity, and requires a correction of $O(10\,\%)$ in the physics data-taking regime. This effect can clearly be seen by studying the ratios between instantaneous luminosities measured with LUCID and with other detectors, and how they change as a function of instantaneous luminosity or μ.

The effects of this non-linearity were corrected by comparing LUCID to track-counting luminosity measurements, which are known from comparisons with EMEC and TILE to have much smaller non-linearity as a function of μ. In 2016 and 2018 data-taking a single long high-luminosity physics fill, near in time to the vdM calibration period, was used to derive a correction of the form

$$\mu_{\text{corr}} = p_0 \mu_{\text{uncorr}} + p_1 \mu_{\text{uncorr}}^2 \,, \tag{6}$$

where μ_{uncorr} is the uncorrected and μ_{corr} the corrected LUCID μ value, and the parameters p_0 and p_1 were obtained from a linear fit to the ratio of μ values measured by track-counting and LUCID, $R = \mu_{\text{track}}/\mu_{\text{uncorr}}$, as a function of μ_{uncorr}. The track-counting luminosity was first normalised to the absolute luminosity measured by LUCID in parts of the vdM fill with stable, almost-constant luminosity, where the beams were colliding head-on in ATLAS (typically while CMS was performing vdM scans), thus ensuring that $R = 1$ at low luminosity with isolated bunches. In practice, the corrections were obtained from Eq. (6) written in terms of $\langle\mu_{\text{uncorr}}\rangle$, averaged over all colliding-bunch pairs at any given time, rather than the per-bunch μ_{uncorr}. This correction also absorbs the $O(1\,\%)$ change in the LUCID calibration caused by the non-zero LHC beam crossing angle used in physics runs. The 2016 correction to the LUCID BiHitOR algorithm is illustrated in Figure 5(a), and the 2018 correction to the C12 single-PMT algorithm in Figure 5(b).

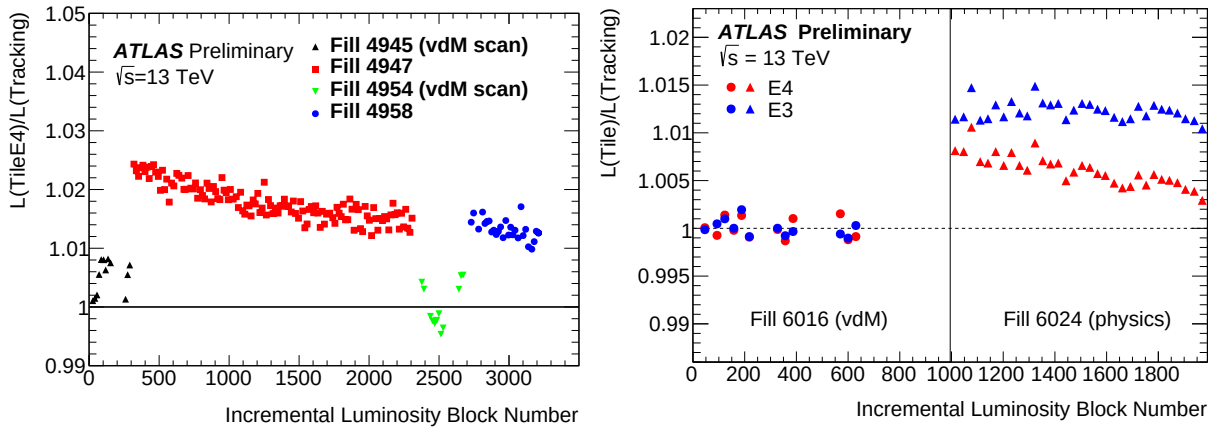
This procedure can be carried out for any long physics fill with a significant reduction in $\langle\mu\rangle$ over time, allowing the stability of the p_0 and p_1 correction parameters to be monitored throughout the data-taking year. Such studies showed that single sets of correction parameters were appropriate for the 2016 and 2018 datasets, but that the 2017 data could be better characterised as two separate periods, with p_0 and p_1

parameters determined from LHC fill 6024 (29th July) being used for data up to fill 6072 (12th August), and a separate set of parameters determined from fill 6259 (30th September) being used thereafter. A simpler procedure was used for the 2015 dataset, where the corrections were smaller due to the lower $\langle\mu\rangle$ values. For this dataset, a constant correction factor of $-2.6\,\%$ was applied based on the ratio of total integrated luminosities measured by track-counting and LUCID in seven physics fills recorded before the vdM fill, and no attempt was made to correct the instantaneous luminosity within fills as a function of μ.

