arXiv:2212.07338v1 [hep-ex] 14 Dec 2022

Studies of the muon momentum calibration and performance of the ATLAS detector with pp collisions at $\sqrt{s}$=13 TeV

The ATLAS Collaboration

This paper presents the muon momentum calibration and performance studies for the ATLAS detector based on the pp collisions data sample produced at $\sqrt{s}$ = 13 TeV at the LHC during Run 2 and corresponding to an integrated luminosity of 139 fb^{-1}. An innovative approach is used to correct for potential charge-dependent momentum biases related to the knowledge of the detector geometry, using the $Z \rightarrow \mu^+\mu^-$ resonance. The muon momentum scale and resolution are measured using samples of $J/\psi \rightarrow \mu^+\mu^-$ and $Z \rightarrow \mu^+\mu^-$ events. A calibration procedure is defined and applied to simulated data to match the performance measured in real data. The calibration is validated using an independent sample of $\Upsilon \rightarrow \mu^+\mu^-$ events. At the Z (J/ψ) peak, the momentum scale is measured with an uncertainty at the 0.05% (0.1%) level, and the resolution is measured with an uncertainty at the 1.5% (2%) level. The charge-dependent bias is removed with a dedicated *in situ* correction for momenta up to 450 GeV with a precision better than 0.03 TeV^{-1}.

Contents

1 Introduction

Muons are a crucial component of the physics programme of the ATLAS experiment [1] at the Large Hadron Collider (LHC) [2]. The discovery of the Higgs boson [3, 4] and the measurement of its properties [5], precision tests of the Standard Model [6], and searches for physics processes beyond the Standard Model [7, 8] al strongly rely on the performance of muon identification and measurement with the ATLAS detector.

During its second data-taking campaign (Run 2; 2015–2018), ATLAS collected 139 fb^{-1} of pp collisions at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV, a data sample of unprecedented size. A recent publication presented the performance of the new and optimised muon reconstruction and identification techniques developed for the analysis of the full Run 2 data sample [9]. The muon momentum measurement was previously published for early Run 2 data, corresponding to 3.2 fb^{-1} collected in 2015 [10].

This paper describes the methods used to calibrate the momentum measurement of muons reconstructed by the ATLAS detector for the full Run 2 data sample. First, the potential charge-dependent bias on the scale of the muon momentum measurement, introduced by the imperfect knowledge of the real detector geometry, is measured in reconstructed collision data with a sample of $Z \rightarrow \mu\mu$ events.[1] A dedicated

---

[1] Charges of decay products are omitted in the following for simplicity.

2

correction is derived and applied to the data. Then, $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays are selected in data and used to measure the resolution and scale of the muon momentum, which are compared with those predicted by simulation. A calibration procedure is applied to simulated events, to improve the agreement between the simulation and the data. Finally, a validation procedure of this calibration is performed using $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events, in addition to an independent sample of $\Upsilon \rightarrow \mu\mu$ decays.

In contrast to previous publications, this paper presents a completely new methodology to determine the charge-dependent bias, sensitive to global shifts in the detectors' positions, which could not be measured with previous techniques. Furthermore, the inclusion of new $J/\psi \rightarrow \mu\mu$ data collected with dedicated trigger strategies, the development of new fitting techniques with better convergence properties, and the significantly larger data sample result in an unprecedented precision of the momentum calibration procedure. In addition to the calibrations developed separately for the tracks measured in the inner detector and in the muon spectrometer of the ATLAS detector, for the first time a dedicated calibration for tracks obtained by combining information from both sub-detectors is derived. The validation of the calibration procedure is performed using $\Upsilon \rightarrow \mu\mu$ events as well as $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events in regions with finer granularity than those used in the calibration procedure.

This paper is structured as follows: the ATLAS detector is described in Section 2; the simulated and real data samples used for the measurements are presented in Section 3; the identification of muon candidates is discussed in Section 4; the methodologies for measuring the muon scale and momentum corrections are presented in Section 5; the results of the measurement and the validation of the data-driven corrections derived for simulated samples are presented in Section 6; final conclusions are provided in Section 7.

2 ATLAS detector

The ATLAS detector [1] at the LHC covers nearly the entire solid angle around the collision point.[2] It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadron calorimeters, and a muon spectrometer incorporating three large superconducting air-core toroidal magnets.

The inner detector system (ID) is immersed in a 2 T axial magnetic field (bending the tracks of charged particles in the transverse plane) and provides charged-particle tracking in the range of $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track. The first hit is normally found in the insertable B-layer (IBL) installed before Run 2 [11, 12]. It is followed by the silicon microstrip tracker (SCT), which typically provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$, typically providing 30 measurements per track. The TRT also provides electron identification information based on the fraction of hits above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. In the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with

---

[2] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadron calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures with $|\eta| < 1.7$, and two copper/LAr hadron endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic energy measurements respectively.

The muon spectrometer (MS) comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by the superconducting air-core toroidal magnets, which bend charged particles tracks in the r–z plane, i.e. in the plane formed by the muon momentum vector and the beam axis. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. Three layers of precision chambers, each consisting of layers of monitored drift tubes, cover the region $|\eta| < 2.7$, complemented by cathode-strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

Events are selected by the first-level trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger [13]. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the high-level trigger further reduces to record events to disk at about 1 kHz.

An extensive software suite [14] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

3 Data and Monte Carlo samples

3.1 Data sample

The analysis is performed using the $\sqrt{s}$ =13 TeV pp collision data sample recorded by the ATLAS detector between 2015 and 2018. Only events collected in stable beam conditions and with all relevant ATLAS magnets and detector subsystems fully operational are used in the analysis. A combination of single-muon trigger algorithms (for $Z \rightarrow \mu\mu$ events) and trigger algorithms dedicated to $J/\psi \rightarrow \mu\mu$ and $\Upsilon \rightarrow \mu\mu$ topologies [15] were used. Starting from the middle of 2016, due to the increasing instantaneous luminosity of the LHC, events associated with $J/\psi \rightarrow \mu\mu$ and $\Upsilon \rightarrow \mu\mu$ triggers were collected with a dedicated data acquisition stream for B-physics and low-mass final states, separated from the main physics stream so that their reconstruction could be delayed if the availability of processing resources was limited [16]. This additional stream is analysed for a muon momentum calibration for the first time in this paper, significantly increasing the number of $J/\psi \rightarrow \mu\mu$ and $\Upsilon \rightarrow \mu\mu$ events available. For the data sample used in this analysis a procedure to calibrate the muon detectors was applied, as described in Ref. [17]. The analysed data corresponds to an integrated luminosity of 139 fb^{-1} after trigger and data-quality requirements, with an average number of pp collisions per bunch crossing $< \mu > = 33.7$.

3.2 Simulated samples

Samples of Monte Carlo (MC) simulated inclusive prompt production of $J/\psi \rightarrow \mu\mu$ and $\Psi(2S) \rightarrow \mu\mu$ events were generated at leading-order (LO) accuracy using PYTHIA 8.186 [18], for matrix element (ME) calculations, and for the modelling of the parton shower, hadronisation, and the underlying event. The

CTEQ6L1 [19] parton density function (PDF) set and the A14 [20] set of tuned generator parameters were used. The effect of QED final-state radiation was simulated with PHOTOS++ [21, 22]. To increase the effective number of events for the analysis, the generated events were filtered, requiring at least one muon in the event to satisfy $p_T > 6$ GeV and $|\eta| < 2.5$.

Samples of $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ simulated events with subsequent decays into two muons were generated with a similar configuration to that of the $J/\psi \to \mu\mu$ samples, though the filter required at least two muons with $p_T > 4$ GeV and $|\eta| < 2.7$ in the event.

The samples of $Z \to \mu\mu$, $Z \to \tau\tau$, $W \to \mu\nu$, and $Z^*/\gamma^*(\to \mu\mu)$ events were simulated using the POWHEG BOX v1 [23–26] generator at next-to-leading-order (NLO) accuracy in perturbative QCD, interfaced to PYTHIA 8.186 with the AZNLO [27] set of tuned generator parameters. The CT10NLO PDF set [28] was used for the hard-scattering processes, whereas the CTEQ6L1 PDF set was used for the modelling of the parton showering, hadronisation, and underlying event. The effect of QED final-state radiation was simulated with PHOTOS++. The EVTGEN 1.2.0 program [29] was used to decay bottom and charm hadrons. To improve the agreement with the data, the simulated Z samples are reweighted to reproduce the muon-pair tranverse momentum and rapidity distributions measured in data.

A configuration similar to the one used for the Z samples described above was used to simulate diboson processes decaying into final states containing muons, (ZZ, WZ, WW), using POWHEG BOX v2 and PYTHIA 8.210 [30] in this case.

The production of $t\bar{t}$ events was modelled using the POWHEG BOX v2 [23–25, 31] generator at NLO with the NNPDF3.0NLO [32] PDF set and the h_{damp} parameter[3] set to $1.5\,m_{\mathrm{top}}$ [33]. The events were interfaced to PYTHIA 8.230 with the A14 set of tuned generator parameters and using the NNPDF2.3LO PDF set [34]. The decays of bottom and charm hadrons were performed by EVTGEN 1.6.0.

An alternative sample of $Z \to \mu\mu$ events was simulated with the SHERPA 2.2.1 [35] generator using NLO ME for up to two partons, and LO ME for up to four partons calculated with Comix [36] and OPENLOOPS [37–39]. They were matched to the SHERPA parton shower [40] using the MEPS@NLO prescription [41–44] using the set of tuned parameters developed by the SHERPA authors. The NNPDF3.0NNLO PDF set [32] was used and the samples were normalised to a next-to-next-to-leading-order (NNLO) prediction [45]. The re-weighting procedure applied to the nominal $Z \to \mu\mu$ events sample is also applied to this alternative sample.

The effect of multiple interactions in the same and neighbouring bunch crossings (pile-up) was modelled by overlaying the simulated hard-scattering event with inelastic proton–proton (pp) events generated with PYTHIA 8.186 [18] using the NNPDF2.3LO PDF set [34] and the A3 set of tuned parameters [46]. The MC events were weighted to reproduce the distribution of the average number of interactions per bunch crossing observed in the data.

The detector response was simulated using the GEANT4 toolkit and reconstructed with the same algorithms as used for real data [47, 48]. The detector simulation assumes perfect alignment for the ID and MS sub-detectors. Dedicated simulated samples with misaligned detector geometry are used for the validation of the calibration procedure. These include samples with a rotation of the ID layers with a linear dependence on the radius [49] and a perfect geometry for the MS. They also include samples with the positions of the middle layer of the MS chambers shifted by amounts corresponding to the residual alignment precision [50].

---

[3] The h_{damp} parameter is a re-summation damping factor and one of the parameters that controls the matching of POWHEG MEs to the parton shower and thus effectively regulates the high-p_T radiation against which the $t\bar{t}$ system recoils.

4 Muon reconstruction, identification and event selection

4.1 Muon reconstruction and identification

The muon reconstruction is described in detail in Ref. [9]; only a brief summary is provided here. Muons are reconstructed using information from the ID and/or MS sub-detectors, which provide two independent measurements of muons crossing the ATLAS detector. In this analysis, the following reconstruction algorithms are used.

- ID tracks: tracks reconstructed using ID hits only.

- Standalone tracks (SA): muon candidates obtained using MS hits only; to define kinematic quantities at the interaction point, SA tracks are extrapolated to the beam-spot taking into account energy loss and multiple scattering in the material upstream of the muon spectrometer.

- Combined tracks (CB): muon candidates obtained by starting from a MS track and matching it to an ID track. A combined fit to the hits, taking into account energy loss in the calorimeters and multiple scattering in the spectrometer, is performed.

Several working points (WPs) are introduced in Ref. [9], defining sets of quality cuts applied to reconstructed muons. The *Medium* WP is the baseline for this analysis; this algorithm only selects CB muons in $|\eta| < 2.5$ and applies a set of requirements to reject poorly reconstructed tracks. The *High-p_T* WP applies tighter cuts on CB muons to ensure an optimal muon reconstruction for analyses interested in the very high p_T regime, with a better momentum resolution. Other working points target either lower background rates, at the expense of a lower efficiency, or higher efficiency but with a lower purity for muon identification; these WPs have a momentum resolution performance similar to the *Medium* one. The individual ID and SA tracks associated with the CB muons are still accessible and used in the analysis, as discussed in the following. For simplicity, muon candidates reconstructed using the ID, MS or combined ID+MS information are referred to as ID, MS or CB muons, respectively.

4.2 Event selection

Proton–proton collision vertices are constructed from reconstructed trajectories of charged particles in the ID with transverse momentum $p_T > 500$ MeV. Events are required to have at least one collision vertex. The vertex with the highest $\sum p_T^2$ of reconstructed tracks is selected as the primary vertex of the hard interaction. The data are subjected to quality requirements to reject events in which detector components were not operating correctly [51].

Events are required to be selected by a combination of the following trigger chains:

- For $Z \rightarrow \mu\mu$ candidate events, unprescaled single-muon trigger chains with the lowest kinematic threshold available in each data sample are used.

  - Data collected in 2015: at least one muon with $p_T > 20$ GeV and passing a loose isolation requirement based on the scalar sum of the p_T of tracks in a cone around the muon candidate track. Events are also retained if selected by a second chain requiring at least one muon with $p_T > 40$ GeV without any isolation requirement.