This calibration transfer procedure implicitly assumes that track-counting suffers from no significant non-linearity between the low-luminosity vdM and high-luminosity physics regimes. This assumption was probed using both 'internal' studies comparing track-counting results with different track selections (working points), and 'external' studies comparing track-counting to independent detectors. For example, Figure 5(a) also shows the ratio of luminosities measured by track-counting and by the TILE D6 cells, located in the extended barrel of the calorimeter [15], for the reference run used to determine the 2016 LUCID μ correction. This ratio shows no significant relative non-linearity over the full range of instantaneous luminosities relevant in normal physics running. However, the lowest instantaneous luminosity probed here is still a factor ~ 1000 larger than that of the vdM regime, where the D6 cells have no sensitivity. The only other detectors with adequate sensitivity in both the vdM and physics regimes are the E1 to E4 scintillators of the tile calorimeter, installed in the gap between the barrel and endcap calorimeter assemblies, and designed primarily to measure energy loss in this region. Of these scintillators, the E4 cells are closest to the beamline and therefore the most sensitive at low luminosity. The calibration transfer uncertainty was evaluated from the difference between the ratio $R_{\mathrm{TILE/trk}} = \mathcal{L}_{\mathrm{TILE-E}}/\mathcal{L}_{\mathrm{track}}$ between TILE E-cell and track-counting luminosity measurements in the low-luminosity vdM run and a close-in-time high-luminosity physics fill. A difference in the $R_{\mathrm{TILE/trk}}$ values observed in the two regimes indicates a non-linearity between their relative responses, which was interpreted as an upper limit on the non-linearity of the track-counting luminosity measurement and thus assigned as a systematic uncertainty on the correction applied to LUCID at high luminosity.

In practice, these studies are made significantly more complicated due to material activation effects and rapid scintillator ageing in the tile calorimeter E cells. The residual activation from a high-luminosity physics fill preceding the vdM fill produces a TILE luminosity signal which is initially of similar magnitude to the typical luminosity in a vdM fill, and which decays with various time constants, the largest of which are longer than ten hours. Care was therefore taken when scheduling the sequence of LHC fills to avoid high-luminosity running immediately before the vdM fills. Extended pedestal runs were taken before and after the vdM fills, allowing exponential fits to be used to subtract the residual activation from under the luminosity signals during the vdM running.

Figure 6 illustrates this procedure for (a) 2016 and (b) 2017 data. In 2016, two vdM scan fills were performed (fills 4945 and 4954), each followed by a high-luminosity fill with 1165 (fill 4947) or 1453 (fill 4958) colliding bunch pairs. A clear difference of around 1.5–2 % is seen between the two regimes. Furthermore, the ratios are not constant within a fill, indicating residual activation effects not fully accounted for by the pedestal subtraction procedure, and scintillator ageing at high luminosity. The ratio also decreases by around 0.5 % from the first to the second vdM fill, due to the short-term ageing in the scintillator tiles. From these studies, a calibration transfer uncertainty of 1.6 % was derived, as the average of the run-integrated shifts between vdM and physics running for the E4 and E3 (not shown) cells and the two run pairs. Figure 6(b) shows the corresponding study in 2017 data, with a single vdM fill (6016) followed soon after by a 2544 colliding-bunch-pair physics fill (6024). Here, the activation effects in the vdM fill were better under control, and an uncertainty of 1.3 % was derived from the average of E1–E4 cell results (only E3 and E4 are shown). In 2018, scheduling constraints resulted in the vdM data being taken directly after a



(a) (b)

Figure 6: Ratios of the instantaneous luminosity measured by TILE E-cell scintillators to that from track-counting, in vdM fills and closely-following physics fills. The luminosity block numbers for each fill are displayed with a fill-dependent offset to present the data as a continuous sequence, with gaps between the fills being suppressed. (a) Ratios using TILE E4 cells for the two vdM fills and following physics fills in 2016 data, averaged over 15 minute intervals and with the integrated ratio normalised to unity in the second vdM fill (green points); (b) Ratios using both TILE E3 and E4 cells in the vdM and following physics fill in 2017 data, averaged over 30 minute intervals and with the integrated ratio normalised to unity in the vdM fill.

2448 colliding-bunch-pair physics fill, resulting in a large TILE pedestal signal from material activation. To mitigate this, a dedicated 140 bunch fill with a low-luminosity $\langle\mu\rangle \approx 0.5$ period mimicking vdM conditions took place five days before the vdM fill, with no high-luminosity running in the week before this fill. Since the data from this campaign is still being analysed, the calibration transfer uncertainty from 2017 is also used for the current 2018 luminosity calibration. The performance of the track-counting measurement (in particular any non-linearity) is expected to be very similar in the two datasets.

A further uncertainty in the LUCID luminosity measurement in physics running arises from the contributions from afterglow and beam-halo. This was estimated using similar procedures to those used at Run 1 [3], and an uncertainty of 0.1 % was assigned.