- Data collected in 2016, 2017, 2018: a similar strategy as the one used in 2015 was employed, but with the first chain requiring a muon with $p_T > 26$ GeV passing a tighter isolation requirement, and the second chain requiring at least one muon with $p_T > 50$ GeV.

- For $J/\psi \to \mu\mu$ and $\Upsilon \to \mu\mu$ candidate events, trigger chains requiring at least two muons in the event were used. The trigger algorithms also performed common vertex fits to pairs of oppositely charged muon candidates, requiring at least one of the fitted vertices to satisfy fit quality criteria and have an invariant mass consistent with that of a J/ψ or Υ resonance [15]. The last requirement is significantly looser than that applied by the offline analysis, to avoid introducing a bias. The trigger algorithms applied the following kinematic requirements on the muons associated with the candidate resonance:

  - Data collected in 2015: both muons must satisfy $p_T > 4$ GeV.

  - Data collected in 2016: both muons must satisfy $p_T > 6$ GeV.

  - Data collected in 2017 and 2018: both muons must satisfy $p_T > 6$ GeV. Additional requirements were also applied on the angular distance between the two muons associated with the vertex and on the displacement of the fitted vertex in the transverse plane. Events were also selected using auxiliary trigger algorithms requiring the leading (sub-leading) muon to have $p_T > 11$ GeV ($p_T > 6$ GeV) but without any additional requirement.

Events are required to have two oppositely charged CB muons satisfying the *Medium* WP requirements and spatially matched with the muon candidates reconstructed by the respective trigger chains. Further kinematic requirements are applied to ensure that selected muons are in the regions where the triggers used for the analysis were fully efficient:

- For $Z \to \mu\mu$ candidate events, the leading muon must satisfy $p_T > 27$ GeV.

- For $J/\psi \to \mu\mu$ and $\Upsilon \to \mu\mu$ candidate events, both muons must satisfy $p_T > 6.25$ GeV. For events selected only by the auxiliary trigger chain in the 2017 and 2018 data samples, the requirement on the leading muon is changed to $p_T > 11.25$ GeV.

For each muon, requirements are also applied on the transverse (d_0) and longitudinal (z_0) impact parameters. The d_0 of a charged-particle track is defined in the plane transverse to the beam as the distance from the primary vertex to the track's point of closest approach. The z_0 is the distance in the z direction between this track point and the primary vertex; this distance is represented by $z_0 \sin\theta$ in the longitudinal projection. The candidate muons are required to satisfy:

- $\frac{|d_0|}{\sigma(d_0)} < 3$, with $\sigma(d_0)$ representing the uncertainty in the measured d_0 value;

- $|z_0 \sin\theta| < 0.5$ mm.

After application of these selections, the non-prompt component of $J/\psi \to \mu\mu$ production is reduced to less than 0.5%.

Finally the invariant mass of the muon pair is required to satisfy:

- $2.6 < m_{\mu\mu} < 3.5$ GeV for $J/\psi \to \mu\mu$ candidate events;

- $7 < m_{\mu\mu} < 14$ GeV for $\Upsilon \to \mu\mu$ candidate events;

- $70 < m_{\mu\mu} < 130$ GeV for $Z \to \mu\mu$ candidate events.

5 Momentum scale corrections

The muon momentum scale and resolution are studied using $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays in data and simulated samples. While dedicated procedures are applied to correct the detector alignment [1, 49, 50], residual misalignments introduce a charge-dependent bias in the momentum measurement. Studies of this effect are discussed in Section 5.1, together with the determination of an appropriate set of corrections for the data. Simulated samples assume ideal detector alignment and thus do not require these corrections.

Although the simulation contains an accurate description of the ATLAS detector, the level of detail is not sufficient to reach the accuracy of 0.1% on the muon momentum scale and the percent level precision on the resolution measured in data. The analysis of this sample improves upon the excellent simulated description of the ATLAS detector and its interaction with muons, encompassing the best knowledge of the detector geometry, material distribution, and physics modelling at the time of event simulation. Section 5.2 details the measurement of momentum scale and resolution in data and simulation, and the determination of corresponding corrections for the simulated samples. Validation studies for the corrections are presented in Section 6.

5.1 Charge-dependent momentum scale calibration in data

After dedicated detector alignment procedures are applied, residual effects can still introduce a bias in the momentum measurement of the muon. These effects arise from both the ID and the MS sub-detectors.

The ID alignment is derived by a global χ^2 minimisation of the track-hit residuals [49]. Residual detector displacements relative to the nominal detector geometry may however still be present, the *weak modes*, that are not fully corrected for by the current alignment procedures. ID alignment studies of these modes are periodically performed and data is reconstructed with an improved description of the detector geometry. Nevertheless, residual effects after this correction procedure still prevent the best possible precision being attained, motivating the study discussed in the following.

For the MS, the alignment procedures are less sensitive to charge-dependent biases, though residual effects may remain due to the limited precision of the procedure. An optical alignment system [1] monitors the position of the muon chambers relative to each other and relative to fiducial marks in the detector. The system compensates for chamber position variations as the data is collected. However, several aspects limit the accuracy of this system, including access shafts for detector operations and construction precision in some detector regions. Additionally, the analysis of dedicated data runs, either collecting cosmic-ray data or in proton–proton collisions with the magnetic field generated by the toroid systems switched off, allows the precision of the relative position of the muon chambers to be further constrained. This leads to a precision at the level of tens of micrometers on the sagitta of the muon track, and up to 120–130 μm in specific detector regions [50]. This residual uncertainty in the alignment system manifests as charge-dependent effects in the muon spectrometer reconstruction procedure.

The precision of the muon momentum reconstruction can be further improved with a dedicated correction procedure accounting for the ID weak modes and MS alignment system uncertainties, which would otherwise degrade the momentum resolution. The charge-dependent bias can be approximated as:

$$\frac{q}{\hat{p}} = \frac{q}{p} + q \cdot \delta_s, \tag{1}$$

8

where $q = \pm 1$ is the charge of the muon, p and $\hat{p}$ are the corrected and uncorrected momentum of the muon, respectively, and δ_s is the strength of the bias. This parameterisation forms the basis for developing a correction to the data, recovering the residual bias and improving the momentum resolution.

To estimate (and later correct) the bias, the large sample of $Z \rightarrow \mu\mu$ decays is used. The mass of the dimuon system, $m_{\mu\mu}$, can be expressed as a function of the positively and negatively charged muons' transverse momenta p_T^+ and p_T^-, respectively:

$$m_{\mu\mu}^2 = 2 \cdot p_T^+ p_T^- \cdot (\cosh(\Delta\eta) - \cos(\Delta\phi)), \tag{2}$$

where $\Delta\eta$ and $\Delta\phi$ are the differences between the pseudorapidity and azimuthal angles of the two muons. The bias is parameterised as a function of a grid of 48×48 equal-sized η–ϕ detector regions, a granularity chosen to ensure that the measurement is not affected by large statistical fluctuations while still correctly reflecting local biases or deformations. This high granularity allows the transverse momentum p_T of the measurement to be used instead of p, which is necessary when comparing ID and MS results, since the two sub-detectors have different bending planes due to their respective magnet systems. Combining the previous two equations, the biased invariant mass $\hat{m}_{\mu\mu}^2$ can be expressed as:

$$\hat{m}_{\mu\mu}^2 = \frac{m_{\mu\mu}^2}{(1 + \delta_s p_T^+)(1 - \delta_s p_T^-)}. \tag{3}$$

Assuming the bias is small, Equation 3 can be approximated as:

$$m_{\mu\mu}^2 = \hat{m}_{\mu\mu}^2 (1 + \delta_s(\eta, \phi) p_T^+ - \delta_s(\eta, \phi) p_T^-). \tag{4}$$

Since $p_T^+ \sim p_T^-$ in Z boson decays, charge-dependent biases largely cancel in the expression of $m_{\mu\mu}^2$, hence do not impact the average dimuon mass over all detector regions, but only broaden the resolution of the $Z \rightarrow \mu\mu$ peak. Therefore, the reconstructed invariant mass $m_{\mu\mu}$ distribution defined in Equation 4 is sensitive to the bias through its impact on the variance of the distribution. The values of the sagitta biases $\hat{\delta}_s(\eta, \phi)$ are evaluated by minimising the variance of the invariant mass distributions; then, the biased momentum of the muon $\hat{p}_T$ is corrected using the following equation:

$$p_T = \frac{\hat{p}_T}{1 - q\hat{\delta}_s(\eta, \phi)\hat{p}_T}. \tag{5}$$

Finally, the updated p_T values are used to recalculate the invariant mass distribution and a new iteration is started. Given the large number of degrees of freedom induced by the dependence of the biases on η and ϕ, a direct solution of the equation becomes impractical. Instead, an iterative approach is chosen. At each iteration the values of $\hat{\delta}_s(\eta, \phi)$ obtained from the previous iteration are used as an input to $\hat{p}_T$ of the next.

A complementary method to estimate this bias is also defined by approximating the average p_T of the muon to half the invariant mass of the dimuon pair. In this case, the bias can be approximated by:

$$\delta_s(\eta, \phi) = 4 \cdot \frac{q(\hat{m}_{\mu\mu}(\eta, \phi) - \langle m_{\mu\mu} \rangle)}{\langle m_{\mu\mu} \rangle^2}, \tag{6}$$

where $\langle m_{\mu\mu} \rangle$ is the average of the invariant mass of the dimuon pairs used to derive the correction, while $\hat{m}_{\mu\mu}(\eta, \phi)$ is the average invariant mass of the dimuon pairs when the muon with the highest transverse momentum is in the given (η, ϕ) region.

Both methods additionally assume that the kinematic properties of positively and negatively charged muons produced by Z decays are symmetrical. In addition to the small asymmetry introduced by the weak mixing angle, this assumption breaks down due to the detector acceptance and selection requirements restricting the phase space of the selected muon pairs. The impact of these acceptance and kinematic restrictions can be estimated by applying the same procedure to events simulated with ideal detector alignment, obtaining a residual bias term δ_s^{MC}. The δ_s^{data} measured in data can be corrected by subtracting the values measured in MC simulation, δ_s^{MC}, to define the final corrections:

$$\delta_s(\eta, \phi) = \delta_s^{\mathrm{data}}(\eta, \phi) - \delta_s^{\mathrm{MC}}(\eta, \phi). \tag{7}$$

Typical values for $\delta_s^{\mathrm{MC}}(\eta, \phi)$ range from up to 0.01 TeV^{-1} in the central pseudorapidity region to up to about 0.05 TeV^{-1} in the forward pseudorapidity region, after event selection. The procedure is applied separately for the muon momentum measured using the CB, ID, and MS information, and dedicated corrections for each measurement are derived. Both procedures target a residual δ_s of less than 10^{-3} TeV^{-1} through simultaneous iterative updates on the ID, MS, and CB momenta. The convergence is achieved once the target is reached for all momenta. For the first method this happens after five iterations on average, while the second typically requires 25 iterative updates. In both methods, the angular correlations between the two muons have a small impact on the estimate of δ_s, given the fine granularity chosen in $\eta \times \phi$. Figure 1 shows the measured sagitta corrections evaluated on CB momenta, with biases up to 0.4 TeV^{-1}. To avoid introducing a bias in the measurement due to the value assumed for m_Z, the unbiased value of $m_{\mu\mu}$ is taken from the data when integrating over all η and ϕ bins for both methods. Results from the two methods, which rely on different assumptions, are compared. For the first method, a comparable bias for the leading and subleading p_{T} is assumed, while for the second method the approximation of $p_{\mathrm{T}}^+ \simeq p_{\mathrm{T}}^-$ is less precise at high η values. The $\delta_s(\eta, \phi)$ corrections obtained with the two methods have a correlation close to 100%, with an agreement between the two methods at the level of 0.01 TeV^{-1} or better, well within the systematic uncertainties of either method. Results for the second method are given in Appendix A.

Using the derived bias results, the transverse momenta of the muon are corrected based on Equation 5 using the $\hat{\delta}_s$ derived after the last iteration. The biases are re-evaluated after the application of the corrections and the results are shown in Figure 2 for CB momenta. Both methods address relative differences between nearby η, ϕ regions, as well as biases that are constant across full sub-detector regions. The biases are reduced to less than $2 \cdot 10^{-4} \text{ TeV}^{-1}$ in all regions of the detector, validating the method ('closure') to several orders of magnitude better than the typical size of the correction.

To validate the correction procedure, simulated samples with misaligned detector configurations, as described in Section 3.2, were used. The injected biases were produced by a rotation of the ID detector layers with a linear dependence on the radial distance from the interaction point. The sizes of these biases in the simulation are the same as those observed in data from $Z \to \mu\mu$ decays before applying the correction procedure. Figure 3 uses these biased detector geometry samples to compare the reconstructed and generated momentum distributions, as evaluated before and after the correction, plotted separately for positive and negative muons. The application of the correction reduces the differences between the positively and negatively charged muons, improving the reconstructed p_{T} resolution. Further closure tests are performed using simulated samples with distorted geometry on $H \to ZZ \to \mu\mu\mu\mu$ and hypothetical narrow resonances decaying into two muons with masses greater than 500 GeV, to estimate the impact of the correction in muon momentum regimes far from that of muons originating from Z decays. Closure is observed at the same level of precision.

Several sources of systematic uncertainty associated with the correction procedure are studied. They relate to both the accuracy of the measurement of δ_s and the assumptions used in the correction procedure.
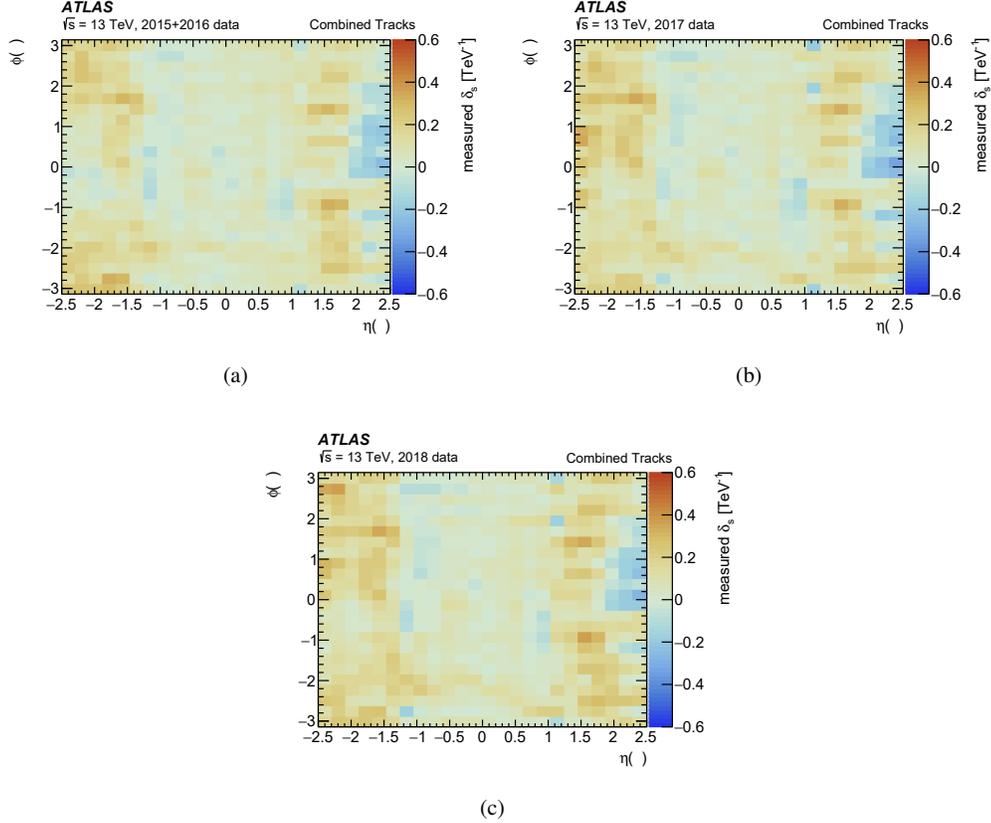
Figure 1: Charge-dependent biases on the muon p_T as evaluated on data for CB momenta after alignment and before applying the dedicated corrections. The biases are shown separately for the three data-taking periods, (a) years 2015 and 2016, (b) 2017, and (c) 2018. The bias is defined in Eq. (1).

The systematic uncertainty originating from the residual non-closure of the correction compared with a perfectly aligned detector simulation is evaluated as follows. First, the δ_s value obtained after the last iteration on data is used to bias the p_T in a perfectly aligned simulation. This is done using Equation 5, by flipping the sign of δ_s, and using the same granularity in η and ϕ as used for deriving the correction. The resulting p_T is compared with the unbiased momentum and the difference is taken as a systematic uncertainty. A second component of the residual non-closure is estimated in simulation by injecting a set of constant known biases and by taking the difference of δ_s between the injected and measured δ_s. The injected biases range from percent fractions up to twice those observed in the data at the first iteration of the correction procedure; the largest resulting δ_s difference is taken as the uncertainty for all bins. Both methods show very good closure for the full range of injected biases. The uncertainties originating from the non-closures amount to about 0.02 TeV^{-1}, depending on the pseudorapidity region.

A second contribution to the systematic uncertainties arises from extrapolations to momenta outside the phase-space region covered by Z decays. In fact, the relative contribution of the ID and MS detectors on the CB momentum measurement may vary in a non-trivial way. This may lead to biases on the estimate of δ_s and of the correction. In this case, the residual bias is evaluated as the difference between simulation and data and as a function of η and p_T using the alternative method of Equation 6. The difference
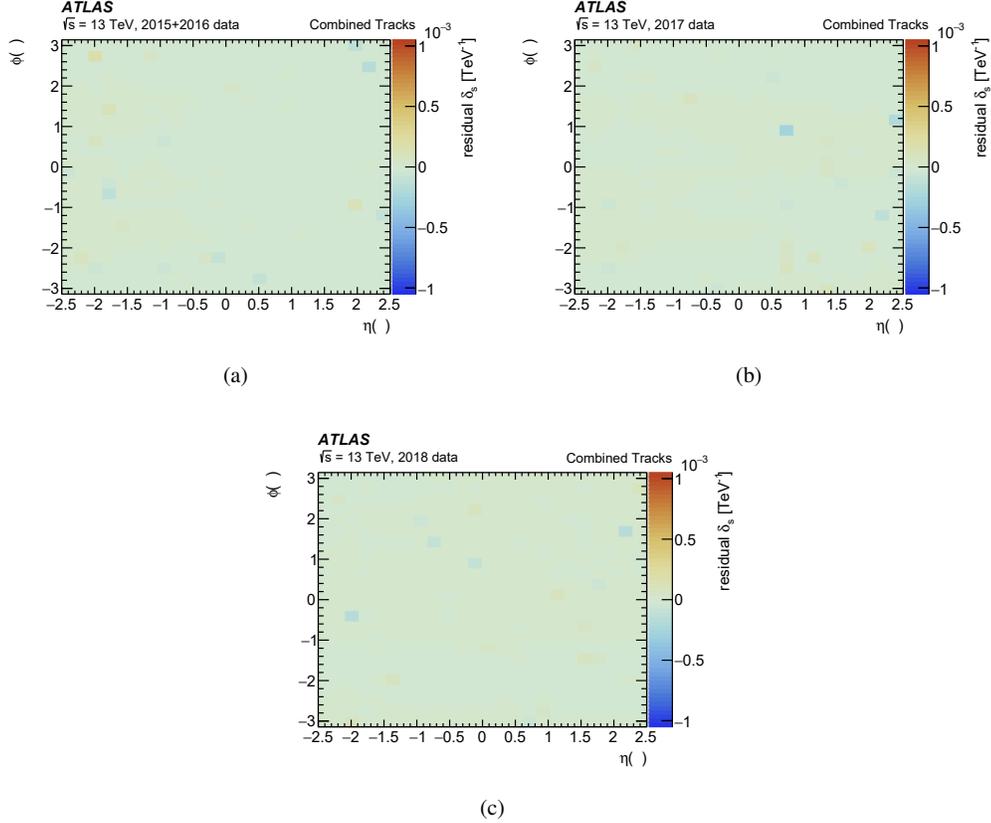
Figure 2: Residual charge-dependent biases on the muon p_T as evaluated on data for CB momenta after alignment and after application of the correction procedure. The residual biases are shown separately for the three data-taking periods, (a) years 2015 and 2016, (b) 2017, and (c) 2018. The bias is defined in Eq. (1).

compared with the simulation is used to account for the kinematic dependencies of the δ_s estimate. The resulting uncertainty is small, compared with the non-closure described previously, in the central regions of pseudorapidity. However, in the forward pseudorapidity regions, it becomes sizeable and evolves as $2 \times 10^{-4} \, \text{TeV}^{-2} \times (p_T - 0.045 \, \text{TeV})$.

The statistical uncertainties due to the observed and simulated number of Z decays are also taken into account. These are approximately an order of magnitude smaller than the first two components, but in the forward region they reach up to $0.02 \, \text{TeV}^{-1}$.

Given the lack of sufficient $Z \rightarrow \mu\mu$ decays above 450 GeV in data to validate the methodology, the correction is not extrapolated beyond this value. Instead, an uncertainty accounting for the charge-dependent biases before any correction is assigned to the simulated momenta. This is estimated by biasing the simulated muon momenta in a perfectly aligned simulation, according to Equation 5, using the values of δ_s as measured before any correction. As done for the non-closure uncertainty, the uncertainty is derived by injecting bias into simulated events and taking the difference between the biased and the non-biased muon momentum values, resulting in an uncertainty of up to $0.4 \, \text{TeV}^{-1}$ as shown in Figure 1.
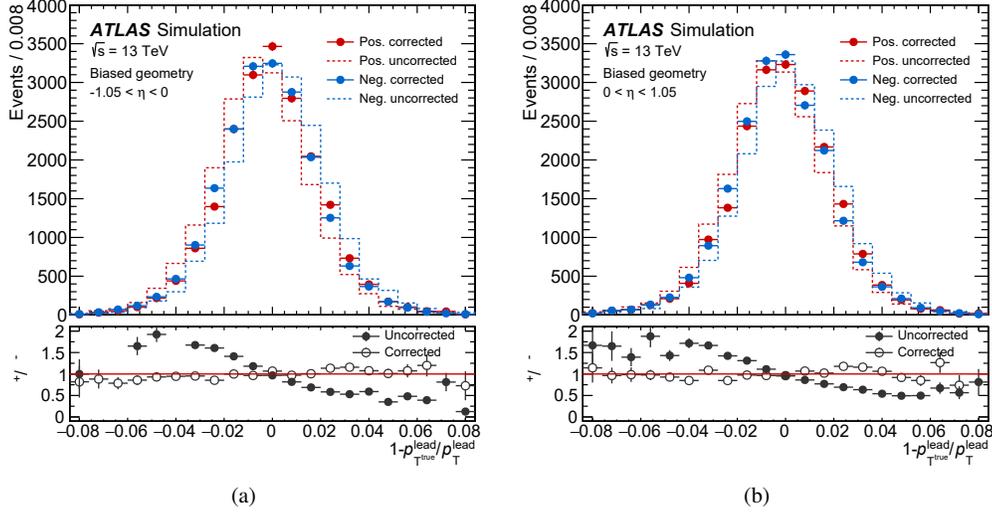
Figure 3: The bias of the transverse momentum measured in simulated data. The upper panels show the leading muon in the event (p_T^{lead}) compared with its true value separately for positive muons and negative muons, (a) for $-1.05 < \eta < 0$ and (b) for $0 < \eta < 1.05$. The dotted lines correspond to the momentum bias as evaluated on a simulation with a biased geometry, for positive and negative muons. The markers correspond to the bias evaluated with the same simulation after the application of the correction procedures. The lower panels show the ratios of positive to negative muons before and after correction.

5.2 Muon momentum calibration procedure in simulation

In the following, the *muon momentum calibration* is defined as the procedure used to identify the corrections to the reconstructed muon transverse momenta in simulation in order to match the measurement of the same quantities in data. This procedure is performed after correcting for the charge-dependent bias discussed in the previous section. The transverse momenta of the ID and MS tracks associated with a CB muon, referred to as p_T^{ID} and p_T^{MS}, respectively, are used in addition to the transverse momentum of the CB track, referred to as p_T^{CB}.

The corrected transverse momentum after applying the calibration procedure, $p_T^{Cor,Det}$ (Det= CB, ID, MS), is described as:

$$p_T^{Cor,Det} = \frac{p_T^{MC,Det} + \sum\limits_{n=0}^{1} s_n^{Det}(\eta, \phi) \left(p_T^{MC,Det}\right)^n}{1 + \sum\limits_{m=0}^{2} \Delta r_m^{Det}(\eta, \phi) \left(p_T^{MC,Det}\right)^{m-1} g_m}, \tag{8}$$

where $p_T^{MC,Det}$ is the uncorrected transverse momentum in simulation, g_m are normally distributed random variables with zero mean and unit width, and the terms $\Delta r_m^{Det}(\eta, \phi)$ and $s_n^{Det}(\eta, \phi)$ describe the momentum resolution smearing and scale corrections respectively, applied in a specific (η, ϕ) detector region. A possible $s_2^{Det}(\eta, \phi)$ term in the numerator is neglected because it would model effects already corrected in data with the procedure described in Section 5.1.

The corrections described in Equation 8 are defined in η–ϕ detector regions homogeneous in detector technology and performance. All corrections are divided into 18 pseudorapidity regions. In addition, the

CB and MS corrections are divided into two ϕ bins separating the two types of MS sectors: those that include the magnet coils (*small sectors*) and those between two coils (*large sectors*). The small and large MS sectors are subjected to independent alignment techniques and cover detector areas with different material distribution, leading to scale and resolution differences.

The numerator of Equation 8 describes the momentum scales. The $s_1^{\mathrm{Det}}(\eta, \phi)$ term corrects for inaccuracy in the description of the magnetic field integral and the dimension of the detector in the direction perpendicular to the magnetic field. The $s_0^{\mathrm{Det}}(\eta, \phi)$ term models the effect on the CB and MS momentum from the inaccuracy in the simulation of the energy loss in the calorimeter and other material between the interaction point and the exit of the MS. As the energy loss between the interaction point and the ID is negligible, $s_0^{\mathrm{ID}}(\eta)$ is set to zero [10].