5 Long-term stability

The vdM calibration, followed by calibration transfer to the physics regime using a close-in-time physics fill, determines the absolute LUCID luminosity scale at one point during the data-taking year. To quantify any possible drifts over the year, the LUCID integrated luminosity estimates (corrected according to Eq. (6)) were compared for each physics run with those from other subdetectors, after renormalising the other luminosity estimates to agree with LUCID in a long 'reference' run close to the vdM scan. Long-term stability was also studied within a single luminosity detector, for example comparing different sets of PMTs and algorithms within LUCID, different track-counting working points, and different subsets of calorimeter cells. As a result of these studies, several corrections were made to the luminosity estimates from individual detectors. The time-dependence of the LUCID μ-correction in 2017 was already discussed in Section 4. In addition, the first 12 runs in 2017 (corresponding to 3 % of the total 2017 integrated luminosity) were corrected upwards by 2 % based on comparison with track-counting and calorimeters to

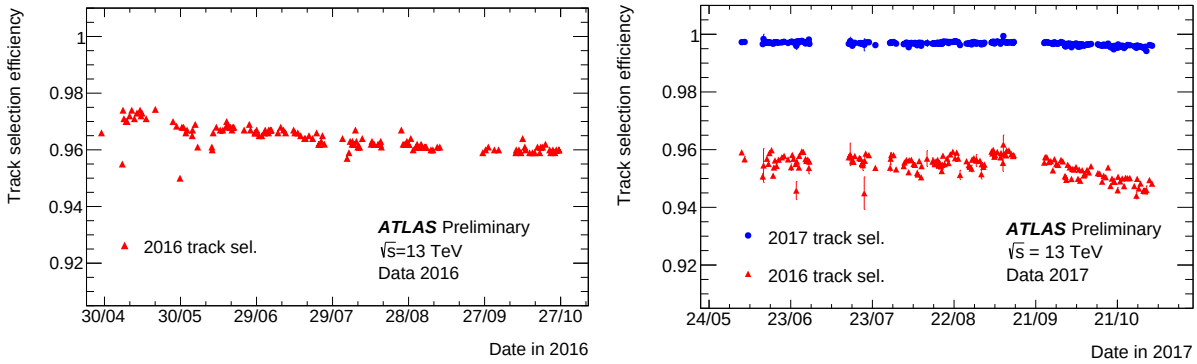
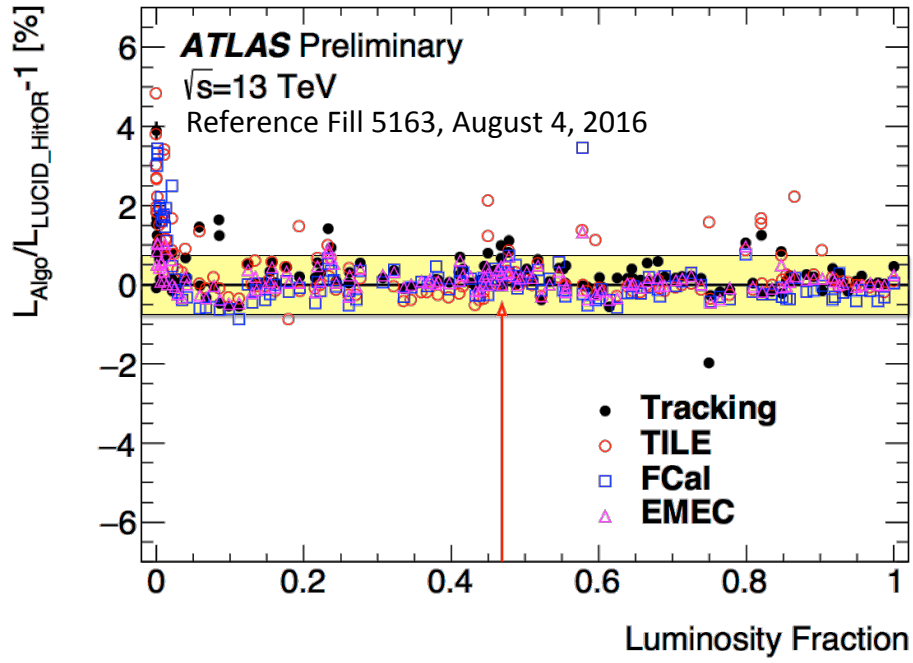


Figure 7: Track-counting efficiencies as a function of time derived from studies of $Z \to \mu\mu$ events: (a) track selection efficiency calculated using 2016 data; (b) track selection efficiency calculated using 2017 data for the track selection used in 2016, and the more robust selection used for the 2017 track-counting luminosity analysis.