The denominator of Equation 8 describes the momentum smearing that broadens the relative p_{T} resolution in simulation, $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$, to properly describe the data. The corrections to the resolution assume that the relative p_{T} resolution can be parameterized as follows:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = r_0/p_{\mathrm{T}} \oplus r_1 \oplus r_2 \cdot p_{\mathrm{T}}, \tag{9}$$

with $\oplus$ denoting a sum in quadrature. In Equation 9, the first term (r_0) mainly accounts for fluctuations of the energy loss in the traversed material; the second term (r_1) mainly accounts for multiple scattering, uncertainties related to, and inhomogeneities in, the modelling of the local magnetic field, and length-scale radial expansions of the detector layers; and the third term (r_2) mainly describes intrinsic resolution effects caused by the spatial resolution of the hit measurements and residual misalignments between the different detector elements. The energy loss term has a negligible impact on the muon resolution in the momentum range considered, and therefore $\Delta r_0^{\mathrm{Det}}$ is set to zero.

In a second step, to cross-check the validity of the corrections obtained directly for the CB tracks, the corrected combined momenta from ID and MS measurements $p_{\mathrm{T}}^{\mathrm{Corr\ ID+MS}}$ is also obtained by combining the ID and MS corrected momenta with a weighted average:

$$p_{\mathrm{T}}^{\mathrm{Cor,ID+MS}} = f \cdot p_{\mathrm{T}}^{\mathrm{Cor,ID}} + (1 - f) \cdot p_{\mathrm{T}}^{\mathrm{Cor,\ MS}}. \tag{10}$$

The weight f is calculated by solving the following linear equation:

$$p_{\mathrm{T}}^{\mathrm{MC,CB}} = f \cdot p_{\mathrm{T}}^{\mathrm{MC,ID}} + (1 - f) \cdot p_{\mathrm{T}}^{\mathrm{MC,\ MS}}, \tag{11}$$

which assumes that the relative contribution of the two sub-detectors to the combined track remains unchanged before and after momentum corrections.

5.2.1 Determination of the p_{T} calibration constants

The CB, ID, and MS correction parameters contained in Equation 8 are extracted from data using a fitting procedure that compares the invariant mass distributions for $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ candidates in data and simulation, selected as discussed in Section 4.

When extracting correction parameters, muons are assigned to η–ϕ regions of fit (ROFs), which are defined separately for the ID and the MS. The values of Δr_0^{ID}, Δr_0^{MS}, Δr_0^{CB}, and s_0^{ID}, are set to zero as previously

discussed, while Δr_2^{MS} is determined from alignment studies using cosmic-ray data and special runs with the toroidal magnetic field turned off [50], and from detector calibration procedures [17].

The corrections are extracted using the distributions of the dimuon invariant mass, $m_{\mu\mu}^{\text{Det}}$. When fitting a specific ROF, one muon is required to belong to the ROF and the other can be anywhere in the detector. To enhance the sensitivity to p_{T}-dependent correction effects, the dimuon pair is classified according to the p_{T} of the muons. For $J/\psi \to \mu\mu$ decays, the fit is performed in two exclusive categories defined by the subleading muon transverse momentum: $6.25 < p_{\text{T}}^{\text{Det,sublead}} < 9$ GeV and $p_{\text{T}}^{\text{Det,sublead}} > 9$ GeV. For $Z \to \mu\mu$ decays the two categories are defined by the leading muon transverse momentum: $20 < p_{\text{T}}^{\text{Det,lead}} < 50$ GeV and $p_{\text{T}}^{\text{Det,lead}} > 50$ GeV.

Templates for the $m_{\mu\mu}^{\text{Det}}$ variables are built using $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ simulated signal samples. The first step in the fitting procedure consists of estimating the contribution of the background processes. For the $Z \to \mu\mu$ selection, the small background component (approximately 0.1%) is extracted from simulation and added to the templates. A much larger (up to 15%) non-resonant background from decays of light and heavy hadrons and from continuum Drell–Yan production is present for events selected in the $J/\psi \to \mu\mu$ channel, estimated with a data-driven approach as this background is difficult to simulate. The dimuon invariant mass distribution in data is fitted in each ROF using a Crystal Ball function [52] with an exponential background distribution. This background model is then combined with the simulated signal templates used in the fit.

To estimate the scale and smearing parameters, a multi-stage procedure is used. A two-step random walk minimisation is first performed where the parameters are sampled within a specified range. The parameter configuration that leads to the best agreement for the invariant mass distribution between simulation and data is kept. In the first step, the best agreement is defined by a simplified metric where the compatibility between the means and standard deviations of the distributions is compared. In the second step, a binned χ^2 is computed that compares the dimuon mass spectrum in simulation to that of data. Around this best-agreement parameter configuration, signal templates are then generated at discrete intervals and interpolated using moment morphing [53]. Using this signal model, a full minimisation using the binned χ^2 is performed. This procedure is iteratively repeated, where the second muon outside the ROF definition is calibrated using results from the previous iteration, until the change of each parameter stays within the root mean squared of its values as estimated from the last five iterations. This happens after 16 iterations. Each calibration parameter's values from the last five iterations are averaged to produce their final value.

The calibration parameters obtained from the fits to the data are summarised in Tables 1, 2, and 3, averaged over three η regions. The sums in quadrature of the statistical and systematic uncertainties are shown, with the latter dominating. All sources of uncertainties are evaluated by varying the parameters of the template fit. The higher uncertainty for Δr_1^{ID} in the second bin is associated with the region of the ID with the largest amount of material, corresponding to the transition between the barrel and endcap region of its subdetectors. The increase in uncertainty as a function of $|\eta|$ for s_1^{ID} is due to the decrease of the magnetic field integral of the solenoid as a function of $|\eta|$. The larger uncertainties for some of the MS terms in the $1.05 \leq |\eta| < 2.0$ region are associated with the low bending power of the magnet system of the MS in part of that region.

The main contributions to the final systematic uncertainty are:

- J/ψ only and Z only fit: The fit is repeated using only the $J/\psi \to \mu\mu$ or $Z \to \mu\mu$ decays with only s_1^{Det} left as a free parameter and the other parameters fixed to their nominal fitted value. The resulting uncertainty is defined by taking the difference relative to the nominal fitted s_1^{Det} and accounts for the

extrapolation from the regions dominated by either the $J/\psi \rightarrow \mu\mu$ or the $Z \rightarrow \mu\mu$ in their respective p_T spectra.

- Z kinematic reweighting: As discussed in Section 3.2, the reweighting maps are derived separately for each simulated sample. To evaluate the impact of the reweighting on the analysis, new templates of $m_{\mu\mu}^{\mathrm{ID}}$, $m_{\mu\mu}^{\mathrm{CB}}$, and $m_{\mu\mu}^{\mathrm{MS}}$ are produced. The results obtained when using the different $Z \rightarrow \mu\mu$ samples with their respective reweighting scheme and the non-reweighted results for the nominal $Z \rightarrow \mu\mu$ samples are used to derive the uncertainty. This systematic uncertainty impacts the overall Z lineshape by changing the relative weight of various momentum regimes and, by extension, the contributions of their dedicated corrections in the final fit.

- Decay modelling and final state radiation modelling: The nominal simulated sample of $Z \rightarrow \mu\mu$ events, of which the QED final-state radiation is handled by PHOTOS++, is compared with $Z \rightarrow \mu\mu$ events simulated with the SHERPA 2.2.1 sample which uses an alternative radiative modelling. The impact of the uncertainties in the decay modelling are significantly smaller than those originating from the components discussed above, partially due to the already discussed kinematic reweighting.

- J/ψ and Z p_T template range: The p_T ranges of the J/ψ and Z templates are varied. For J/ψ, the ranges of the two nominal p_T templates are $p_\mathrm{T}^{J/\psi,\mathrm{nom}} \in [6.25, 9]$ GeV and $p_\mathrm{T}^{J/\psi,\mathrm{nom}} \in [9, 20]$ GeV. For Z, the ranges of two templates are $p_\mathrm{T}^{Z,\mathrm{nom}} \in [20, 50]$ GeV and $p_\mathrm{T}^{Z,\mathrm{nom}} \in [50, 300]$ GeV. The boundary of the two templates for J/ψ is varied from 9 to 8 and 12 GeV, and for Z is varied from 50 to 40 and 80 GeV. This preserves the numbers of events, and covers any additional p_T dependencies.

- Mass binning of J/ψ and Z: The number of bins for the $m_{\mu\mu}$ templates are reduced from 200 to 150 for Z, and from 90 to 60 for J/ψ. This systematic uncertainty covers any binning effect in generating templates.

- Z mass window: New templates of $m_{\mu\mu}$ are generated for the $Z \rightarrow \mu\mu$ sample changing the m_Z window selection to $75 < m_{\mu\mu} < 115$ GeV for all track types. The high and low mass regions away from the resonant pole of the Z are more sensitive to initial- and final-state radiation contributions, the running of α_Z^{EM}, and other specific choices of the MC generators used in the $Z \rightarrow \mu\mu$ simulation.

- J/ψ mass window: The same approach as in the case of the Z is applied, changing the invariant mass window selection to $2.75 < m_{\mu\mu} < 3.4$ GeV for all track types. The region furthest from the $m_{J/\psi}$ peak is more sensitive to the shape variation of the combinatorial background, and the specific choices of the MC generators used in the $J/\psi \rightarrow \mu\mu$ simulation.

- J/ψ background: New $m_{\mu\mu}$ templates are generated for the $J/\psi \rightarrow \mu\mu$ samples using a background parameterisation different from the exponential function. For this systematic uncertainty, background parameterisations including a combination of two Crystal ball functions and a Chebyshev function are used.

- Closure test and statistical uncertainty: The root mean squared of the last five iterations of each calibration parameter is used as the statistical uncertainty of the method. This accounts for any instabilities in the fit, and the correlations between different parameters.

For muons with transverse momenta of about 45 GeV, the largest component of the total uncertainty in the momentum scale originates from the comparison of the J/ψ and Z only fits. The remaining components are smaller by about a factor of two. For the resolution smearing terms, in the same kinematic regime, the largest uncertainty comes from the J/ψ and Z p_T template fit range. At significantly higher momenta the uncertainty in the J/ψ and Z p_T template fit range still dominates the resolution uncertainty in $p_\mathrm{T}^{\mathrm{Cor,CB}}$.

Table 1: Summary of ID muon momentum resolution and scale corrections used in Equation 8, averaged over the three main detector regions. The corrections are derived in 18 pseudorapidity regions, as described in Section 5, and averaged, with each region assigned a weight proportional to its η width. The uncertainties represent the sum in quadrature of the statistical and systematic uncertainties.

Region	$\Delta r_1^{\mathrm{ID}}(\times 10^{-3})$	Δr_2^{ID} [TeV^{-1}]	$s_1^{\mathrm{ID}}(\times 10^{-3})$
$|\eta| < 1.05$	$4.4^{+3.0}_{-2.7}$	$0.12^{+0.03}_{-0.03}$	$-0.82^{+0.19}_{-0.06}$
$1.05 \leq |\eta| < 2.0$	$6.7^{+19.5}_{-3.1}$	$0.31^{+0.03}_{-0.03}$	$-0.86^{+0.36}_{-0.21}$
$|\eta| \geq 2.0$	$9.4^{+3.6}_{-5.3}$	$0.08^{+0.02}_{-0.02}$	$-0.45^{+1.19}_{-0.55}$

Table 2: Summary of MS momentum resolution and scale corrections for small and large MS sectors, averaged over three main detector regions. The corrections for large and small MS sectors are derived in 18 pseudorapidity regions, as described in Section 5, and averaged, with each region assigned a weight proportional to its η width. The uncertainties represent the sum in quadrature of the statistical and systematic uncertainties. The larger uncertainties for some terms in the $1.05 \leq |\eta| < 2.0$ region are associated with the low bending power of the magnet system of the MS in part of that region.