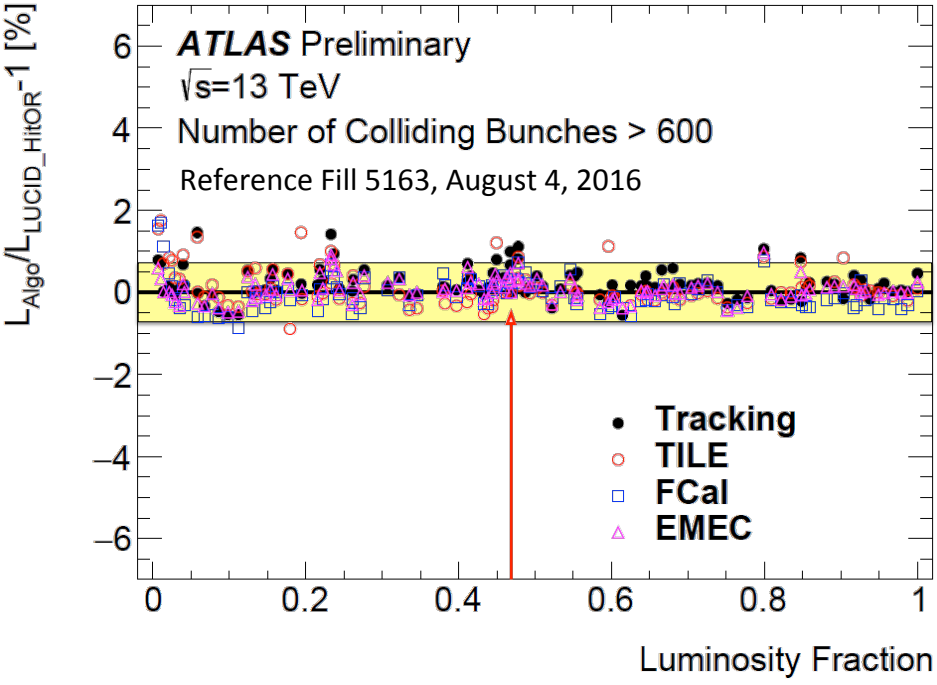
compensate for start-of-year effects (discussed further below). In 2016, a $-0.7\,\%$ correction was applied to data in the first half of the year to compensate for drifts in the LUCID calibration.

The track-counting luminosity measurement also suffers from long-term drifts if the track selection efficiency changes with time. This was monitored using $Z \to \mu\mu$ events, where both muons were reconstructed as isolated 'combined' muons, with matched tracks in the ATLAS inner detector and muon spectrometer [16]. The fraction of the inner detector tracks associated with these muons that pass the track-counting working points, which only rely on selection criteria related to the inner detector, gives a measure of the inner detector track selection efficiency, allowing an efficiency correction to be calculated. Figure 7(a) shows the resulting efficiency for 2016 data; a progressive decrease of more than 1 % over the course of the year is visible, and the corresponding correction was applied to the track-counting luminosity results. The decrease is attributed to the gradually increasing peak instantaneous luminosity and pileup (and hence detector occupancy) through the year. As discussed in Section 2, a more robust track selection with a smaller sensitivity to the amount of pileup was used for 2017–18 data. Figure 7(b) shows the track selection efficiency for both selections measured throughout the 2017 data-taking period. The new selection is much more stable, showing only a small ($\sim 0.2\,\%$) decrease towards the end of the year during the 8b4e running period with luminosity levelling and sustained high pileup levels.

Figures 8, 9 and 10 compare the final luminosity estimates from each detector, after taking into account all corrections for known time-dependent effects. The figures show fractional differences between the run-integrated luminosities measured from track-counting, EMEC, FCal and TILE calorimeters (and TPX for 2017 data) with respect to the baseline LUCID algorithm chosen for each of the years 2016–2018. The values for each run are plotted as a function of the cumulative delivered luminosity fraction, ranging from zero at the start of the year to one at the end, giving an axis monotonically increasing in time. Short runs with less than about two hours data-taking are not shown. In these runs, activation effects in the calorimeters (with time constants of order tens of minutes) can affect the ratios, as the activation build up and decay do not have time to reach equilibrium. The statistical uncertainties for the remaining runs are negligible for all luminosity algorithms, and the remaining differences are attributed to systematic effects. The run-to-run agreement between the various luminosity measurements is generally at the percent level or better for the bulk of the data, with various short- and long-term trends being visible. As shown for 2016 data in Figure 8(a) and (b), many of the outliers are due to LHC fills with reduced numbers of



(a)



(b)

Figure 8: Fractional differences in run-integrated luminosity between the LUCID BiHitOR algorithm and the track-counting, TILE, EMEC and FCal measurements, plotted as a function of the cumulative delivered luminosity normalised to the 2016 total. The luminosity measurements from the other detectors have been normalised to that of LUCID in the reference run indicated by the arrow. Figure (a) shows all runs, whereas figure (b) shows only those runs with more than 600 colliding bunch pairs. The assigned ±0.7 % long-term stability uncertainty is shown by the yellow band.