Region	Δr_0^{MS} [MeV]	$\Delta r_1^{\mathrm{MS}}(\times 10^{-3})$	Δr_2^{MS} [TeV^{-1}]	s_0^{MS} [MeV]	$s_1^{\mathrm{MS}}(\times 10^{-3})$
$|\eta| < 1.05$ (large)	71^{+41}_{-10}	$13.5^{+3.1}_{-2.6}$	$0.11^{+0.02}_{-0.02}$	-27^{+18}_{-14}	$2.33^{+0.61}_{-0.61}$
$|\eta| < 1.05$ (small)	63^{+35}_{-23}	$17.2^{+2.3}_{-3.6}$	$0.12^{+0.02}_{-0.02}$	-18^{+14}_{-24}	$-1.51^{+1.15}_{-0.45}$
$1.05 \leq |\eta| < 2.0$ (large)	42^{+401}_{-53}	$26.7^{+19.5}_{-2.8}$	$0.14^{+0.02}_{-0.02}$	-29^{+43}_{-57}	$-0.9^{+2.8}_{-1.9}$
$1.05 \leq |\eta| < 2.0$ (small)	44^{+437}_{-59}	$29.2^{+8.2}_{-1.5}$	$0.18^{+0.02}_{-0.02}$	-28^{+39}_{-43}	$-1.3^{+2.0}_{-1.6}$
$|\eta| \geq 2.0$ (large)	61^{+48}_{-27}	$16.8^{+3.2}_{-4.3}$	$0.11^{+0.02}_{-0.02}$	-1^{+43}_{-29}	$1.7^{+1.3}_{-1.4}$
$|\eta| \geq 2.0$ (small)	47^{+32}_{-40}	$17.8^{+1}_{-1.9}$	$0.16^{+0.02}_{-0.02}$	-8^{+14}_{-12}	$1.94^{+0.76}_{-0.54}$

As an additional check, the value of $p_{\mathrm{T}}^{\mathrm{Cor,ID+MS}}$ is compared with $p_{\mathrm{T}}^{\mathrm{Cor,CB}}$. Results from both methods are in agreement within their respective systematic uncertainties. The comparison is performed on simulated samples with a biased alignment of the MS on muons with momenta ranging from few tens of GeV to about 1 TeV. The uncertainties in the momentum scale of $p_{\mathrm{T}}^{\mathrm{Cor,CB}}$ are found to be smaller than those of $p_{\mathrm{T}}^{\mathrm{Cor, ID+MS}}$ with a slightly larger uncertainty increase in the resolution as a function of the transverse momentum, as shown in Appendix B. At high momenta the dominant uncertainty in $p_{\mathrm{T}}^{\mathrm{Cor,ID+MS}}$ originates from the residual muon spectrometer misalignment.

The resolution as a function of the muon momentum is shown in Figure 4, (a) for muons in the barrel region ($|\eta| < 1.05$) and (b) for all other muons ($|\eta| > 1.05$). The resolution is estimated by taking half of the smallest interval containing the 68% of the distribution of the relative difference between the corrected and the generated single muon momentum in simulation for muons satisfying the *High-p_T* WP. The resolution is approximately 2% (3%) at p_{T} values of 45 GeV in the $|\eta| < 1.05$ ($|\eta| > 1.05$) region and increases with p_{T}. The resulting CB track momentum resolution is always better than that of the individual measurements in the ID or the MS. At low p_{T} the ID dominates while in the intermediate regime of few tens of GeV to a few hundreds of GeV both detector systems contribute equally to the measurement. At p_{T} values higher than few hundred GeV the resolution is dominated by the MS. The resolution is approximately 14% (19%)

Table 3: Summary of CB momentum resolution and scale corrections for small and large MS sectors, averaged over three main detector regions. The corrections for large and small MS sectors are derived in 18 pseudorapidity regions, as described in Section 5, and averaged, with each region assigned a weight proportional to its η width. The energy loss term Δr_0^{CB} is negligible and therefore fixed to zero in the fit for all η. The uncertainties represent the sum in quadrature of the statistical and systematic uncertainties.

Region	$\Delta r_1^{CB}(\times 10^{-3})$	Δr_2^{CB} [TeV^{-1}]	s_0^{CB} [MeV]	$s_1^{CB}(\times 10^{-3})$
$|\eta| < 1.05$ (large)	$6.7^{+1.4}_{-0.9}$	$0.08^{+0.04}_{-0.05}$	$-5.0^{+2.9}_{-4.0}$	$0.35^{+0.24}_{-0.22}$
$|\eta| < 1.05$ (small)	$6.5^{+1.3}_{-1.0}$	$0.11^{+0.05}_{-0.05}$	$-0.9^{+2.5}_{-3.6}$	$-0.83^{+0.25}_{-0.14}$
$1.05 \leq |\eta| < 2.0$ (large)	$10.3^{+2.6}_{-2.7}$	$0.24^{+0.10}_{-0.07}$	$-2.0^{+5.7}_{-6.7}$	$-0.83^{+0.39}_{-0.30}$
$1.05 \leq |\eta| < 2.0$ (small)	$8.9^{+1.7}_{-2.7}$	$0.29^{+0.08}_{-0.03}$	$-3.0^{+3.3}_{-4.0}$	$-0.80^{+0.26}_{-0.21}$
$|\eta| \geq 2.0$ (large)	$10.6^{+2.2}_{-2.7}$	$0.21^{+0.10}_{-0.07}$	$2.3^{+13}_{-9.3}$	$0.80^{+1.09}_{-0.42}$
$|\eta| \geq 2.0$ (small)	$11.5^{+2.2}_{-2.1}$	$0.26^{+0.08}_{-0.06}$	$-12.6^{+8.2}_{-9.7}$	$1.59^{+0.47}_{-0.43}$

(a) (b)
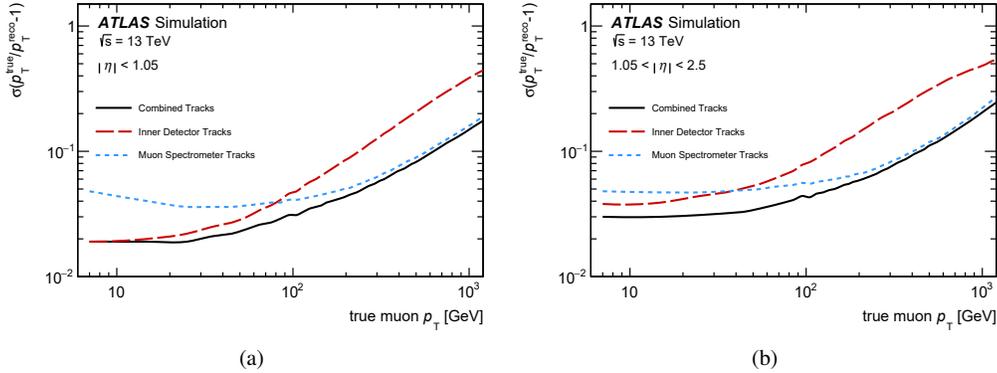
Figure 4: Resolution of the muon p_T as obtained from simulation after derivation and application of all correction constants. Muons are selected using the *High-p_T* WP. The resolution is shown as a function of the true p_T of the muon for a range from 1 GeV to 2.5 TeV, (a) for muons with $|\eta| < 1.05$ and (b) for muons with $|\eta| > 1.05$. The resolution lines are obtained by interpolating between points sampled in steps of p_T. The continuous lines correspond to the CB momentum, while the dashed lines to that from the ID, and the dotted lines to that of the MS. Statistical fluctuations, due to the simulated sample size result in small fluctuations on the interpolation between measured values.

at 1 TeV in the $|\eta| < 1.05$ ($|\eta| > 1.05$) region.

6 Performance studies

The samples of $J/\psi \rightarrow \mu\mu$, $\Upsilon \rightarrow \mu\mu$, and $Z \rightarrow \mu\mu$ decays are used to validate the momentum corrections obtained with the template fit method described in the previous sections and measure the muon momentum reconstruction performance. The detector segmentation applied to $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays is

chosen to be at least twice as fine as that used for deriving the simulation corrections. Furthermore, the momentum requirements used in both cases are looser than those used for deriving the corrections. The combination of the finer segmentation and looser requirements allows the data-driven validation of the correction methods both within the bins assumed in the template fits and extrapolated in p_T beyond the range of the calibration procedure. To complete the data-driven validation and performance measurements, an additional study is performed using $\Upsilon \rightarrow \mu\mu$ decays, which are statistically fully independent of the samples used to derived the calibration corrections.

The invariant mass distributions for the $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ candidates are shown in Figure 5 and compared with corrected simulation. The lineshapes of the two resonances in simulation agree with the data within the systematic uncertainties, demonstrating the overall effectiveness of the p_T calibration.

(a)   (b)
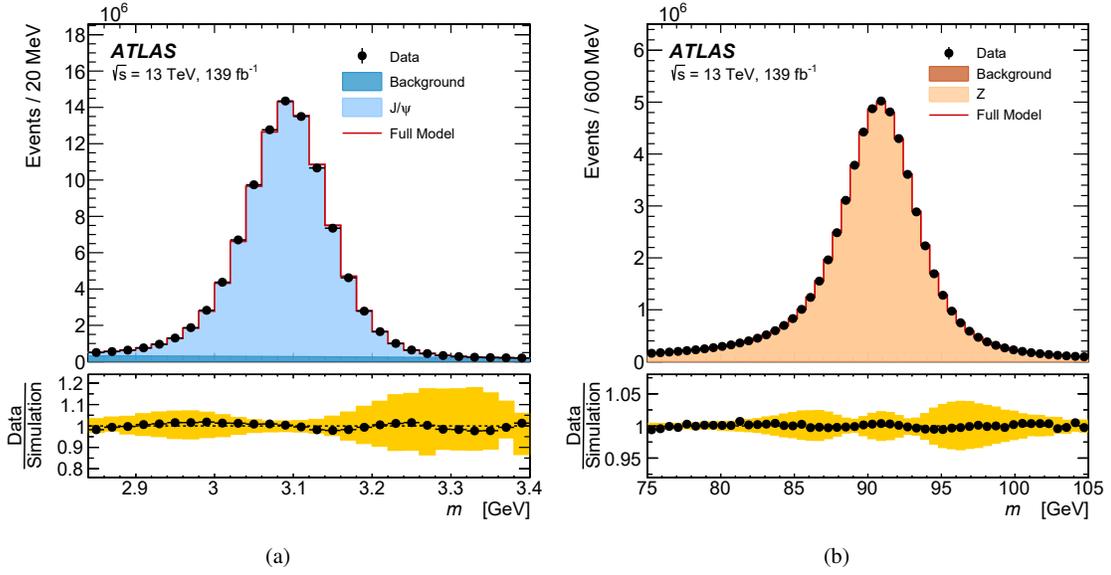
Figure 5: Dimuon invariant mass distribution of (a) $J/\psi \rightarrow \mu\mu$ and (b) $Z \rightarrow \mu\mu$ candidate events reconstructed with CB muons. The upper panels show the invariant mass distribution for data and for the signal simulation plus the background estimate. The points show the data. The continuous line corresponds to the simulation with the MC momentum corrections applied. Background estimates are added to the signal simulation. The bands represent the effect of the systematic uncertainties in the MC momentum corrections. The lower panels show the ratios of the data to the MC simulation. In the Z sample, the MC background samples are added to the signal sample according to their expected cross-sections. In the J/ψ sample, the background is estimated from a fit to the data as described in the text. The sum of background and signal MC distributions is normalised to the data.

When the two muons have similar momentum resolution and angular effects are neglected, the relative mass resolution, $\sigma_{\mu\mu}/m_{\mu\mu}$, is directly proportional to the relative muon momentum resolution, σ_{p_μ}/p_μ:

$$\frac{\sigma_{\mu\mu}}{m_{\mu\mu}} = \frac{1}{\sqrt{2}} \frac{\sigma_{p_\mu}}{p_\mu}. \tag{12}$$

Similarly, the total muon momentum scale, defined as $s = \langle (p^{\text{meas}} - p^{\text{true}})/p^{\text{true}} \rangle$, is directly related to the dimuon mass scale, defined as $s_{\mu\mu} = \langle (m_{\mu\mu}^{\text{meas}} - m_{\mu\mu}^{\text{true}})/m_{\mu\mu}^{\text{true}} \rangle$:

$$s_{\mu\mu} = \sqrt{s_{\mu_1} s_{\mu_2}}, \tag{13}$$

where s_{μ_1} and s_{μ_2} are the momentum scales of the two muons. The effectiveness of the momentum calibration is also measured by comparing the mean value $m_{\mu\mu}$, and resolution $\sigma(m_{\mu\mu})$ of the dimuon mass resonances. To measure such quantities, fully analytical fit functions are created for each resonance modelling both the signal and the background. Using the same definition of the resolution function for all resonances, this methodology allows a direct comparison of the resolution quantities as extracted separately for each resonance. In turn, this allows a precise measurement of the momentum reconstruction performance across a wide kinematic regime. The fitting procedures used to measure the mean value and resolution quantities are optimised, subsequently, for each resonance.

In $J/\psi \rightarrow \mu\mu$ decays, the intrinsic width of the resonance is negligible compared with the experimental resolution. In contrast to Section 5.2, the mass window for the $J/\psi \rightarrow \mu\mu$ validation fits is modified to $2.8 < m_{\mu\mu} < 3.9$ GeV. The lower bound is raised to remove the abrupt spectrum sculpting due to the trigger criteria, which is difficult to model well analytically. The upper bound is extended to include the $\Psi(2S)$ resonance and add a lever arm to constrain the estimation of the resolution.

The resonant peak of the J/ψ is modelled by a double-sided Crystal Ball function; the four parameters modelling the tails are extracted from simulation, and subsequently fixed in the fits to data. A secondary double-sided Crystal Ball function models the $\Psi(2S)$ resonance that follows the $1s$ resonance in the J/ψ validation mass window. Similarly to the $1S$ resonance, the parameters modelling the tails are extracted from simulation and kept fixed when fitting to data. The mean value parameter of the resolution function is kept free floating for the $\Psi(2S)$ resonance for fits on simulated and real data, while the resolution parameter is constrained to scale linearly with the resolution parameter of the $1S$ resonance using the ratio of the mean parameters. The non-resonant background or the background from mis-identified muons is described by an exponential function.