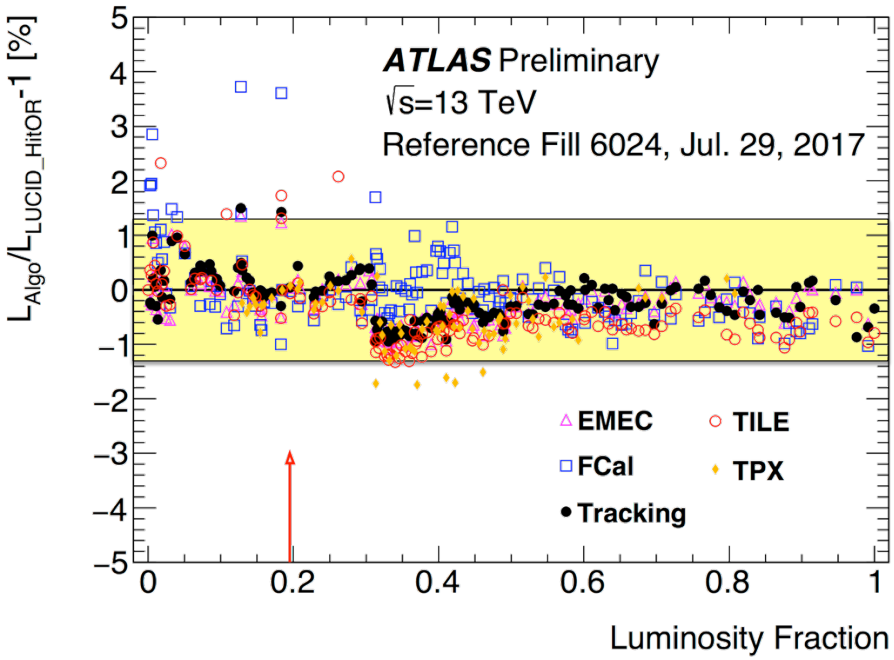


Figure 9: Fractional differences in run-integrated luminosity between the LUCID BiHitOR algorithm and the track-counting, TILE, EMEC, FCal and TPX measurements, plotted as a function of the cumulative delivered luminosity normalised to the 2017 total. The luminosity measurements from the other detectors have been normalised to that of LUCID in the reference run indicated by the arrow. The assigned ±1.3 % long-term stability uncertainty is shown by the yellow band.

bunches (and hence low integrated luminosity), typically in intensity ramp-up periods after LHC stops or operational difficulties. These lead to larger differences between luminosity measurements due to residual luminosity or bunch-pattern dependencies in some of the algorithms and detectors that are not addressed by the calibration and correction procedures. The start of each data-taking year also shows a larger spread, with LUCID systematically underestimating the luminosity compared to other detectors. Part of this is attributed to dependencies on the number of colliding-bunch pairs, as the LHC commissioning periods at the beginning of each year featured a gradual step-by-step increase in the number of bunches over a period of weeks. These periods also show evidence for 'burn-in' effects in LUCID as the performance stabilises over time, particularly as several LUCID PMTs were exchanged during each winter shutdown. However, the integrated luminosity fraction affected by such problems in each year is typically only a few percent of the total.

A long-term stability uncertainty was evaluated conservatively from these studies by considering a 'stability band', a range which encloses the bulk of the differences between LUCID and any of the other luminosity measurements, excluding the outliers for the reasons discussed above. This procedure led to uncertainties of ±1.0 %, ±0.7 %, ±1.3 % and ±0.8 % for the four data-taking years 2015–2018.

The time-dependence of the track-counting luminosity calibration introduces an extra uncertainty in the transfer of the LUCID calibration from the vdM to the physics regime, since the procedure discussed in Section 4 assumes the track-counting measurement is stable between the times of the vdM and reference physics runs. Unfortunately, the $Z \rightarrow \mu\mu$ rate in low-luminosity vdM runs is insufficient to perform a

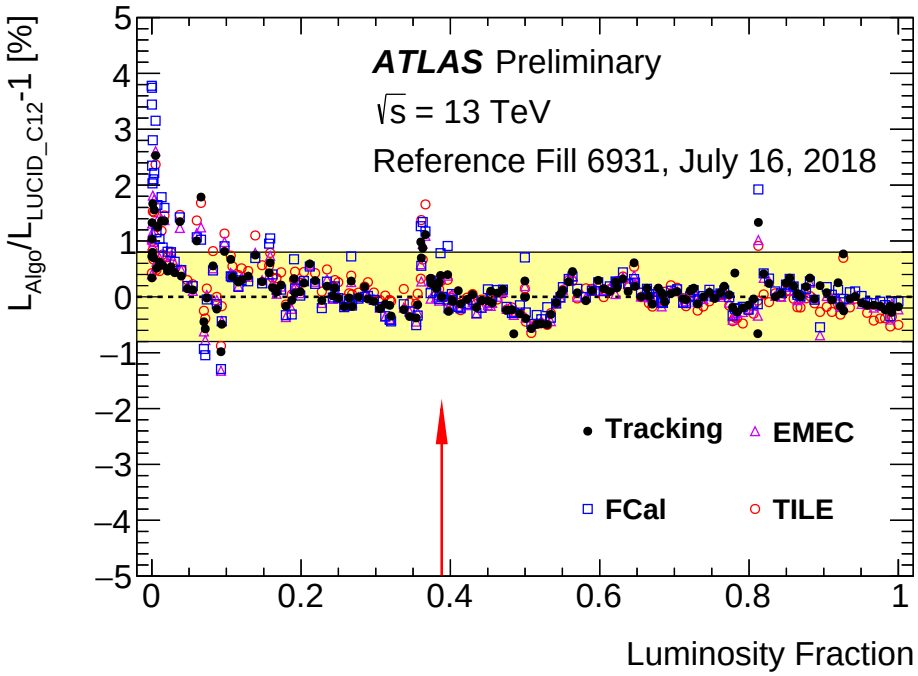
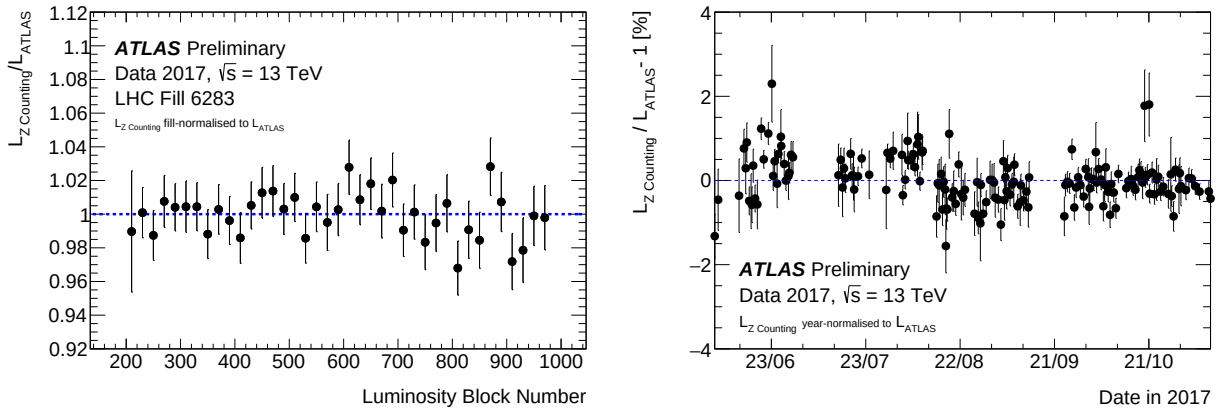