In $Z \rightarrow \mu\mu$ decays, the fits use a convolution of the true lineshape (modelled by a Breit–Wigner function) with an experimental resolution function (a double-sided Crystal Ball). A fit range of $75 < m_{\mu\mu} < 105$ GeV is used. Similarly to the J/ψ, the non-resonant background is described by an exponential function. The peak position and width of the Crystal Ball function are used as estimators for the $m_{\mu\mu}$ and $\sigma(m_{\mu\mu})$ variables in each of the η and p_T bins.

Figure 6 shows the position of the mean value of the invariant mass distribution, $m_{\mu\mu}$, obtained from the fits to the Z boson and J/ψ samples as a function of the pseudorapidity of the highest-p_T muon for each decay. The distributions are shown for data and corrected simulation, with the ratios of the two in the lower panels. The simulation is in good agreement with the data. Minor deviations are within the momentum scale systematic uncertainties of 0.05% in the barrel region increasing with $|\eta|$ to 0.15% for $J/\psi \rightarrow \mu\mu$ decays, and to 0.05% in the barrel region increasing with $|\eta|$ to 0.1% for $Z \rightarrow \mu\mu$ decays. The systematic uncertainties shown in the plots include the effects of the uncertainties in the calibration constants described in Section 5.2. The observed level of agreement demonstrates that the p_T calibration for combined muon tracks described above provides an accurate description of the momentum scale in all η regions, over a wide p_T range. Similar levels of data/MC agreement are observed for the ID and MS components of the combined tracks.

Figure 7 displays the dimuon mass resolution $\sigma(m_{\mu\mu})$ as a function of the leading-muon η for the two resonances. The dimuon mass resolution is about 1.3% and 1.6% at small η values for the J/ψ and Z bosons, respectively, and increases to 2.1% and 2.4% in the endcaps. This corresponds to a relative muon p_T resolution of 1.8% and 2.3% in the centre of the detector and 3.0% and 3.4% in the endcaps for J/ψ and
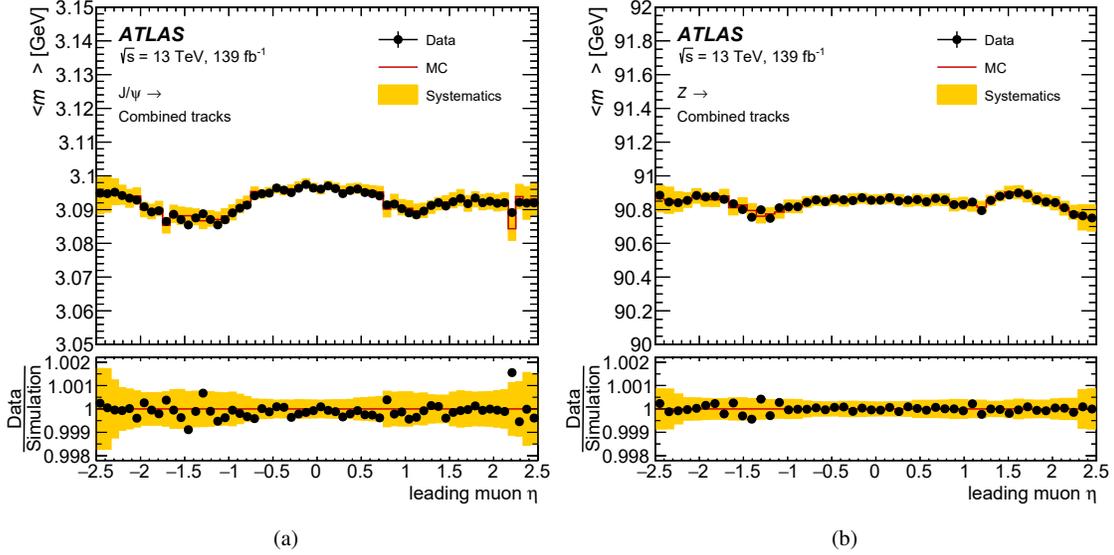
Figure 6: Fitted mean mass of the dimuon system for CB muons for (a) $J/\psi \to \mu\mu$ and (b) $Z \to \mu\mu$ events for data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. The upper panels show the fitted mean mass value for data and corrected simulation. The small variations of the invariant mass estimator as a function of pseudorapidity are due to imperfect energy loss corrections and magnetic field description in the muon reconstruction. Both effects are well reproduced in the simulation. The lower panels show the ratios of the data to the MC simulations. The error bars represent the statistical uncertainty; the shaded bands represent the systematic uncertainty in the correction and the systematic uncertainty in the extraction method added in quadrature.

Z boson decays, respectively. Uncertainties in the dimuon mass resolution range between 2% and 5% for the J/ψ and between 3% and about 6% for the Z boson, depending on the detector region.

Using the same methodologies as for the $J/\psi \to \mu\mu$ decays, the measurement of the scale dependence and resolution is repeated on the fully independent set of data from $\Upsilon \to \mu\mu$ decays. The same approach as for the J/ψ resonances models the 2S and 3S resonances of the Υ. This entails constraining the resolution parameter of the 3S resonance to scale linearly with that of the 2S resonance and that of the 2S resonance to scale linearly with that of the 1S resonance. The mean parameters of the resolution models of the 2S and 3S are kept free independently in data and simulation, similarly to what is done for the J/ψ resonance. The background is, however, modelled using a polynomial function derived from data. Figure 8 shows the data and simulation agreement after applying the momentum corrections on simulation for the invariant mass distributions of the three resonances of the $\Upsilon \to \mu\mu$. As done for the $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ decays, the momentum scale dependency and the resolution are extracted as a function of the pseudorapidity of the leading muon in the event. Figure 9 compares the data and simulation after all corrections. The data agree with the simulation within the systematic uncertainties. The uncertainty in the momentum scale is up to 0.05% in the central region and up to 0.1% in the forward region. Similarly, the extracted resolution is in agreement within about 3%. Good agreement between the dimuon mass resolution measured in data and simulation is also observed when deriving the corrections independently for the ID and MS components of the combined tracks, as shown in Appendix B.

The relative dimuon mass resolution $\sigma_{\mu\mu}/m_{\mu\mu}$ is approximately proportional to the average momentum of
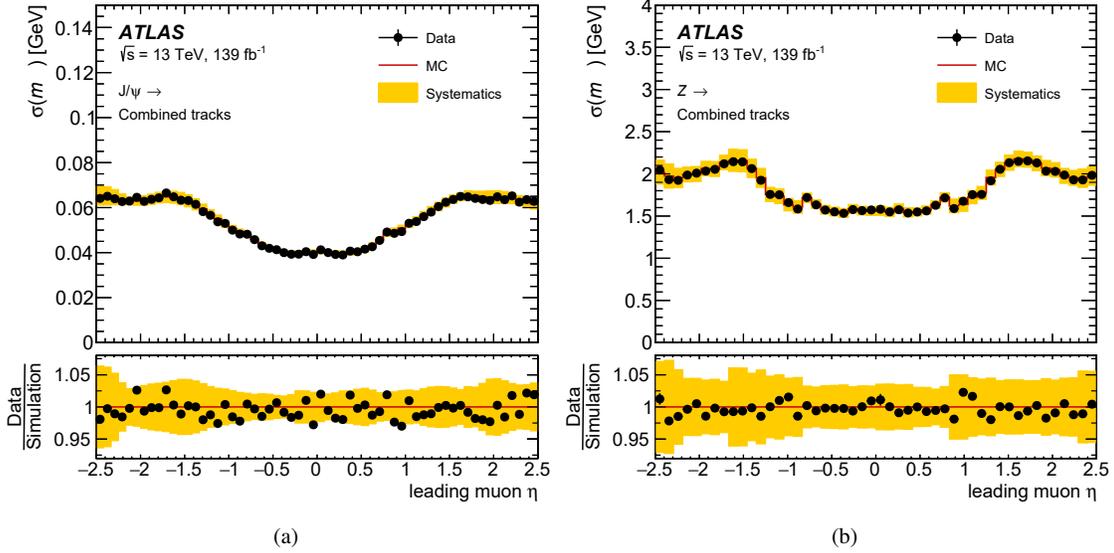
Figure 7: Dimuon invariant mass resolution for CB muons for (a) $J/\psi \to \mu\mu$ and (b) $Z \to \mu\mu$ events for data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. The upper panels show the fitted resolution value for data and corrected simulation. The lower panels show the ratios of the data to the MC simulations. The error bars represent the statistical uncertainty; the shaded bands represent the systematic uncertainty in the correction and the systematic uncertainty in the extraction method added in quadrature. For the $Z \to \mu\mu$ the estimator shown in the plot is only sensitive to the detector resolution and not to the natural width of the Z boson, as discussed in the text.

the muons, as shown in Equation 12. A direct comparison of the momentum resolution function determined with J/ψ, Υ, and Z boson decays can therefore be performed. To remove the effect of the correlation between the measurement of the dimuon mass resolution and the p_T of the muons, the following definition of transverse momentum is used:

$$p_T^* = \hat{m}\sqrt{\frac{\sin\theta_1 \sin\theta_2}{2(1 - \cos\alpha_{12})}},$$

(14)

where $\hat{m}$ is a fixed value, corresponding to the best known value of the mass of the Z boson, the J/ψ, and the Υ, respectively. The variables θ_1 and θ_2 are the polar angles of the two muons, and α_{12} is the opening angle of the muon pair. The relative dimuon mass resolution from $J/\psi \to \mu\mu$, $\Upsilon \to \mu\mu$, and $Z \to \mu\mu$ events measured in data is compared with that obtained from calibrated simulation as a function of of p_T^* in Figure 10. The resolutions are in good agreement. Due to the larger number of events, the $J/\psi \to \mu\mu$ measurement extends to higher momenta than the $\Upsilon \to \mu\mu$ one.
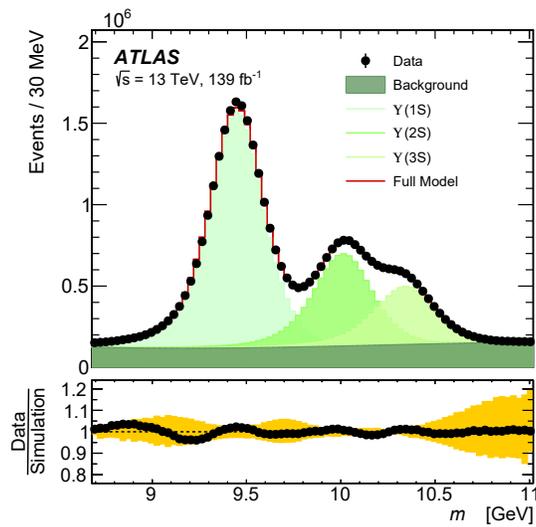
Figure 8: Dimuon invariant mass distribution of $\Upsilon \to \mu\mu$ candidate events reconstructed with CB muons. The upper panels show the invariant mass distribution for data and for the signal simulation plus the background estimate. The points show the data. The continuous line corresponds to the simulation with the MC momentum corrections applied. The band represents the effect of the systematic uncertainties in the MC momentum corrections. The lower panel shows ratio of the data to the MC simulations. The background is estimated from a fit to the data as described in the text. The sum of background and signal MC distributions is normalised to the data.
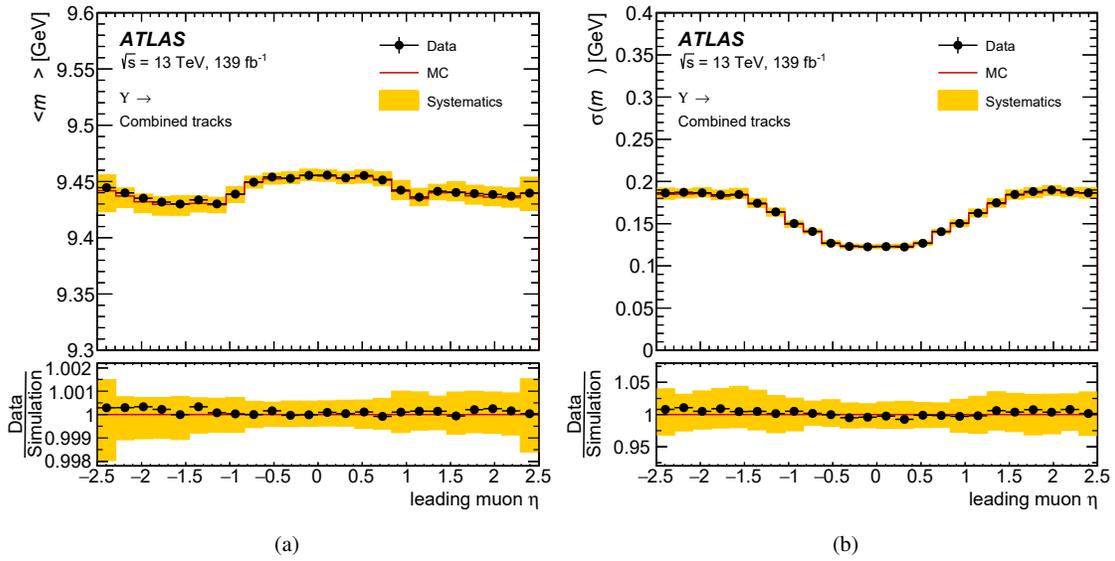
Figure 9: (a) Fitted mean mass of the dimuon system for CB muons for $\Upsilon \rightarrow \mu\mu$ events in data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. (b) Dimuon invariant mass resolution for CB muons for $\Upsilon \rightarrow \mu\mu$ events for data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. The upper panels show the results obtained on data and simulation after all corrections. The lower panels show the ratios of the data to the MC simulation. The error bars represent the statistical uncertainty; the shaded bands represent the systematic uncertainty in the correction and the systematic uncertainty in the extraction method added in quadrature.
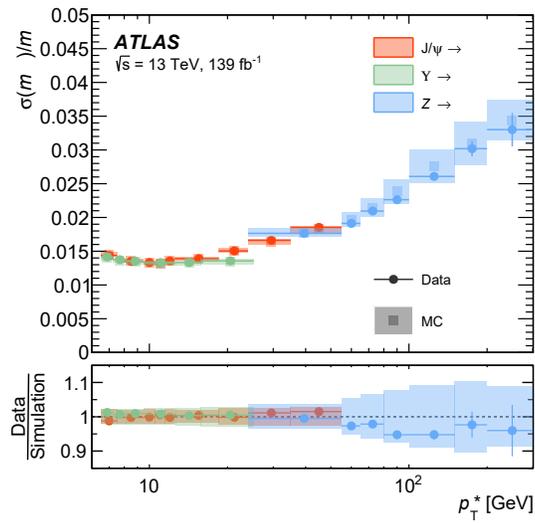
Figure 10: The upper panel shows the dimuon invariant mass resolution divided by the dimuon invariant mass for CB muons measured from $J/\psi \to \mu\mu$, $\Upsilon \to \mu\mu$ and $Z \to \mu\mu$ events as a function of the p_T^* variable defined in the text. To account for the resolution differences induced by the different pseudorapidity distributions among muons from the three resonances, the data and simulations are re-weighted according to the pseudorapidity distribution of the muon with the highest momentum from $\Upsilon \to \mu\mu$ decays. The lower panel shows the ratio of the data to the MC simulation. The error bars represent the statistical uncertainty, while the bands show the systematic uncertainties.