Figure 10: Fractional differences in run-integrated luminosity between the LUCID C12 single-PMT algorithm and the track-counting, TILE, EMEC and FCal measurements, plotted as a function of the cumulative delivered luminosity normalised to the 2018 total. The luminosity measurements from the other detectors have been normalised to that of LUCID in the reference run indicated by the arrow. The assigned ±0.8 % long-term stability uncertainty is shown by the yellow band.

tag-and-probe calibration. Instead, a more indirect method was used, based on the comparison of track $\eta - \phi$ maps to probe module inefficiencies and variations across the acceptance between the vdM and physics runs. Based on these studies, an additional conservative uncertainty of 0.6 % was assigned to the 2016 luminosity calibration, shown as the 'Tracking efficiency time-dependence uncertainty' in Table 2. As the new track-counting working point used in 2017 and 2018 showed much smaller time-dependence, no corresponding uncertainty was assigned for these years, nor for 2015 where no μ-correction was made.

Counting the rate of reconstructed $Z \to \mu\mu$ events in a well-defined fiducial region ('Z-counting') provides a further independent measurement of the luminosity in physics runs, exploited throughout the Run 2 data-taking period. Events were selected with two reconstructed muons each having $p_T > 27$ GeV and $|\eta| < 2.4$, and a dimuon invariant mass satisfying $66 < m_{\mu\mu} < 116$ GeV. The tag-and-probe technique was employed to measure the muon trigger and reconstruction efficiency in situ, with residual corrections taken from simulation. Results from this method are shown in Figure 11. The left plot shows the ratio of relative luminosities measured by Z-counting and the baseline LUCID algorithm as a function of time in LHC fill 6283. The counts are averaged over 20 consecutive luminosity blocks (typically corresponding to 20 minutes), and the Z-counting result is normalised to give the same run-integrated luminosity as LUCID. The right plot shows the fractional differences between Z-counting and LUCID in each run, normalising the two measurements to give the same total integrated luminosity over the entire year. Normalising the Z-counting results to LUCID allows Z-counting to be used to study relative luminosity variations, without needing to assume a $pp \to Z$ production cross-section value, which would be required to make an absolute luminosity measurement independent of the vdM calibration. The results in Figure 11 show good



(a) (b)

Figure 11: Comparison of relative luminosity measurements from Z-counting and LUCID in 2017 data: (a) Ratio of Z-counting to LUCID measurements as a function of time (luminosity block number) within LHC fill 6283. The measurements are averaged over 20 consecutive luminosity blocks (typically 20 minutes), and normalised so that the run-integrated ratio is unity. (b) Fractional differences between per-run Z-counting and LUCID integrated luminosities over the entire 2017 data-taking period. Only runs with at least 10 000 reconstructed Z bosons and a minimum length of about 40 minutes are included. The measurements are normalised to give the same integrated luminosity over the entire 2017 data-taking period.

agreement at the sub-percent level in both the intra-run and inter-run comparisons, providing a further cross-check of the stability of the baseline luminosity determination.

6 Uncertainties and inter-year correlations

The various sources of uncertainty on the luminosity calibration have been discussed in Sections 3–5 above. The numerical values for the different years are summarised in Table 2, where the 2015 and 2016 data have been combined. The uncertainties for 2015–2016 and 2018 are both around 2 %, while the 2017 uncertainty of 2.4 % is slightly larger due to the poorer scan-to-scan reproducibility in the vdM calibration, and the larger variations seen between different algorithms over the course of the year (see Figure 9). Except for this last uncertainty in 2017, the largest single uncertainty in each individual year comes from the calibration transfer. Table 2 also shows the total integrated luminosity in each year, after applying the data quality selections of a typical ATLAS physics analysis. The integrated luminosities for 2016, 2017 and 2018 are comparable (within a factor of two), whilst 2015 contributes only 2 % of the full Run 2 data sample.