7 Conclusion

The momentum performance of the ATLAS muon reconstruction is measured using 139 fb^{-1} of data from pp collisions at $\sqrt{s} = 13$ TeV recorded during the second run of LHC between 2015 and 2018.

Events from the $Z \rightarrow \mu\mu$ process are used to correct the reconstructed data for a charge-dependent bias in the muon momentum measurement associated with detector alignment effects. This bias is reduced on average from up to 0.4 TeV^{-1} to $2 \cdot 10^{-4}$ TeV^{-1} for muons with $p_T < 450$ GeV after the corrections are applied, with associated uncertainty at the level of 0.03 TeV^{-1} for muons with transverse momentum of about 45 GeV.

The scale and resolution of the muon momentum measurement is studied in detail using $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays. These studies are used to correct the simulation, improving the agreement with the data and reducing the systematic uncertainties related to the muon calibration in physics analyses. The improvements in the p_T correction methods described in this paper and the substantial number of $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays collected in the Run 2 data sample together improve the precision of the momentum scale by up to a factor of two relative to the previous publication based on 3.2 fb^{-1} of collected data [10]. For $Z \rightarrow \mu\mu$ decays, the uncertainty in the momentum scale varies from a minimum of 0.05% for $|\eta| < 1$ to a maximum of 0.15% for $|\eta| \sim 2.5$.

The dimuon mass resolution is about 1.3% (1.6%) at small values of pseudorapidity for J/ψ (Z) decays, and increases up to 2.1% (2.4%) in the endcaps. This corresponds to a relative muon p_T resolution of 1.8% (2.3%) at small values of pseudorapidity and 3.0% (3.4%) in the endcaps for J/ψ (Z) decays. After applying momentum corrections, the p_T resolution in data and simulation agree within the quoted uncertainties, which are at the level of, or better than, 5% (6%) for J/ψ (Z) decays depending on the η range.

Validation studies performed with $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ show that the corrections applied to the simulation bring the agreement with the data to the level of their estimated uncertainties. An additional, statistically independent validation performed using $\Upsilon \rightarrow \mu\mu$ decays confirms the correctness of the correction procedure.

Appendix

This appendix complements the results presented in the paper with some additional material. Appendix A provides the measured charge-dependent biases before and after application of the correction procedure for the alternative method discussed in Section 5.1. Appendix B documents the measured performance when applying the corrections for CB tracks as obtained by the combination of corrections to ID and MS measurements.
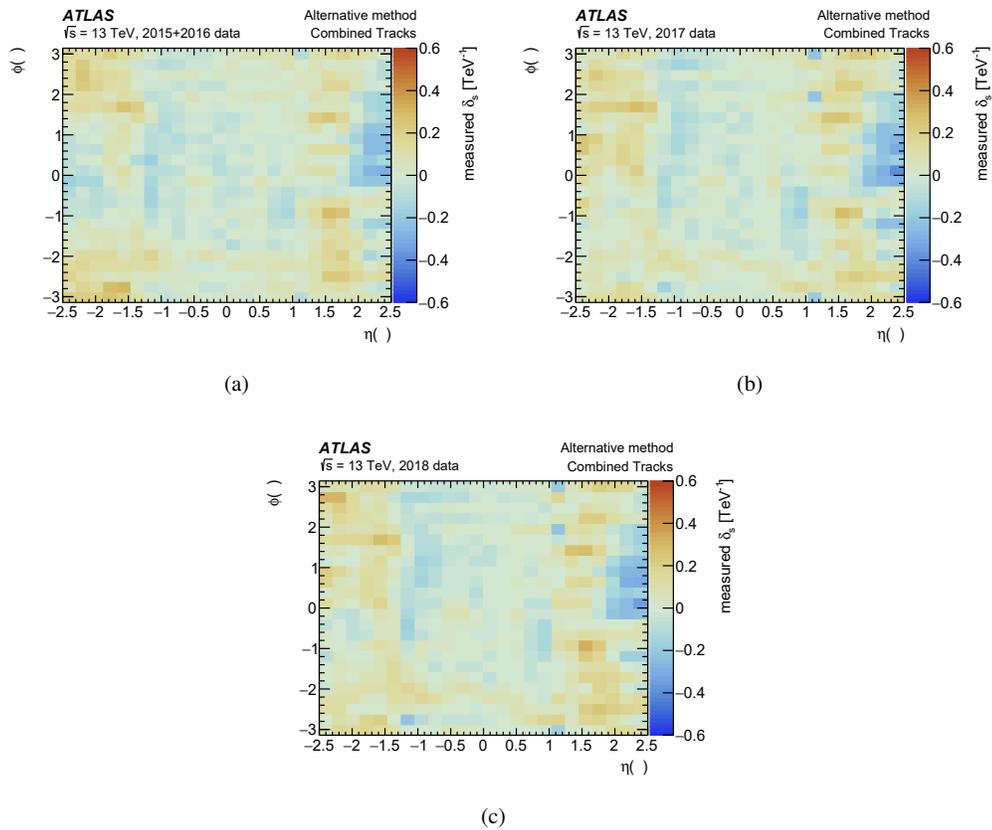
A Charge-dependent momentum scale calibration in data

(a)

(b)

(c)

Figure 11: Charge-dependent biases on the muon p_T as evaluated on data for CB momenta after alignment and before applying the dedicated corrections using the alternative method described in Section 5.1. The biases are shown separately for the three data-taking periods of (a) years 2015 and 2016, (b) 2017 and (c) 2018. The bias is defined in Eq. (1).

As described in the main body of the paper, a complementary method to estimate the charge-dependent bias was also developed. The details of the method are explained in Section 5.1. In this Appendix the results obtained with this method are presented. Similarly to Figure 1, Figure 11 shows the measured

sagitta corrections as evaluated on CB momenta. It can be seen that the biases also extend to about 0.4 TeV^{-1}. After using the correction procedure with the iterative method explained in Section 5.1, using 25 iterations, the residual biases are re-evaluated with the same methodology. The resulting values, as shown in Figure 12, indicate the same reduction of the measured biases to values of less than 10^{-3}TeV^{-1}.
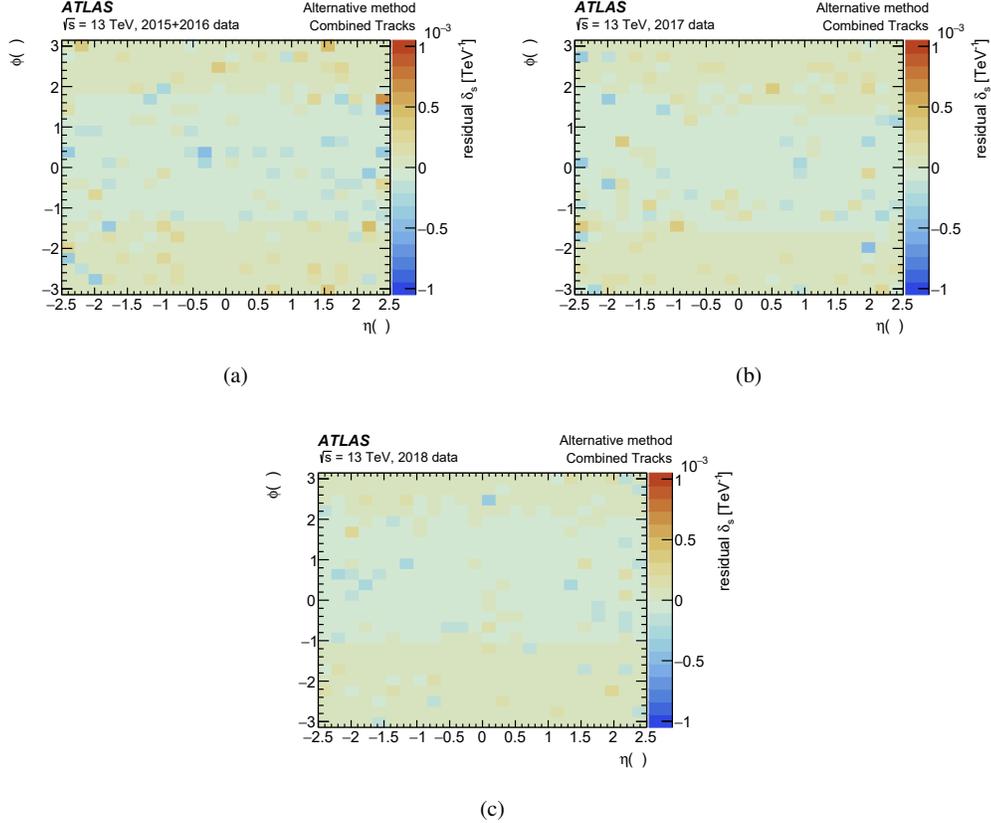

(a)

(b)

(c)

Figure 12: Residual charge-dependent biases on the muon p_T as evaluated on data for CB momenta after alignment and after application of the alternative correction procedure. The residual biases are shown separately for the three data-taking periods of (a) years 2015 and 2016, (b) 2017 and (c) 2018. The bias is defined in Eq. (1).

B Additional performance studies

A second step, used to cross-check the validity of the corrections obtained directly for CB tracks, is introduced in Section 5.2. Because of the different parameterisation, a one-to-one comparison of the results obtained from the ID+MS method, as shown in Tables 1 and 2, and that of the CB method, as shown in Table 3, is not possible. Therefore, the methods used to estimate the performance in Section 6 are extended to the $p_T^{\mathrm{Cor,ID+MS}}$ momenta as obtained from the ID+MS correction procedure. Using the same resonant decays from $J/\psi \rightarrow \mu\mu$, $\Upsilon \rightarrow \mu\mu$, and $Z \rightarrow \mu\mu$ events in simulation, the dimuon momentum mass scale and the dimuon mass resolution are compared with the measurements obtained from data. The results are obtained with the same fitting methods detailed in Section 6. For the $J/\psi \rightarrow \mu\mu$ decays the results are

presented in Figure 13, and they show a good agreement between data and simulation for both the mass scale and the mass resolution. Similarly, Figure 14 and Figure 15 show the results obtained for $\Upsilon \to \mu\mu$ and $Z \to \mu\mu$ decays, respectively. As expected from the systematic uncertainty study detailed in Section 5.2, the ID+MS methodology results in a reduced smearing uncertainty to a higher scale uncertainty than the results of the CB methodology. All results show an agreement within quoted uncertainties of the measured values between data and simulation.
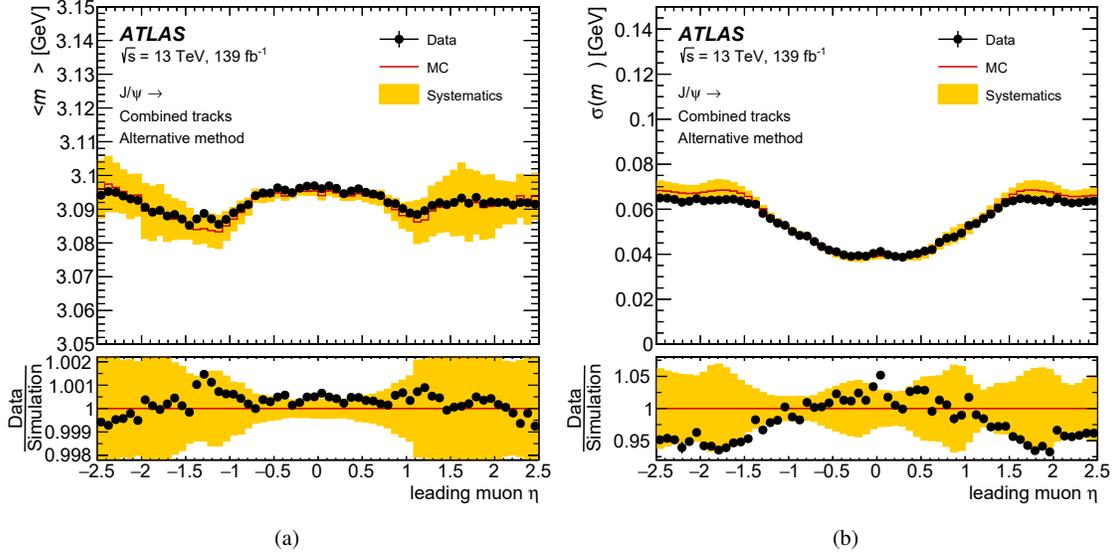
Figure 13: (a) Fitted mean mass of the dimuon system for CB muons for $J/\psi \to \mu\mu$ events data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. (b) Dimuon invariant mass resolution for CB muons for $J/\psi \to \mu\mu$ events for data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. The upper panels show the results obtained on data and simulation after all corrections using the alternative ID+MS correction method for CB tracks. The lower panels show the ratios of the data to the MC simulation. The error bars represent the statistical uncertainty; the shaded bands represent the systematic uncertainty in the correction and the systematic uncertainty in the extraction method added in quadrature.
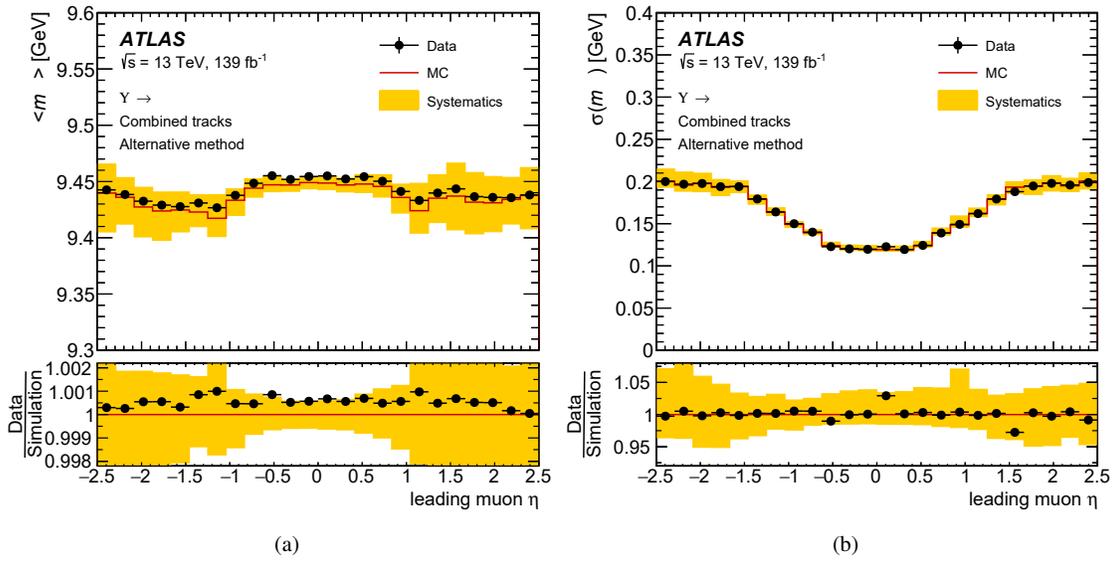
Figure 14: (a) Fitted mean mass of the dimuon system for CB muons for $\Upsilon \rightarrow \mu\mu$ events data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. (b) Dimuon invariant mass resolution for CB muons for $\Upsilon \rightarrow \mu\mu$ events for data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. The upper panels show the results obtained on data and simulation after all corrections using the alternative ID+MS correction method for CB tracks. The lower panels show ratios of the data to the MC simulation. The error bars represent the statistical uncertainty; the shaded bands represent the systematic uncertainty in the correction and the systematic uncertainty in the extraction method added in quadrature.
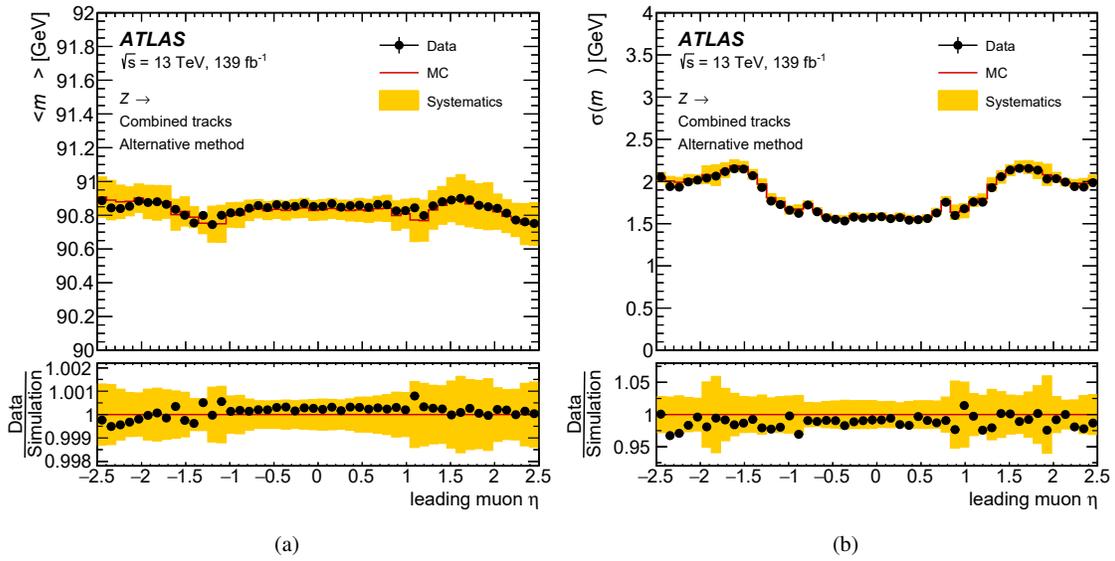
(a)

(b)

Figure 15: (a) Fitted mean mass of the dimuon system for CB muons for $Z \to \mu\mu$ events data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. (b) Dimuon invariant mass resolution for CB muons for $Z \to \mu\mu$ events for data and corrected simulation as a function of the pseudorapidity of the highest-p_T muon. The upper panels show the results obtained on data and simulation after all corrections using the alternative ID+MS correction method for CB tracks. The lower panels show the ratios of the data to the MC simulations. The error bars represent the statistical uncertainty; the shaded bands represent the systematic uncertainty in the correction and the systematic uncertainty in the extraction method added in quadrature.

References

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

[2] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[3] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs
boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1,
arXiv: 1207.7214 [hep-ex].

[4] CMS Collaboration,
Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,
Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex].

[5] ATLAS Collaboration,
A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery,
Nature **607** (2022) 52, arXiv: 2207.00092 [hep-ex].

[6] ATLAS Collaboration,
Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector,
Eur. Phys. J. C **78** (2018) 110, arXiv: 1701.07240 [hep-ex],
Erratum: Eur. Phys. J. C **78** (2018) 898.

[7] ATLAS Collaboration, *Search for high-mass dilepton resonances using $139\,fb^{-1}$ of pp collision
data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **796** (2019) 68,
arXiv: 1903.06248 [hep-ex].

[8] ATLAS Collaboration, *Search for a heavy charged boson in events with a charged lepton and
missing transverse momentum from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
Phys. Rev. D **100** (2019) 052013, arXiv: 1906.05609 [hep-ex].

[9] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full
Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **81** (2021) 578,
arXiv: 2012.00578 [hep-ex].

[10] 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].

[11] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*,
ATLAS-TDR-19; CERN-LHCC-2010-013, 2010,
URL: https://cds.cern.ch/record/1291633, Addendum: ATLAS-TDR-19-ADD-1;
CERN-LHCC-2012-009, 2012, URL: https://cds.cern.ch/record/1451888.

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

[13] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*,
Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[14] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*,
ATL-SOFT-PUB-2021-001, 2021, URL: https://cds.cern.ch/record/2767187.

[15] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*,
JINST **15** (2020) P09015, arXiv: 2004.13447 [hep-ex].

[16] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004,
arXiv: 2007.12539 [hep-ex].

[17] ATLAS Collaboration,
      Resolution of the ATLAS muon spectrometer monitored drift tubes in LHC Run 2,
      JINST **14** (2019) P09011, arXiv: 1906.12226 [hep-ex].

[18] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*,
      Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

[19] J. Pumplin et al.,
      New Generation of Parton Distributions with Uncertainties from Global QCD Analysis,
      JHEP **07** (2002) 012, arXiv: hep-ph/0201195.

[20] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014,
      URL: https://cds.cern.ch/record/1966419.

[21] P. Golonka and Z. Was,
      PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays,
      Eur. Phys. J. C **45** (2006) 97, arXiv: hep-ph/0506026.

[22] N. Davidson, T. Przedzinski and Z. Was,
      PHOTOS Interface in C++: Technical and physics documentation,
      Comput. Phys. Commun. **199** (2016) 86, arXiv: 1011.0937 [hep-ph].

[23] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*,
      JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[24] S. Frixione, P. Nason and C. Oleari,
      Matching NLO QCD computations with parton shower simulations: the POWHEG method,
      JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[25] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043,
      arXiv: 1002.2581 [hep-ph].

[26] S. Alioli, P. Nason, C. Oleari and E. Re,
      NLO vector-boson production matched with shower in POWHEG, JHEP **07** (2008) 060,
      arXiv: 0805.4802 [hep-ph].

[27] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014) 145,
      arXiv: 1406.3660 [hep-ex].

[28] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024,
      arXiv: 1007.2241 [hep-ph].

[29] D. J. Lange, *The EvtGen particle decay simulation package*,
      Nucl. Instrum. Meth. A **462** (2001) 152.

[30] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159,
      arXiv: 1410.3012 [hep-ph].

[31] S. Frixione, G. Ridolfi and P. Nason,
      A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction,
      JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph].

[32] R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040,
      arXiv: 1410.8849 [hep-ph].

[33] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*,
ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168.

[34] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244,
arXiv: 1207.1303 [hep-ph].

[35] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034,
arXiv: 1905.09127 [hep-ph].

[36] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039,
arXiv: 0808.3674 [hep-ph].

[37] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019) 866, arXiv: 1907.13071 [hep-ph].

[38] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*,
Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph].

[39] A. Denner, S. Dittmaier and L. Hofer,
COLLIER: A fortran-based complex one-loop library in extended regularizations,
Comput. Phys. Commun. **212** (2017) 220, arXiv: 1604.06792 [hep-ph].

[40] S. Schumann and F. Krauss,
A parton shower algorithm based on Catani–Seymour dipole factorisation, JHEP **03** (2008) 038,
arXiv: 0709.1027 [hep-ph].

[41] S. Höche, F. Krauss, M. Schönherr and F. Siegert,
A critical appraisal of NLO+PS matching methods, JHEP **09** (2012) 049,
arXiv: 1111.1220 [hep-ph].

[42] S. Höche, F. Krauss, M. Schönherr and F. Siegert,
QCD matrix elements + parton showers. The NLO case, JHEP **04** (2013) 027,
arXiv: 1207.5030 [hep-ph].

[43] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*,
JHEP **11** (2001) 063, arXiv: hep-ph/0109231.

[44] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*,
JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph].

[45] C. Anastasiou, L. Dixon, K. Melnikov and F. Petriello, *High-precision QCD at hadron colliders:
Electroweak gauge boson rapidity distributions at next-to-next-to leading order*,
Phys. Rev. D **69** (2004) 094008, arXiv: hep-ph/0312266.

[46] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic
measurements incorporating the Donnachie–Landshoff diffractive model*,
ATL-PHYS-PUB-2016-017, 2016, URL: https://cds.cern.ch/record/2206965.

[47] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 – a simulation toolkit*,
Nucl. Instrum. Meth. A **506** (2003) 250.

[48] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823,
arXiv: 1005.4568 [physics.ins-det].

[49] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector in Run-2*,
Eur. Phys. J. C **80** (2020) 1194, arXiv: 2007.07624 [hep-ex].

[50] ATLAS Collaboration, *Performance of the muon spectrometer alignment in 2017 and 2018 data*,
ATL-MUON-PUB-2021-002, 2021, URL: https://cds.cern.ch/record/2753329.

[51] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det].

[52] M. J. Oreglia, *A study of reactions $\psi\prime \rightarrow \gamma\gamma\psi$*, PhD thesis, Univ. Stanford, 1980. SLAC-R-0236, Appendix D.

[53] M. Baak, S. Gadatsch, R. Harrington and W. Verkerke, *Interpolation between multi-dimensional histograms using a new non-linear moment morphing method*, Nucl. Instrum. Meth. A **771** (2015) 39, arXiv: 1410.7388 [physics.data-an].