Most ATLAS physics analyses treat the full Run 2 $\sqrt{s} = 13$ TeV pp collision data sample as a single dataset, and therefore require the luminosity uncertainty on the entire dataset. Since some of the uncertainty sources are correlated (fully or partially) between all years, some are correlated between a subset of years, and some are uncorrelated, the relative error on the total is not simply a weighted sum of the relative errors on the individual years. Instead, expressing the total Run 2 integrated luminosity $\mathcal{L}_{\text{tot}}$ as the sum of the integrated luminosities $\mathcal{L}_i$ in each year i:

$$\mathcal{L}_{\text{tot}} = \Sigma_i \mathcal{L}_i \,,$$

Data sample	2015+16	2017	2018	Comb.
Integrated luminosity (fb^{-1})	36.2	44.3	58.5	139.0
Total uncertainty (fb^{-1})	0.8	1.0	1.2	2.4
Uncertainty contributions (%):				
DCCT calibration†	0.2	0.2	0.2	0.1
FBCT bunch-by-bunch fractions	0.1	0.1	0.1	0.1
Ghost-charge correction*	0.0	0.0	0.0	0.0
Satellite correction†	0.0	0.0	0.0	0.0
Scan curve fit model†	0.5	0.4	0.5	0.4
Background subtraction	0.2	0.2	0.2	0.1
Orbit-drift correction	0.1	0.2	0.1	0.1
Beam position jitter†	0.3	0.3	0.2	0.2
Beam-beam effects*	0.3	0.3	0.2	0.3
Emittance growth correction*	0.2	0.2	0.2	0.2
Non-factorization effects*	0.4	0.2	0.5	0.4
Length-scale calibration	0.3	0.3	0.4	0.2
ID length scale*	0.1	0.1	0.1	0.1
Bunch-by-bunch σ_{vis} consistency	0.2	0.2	0.4	0.2
Scan-to-scan reproducibility	0.5	1.2	0.6	0.5
Reference specific luminosity	0.2	0.2	0.4	0.2
Subtotal for absolute vdM calibration	1.1	1.5	1.2	-
Calibration transfer†	1.6	1.3	1.3	1.3
Afterglow and beam-halo subtraction*	0.1	0.1	0.1	0.1
Long-term stability	0.7	1.3	0.8	0.6
Tracking efficiency time-dependence	0.6	0.0	0.0	0.2
Total uncertainty (%)	2.1	2.4	2.0	1.7

Table 2: Summary of the integrated luminosities (after data quality requirements) and uncertainties for the preliminary calibration of each individual year of the Run 2 pp data sample at $\sqrt{s} = 13$ TeV (already combining 2015 and 2016) and the full combined sample. As well as the integrated luminosities and total uncertainties, the table gives the breakdown and total of contributions to the absolute vdM calibration, the additional uncertainties for the physics data sample, and the total relative uncertainty in percent. Contributions marked * are considered fully correlated between years, those marked † are considered partially correlated, and the other uncertainties are considered uncorrelated.

the absolute uncertainty on the total, $\sigma_{\mathcal{L}_{\text{tot}}}$, is given by standard error propagation as

$$\sigma^2_{\mathcal{L}_{\text{tot}}} = \mathbf{G}\mathbf{V}_{\mathbf{L}}\tilde{\mathbf{G}}.$$

Here, $\mathbf{V}_{\mathbf{L}}$ is the covariance matrix of the absolute luminosity uncertainties for the different years, and $\mathbf{G}$ is a unit vector.[3] The covariance matrix is made up of the sum of terms corresponding to each uncertainty source in Table 2; uncorrelated uncertainties give rise to terms corresponding to diagonal matrices, whilst correlated sources are represented by terms with non-zero off-diagonal entries.

Since the absolute luminosity scale was calibrated separately for each year with an independent vdM scan session, a large part of the associated uncertainties are uncorrelated between years. The bunch-to-bunch, scan-to-scan and reference specific luminosity uncertainties are driven by the internal consistency of each

[3] In general, $\mathbf{G} = (\mathrm{d}\mathcal{L}_{\text{tot}}/\mathrm{d}\mathcal{L}_i, i = 1...n)$, but as $\mathcal{L}_{\text{tot}} = \Sigma_i \mathcal{L}_i$, $\mathbf{G}$ reduces to a unit vector.

set of vdM scans, and were considered uncorrelated, as were the background subtraction uncertainties, and the uncertainties on the orbit drift corrections, which depend on the specifics of the drifts observed in each year. In contrast, the non-factorisation uncertainty was taken to be correlated, as the same methodology was used to evaluate it each year, and the underlying cause is likely to be correlated. The beam-beam effects and fit model uncertainties were treated in the same way. Effects which were evaluated only once during Run 2, or taken to be the same as in Run 1, such as the ID length scale and the emittance growth correction, were considered correlated. The beam position jitter uncertainties used for 2015–17 were taken from a Run 1 study, but were considered uncorrelated with those for 2018 data, where a new evaluation was made. The length scale was determined from dedicated calibration runs in each year, and considered uncorrelated. The uncertainties on the DCCT calibration were taken to be uncorrelated, apart from the minor contributions related to the calibration pulse generator and a potential bunch-pattern dependence of the response. The bunch-by-bunch fraction uncertainties are dominated by electronic noise and were taken to be uncorrelated between years. Finally, the residual uncertainties on the ghost and satellite corrections were taken to be correlated, as the instrumentation and methodology are common to all years.

The calibration transfer uncertainties were based on similar comparisons of TILE and track-counting luminosity measurements in all years. A correlated uncertainty of 1.3 % was considered, with an additional uncorrelated component of 0.9 % in 2016 where the overall uncertainty is larger. The additional uncertainty on the possible time-dependence of the tracking efficiency applies only to 2015–16 data, as a more stable working point was used in 2017–18 (see Figure 7). The long-term stability uncertainties derived from the comparisons shown in Figure 8–10 were considered uncorrelated between years, as the trends seen in the evolution of ratios appear different in each year (apart from the initial intensity ramp-up at the start of each year, which affects a very small fraction of the total integrated luminosity). Finally, the afterglow and beam-halo related uncertainty was considered correlated, as it was derived from a Run 1 study.

The 'Comb.' column of Table 2 shows the total uncertainty of 1.7 % on the full Run 2 integrated luminosity estimate, and the breakdown in terms of individual components. The largest uncertainties come from the calibration transfer of the luminosity scale from the vdM to the physics regime (1.3 %), followed by the long-term stability (0.6 %). The subtotals for the uncertainties on the absolute luminosity scale from the vdM calibration are larger than 1 % in all years, but make a smaller contribution to the total Run 2 uncertainty as they are only partially correlated between years.

7 Conclusion

The luminosity scale for the Run 2 $\sqrt{s} = 13$ TeV pp collision data recorded by the ATLAS experiment at the LHC during the years 2015–18 has been calibrated using dedicated van der Meer scans in each year, and extrapolated to the physics regime using complementary measurements from several luminosity-sensitive detectors. For this preliminary calibration, the total uncertainties on the integrated luminosities for each individual year vary from 2.0–2.4 %, and the uncertainty on the combined Run 2 dataset is $\delta\mathcal{L}/\mathcal{L} = \pm 1.7$ %. The largest contribution to the uncertainty comes from the extrapolation of the calibration from the low-luminosity vdM scans to the high-luminosity physics data-taking regime. This calibration is applicable to the full Run 2 dataset recorded at high instantaneous luminosity and pileup, but not to the special runs recorded under low pileup conditions for some physics studies (e.g. W/Z production) and at high β^* for the ALFA forward physics programme, which require dedicated analyses appropriate to their specific experimental conditions.

References

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[2] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **71** (2011) 1630, arXiv: 1101.2185 [hep-ex].

[3] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **73** (2013) 2518, arXiv: 1302.4393 [hep-ex].

[4] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653, arXiv: 1608.03953 [hep-ex].

[5] S. van der Meer, *Calibration of the effective beam height in the ISR*, CERN-ISR-PO-68-31, 1968, URL: https://cds.cern.ch/record/296752.

[6] P. Grafström and W. Kozanecki, *Luminosity determination at proton colliders*, Prog. Part. Nucl. Phys. **81** (2015) 97.

[7] H. Wiedemann, *Particle Accelerator Physics*, Graduate Texts in Physics, Springer, 2015, URL: https://www.springer.com/us/book/9783319183169?wt_mc=ThirdParty.SpringerLink.3.EPR653.About_eBook.

[8] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017.

[9] A. Sopczak et al., *Precision Luminosity of LHC Proton–Proton Collisions at 13 TeV using Hit Counting with TPX Pixel Devices*, IEEE Trans. Nucl. Sci. **64** (2017) 915, arXiv: 1702.00711 [physics.ins-det].

[10] B. Abbott et al., *Production and Integration of the ATLAS Insertable B-Layer*, JINST **13** (2018) T05008, arXiv: 1803.00844 [physics.ins-det].

[11] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19, 2010, URL: https://cds.cern.ch/record/1291633, *ATLAS Insertable B-Layer Technical Design Report Addendum*, ATLAS-TDR-19-ADD-1, 2012, URL: https://cds.cern.ch/record/1451888.

[12] ATLAS Collaboration, *Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-051, 2015, URL: https://cds.cern.ch/record/2110140.

[13] C. Barschel and M. Ferro-Luzzi, *Precision luminosity measurement at LHCb with beam-gas imaging*, CERN-THESIS-2013-301, 2014, URL: https://cds.cern.ch/record/1693671.

[14] W. Herr, *Beam-beam effects and dynamic β^**, Proc. LHC Lumi Days 2012, URL: https://indico.cern.ch/event/162948/contributions/1417430/attachments/191879/269237/S3_WH.pdf.

[15] ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, Eur. Phys. J. C **70** (2010) 1193, arXiv: 1007.5423 [hep-ex].

[16] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex].