Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector

The ATLAS Collaboration

DOI
10.1140/epjc/s10052-017-5475-4

Publication date
2018

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date: 29 Mar 2024
Measurement of the $W$-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector

ATLAS Collaboration

CERN, 1211 Geneva 23, Switzerland

Received: 26 January 2017 / Accepted: 18 December 2017 / Published online: 6 February 2018

© CERN for the benefit of the ATLAS collaboration 2018. This article is an open access publication

Abstract A measurement of the mass of the $W$ boson is presented based on proton–proton collision data recorded in 2011 at a centre-of-mass energy of 7 TeV with the ATLAS detector at the LHC, and corresponding to 4.6 fb$^{-1}$ of integrated luminosity. The selected data sample consists of $7.8 \times 10^6$ candidates in the $W \rightarrow \mu \nu$ channel and $5.9 \times 10^6$ candidates in the $W \rightarrow e\nu$ channel. The $W$-boson mass is obtained from template fits to the reconstructed distributions of the charged lepton transverse momentum and of the $W$ boson transverse mass in the electron and muon decay channels, yielding

$$m_W = 80370 \pm 7 \text{ (stat.)} \pm 11\text{ (exp. syst.)} \pm 14 \text{ (mod. syst.)} \text{ MeV}$$

$$= 80370 \pm 19 \text{ MeV},$$

where the first uncertainty is statistical, the second corresponds to the experimental systematic uncertainty, and the third to the physics-modelling systematic uncertainty. A measurement of the mass difference between the $W^+$ and $W^-$ bosons yields $m_{W^+} - m_{W^-} = -29 \pm 28$ MeV.

1 Introduction

The Standard Model (SM) of particle physics describes the electroweak interactions as being mediated by the $W$ boson, the $Z$ boson, and the photon, in a gauge theory based on the $SU(2)_L \times U(1)_Y$ symmetry [1–3]. The theory incorporates the observed masses of the $W$ and $Z$ bosons through a symmetry-breaking mechanism. In the SM, this mechanism relies on the interaction of the gauge bosons with a scalar doublet field and implies the existence of an additional physical state known as the Higgs boson [4–7]. The existence of the $W$ and $Z$ bosons was first established at the CERN SPS in 1983 [8–11], and the LHC collaborations ATLAS and CMS reported the discovery of the Higgs boson in 2012 [12,13].

At lowest order in the electroweak theory, the $W$-boson mass, $m_W$, can be expressed solely as a function of the $Z$-boson mass, $m_Z$, the fine-structure constant, $\alpha$, and the Fermi constant, $G_F$. Higher-order corrections introduce an additional dependence of the $W$-boson mass on the gauge couplings and the masses of the heavy particles of the SM. The mass of the $W$ boson can be expressed in terms of the other SM parameters as follows:

$$m_W^2 \left(1 - \frac{m_W^2}{m_Z^2}\right) = \frac{\pi \alpha}{\sqrt{2}G_F}(1 + \Delta r),$$

where $\Delta r$ incorporates the effect of higher-order corrections [14,15]. In the SM, $\Delta r$ is in particular sensitive to the top-quark and Higgs-boson masses; in extended theories, $\Delta r$ receives contributions from additional particles and interactions. These effects can be probed by comparing the measured and predicted values of $m_W$. In the context of global fits to the SM parameters, constraints on physics beyond the SM are currently limited by the $W$-boson mass measurement precision [16]. Improving the precision of the measurement of $m_W$ is therefore of high importance for testing the overall consistency of the SM.

Previous measurements of the mass of the $W$ boson were performed at the CERN SPS proton–antiproton ($p \bar{p}$) collider with the UA1 and UA2 experiments [17,18] at centre-of-mass energies of $\sqrt{s} = 546$ GeV and $\sqrt{s} = 630$ GeV, at the Tevatron $p \bar{p}$ collider with the CDF and D0 detectors at $\sqrt{s} = 1.8$ TeV [19–21] and $\sqrt{s} = 1.96$ TeV [22–24], and at the LEP electron–positron collider by the ALEPH, DELPHI, L3, and OPAL collaborations at $\sqrt{s} = 161$–209 GeV [25–28]. The current Particle Data Group world average value of $m_W = 80385 \pm 15$ MeV [29] is dominated by the CDF and D0 measurements performed at $\sqrt{s} = 1.96$ TeV. Given the precisely measured values of $\alpha$, $G_F$, and $m_Z$, and taking recent top-quark and Higgs-boson mass measurements, the SM prediction of $m_W$ is $m_W = 80358 \pm 8$ MeV in Ref. [16] and $m_W = 80362 \pm 8$ MeV in Ref. [30]. The SM prediction uncertainty of 8 MeV represents a target for the precision of future measurements of $m_W$. 

* e-mail: atlas.publications@cern.ch
At hadron colliders, the $W$-boson mass can be determined in Drell–Yan production \cite{31} from $W \to \ell\nu$ decays, where $\ell$ is an electron or muon. The mass of the $W$ boson is extracted from the Jacobian edges of the final-state kinematic distributions, measured in the plane perpendicular to the beam direction. Sensitive observables include the transverse momenta of the charged lepton and neutrino and the $W$-boson transverse mass.

The ATLAS and CMS experiments benefit from large signal and calibration samples. The numbers of selected $W$- and $Z$-boson events, collected in a sample corresponding to approximately $4.6 \text{ fb}^{-1}$ of integrated luminosity at a centre-of-mass energy of 7 TeV, are of the order of $10^7$ for the $W \to \ell\nu$, and of the order of $10^6$ for the $Z \to \ell\ell$ processes. The available data sample is therefore larger by an order of magnitude compared to the corresponding samples measured at the Tevatron. Other important aspects of the $W$- and $Z$-boson mass are the theoretical predictions, measured in the plane perpendicular to the beam direction. 

2 Measurement overview

This section provides the definition of the observables used in the analysis, an overview of the measurement strategy for the determination of the mass of the $W$ boson, and a description of the methodology used to estimate the systematic uncertainties.

2.1 Observable definitions

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 upward. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane. The $z$-axis is defined as the polar angle $\theta$ as $\theta = -\ln \tan(\phi/2)$.

The kinematic properties of charged leptons from $W$- and $Z$-boson decays are characterised by the measured transverse momentum, $p_T^\ell$, pseudorapidity, $\eta^\ell$, and azimuth, $\phi^\ell$. The mass of the lepton, $m^\ell$, completes the four-vector. For $Z$-boson events, the invariant mass, $m^\ell\ell$, the rapidity, $y^\ell\ell$, and the transverse momentum, $p_T^\ell\ell$, are obtained by combining the four-momenta of the decay-lepton pair.

The recoil in the transverse plane, $\vec{u}_T$, is reconstructed from the vector sum of the transverse energy of all clusters reconstructed in the calorimeters (Sect. 3), excluding energy deposits associated with the decay leptons. It is defined as:

$$\vec{u}_T = \sum_i \vec{E}_{T,i},$$

where $\vec{E}_{T,i}$ is the vector of the transverse energy of cluster $i$. The transverse-energy vector of a cluster has magnitude $E_T = E / \cosh \eta$, with the energy deposit of the cluster $E$ and its pseudorapidity $\eta$. The azimuth $\phi$ of the transverse-energy vector is defined from the coordinates of the cluster in the transverse plane. In $W$- and $Z$-boson events, $-\vec{u}_T$ provides an estimate of the boson transverse momentum. The related quantities $u_x$ and $u_y$ are the projections of the recoil onto the axes of the transverse plane in the ATLAS coordinate system. In $Z$-boson events, $u_x^\ell\ell$ and $u_y^\ell\ell$ represent the projections of the recoil onto the axes parallel and perpendicular to the $Z$-boson transverse momentum reconstructed from the decay-lepton pair. Whereas $u_x^\ell\ell$ can be compared to $-p_T^\ell\ell$ and probes the detector response to the recoil in terms of linearity and resolution, the $u_y^\ell\ell$ distribution satisfies $|u_y^\ell\ell| = 0$ and its width provides an estimate of the recoil resolution. In $W$-boson events, $u_x^\ell$ and $u_y^\ell$ are the projections of the recoil onto the
axes parallel and perpendicular to the reconstructed charged-lepton transverse momentum.

The resolution of the recoil is affected by additional event properties, namely the per-event number of $pp$ interactions per bunch crossing (pile-up) $\mu$, the average number of $pp$ interactions per bunch crossing $\langle \mu \rangle$, the total reconstructed transverse energy, defined as the scalar sum of the transverse energy of all calorimeter clusters, $\Sigma E_T \equiv \sum_i E_{T,i}$, and the quantity $\Sigma E_T^u \equiv \Sigma E_T - | u_T |$. The latter is less correlated with the recoil than $\Sigma E_T$, and better represents the event activity related to the pile-up and to the underlying event.

The magnitude and direction of the transverse-momentum vector of the decay neutrino, $\vec{p}_T^\nu$, are inferred from the vector of the missing transverse momentum, $\vec{p}_T^{\text{miss}}$, which corresponds to the momentum imbalance in the transverse plane and is defined as:

$$\vec{p}_T^{\text{miss}} = -\left( \vec{p}_T^\ell + u_T \right).$$

The $W$-boson transverse mass, $m_T$, is derived from $p_T^{\text{miss}}$ and from the transverse momentum of the charged lepton as follows:

$$m_T = \sqrt{2p_T^\ell p_T^{\text{miss}} (1 - \cos \Delta \phi)},$$

where $\Delta \phi$ is the azimuthal opening angle between the charged lepton and the missing transverse momentum.

All vector-boson masses and widths are defined in the running-width scheme. Resonances are expressed by the relativistic Breit–Wigner mass distribution:

$$\frac{d\sigma}{dm} \propto \frac{m^2}{(m^2 - m_V^2)^2 + m^4 \Gamma_V^2 / m_V^2},$$

where $m$ is the invariant mass of the vector-boson decay products, and $m_V$ and $\Gamma_V$, with $V = W, Z$, are the vector-boson masses and widths, respectively. This scheme was introduced in Ref. [35], and is consistent with earlier measurements of the $W$- and $Z$-boson resonance parameters [24,32].

### 2.2 Analysis strategy

The mass of the $W$ boson is determined from fits to the transverse momentum of the charged lepton, $p_T^\ell$, and to the transverse mass of the $W$ boson, $m_T$. For $W$ bosons at rest, the transverse-momentum distributions of the $W$ decay lepton have a Jacobian edge at a value of $m/2$, whereas the distribution of the transverse mass has an endpoint at the value of $m$ [36], where $m$ is the invariant mass of the charged-lepton and neutrino system, which is related to $m_W$ through the Breit–Wigner distribution of Eq. (1).

The expected final-state distributions, referred to as templates, are simulated for several values of $m_W$ and include signal and background contributions. The templates are compared to the observed distribution by means of a $\chi^2$ compatibility test. The $\chi^2$ as a function of $m_W$ is interpolated, and the measured value is determined by analytical minimisation of the $\chi^2$ function. Predictions for different values of $m_W$ are obtained from a single simulated reference sample, by reweighting the $W$-boson invariant mass distribution according to the Breit–Wigner parameterisation of Eq. (1). The $W$-boson width is scaled accordingly, following the SM relation $\Gamma_W \propto m_W^3$.

Experimentally, the $p_T^\ell$ and $p_T^{\text{miss}}$ distributions are affected by the lepton energy calibration. The latter is also affected by the calibration of the recoil. The $p_T^\ell$ and $p_T^{\text{miss}}$ distributions are broadened by the $W$-boson transverse-momentum distribution, and are sensitive to the $W$-boson helicity states, which are influenced by the proton PDFs [37]. Compared to $p_T^\ell$, the $m_T$ distribution has larger uncertainties due to the recoil, but smaller sensitivity to such physics-modelling effects. Imperfect modelling of these effects can distort the template distributions, and constitutes a significant source of uncertainties for the determination of $m_W$.

The calibration procedures described in this paper rely mainly on methods and results published earlier by ATLAS [38–40], and based on $W$ and $Z$ samples at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The $Z \rightarrow \ell \ell$ event samples are used to calibrate the detector response. Lepton momentum corrections are derived exploiting the precisely measured value of the $Z$-boson mass, $m_Z$ [32], and the recoil response is calibrated using the expected momentum balance with $p_T^{\ell\ell}$. Identification and reconstruction efficiency corrections are determined from $W$- and $Z$-boson events using the tag-and-probe method [38,40]. The dependence of these corrections on $p_T^\ell$ is important for the measurement of $m_W$, as it affects the shape of the template distributions.

The detector response corrections and the physics modelling are verified in $Z$-boson events by performing measurements of the $Z$-boson mass with the same method used to determine the $W$-boson mass, and comparing the results to the LEP combined value of $m_Z$, which is used as input for the lepton calibration. The determination of $m_Z$ from the lepton-pair invariant mass provides a first closure test of the lepton energy calibration. In addition, the extraction of $m_Z$ from the $p_T^{\ell\ell}$ distribution tests the $p_T^{\ell\ell}$-dependence of the efficiency corrections, and the modelling of the $Z$-boson transverse-momentum distribution and of the relative fractions of $Z$-boson helicity states. The $p_T^{\text{miss}}$ and $m_T$ variables are defined in $Z$-boson events by treating one of the reconstructed decay leptons as a neutrino. The extraction of $m_Z$ from the $m_T$ distribution provides a test of the recoil calibration. The combination of the extraction of $m_Z$ from the $m_{\ell\ell}$, $p_T^\ell$ and $m_T$ distributions provides a closure test of the measurement procedure. The precision of this validation procedure is limited by the finite size of the $Z$-boson sample, which is approximately ten times smaller than the $W$-boson sample.
The analysis of the Z-boson sample does not probe differences in the modelling of W- and Z-boson production processes. Whereas W-boson production at the Tevatron is charge symmetric and dominated by interactions with at least one valence quark, the sea-quark PDFs play a larger role at the LHC, and contributions from processes with heavy quarks in the initial state have to be modelled properly. The W+ -boson production rate exceeds that of W- bosons by about 40%, with a broader rapidity distribution and a softer transverse-momentum distribution. Uncertainties in the modelling of these distributions and in the relative fractions of the W-boson helicity states are constrained using measurements of W- and Z-boson production performed with the ATLAS experiment at √s = 7 TeV and √s = 8 TeV [41–45].

The final measured value of the W-boson mass is obtained from the combination of various measurements performed in the electron and muon decay channels, and in charge- and |ηℓ|-dependent categories, as defined in Table 1. The boundaries of the |ηℓ| categories are driven mainly by experimental and statistical constraints. The measurements of mW used in the combination are based on the observed distributions of pT and mT, which are only partially correlated. Measurements of mW based on the pTmiss distributions are performed as consistency tests, but they are not used in the combination due to their significantly lower precision. The consistency of the results in the electron and muon channels provide a further test of the experimental calibrations, whereas the consistency of the results for the different charge and |ηℓ| categories tests the W-boson production model.

Further consistency tests are performed by repeating the measurement in three intervals of (|μ|), in two intervals of uT and uTmiss, and by removing the pTmiss selection requirement, which is applied in the nominal signal selection. The consistency of the values of mW in these additional categories probes the modelling of the recoil response, and the modelling of the transverse-momentum spectrum of the W boson. Finally, the stability of the result with respect to the charged-lepton azimuth, and upon variations of the fitting ranges is verified.

Systematic uncertainties in the determination of mW are evaluated using pseudodata samples produced from the nominal simulated event samples by varying the parameters corresponding to each source of uncertainty in turn. The differences between the values of mW extracted from the pseudodata and nominal samples are used to estimate the uncertainty. When relevant, these variations are applied simultaneously in the W-boson signal samples and in the background contributions. The systematic uncertainties are estimated separately for each source and for fit ranges of 32 < pT < 45 GeV and 66 < mT < 99 GeV. These fit ranges minimise the total expected measurement uncertainty, and are used for the final result as discussed in Sect. 11.

In Sects. 6, 7, 8, and 10, which discuss the systematic uncertainties of the mW measurement, the uncertainties are also given for combinations of measurement categories. This provides information showing the reduction of the systematic uncertainty obtained from the measurement categorisation. For these cases, the combined uncertainties are evaluated including only the expected statistical uncertainty in addition to the systematic uncertainty being considered. However, the total measurement uncertainty is estimated by adding all uncertainty contributions in quadrature for each measurement category, and combining the results accounting for correlations across categories.

During the analysis, an unknown offset was added to the value of mW used to produce the templates. The offset was randomly selected from a uniform distribution in the range [−100, 100] MeV, and the same value was used for the W+ and W- templates. The offset was removed after the mW measurements performed in all categories were found to be compatible and the analysis procedure was finalised.

<table>
<thead>
<tr>
<th>Decay channel</th>
<th>W → eν</th>
<th>W → μν</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinematic distributions</td>
<td>pT, mT</td>
<td>pT, mT</td>
</tr>
<tr>
<td>Charge categories</td>
<td>W+, W−</td>
<td>W+, W−</td>
</tr>
<tr>
<td></td>
<td>[0, 0.6], [0.6, 1.2], [1.8, 2.4]</td>
<td>[0, 0.8], [0.8, 1.4], [1.4, 2.0], [2.0, 2.4]</td>
</tr>
</tbody>
</table>

3 The ATLAS detector

The ATLAS experiment [46] is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroid magnets.

The inner-detector system (ID) is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range |η| < 2.5. At small radii, a high-granularity silicon pixel detector covers the vertex region and typically provides three measurements per track. It is followed by the silicon microstrip tracker, which usually provides eight measurement points per track. These silicon detectors are complemented by a gas-filled straw-tube transition radiation tracker, which enables radially extended track reconstruction up to
|η| = 2.0. The transition radiation tracker also provides electron identification information based on the fraction of hits (typically 35 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range |η| < 4.9. Within the region |η| < 3.2, electromagnetic (EM) calorimetry is provided by high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler covering |η| < 1.8 to correct for upstream energy-loss fluctuations. The EM calorimeter is divided into a barrel section covering |η| < 1.475 and two endcap sections covering 1.375 < |η| < 3.2. For |η| < 2.5 it is divided into three layers in depth, which are finely segmented in η and φ. Hadronic calorimetry is provided by a steel/scintillator-tile calorimeter, segmented into three barrel structures within |η| < 1.7 and two copper/LAr hadronic endcap calorimeters covering 1.5 < |η| < 3.2. The solid-angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules in 3.1 < |η| < 4.9, optimised for electromagnetic and hadronic 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 superconducting air-core toroids. The precision chamber system covers the region |η| < 2.7 with three layers of monitored drift tubes, complemented by cathode strip chambers in the forward region. The muon trigger system covers the range |η| < 2.4 with resistive plate chambers in the barrel, and thin gap chambers in the endcap regions.

A three-level trigger system is used to select events for offline analysis [47]. The level-1 trigger is implemented in hardware and uses a subset of detector information to reduce the event rate to a design value of at most 75 kHz. This is followed by two software-based trigger levels which together reduce the event rate to about 300 Hz.

4 Data samples and event simulation

The data sample used in this analysis consists of W- and Z-boson candidate events, collected in 2011 with the ATLAS detector in proton–proton collisions at the LHC, at a centre-of-mass energy of √s = 7 TeV. The sample for the electron channel, with all relevant detector systems operational, corresponds to approximately 4.6 fb−1 of integrated luminosity. A smaller integrated luminosity of approximately 4.1 fb−1 is used in the muon channel, as part of the data was discarded due to a timing problem in the resistive plate chambers, which affected the muon trigger efficiency. The relative uncertainty of the integrated luminosity is 1.8% [48]. This data set provides approximately 1.4 × 10^5 reconstructed W-boson events and 1.8 × 10^6 Z-boson events, after all selection criteria have been applied.

The POWHEG MC generator [49–51] (v1/r1556) is used for the simulation of the hard-scattering processes of W- and Z-boson production and decay in the electron, muon, and tau channels, and is interfaced to PYTHIA 8 (v8.170) for the modelling of the parton shower, hadronisation, and underlying event [52,53], with parameters set according to the A2NLO tune [44]. The CT10 PDF set [54] is used for the hard-scattering processes, whereas the CTEQ6L1 PDF set [55] is used for the parton shower. In the Z-boson samples, the effect of virtual photon production (γ*) and Z/γ* interference is included. The effect of QED final-state radiation (FSR) is simulated with PHOTOS (v2.154) [56]. Tau lepton decays are handled by PYTHIA 8, taking into account polarisation effects. An alternative set of samples for W- and Z-boson production is generated with POWHEG interfaced to HERWIG (v6.520) for the modelling of the parton shower [57], and to JIMMY (v4.31) for the underlying event [58]. The W- and Z-boson masses are set to mw = 80.399 GeV and mz = 91.1875 GeV, respectively. During the analysis, the value of the W-boson mass in the W → ℓν and W → τν samples was blinded using the reweighting procedure described in Sect. 2.

Top-quark pair production and the single-top-quark processes are modelled using the MC@NLO MC generator (v4.01) [59–61], interfaced to HERWIG and JIMMY. Gauge-boson pair production (WW, WZ, ZZ) is simulated with HERWIG (v6.520). In all the samples, the CT10 PDF set is used. Samples of heavy-flavour multijet events (pp → b̄b + X and pp → c̄c + X) are simulated with PYTHIA 8 to validate the data-driven methods used to estimate backgrounds with non-prompt leptons in the final state.

 Whereas the extraction of mw is based on the shape of distributions, and is not sensitive to the overall normalisation of the predicted distributions, it is affected by theoretical uncertainties in the relative fractions of background and signal. The W- and Z-boson event yields are normalised according to their measured cross sections, and uncertainties of 1.8% and 2.3% are assigned to the W+ /Z and W- /Z production cross-section ratios, respectively [41]. The tt sample is normalised to its measured cross section [62] with an uncertainty of 3.9%, whereas the cross-section predictions for the single-top production processes of Refs. [63–65] are used for the normalisation of the corresponding sample, with an uncertainty of 7%. The samples of events with massive gauge-boson pair production are normalised to the NLO predictions calculated with MCFM [66], with an uncertainty of 10% to cover the differences to the NNLO predictions [67].

The response of the ATLAS detector is simulated using a program [68] based on GEANT 4 [69]. The ID and the MS were simulated assuming an ideal detector geometry; alignment corrections are applied to the data during event reconstruction. The description of the detector material incorporates the results of extensive studies of the electron and photon calibration [39]. The simulated hard-scattering process...
is overlaid with additional proton–proton interactions, simulated with PYTHIA 8 (v8.165) using the A2 tune [70]. The distribution of the average number of interactions per bunch crossing ($\langle \mu \rangle$) spans the range 2.5–16.0, with a mean value of approximately 9.0.

Simulation inaccuracies affecting the distributions of the signal, the response of the detector, and the underlying-event modelling, are corrected as described in the following sections. Physics-modelling corrections, such as those affecting the $W$-boson transverse-momentum distribution and the angular decay coefficients, are discussed in Sect. 6. Calibration and detector response corrections are presented in Sects. 7 and 8.

5 Particle reconstruction and event selection

This section describes the reconstruction and identification of electrons and muons, the reconstruction of the recoil, and the requirements used to select $W$- and $Z$-boson candidate events. The recoil provides an event-by-event estimate of the $W$-boson transverse momentum. The reconstructed kinematic properties of the leptons and of the recoil are used to infer the transverse momentum of the neutrino and the transverse-mass kinematic variables.

5.1 Reconstruction of electrons, muons and the recoil

Electron candidates are reconstructed from clusters of energy deposited in the electromagnetic calorimeter and associated with at least one track in the ID [38,39]. Quality requirements are applied to the associated tracks in order to reject poorly reconstructed charged-particle trajectories. The energy of the electron is reconstructed from the energy collected in calorimeter cells within an area of size $\Delta \eta \times \Delta \phi = 0.075 \times 0.175$ in the barrel, and $0.125 \times 0.125$ in the endcaps. A multivariate regression algorithm, developed and optimised on simulated events, is used to calibrate the energy reconstruction. The reconstructed electron energy is corrected to account for the energy deposited in front of the calorimeter and outside the cluster, as well as for variations of the energy response as a function of the impact point of the electron in the calorimeter. The energy calibration algorithm takes as inputs the energy collected by each calorimeter layer, including the presampler, the pseudorapidity of the cluster, and the local position of the shower within the cell of the second layer, which corresponds to the cluster centroid. The kinematic properties of the reconstructed electron are inferred from the energy measured in the EM calorimeter, and from the pseudorapidity and azimuth of the associated track. Electron candidates are required to have $p_T > 15$ GeV and $|\eta| < 2.4$ and to fulfil a set of tight identification requirements [38]. The pseudorapidity range $1.2 < |\eta| < 1.82$ is excluded from the measurement, as the amount of passive material in front of the calorimeter and its uncertainty are largest in this region [39], preventing a sufficiently accurate description of non-Gaussian tails in the electron energy response. Additional isolation requirements on the nearby activity in the ID and calorimeter are applied to improve the background rejection. These isolation requirements are implemented by requiring the scalar sum of the $p_T$ of tracks in a cone of size $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ around the electron, $p_T^{\text{cone}}$, and the transverse energy deposited in the calorimeter within a cone of size $\Delta R < 0.2$ around the electron, $E_T^{\text{cone}}$, to be small. The contribution from the electron candidate itself is excluded. The specific criteria are optimised as a function of electron $\eta$ and $p_T$ to have a combined efficiency of about 95% in the simulation for isolated electrons from the decay of a $W$ or $Z$ boson.

The muon reconstruction is performed independently in the ID and in the MS, and a combined muon candidate is formed from the combination of a MS track with an ID track, based on the statistical combination of the track parameters [40]. The kinematic properties of the reconstructed muon are defined using the ID track parameters alone, which allows a simpler calibration procedure. The loss of resolution is small (10–15%) in the transverse-momentum range relevant for the measurement of the $W$-boson mass. The ID tracks associated with the muons must satisfy quality requirements on the number of hits recorded by each subdetector [40]. In order to reject muons from cosmic rays, the longitudinal coordinate of the point of closest approach of the track to the beamline is required to be within 10 mm of the collision vertex. Muon candidates are required to have $p_T > 20$ GeV and $|\eta| < 2.4$. Similarly to the electrons, the rejection of multijet background is increased by applying an isolation requirement: the scalar sum of the $p_T$ of tracks in a cone of size $\Delta R < 0.2$ around the muon candidate, $p_T^{\text{cone}}$, is required to be less than 10% of the muon $p_T$.

The recoil, $\vec{u}_T$, is reconstructed from the vector sum of the transverse energy of all clusters measured in the calorimeters, as defined in Sect. 2.1. The ATLAS calorimeters measure energy depositions in the range $|\eta| < 4.9$ with a topological clustering algorithm [71], which starts from cells with an energy of at least four times the expected noise from electronics and pile-up. The momentum vector of each cluster is determined by the magnitude and coordinates of the energy deposition. Cluster energies are initially measured assuming that the energy deposition occurs only through electromagnetic interactions, and are then corrected for the different calorimeter responses to hadrons and electromagnetic particles, for losses due to dead material, and for energy which is not captured by the clustering process. The definition of $\vec{u}_T$ and the inferred quantities $m_T^{\text{miss}}$ and $m_T$ do not involve the explicit reconstruction of particle jets, to avoid possible threshold effects.

Springer
Clusters located a distance $\Delta R < 0.2$ from the reconstructed electron or muon candidates are not used for the reconstruction of $\vec{u}_T$. This ensures that energy deposits originating from the lepton itself or from accompanying photons (from FSR or Bremsstrahlung) do not contribute to the recoil measurement. The energy of any soft particles removed along with the lepton is compensated for using the total transverse energy measured in a cone of the same size $\Delta R = 0.2$, placed at the same absolute pseudorapidity as the lepton with randomly chosen sign, and at different $\phi$. The total transverse momentum measured in this cone is rotated to the position of the lepton and added to $\vec{u}_T$.

5.2 Event selection

The $W$-boson sample is collected during data-taking with triggers requiring at least one muon candidate with transverse momentum larger than 18 GeV or at least one electron candidate with transverse momentum larger than 20 GeV. The transverse-momentum requirement for the electron candidate was raised to 22 GeV in later data-taking periods to cope with the increased instantaneous luminosity delivered by the LHC. Selected events are required to have a reconstructed primary vertex with at least three associated tracks.

$W$-boson candidate events are selected by requiring exactly one reconstructed electron or muon with $p_T^f > 30$ GeV. The leptons are required to match the corresponding trigger object. In addition, the reconstructed recoil is required to be $u_T < 30$ GeV, the missing transverse momentum $p_T^{\text{miss}} > 30$ GeV and the transverse mass $m_T > 60$ GeV. These selection requirements are optimised to reduce the multijet background contribution, and to minimise model uncertainties from $W$ bosons produced at high transverse momentum. A total of $5.89 \times 10^6$ $W$-boson candidate events are selected in the $W \rightarrow e\nu$ channel, and $7.84 \times 10^6$ events in the $W \rightarrow \mu\nu$ channel.

As mentioned in Sect. 2, $Z$-boson events are extensively used to calibrate the response of the detector to electrons and muons, and to derive recoil corrections. In addition, $Z$-boson events are used to test several aspects of the modelling of vector-boson production. $Z$-boson candidate events are collected with the same trigger selection used for the $W$-boson sample. The analysis selection requires exactly two reconstructed leptons with $p_T^f > 25$ GeV, having the same flavour and opposite charges. The events are required to have an invariant mass of the dilepton system in the range $80 < m_{\ell\ell} < 100$ GeV. In both channels, selected leptons are required to be isolated in the same way as in the $W$-boson event selection. In total, $0.58 \times 10^6$ and $1.23 \times 10^6$ $Z$-boson candidate events are selected in the electron and muon decay channels, respectively.

6 Vector-boson production and decay

Samples of inclusive vector-boson production are produced using the POWHEG MC generator interfaced to PYTHIA 8, henceforth referred to as POWHEG+PYTHIA 8. The $W$- and $Z$-boson samples are reweighted to include the effects of higher-order QCD and electroweak (EW) corrections, as well as the results of fits to measured distributions which improve the agreement of the simulated lepton kinematic distributions with the data. The effect of virtual photon production and $Z/\gamma^*$ interference is included in both the predictions and the POWHEG+PYTHIA 8 simulated $Z$-boson samples. The reweighting procedure used to include the corrections in the simulated event samples is detailed in Sect. 6.4.

The correction procedure is based on the factorisation of the fully differential leptonic Drell–Yan cross section [31] into four terms:

$$
\frac{d\sigma}{dp_1 \cdot dp_2} = \left[ \frac{d\sigma(m)}{dm} \right] \left[ \frac{d\sigma(y)}{dy} \right] \left[ \frac{d\sigma(p_{T, y})}{dp_T \cdot dy} \right] \left[ \frac{d\sigma(y)}{dy} \right]^{-1} \times \left[ (1 + \cos^2 \theta) + \sum_{i=0}^{7} A_i(p_{T, y}) \cdot P_i(\cos \theta, \phi) \right],
$$

where $p_1$ and $p_2$ are the lepton and anti-lepton four-momenta; $m$, $p_T$, and $y$ are the invariant mass, transverse momentum, and rapidity of the dilepton system; $\theta$ and $\phi$ are the polar angle and azimuth of the lepton$^1$ in any given rest frame of the dilepton system; $A_i$ are numerical coefficients, and $P_i$ are spherical harmonics of order zero, one and two.

The differential cross section as a function of the invariant mass, $\sigma(m)/dm$, is modelled with a Breit–Wigner parameterisation according to Eq. (1). In the case of the $Z$-boson samples, the photon propagator is included using the running electromagnetic coupling constant; further electroweak corrections are discussed in Sect. 6.1. The differential cross section as a function of boson rapidity, $\sigma(y)/dy$, and the coefficients $A_i$ are modelled with perturbative QCD fixed-order predictions, as described in Sect. 6.2. The transverse-momentum spectrum at a given rapidity, $\sigma(p_{T, y})/(dp_T \cdot dy) \cdot (\sigma(y)/dy)^{-1}$, is modelled with predictions based on the PYTHIA 8 MC generator, as discussed in Sect. 6.3. An exhaustive review of available predictions for $W$- and $Z$-boson production at the LHC is given in Ref. [72].

Measurements of $W$- and $Z$-boson production are used to validate and constrain the modelling of the fully differential leptonic Drell–Yan cross section. The PDF central values and uncertainties, as well as the modelling of the differential cross section as a function of boson rapidity, are validated

$^1$ Here, lepton refers to the negatively charged lepton from a $W^-$ or $Z$ boson, and the neutrino from a $W^+$ boson.
by comparing to the 7 TeV W- and Z-boson rapidity measurements [41], based on the same data sample. The QCD parameters of the parton shower model were determined by fits to the transverse-momentum distribution of the Z boson measured at 7 TeV [44]. The modelling of the $A_1$ coefficients is validated by comparing the theoretical predictions to the 8 TeV measurement of the angular coefficients in Z-boson decays [42].

6.1 Electroweak corrections and uncertainties

The dominant source of electroweak corrections to W- and Z-boson production originates from QED final-state radiation, and is simulated with PHOTOS. The effect of QED initial-state radiation (ISR) is also included through the PYTHIA 8 parton shower. The uncertainty in the modelling of QED FSR is evaluated by comparing distributions obtained using the default leading-order photon emission matrix elements with predictions obtained using NLO matrix elements, as well as by comparing PHOTOS with an alternative implementation based on the Yennie–Frautschi–Suura formalism [73], which is available in WINHAC [74]. The differences are small in both cases, and the associated uncertainty is considered negligible.

Other sources of electroweak corrections are not included in the simulated event samples, and their full effects are considered as systematic uncertainties. They include the interference between ISR and FSR QED corrections (IFI), pure weak corrections due to virtual-loop and box diagrams, and final-state emission of lepton pairs. Complete $O(\alpha)$ electroweak corrections to the $pp \rightarrow W + X$, $W \rightarrow \ell\nu$ process were initially calculated in Refs. [75,76]. Combined QCD and EW corrections are however necessary to evaluate the effect of the latter in presence of a realistic $p_T^W$ distribution. Approximate $O(\alpha,\alpha)$ corrections including parton shower effects are available from WINHAC, SANC [77] and in the POWHEG framework [78–80]. A complete, fixed-order calculation of $O(\alpha,\alpha)$ corrections in the resonance region appeared in Ref. [81].

In the present work the effect of the NLO EW corrections are estimated using WINHAC, which employs the PYTHIA 6 MC generator for the simulation of QCD and QED ISR. The corresponding uncertainties are evaluated comparing the final state distributions obtained including QED FSR only with predictions using the complete NLO EW corrections in the $\alpha(0)$ and $G_{\mu}$ renormalisation schemes [82]. The latter predicts the larger correction and is used to assign the systematic uncertainty.

Final-state lepton pair production, through $\gamma^* \rightarrow \ell\ell$ radiation, is formally a higher-order correction but constitutes an significant additional source of energy loss for the W-boson decay products. This process is not included in the event simulation, and the impact on the determination of $m_W$ is evaluated using PHOTOS and SANC.

Table 2 summarises the effect of the uncertainties associated with the electroweak corrections on the $m_W$ measurements. All comparisons described above were performed at particle level. The impact is larger for the $p_T^{\ell\ell}$ distribution than for the $m_T$ distribution, and similar between the electron and muon decay channels. A detailed evaluation of these uncertainties was performed in Ref. [83] using POWHEG [78], and the results are in fair agreement with Table 2. The study of Ref. [83] also compares, at fixed order, the effect of the approximate $O(\alpha,\alpha)$ corrections with the full calculation of Ref. [81], and good agreement is found. The same sources of uncertainty affect the lepton momentum calibration through their impact on the $m_{\ell\ell}$ distribution in Z-boson events, as discussed in Sect. 7.

6.2 Rapidity distribution and angular coefficients

At leading order, W and Z bosons are produced with zero transverse momentum, and the angular distribution of the decay leptons depends solely on the polar angle of the lepton in the boson rest frame. Higher-order corrections give rise to sizeable boson transverse momentum, and to azimuthal asymmetries in the angular distribution of the decay leptons. The angular distribution of the W- and Z-boson decay leptons is determined by the relative fractions of helicity cross sections for the vector-boson production. The fully differential leptonic Drell–Yan cross section can be decomposed as a weighted sum of nine harmonic polynomials, with weights given by the helicity cross sections. The harmonic polynomials are estimated using PHOTOS and SANC. The full calculation of Ref. [81] also compares, at fixed order, the effect of the approximate $O(\alpha,\alpha)$ corrections with the full calculation of Ref. [81], and good agreement is found. The same sources of uncertainty affect the lepton momentum calibration through their impact on the $m_{\ell\ell}$ distribution in Z-boson events, as discussed in Sect. 7.

Table 2 Impact on the $m_W$ measurement of systematic uncertainties from higher-order electroweak corrections, for the $p_T^\ell$ and $m_T$ distributions in the electron and muon decay channels

<table>
<thead>
<tr>
<th>Decay channel</th>
<th>$W \rightarrow e\nu$</th>
<th>$W \rightarrow \mu\nu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinematic distribution</td>
<td>$p_T^\ell$</td>
<td>$m_T$</td>
</tr>
<tr>
<td>$\delta m_W$ [MeV]</td>
<td>$&lt;0.1$</td>
<td>$&lt;0.1$</td>
</tr>
<tr>
<td>FSR (real)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pure weak and IFI corrections</td>
<td>$3.3$</td>
<td>$2.5$</td>
</tr>
<tr>
<td>FSR (pair production)</td>
<td>$3.6$</td>
<td>$0.8$</td>
</tr>
<tr>
<td>Total</td>
<td>$4.9$</td>
<td>$2.6$</td>
</tr>
</tbody>
</table>
mials depend on the polar angle, \( \theta \), and the azimuth, \( \phi \), of the lepton in a given rest frame of the boson. The helicity cross sections depend, in their most general expression, on the transverse momentum, \( p_T \), rapidity, \( y \), and invariant mass, \( m \), of the boson. It is customary to factorise the unpolarised, or angular-integrated, cross section, \( d\sigma/(dp_T^2 \, dy \, dm) \), and express the decomposition in terms of dimensionless angular coefficients, \( A_i \), which represent the ratios of the helicity cross sections with respect to the unpolarised cross section [34], leading to the following expression for the fully differential Drell–Yan cross section:

\[
\frac{d\sigma}{dp_T^2 \, dy \, dm \, d\cos \theta \, d\phi} = \frac{3}{16\pi} \frac{d\sigma}{dp_T^2 \, dy \, dm} \times \left[ (1 + \cos^2 \theta) + A_0 \frac{1}{2} (1 - 3 \cos^2 \theta) + A_1 \sin 2\theta \cos \phi + A_2 \frac{1}{2} \sin^2 \theta \cos 2\phi + A_3 \sin \theta \cos \phi + A_4 \cos \theta + A_5 \sin^2 \theta \sin 2\phi + A_6 \sin 2\theta \sin \phi + A_7 \sin \theta \sin \phi \right].
\]

The angular coefficients depend in general on \( p_T, y \) and \( m \). The \( A_5-A_7 \) coefficients are non-zero only at order \( O(\alpha_s^2) \) and above. They are small in the \( p_T \) region relevant for the present analysis, and are not considered further. The angles \( \theta \) and \( \phi \) are defined in the Collins–Soper (CS) frame [84].

The differential cross section as a function of boson rapidity, \( d\sigma(y)/dy \), and the angular coefficients, \( A_i \), are modelled with fixed-order perturbative QCD predictions, at \( O(\alpha_s^2) \) in the perturbative expansion of the strong coupling constant and using the CT10nnlo PDF set [85]. The dependence of the angular coefficients on \( m \) is neglected; the effect of this approximation on the measurement of \( m_W \) is discussed in Sect. 6.4. For the calculation of the predictions, an optimised version of DYNLO [86] is used, which explicitly decomposes the calculation of the cross section into the different pieces of the \( q_T \)-subtraction formalism, and allows the computation of statistically correlated PDF variations. In this optimised version of DYNLO, the Cuba library [87] is used for the numerical integration.

The values of the angular coefficients predicted by the POWHEG+PYTHIA 8 samples differ significantly from the corresponding NNLO predictions. In particular, large differences are observed in the predictions of \( A_0 \) at low values of \( p_T^{W,Z} \). Other coefficients, such as \( A_1 \) and \( A_2 \), are affected by significant NNLO corrections at high \( p_T^{W,Z} \). In \( Z \)-boson production, \( A_3 \) and \( A_4 \) are sensitive to the vector couplings between the \( Z \) boson and the fermions, and are predicted assuming the measured value of the effective weak mixing angle \( \sin^2 \theta_{\text{eff}}^L \) [32].

6.3 Transverse-momentum distribution

Predictions of the vector-boson transverse-momentum spectrum cannot rely solely on fixed-order perturbative QCD. Most \( W \)-boson events used for the analysis have a low transverse-momentum value, in the kinematic region \( p_T^W < 30 \text{ GeV} \), where large logarithmic terms of the type \( \log (m_W/p_T^W) \) need to be resummed, and non-perturbative effects must be included, either with parton showers or with predictions based on analytic resummation [88–92]. The modelling of the transverse-momentum spectrum of vector bosons at a given rapidity, expressed by the term \( d\sigma(p_T, y)/(dp_T \, dy) \cdot (d\sigma(y)/dy)^{-1} \) in Eq. (2), is based on the PYTHIA 8 parton shower MC generator. The predictions of vector-boson production in the PYTHIA 8 MC generator employ leading-order matrix elements for the \( q\bar{q}' \to W, Z \) processes and include a reweighting of the first parton shower emission to the leading-order \( V+\text{jet} \) cross section [93]. The resulting prediction of the boson \( p_T \) spectrum is comparable in accuracy to those of an NLO plus parton shower generator setup such as POWHEG+PYTHIA 8, and of resummed predictions at next-to-leading logarithmic order [94].

The values of the QCD parameters used in PYTHIA 8 were determined from fits to the \( Z \)-boson transverse momentum distribution measured with the ATLAS detector at a centre-of-mass energy of \( \sqrt{s} = 7 \text{ TeV} \) [44]. Three QCD parameters were considered in the fit: the intrinsic transverse momentum of the incoming partons, the value of \( \alpha_s(m_Z) \) used for the QCD ISR, and the value of the ISR infrared cut-off. The resulting values of the PYTHIA 8 parameters constitute the \( AZ \) tune. The PYTHIA 8 AZ prediction was found to provide a satisfactory description of the \( p_T^Z \) distribution as a function of rapidity, contrarily to POWHEG+PYTHIA 8 AZNLO; hence the former is chosen to predict the \( p_T^Z \) distribution. The good consistency of the \( m_W \) measurement results in \( |\eta| \) categories, presented in Sect. 11, is also a consequence of this choice.

To illustrate the results of the parameters optimisation, the PYTHIA 8 AZ and 4C [95] predictions of the \( p_T^Z \) distribution are compared in Fig. 1a to the measurement used to determine the \( AZ \) tune. Kinematic requirements on the decay leptons are applied according to the experimental acceptance. For further validation, the predicted differential cross-section ratio,

\[
R_{W/Z}(p_T) = \left( \frac{1}{\sigma_W} \cdot \frac{d\sigma_W(p_T)}{dp_T} \right)^{-1} \left( \frac{1}{\sigma_Z} \cdot \frac{d\sigma_Z(p_T)}{dp_T} \right)^{-1},
\]

is compared to the corresponding ratio of ATLAS measurements of vector-boson transverse momentum [44,45]. The comparison is shown in Fig. 1b, where kinematic requirements on the decay leptons are applied according to the experimental acceptance. The measured \( Z \)-boson \( p_T \) distribution is
Fig. 1 a Normalised differential cross section as a function of $p_T^{\ell\ell}$ in Z-boson events [44] and b differential cross-section ratio $R_{W/Z}(p_T)$ as a function of the boson $p_T$ [44,45]. The measured cross sections are compared to the predictions of the PYTHIA 8 AZ tune and, in a, of the PYTHIA 8 4C tune. The shaded bands show the total experimental uncertainties.

Fig. 2 Ratios of the reconstruction-level a $p_T^{\ell\ell}$ and b $m_T$ normalised distributions obtained using POWHEG+PYTHIA 8 AZNLO, DYRes and POWHEG MiNLO+PYTHIA 8 to the baseline normalised distributions obtained using PYTHIA 8 AZ rebinned to match the coarser bins of the $W$-boson $p_T$ distribution, which was measured using only 30 pb$^{-1}$ of data. The theoretical prediction is in agreement with the experimental measurements for the region with $p_T < 30$ GeV, which is relevant for the measurement of the $W$-boson mass.

The predictions of RESBOS [89,90], DYRes [91] and POWHEG MiNLO+PYTHIA 8 [96,97] are also considered. All predict a harder $p_T^W$ distribution for a given $p_T^{\ell\ell}$ distribution, compared to PYTHIA 8 AZ. Assuming the latter can be adjusted to match the measurement of Ref. [44], the corresponding $p_T^W$ distribution induces a discrepancy with the detector-level $m_T$ and $u_T^{\ell\ell}$ distributions observed in the $W$-boson data, as discussed in Sect. 11.2. This behaviour is observed using default values for the non-perturbative parameters of these programs, but is not expected to change significantly under variations of these parameters. These predictions are therefore not used in the determination of $m_W$ or its uncertainty.

Figure 2 compares the reconstruction-level $p_T^{\ell\ell}$ and $m_T$ distributions obtained with POWHEG+PYTHIA 8 AZNLO, DYRes and POWHEG MiNLO+PYTHIA 8 to those of PYTHIA 8 AZ. The effect of varying the $p_T^W$ distribution is largest at high $p_T^{\ell\ell}$, which explains why the uncertainty due to the $p_T^W$ modelling is reduced when limiting the $p_T^{\ell\ell}$ fitting range as described in Sect. 11.3.

Reconstruction-level distributions are obtained from the POWHEG+PYTHIA 8 signal sample by reweighting the particle-level $p_T^W$ distribution according to the product of the $p_T^\ell\ell$ distribution in PYTHIA 8 AZ, and of $R_{W/Z}(p_T)$ as predicted by POWHEG+PYTHIA 8 AZNLO, DYRes and POWHEG MiNLO+PYTHIA 8.
6.4 Reweighting procedure

The $W$ and $Z$ production and decay model described above is applied to the POWHEG+PYTHIA 8 samples through an event-by-event reweighting. Equation (3) expresses the factorisation of the cross section into the three-dimensional boson production phase space, defined by the variables $m$, $p_T$, and $y$, and the two-dimensional boson decay phase space, defined by the variables $\theta$ and $\phi$. Accordingly, a prediction of the kinematic distributions of vector bosons and their decay products can be transformed into another prediction by applying separate reweighting of the three-dimensional boson production phase-space distributions, followed by a reweighting of the angular decay distributions.

The reweighting is performed in several steps. First, the inclusive rapidity distribution is reweighted according to the NNLO QCD predictions evaluated with DYNNLO. Then, at a given rapidity, the vector-boson transverse-momentum shape is reweighted to the PYTHIA 8 prediction with the AZ tune. This procedure provides the transverse-momentum distribution of vector bosons predicted by PYTHIA 8, preserving the rapidity distribution at NNLO. Finally, at given rapidity and transverse momentum, the angular variables are reweighted according to:

$$ w(\cos \theta, \phi, p_T, y) = \frac{1 + \cos^2 \theta + \sum_i A_i'(p_T, y) P_i(\cos \theta, \phi)}{1 + \cos^2 \theta + \sum_i A_i(p_T, y) P_i(\cos \theta, \phi)}, $$

where $A_i'$ are the angular coefficients evaluated at $O(\alpha_s^2)$, and $A_i$ are the angular coefficients of the POWHEG+PYTHIA 8 samples. This reweighting procedure neglects the small dependence of the two-dimensional ($p_T, y$) distribution and of the angular coefficients on the final state invariant mass. The procedure is used to include the corrections described in Sects. 6.2 and 6.3, as well as to estimate the impact of the QCD modelling uncertainties described in Sect. 6.5.

The validity of the reweighting procedure is tested at particle level by generating independent $W$-boson samples using the CT10nnlo and NNPDF3.0 [98] NNLO PDF sets, and the same value of $m_W$. The relevant kinematic distributions are calculated for both samples and used to reweight the CT10nnlo sample to the NNPDF3.0 one. The procedure described in Sect. 2.2 is then used to determine the value of $m_W$ by fitting the NNPDF3.0 sample using templates from the reweighted CT10nnlo sample. The fitted value agrees with the input value within $1.5 \pm 2.0$ MeV. The statistical precision of this test is used to assign the associated systematic uncertainty.

The resulting model is tested by comparing the predicted $Z$-boson differential cross section as a function of rapidity, the $W$-boson differential cross section as a function of lepton pseudorapidity, and the angular coefficients in $Z$-boson events, to the corresponding ATLAS measurements [41,42]. The comparison with the measured $W$ and $Z$ cross sections is shown in Fig. 3. Satisfactory agreement between the measurements and the theoretical predictions is observed. A $\chi^2$ compatibility test is performed for the three distributions simultaneously, including the correlations between the uncertainties. The compatibility test yields a $\chi^2$/dof value of 45/34. Other NNLO PDF sets such as NNPDF3.0, CT14 [99], MMHT2014 [100], and ABM12 [101] are in worse agreement with these distributions. Based on the quantitative comparisons performed in

![Fig. 3 a](https://example.com/fig1a.png)  ![Fig. 3 b](https://example.com/fig1b.png)

Fig. 3 a Differential $Z$-boson cross section as a function of boson rapidity, and b differential $W^+$ and $W^-$ cross sections as a function of charged decay-lepton pseudorapidity at $\sqrt{s} = 7$ TeV [41]. The measured cross sections are compared to the POWHEG+PYTHIA 8 predictions, corrected to NNLO using DYNNLO with the CT10nnlo PDF set. The error bars show the total experimental uncertainties, including luminosity uncertainty, and the bands show the PDF uncertainties of the predictions.
Ref. [41], only CT10nnlo, CT14 and MMHT2014 are considered further. The better agreement obtained with CT10nnlo can be ascribed to the weaker suppression of the strange quark density compared to the u- and d-quark sea densities in this PDF set.

The predictions of the angular coefficients in Z-boson events are compared to the ATLAS measurement at $\sqrt{s} = 8$ TeV [42]. Good agreement between the measurements and DYNNLO is observed for the relevant coefficients, except for $A_2$, where the measurement is significantly below the prediction. As an example, Fig. 4 shows the comparison for $A_0$ and $A_2$ as a function of $p_T^{\ell\ell}$. For $A_2$, an additional source of uncertainty in the theoretical prediction is considered to account for the observed disagreement with data, as discussed in Sect. 6.5.3.

6.5 Uncertainties in the QCD modelling

Several sources of uncertainty related to the perturbative and non-perturbative modelling of the strong interaction affect the dynamics of the vector-boson production and decay [33,102–104]. Their impact on the measurement of $m_W$ is assessed through variations of the model parameters of the predictions for the differential cross sections as functions of the boson rapidity, transverse-momentum spectrum at a given rapidity, and angular coefficients, which correspond to the second, third, and fourth terms of the decomposition of Eq. (2), respectively. The parameter variations used to estimate the uncertainties are propagated to the simulated event samples by means of the reweighting procedure described in Sect. 6.4. Table 3 shows an overview of the uncertainties due to the QCD modelling which are discussed below.

**Table 3** Systematic uncertainties in the $m_W$ measurement due to QCD modelling, for the different kinematic distributions and $W$-boson charges. Except for the case of PDFs, the same uncertainties apply to $W^+$ and $W^-$. The fixed-order PDF uncertainty given for the separate $W^+$ and $W^-$ final states corresponds to the quadrature sum of the CT10nnlo uncertainty variations; the charge-combined uncertainty also contains a 3.8 MeV contribution from comparing CT10nnlo to CT14 and MMHT2014.

<table>
<thead>
<tr>
<th>$W$-boson charge</th>
<th>Kinematic distribution</th>
<th>$W^+$</th>
<th>$W^-$</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p_T^{\ell\ell}$</td>
<td>$m_T$</td>
<td>$p_T^{\ell\ell}$</td>
<td>$m_T$</td>
</tr>
<tr>
<td>$\delta m_W$ [MeV]</td>
<td>Fixed-order PDF uncertainty</td>
<td>13.1</td>
<td>14.9</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td>AZ tune</td>
<td>3.0</td>
<td>3.4</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>Charm-quark mass</td>
<td>1.2</td>
<td>1.5</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Parton shower $\mu F$ with heavy-flavour decorrelation</td>
<td>5.0</td>
<td>6.9</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>Parton shower PDF uncertainty</td>
<td>3.6</td>
<td>4.0</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>Angular coefficients</td>
<td>5.8</td>
<td>5.3</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>15.9</td>
<td>18.1</td>
<td>14.8</td>
</tr>
</tbody>
</table>
6.5.1 Uncertainties in the fixed-order predictions

The imperfect knowledge of the PDFs affects the differential cross section as a function of boson rapidity, the angular coefficients, and the \( p_W^+ \) distribution. The PDF contribution to the prediction uncertainty is estimated with the CT10nnlo PDF set by using the Hessian method [105]. There are 25 error eigenvectors, and a pair of PDF variations associated with each eigenvector. Each pair corresponds to positive and negative 90% CL excursions along the corresponding eigenvector. Symmetric PDF uncertainties are defined as the mean value of the absolute positive and negative excursions corresponding to each pair of PDF variations. The overall uncertainty of the CT10nnlo PDF set is scaled to 68% CL by applying a multiplicative factor of 1/1.645.

The effect of PDF variations on the rapidity distributions and angular coefficients are evaluated with DYNNLO, while their impact on the W-boson \( p_T^\ell \) distribution is evaluated using Pythia 8 and by reweighting event-by-event the PDFs of the hard-scattering process, which are convolved with the LO matrix elements. Similarly to other uncertainties which affect the \( p_W^+ \) distribution (Sect. 6.5.2), only relative variations of the \( p_W^+ \) and \( p_T^\ell \) distributions induced by the PDFs are considered. The PDF variations are applied simultaneously to the boson rapidity, angular coefficients, and transverse-momentum distributions, and the overall PDF uncertainty is evaluated with the Hessian method as described above.

Uncertainties in the PDFs are the dominant source of physics-modelling uncertainty, contributing about 14 and 13 MeV when averaging \( p_T^\ell \) and \( m_T \) fits for \( W^+ \) and \( W^- \), respectively. The PDF uncertainties are very similar when using \( p_T^\ell \) or \( m_T \) for the measurement. They are strongly anti-correlated between positively and negatively charged W bosons, and the uncertainty is reduced to 7.4 MeV on average for \( p_T^\ell \) and \( m_T \) fits, when combining opposite-charge categories. The anti-correlation of the PDF uncertainties is due to the fact that the total light-quark sea PDF is well constrained by deep inelastic scattering data, whereas the \( u-, d-, \) and s-quark decomposition of the sea is less precisely known [106]. An increase in the \( \bar{u} \) PDF is at the expense of the \( d \) PDF, which produces opposite effects in the longitudinal polarisation of positively and negatively charged W bosons [37].

Other PDF sets are considered as alternative choices. The envelope of values of \( m_W \) extracted with the MMHT2014 and CT14 NNLO PDF sets is considered as an additional PDF uncertainty of 3.8 MeV, which is added in quadrature after combining the \( W^+ \) and \( W^- \) categories, leading to overall PDF uncertainties of 8.0 MeV and 8.7 MeV for \( p_T^\ell \) and \( m_T \) fits, respectively.

The effect of missing higher-order corrections on the NNLO predictions of the rapidity distributions of Z bosons, and the pseudorapidity distributions of the decay leptons of W bosons, is estimated by varying the renormalisation and factorisation scales by factors of 0.5 and 2.0 with respect to their nominal value \( \mu_R = \mu_F = m_T \) in the DYNNLO predictions. The corresponding relative uncertainty in the normalised distributions is of the order of 0.1–0.3%, and significantly smaller than the PDF uncertainties. These uncertainties are expected to have a negligible impact on the measurement of \( m_W \), and are not considered further.

The effect of the LHC beam-energy uncertainty of 0.65% [107] on the fixed-order predictions is studied. Relative variations of 0.65% around the nominal value of 3.5 TeV are considered, yielding variations of the inclusive \( W^+ \) and \( W^- \) cross sections of 0.6 and 0.5%, respectively. No significant dependence as a function of lepton pseudorapidity is observed in the kinematic region used for the measurement, and the dependence as a function of \( p_T^\ell \) and \( m_T \) is expected to be even smaller. This uncertainty is not considered further.

6.5.2 Uncertainties in the parton shower predictions

Several sources of uncertainty affect the Pythia 8 parton shower model used to predict the transverse momentum of the W boson. The values of the AZ tune parameters, determined by fits to the measurement of the Z-boson transverse momentum, are affected by the experimental uncertainty of the measurement. The corresponding uncertainties are propagated to the \( p_T^\ell \) predictions through variations of the orthogonal eigenvector components of the parameters error matrix [44]. The resulting uncertainty in \( m_W \) is 3.0 MeV for the \( p_T^\ell \) distribution, and 3.4 MeV for the \( m_T \) distribution. In the present analysis, the impact of \( p_T^\ell \) distribution uncertainties is in general smaller when using \( p_T^\ell \) than when using \( m_T \), as a result of the comparatively narrow range used for the \( p_T^\ell \) distribution fits.

Other uncertainties affecting predictions of the transverse-momentum spectrum of the W boson at a given rapidity, are propagated by considering relative variations of the \( p_W^+ \) and \( p_T^\ell \) distributions. The procedure is based on the assumption that model variations, when applied to \( p_T^\ell \), can be largely reabsorbed into new values of the AZ tune parameters fitted to the \( p_T^\ell \) data. Variations that cannot be reabsorbed by the fit are excluded, since they would lead to a significant disagreement of the prediction with the measurement of \( p_T^\ell \). The uncertainties due to model variations which are largely correlated between \( p_W^+ \) and \( p_T^\ell \) cancel in this procedure. In contrast, the procedure allows a correct estimation of the uncertainties due to model variations which are uncorrelated between \( p_W^+ \) and \( p_T^\ell \), and which represent the only relevant sources of theoretical uncertainties in the propagation of the QCD modelling from \( p_T^\ell \) to \( p_T^\ell \).

Uncertainties due to variations of parton shower parameters that are not fitted to the \( p_T^\ell \) measurement include variations of the masses of the charm and bottom quarks, and variations of the factorisation scale used for the QCD ISR.
The mass of the charm quark is varied in Pythia 8, conserva-
tively, by ± 0.5 GeV around its nominal value of 1.5 GeV. The
resulting uncertainty contributes 1.2 MeV for the $p_T$ fits,
and 1.5 MeV for the $m_T$ fits. The mass of the bottom quark is
varied in Pythia 8, conservatively, by ± 0.8 GeV around
its nominal value of 4.8 GeV. The resulting variations have a
negligible impact on the transverse-momentum distributions of
Z and W bosons, and are not considered further.

The uncertainty due to higher-order QCD corrections to
the parton shower is estimated through variations of the fac-
torisation scale, $\mu_F$, in the QCD ISR by factors of 0.5 and 2.0
with respect to the central choice $\mu_F = p_{T,0}^2 + p_T^2$, where $p_{T,0}$
is an infrared cut-off, and $p_T$ is the evolution variable of the
parton shower [108]. Variations of the renormalisation scale in
the QCD ISR are equivalent to a redefinition of $\alpha_s(m_Z)$
used for the QCD ISR, which is fixed from the fits to the $p_T^Z$
data. As a consequence, variations of the ISR renormalisation
scale do not apply when estimating the uncertainty in the
predicted $p_T^W$ distribution.

Higher-order QCD corrections are expected to be largely
correlated between W-boson and Z-boson production induced
by the light quarks, $u$, $d$, and $s$, in the initial state. How-
ever, a certain degree of decorrelation between $W$- and $Z$
-boson transverse-momentum distributions is expected, due
to the different amounts of heavy-quark-initiated production,
where heavy refers to charm and bottom flavours. The physi-
cal origin of this decorrelation can be ascribed to the presence
of independent QCD scales corresponding to the three-to-
four flavours and four-to-five flavours matching scales $\mu_c$
and $\mu_b$ in the variable-flavour-number scheme PDF evolu-
tion [109], which are of the order of the charm- and bottom-
quark masses, respectively. To assess this effect, the vari-
ations of $\mu_F$ in the QCD ISR are performed simultaneously
for all light-quark $q\bar{q} \rightarrow W, Z$ processes, with $q = u, d, s,$
but independently for each of the $c\bar{c} \rightarrow Z, b\bar{b} \rightarrow Z,$
and $c\bar{q} \rightarrow W$ processes, with $q = d, s$. The effect of
the $c\bar{q} \rightarrow W$ variations on the determination of $m_W$
is reduced by a factor of two, to account for the presence of
only one heavy-flavour quark in the initial state. The result-
ing uncertainty in $m_W$ is 5.0 MeV for the $p_T^W$ distribution,
and 6.9 MeV for the $m_T$ distribution. Since the $\mu_F$
variations affect all the branchings of the shower evolution and
not only vertices involving heavy quarks, this procedure is
expected to yield a sufficient estimate of the $\mu_{c,b}$-induced
decorrelation between the $W$- and $Z$-boson $p_T$ distributions.
Treating the $\mu_F$ variations as correlated between all quark
flavours, but uncorrelated between $W$- and $Z$-boson produc-
tion, would yield a systematic uncertainty in $m_W$ of approx-
imately 30 MeV.

The predictions of the Pythia 8 MC generator include a
reweighting of the first parton shower emission to the leading-
order $W$+jet cross section, and do not include matching cor-
tections to the higher-order $W$+jet cross section. As discussed
in Sect. 11.2, predictions matched to the NLO $W$+jet cross
section, such as Powheg MiNLO+Pythia 8 and DYres, are
in disagreement with the observed $u_T^Z$ distribution and cannot
be used to provide a reliable estimate of the associated uncer-
tainty. The $u_T^Z$ distribution, on the other hand, validates the
Pythia 8 AZ prediction and its uncertainty, which gives con-
fidence that missing higher-order corrections to the $W$-boson
$p_T$ distribution are small in comparison to the uncertainties
that are already included, and can be neglected at the present
level of precision.

The sum in quadrature of the experimental uncertainties of
the AZ tune parameters, the variations of the mass of the
charm quark, and the factorisation scale variations, leads to
uncertainties on $m_W$ of 6.0 and 7.8 MeV when using the
$p_T^Z$ distribution and the $m_T$ distribution, respectively. These
sources of uncertainty are taken as fully correlated between
the electron and muon channels, the positively and negatively
charged W-boson production, and the $|p_T|$ bins.

The Pythia 8 parton shower simulation employs the
CTEQ6L1 leading-order PDF set. An additional independent
source of PDF-induced uncertainty in the $p_T^W$ distribution is
estimated by comparing several choices of the leading-
order PDF used in the parton shower, corresponding to the
CT14lo, MMHT2014lo and NNPDF2.3lo [110] PDF sets. The
PDFs which give the largest deviation from the nominal
ratio of the $p_T^W$ and $p_T^Z$ distributions are used to estimate the
uncertainty. This procedure yields an uncertainty of about
4 MeV for $W^+$, and of about 2.5 MeV for $W^-$. Similarly to
the case of fixed-order PDF uncertainties, there is a strong
anti-correlation between positively and negatively charged
W bosons, and the uncertainty is reduced to about 1.5 MeV
when combining positive- and negative-charge categories.

The prediction of the $p_T^W$ distribution relies on the $p_T$
ordered parton shower model of the Pythia 8 MC generator.
In order to assess the impact of the choice of parton shower
model on the determination of $m_W$, the Pythia 8 prediction
of the ratio of the $p_T^W$ and $p_T^Z$ distributions is compared to
the corresponding prediction of the Herwig 7 MC genera-
tor [111,112], which implements an angular-ordered parton
shower model. Differences between the Pythia 8 and Her-
wig 7 predictions are smaller than the uncertainties in the
Pythia 8 prediction, and no additional uncertainty is con-
idered.

6.5.3 Uncertainties in the angular coefficients

The full set of angular coefficients can only be measured pre-
cisely for the production of Z bosons. The accuracy of the
NNLO predictions of the angular coefficients is validated by
comparison to the $Z$-boson measurement, and extrapolated to
$W$-boson production assuming that NNLO predictions have
similar accuracy for the $W$- and $Z$-boson processes. The
ATLAS measurement of the angular coefficients in Z-boson
production at a centre-of-mass energy of $\sqrt{s} = 8\text{ TeV}$ [42] is used for this validation. The $O(\alpha_s^2)$ predictions, evaluated with DYNNLO, are in agreement with the measurements of the angular coefficients within the experimental uncertainties, except for the measurement of $A_2$ as a function of Z-boson $p_T$. Two sources of uncertainty affecting the modelling of the angular coefficients are considered, and propagated to the $W$-boson predictions. One source is defined from the experimental uncertainty of the $Z$-boson measurement of the angular coefficients which is used to validate the NNLO predictions. The uncertainty in the corresponding $W$-boson predictions is estimated by propagating the experimental uncertainty of the $Z$-boson measurement as follows. A set of pseudodata distributions are obtained by fluctuating the angular coefficients within the experimental uncertainties, preserving the correlations between the different measurement bins for the different coefficients. For each pseudoexperiment, the differences in the $A_i$ coefficients between fluctuated and nominal $Z$-boson measurement results are propagated to the corresponding coefficient in $W$-boson production. The corresponding uncertainty is defined from the standard deviation of the $m_W$ values as estimated from the pseudodata distributions.

The other source of uncertainty is considered to account for the disagreement between the measurement and the NNLO QCD predictions observed for the $A_2$ angular coefficient as a function of the $Z$-boson $p_T$ (Fig. 4). The corresponding uncertainty in $m_W$ is estimated by propagating the difference in $A_2$ between the $Z$-boson measurement and the theoretical prediction to the corresponding coefficient in $W$-boson production. The corresponding uncertainty in the measurement of $m_W$ is 1.6 MeV for the extraction from the $p_T^1$ distribution. Including this contribution, total uncertainties of 5.8 and 5.3 MeV due to the modelling of the angular coefficients are estimated in the determination of the $W$-boson mass from the $p_T^1$ and $m_T$ distributions, respectively. The uncertainty is dominated by the experimental uncertainty of the $Z$-boson measurement used to validate the theoretical predictions.

7 Calibration of electrons and muons

Any imperfect calibration of the detector response to electrons and muons impacts the measurement of the $W$-boson mass, as it affects the position and shape of the Jacobian edges reflecting the value of $m_W$. In addition, the $p_T^1$ and $m_T$ distributions are broadened by the electron-energy and muon-momentum resolutions. Finally, the lepton-selection efficiencies depend on the lepton pseudorapidity and transverse momentum, further modifying these distributions. Corrections to the detector response are derived from the data, and presented below. In most cases, the corrections are applied to the simulation, with the exception of the muon sagitta bias corrections and electron energy response corrections, which are applied to the data. Backgrounds to the selected $Z \rightarrow \ell\ell$ samples are taken into account using the same procedures as discussed in Sect. 9. Since the $Z$ samples are used separately for momentum calibration and efficiency measurements, as well as for the recoil response corrections discussed in Sect. 8, correlations among the corresponding uncertainties can appear. These correlations were investigated and found to be negligible.

7.1 Muon momentum calibration

As described in Sect. 5.1, the kinematic parameters of selected muons are determined from the associated inner-detector tracks. The accuracy of the momentum measurement is limited by imperfect knowledge of the detector alignment and resolution, of the magnetic field, and of the amount of passive material in the detector.

Biases in the reconstructed muon track momenta are classified as radial or sagitta biases. The former originate from detector movements along the particle trajectory and can be corrected by an $\eta$-dependent, charge-independent momentum-scale correction. The latter typically originate from curl distortions or linear twists of the detector around the $z$-axis [113], and can be corrected with $\eta$-dependent correction factors proportional to $q \times p_T^1$, where $q$ is the charge of the muon. The momentum scale and resolution corrections are applied to the simulation, while the sagitta bias correction is applied to the data:

$$p_T^{\text{MC,corr}} = p_T^{\text{MC}} \times [1 + \alpha(\eta, \phi)] \times \left[1 + \beta_{\text{curv}}(\eta) \cdot G(0, 1) \cdot p_T^{\text{MC}}\right],$$

$$p_T^{\text{data,corr}} = \frac{p_T^{\text{data}}}{1 + q \cdot \delta(\eta, \phi) \cdot p_T^{\text{data}}},$$

where $p_T^{\text{data,MC}}$ is the uncorrected muon transverse momentum in data and simulation, $G(0, 1)$ are normally distributed random variables with mean zero and unit width, and $\alpha, \beta_{\text{curv}},$ and $\delta$ represent the momentum scale, intrinsic resolution and sagitta bias corrections, respectively. Multiple-scattering contributions to the resolution are relevant at low $p_T$, and the corresponding corrections are neglected.

Momentum scale and resolution corrections are derived using $Z \rightarrow \mu\mu$ decays, following the method described in Ref. [40]. Template histograms of the dimuon invariant mass are constructed from the simulated event samples, including momentum scale and resolution corrections in narrow steps within a range covering the expected uncertainty. The optimal values of $\alpha$ and $\beta_{\text{curv}}$ are determined by means of a $\chi^2$ minimisation, comparing data and simulation in the
range of twice the standard deviation on each side of the mean value of the invariant mass distribution. In the first step, the corrections are derived by averaging over $\phi$, and for 24 pseudorapidity bins in the range $-2.4 < \eta_L < 2.4$. In the second iteration, $\phi$-dependent correction factors are evaluated in coarser bins of $\eta_L$. The typical size of $\alpha$ varies from $-0.0005$ to $-0.0015$ depending on $\eta_L$, while $\beta_{\text{curv}}$ values increase from 0.2 TeV$^{-1}$ in the barrel to 0.6 TeV$^{-1}$ in the high $\eta_L$ region. Before the correction, the $\phi$-dependence has an amplitude at the level of 0.1%.

The $\alpha$ and $\beta_{\text{curv}}$ corrections are sensitive to the following aspects of the calibration procedure, which are considered for the systematic uncertainty: the choice of the fitting range, methodological biases, background contributions, theoretical modelling of $Z$-boson production, non-linearity of the corrections, and material distribution in the ID. The uncertainty due to the choice of fitting range is estimated by varying the range by $\pm 10\%$, and repeating the procedure. The uncertainty due to the fit methodology is estimated by comparing the template fit results with an alternative approach, based on an iterative $\chi^2$ minimisation. Background contributions from gauge-boson pair and top-quark pair production are estimated using the simulation. The uncertainty in these background contributions is evaluated by varying their normalisation within the theoretical uncertainties on the production cross sections. The uncertainty in the theoretical modelling of $Z$-boson production is evaluated by propagating the effect of electroweak corrections to QED FSR, QED radiation, and other NLO electroweak corrections described in Sect. 6.1. The experimental uncertainty in the value of the $Z$-boson mass used as input is also accounted for. These sources of uncertainty are summed in quadrature, yielding an uncertainty $\delta \alpha$ in the muon momentum scale correction of approximately $0.5 \times 10^{-4}$; these sources are considered fully correlated across muon pseudorapidity.

The systematic uncertainty in the muon momentum scale due to the extrapolation from the $Z \rightarrow \mu\mu$ momentum range to the $W \rightarrow \mu\nu$ momentum range is estimated by evaluating momentum-scale corrections as a function of $1/p_T$ for muons in various $|\eta|$ ranges. The extrapolation uncertainty $\delta \alpha$ is parameterised as follows:

$$\delta \alpha = p_0 + \frac{p_1}{\langle p_T(W) \rangle},$$

where $\langle p_T(W) \rangle$ is the average $p_T$ of muons in $W$-boson events, and $p_0$ and $p_1$ are free parameters. If the momentum-scale corrections are independent of $1/p_T$, the fitting parameters are expected to be $p_0 = 1$ and $p_1 = 0$. Deviations of $p_1$ from zero indicate a possible momentum dependence. The fitted values of $\delta \alpha$ are shown in Fig. 5a, and are consistent with one, within two standard deviations of the statistical error. The corresponding systematic uncertainty in $m_W$ is defined assuming, in each bin of $|\eta|$, a momentum non-linearity given by the larger of the fitted value of $p_1$ and its uncertainty. This source of uncertainty is considered uncorrelated across muon pseudorapidity given that $p_1$ is dominated by statistical fluctuations. The effect of the imperfect knowledge of the material in the ID is studied using simulated event samples including an increase of the ID material by 10%, according to the uncertainty estimated in Ref. [114]. The impact of this variation is found to be negligible in comparison with the uncertainties discussed above.

Two methods are used for the determination of the sagitta bias $\delta$. The first method exploits $Z \rightarrow \mu\mu$ events. Muons are categorised according to their charge and pseudorapidity, and for each of these categories, the position of the peak in the dimuon invariant mass distribution is determined for data and simulation. The procedure allows the determination of the charge dependence of the momentum scale for $p_T$ values of approximately 42 GeV, which corresponds to the average transverse momentum of muons from $Z$-boson decays. The second method exploits identified electrons in a sample of $W \rightarrow e\nu$ decays. It is based on the ratio of the measured electron energy deposited in the calorimeter, $E$, to the electron momentum, $p$, measured in the ID. A clean sample of $W \rightarrow e\nu$ events with tightly identified electrons [38] is selected. Assuming that the response of the electromagnetic calorimeter is independent of the charge of the incoming particle, charge-dependent ID track momentum biases are extracted from the average differences in $E/p$ for electrons and positrons [113]. This method benefits from a larger event sample compared to the first method, and allows the determination of charge-dependent corrections for $p_T$ values of approximately 38 GeV, which corresponds to the average transverse momentum of muons in $W$-boson decays. The sagitta bias correction factors are derived using both methods separately in 40 $\eta$ bins and 40 $\phi$ bins. The results are found to agree within uncertainties and are combined, as illustrated in Fig. 5b. The combined correction uncertainty is dominated by the finite size of the event samples.

Figure 6 shows the dimuon invariant mass distribution of $Z \rightarrow \mu\mu$ decays in data and simulation, after applying all corrections. Table 4 summarises the effect of the muon momentum scale and resolution uncertainties on the determination of $m_W$. The dominant systematic uncertainty in the momentum scale is due to the extrapolation of the correction from the $Z$-boson momentum range to the $W$-boson momentum range. The extrapolation uncertainty $\delta \alpha$ is $(2-5) \times 10^{-5}$ for $|\eta_L| < 2.0$, and $(4-7) \times 10^{-4}$ for $|\eta_L| > 2.0$. Systematic uncertainties from other sources are relatively small. The systematic uncertainty of the resolution corrections is dominated by the statistical uncertainty of the $Z$-boson event sample, and includes a contribution from the imperfect closure of the method. The latter is defined from the residual difference between the standard deviations of the dimuon invariant mass in data and simulation, after applying resolution corrections.
7.2 Muon selection efficiency

The selection of muon candidates in $W \rightarrow \mu \nu$ and $Z \rightarrow \mu \mu$ events requires an isolated track reconstructed in the inner detector and in the muon spectrometer. In addition, the events are required to pass the muon trigger selection. Differences in the efficiency of the reconstruction and selection requirements between data and simulation can introduce a systematic shift in the measurement of the $W$-boson mass, and have to be corrected. In particular, the extraction of $m_W$ is sensitive to the dependence of the trigger, reconstruction and isolation efficiencies on the muon $p_T$ and on the projection of the recoil on the lepton transverse momentum, $u_\ell^T$. For muons with $p_T$ larger than approximately 15 GeV the detector simulation predicts constant efficiency as a function of $p_T^\mu$, both for the muon trigger selection and the track reconstruction. In contrast, the efficiency of the isolation requirement is expected to vary as a function of $p_T^\ell$ and $u_\ell^T$. The efficiency corrections also affect the muon selection inefficiency, and hence the estimation of the $Z \rightarrow \mu \mu$ background, which contributes to the $W \rightarrow \mu \nu$ selection when one of the decay muons fails the muon reconstruction or kinematic selection requirements.

Corrections to the muon reconstruction, trigger and isolation efficiencies are estimated by applying the tag-and-probe method [40] to $Z \rightarrow \mu \mu$ events in data and simulation. Efficiency corrections are defined as the ratio of efficiencies evaluated in data to efficiencies evaluated in simulated events. The corrections are evaluated as functions of two variables, $p_T^\mu$ and $u_\ell^T$, and in various regions of the detector. The detector is segmented into regions corresponding to the $\eta$ and $\phi$ coverage of the muon spectrometer. The subdivision accounts for the geometrical characteristics of the detector, such as the presence of uninstrumented or transition regions. The dependence of the efficiencies on $u_\ell^T$ agree in data and simulation. Therefore, the muon efficiency corrections are evaluated only as a function of $p_T^\mu$ and $\eta_\ell$, separately for positive and negative muon charges. The final efficiency correction factors are linearly interpolated as a function of muon $p_T$. No significant
Table 4 Systematic uncertainties in the $m_{WW}$ measurement from muon calibration and efficiency corrections, for the different kinematic distributions and $|\eta_\ell|$ categories, averaged over lepton charge. The momentum-scale uncertainties include the effects of both the momentum scale and linearity corrections. Combined uncertainties are evaluated as described in Sect. 2.2.

| $|\eta_\ell|$ range | Kinematic distribution | $p_T^{\ell}$ | $m_T$ | $p_T^{\ell}$ | $m_T$ | $p_T^{\ell}$ | $m_T$ | $p_T^{\ell}$ | $m_T$ | Combined $p_T^{\ell}$ | $m_T$ |
|-----------------|------------------------|-------------|-------|-------------|-------|-------------|-------|-------------|-------|---------------------|-------|
|                 |                        | $[0.0, 0.8]$ |       | $[0.8, 1.4]$ |       | $[1.4, 2.0]$ |       | $[2.0, 2.4]$ |       |                     |       |
| $\delta m_{WW}$ [MeV] | Momentum scale | 8.9 | 9.3 | 14.2 | 15.6 | 27.4 | 29.2 | 111.0 | 115.4 | 8.4 | 8.8 |
|                  | Momentum resolution | 1.8 | 2.0 | 1.9 | 1.7 | 1.5 | 2.2 | 3.4 | 3.8 | 1.0 | 1.2 |
|                  | Sagitta bias | 0.7 | 0.8 | 1.7 | 1.7 | 3.1 | 3.1 | 4.5 | 4.3 | 0.6 | 0.6 |
|                  | Reconstruction and isolation efficiencies | 4.0 | 3.6 | 5.1 | 3.7 | 4.7 | 3.5 | 6.4 | 5.5 | 2.7 | 2.2 |
|                  | Trigger efficiency | 5.6 | 5.0 | 7.1 | 5.0 | 11.8 | 9.1 | 12.1 | 9.9 | 4.1 | 3.2 |
|                  | Total | 11.4 | 11.4 | 16.9 | 17.0 | 30.4 | 31.0 | 112.0 | 116.1 | 9.8 | 9.7 |

Fig. 7 a Scale factors for the muon reconstruction, trigger and isolation efficiency obtained with the tag and probe method as a function of the muon $p_T$. Scale factors for the trigger efficiency are averaged over two data-taking periods as explained in the text. The error bars on the points show statistical uncertainties only. b Distribution of the reconstructed muons $\eta$ in $Z \rightarrow \mu\mu$ events. The data are compared to the simulation including signal and background contributions. Corrections for momentum scale and resolution, and for reconstruction, isolation, and trigger efficiencies are applied to the muons in the simulated events. Background events contribute less than 0.2% of the observed distribution. The lower panel shows the data-to-prediction ratio, with the error bars showing the statistical uncertainty.

$p_T$-dependence of the corrections is observed in any of the detector regions.

The selection of tag-and-probe pairs from $Z \rightarrow \mu\mu$ events is based on the kinematic requirements described in Sect. 5.2. The tag muon is required to be a combined and energy-isolated muon candidate (see Sect. 5.1) which fulfils the muon trigger requirements. The selection requirements applied to the probe muon candidate differ for each efficiency determination: the selection requirement for which the efficiency is determined is removed from the set of requirements applied to the probe muon. All the efficiency corrections are derived inclusively for the full data set, with the exception of the trigger, for which they are derived separately for two different data-taking periods. The resulting scale factors are shown as a function of $p_T^{\ell}$ and averaged over $\eta_\ell$ in Fig. 7a. The trigger and isolation efficiency corrections are typically below 0.3%, while the reconstruction efficiency correction is on average about 1.1%. The corresponding impact on muon selection inefficiency reaches up to about 20%.

The quality of the efficiency corrections is evaluated by applying the corrections to the $Z \rightarrow \mu\mu$ simulated sample, and comparing the simulated kinematic distributions to the corresponding distributions in data. Figure 7b illustrates this procedure for the $\eta_\ell$ distribution. Further distributions are shown in Sect. 9.

The dominant source of uncertainty in the determination of the muon efficiency corrections is the statistical uncertainty of the $Z$-boson data sample. The largest sources of systematic uncertainty are the multijet background contribution and the momentum-scale uncertainty. The correspond-
ing uncertainty in the measurement of $m_W$ is approximately 5 MeV. The ID tracking efficiencies for muon candidates are above 99.5% without any significant $p_T$ dependence, and the associated uncertainties are not considered further. An overview of the uncertainties associated with the muon efficiency corrections is shown in Table 4.

### 7.3 Electron energy response

The electron-energy corrections and uncertainties are largely based on the ATLAS Run 1 electron and photon calibration results [39]. The correction procedure starts with the intercalibration of the first and second layers of the EM calorimeter for minimum-ionising particles, using the energy deposits of muons in $Z \rightarrow \mu\mu$ decays. After the intercalibration of the calorimeter layers, the longitudinal shower-energy profiles of electrons and photons are used to determine the presampler energy scale and probe the passive material in front of the EM calorimeter, leading to an improved description of the detector material distribution and providing estimates of the residual passive material uncertainty. Finally, a dependence of the cell-level energy measurement on the read-out gain is observed in the second layer and corrected for. After these preliminary corrections, an overall energy-scale correction is determined as a function of $\eta_\ell$ from $Z \rightarrow ee$ decays, by comparing the reconstructed mass distributions in data and simulation. Simultaneously, an effective constant term for the calorimeter energy resolution is extracted by adjusting the width of the reconstructed dielectron invariant mass distribution in simulation to match the distribution in data.

Uncertainties in the energy-response corrections arise from the limited size of the $Z \rightarrow ee$ sample, from the physics modelling of the resonance and from the calibration algorithm itself. Physics-modelling uncertainties include uncertainties from missing higher-order electroweak corrections (dominated by the absence of lepton-pair emissions in the simulation) and from the experimental uncertainty in $m_Z$; these effects are taken fully correlated with the muon channel. Background contributions are small and the associated uncertainty is considered to be negligible. Uncertainties related to the calibration procedure are estimated by varying the invariant mass range used for the calibration, and with a closure test. For the closure test, a pseudodata sample of $Z \rightarrow ee$ events is obtained from the nominal sample by rescaling the electron energies by known $\eta$-dependent factors; the calibration algorithm is then applied, and the measured energy corrections are compared with the input rescaling factors.

These sources of uncertainty constitute a subset of those listed in Ref. [39], where additional variations were considered in order to generalise the applicability of the $Z$-boson calibration results to electrons and photons spanning a wide energy range. The effect of these uncertainties is averaged within the different $\eta_\ell$ categories. The overall relative energy-scale uncertainty, averaged over $\eta_\ell$, is $9.4 \times 10^{-5}$ for electrons from $Z$-boson decays.

In addition to the uncertainties in the energy-scale corrections arising from the $Z$-boson calibration procedure, possible differences in the energy response between electrons from $Z$-boson and $W$-boson decays constitute a significant source of uncertainty. The linearity of the response is affected by uncertainties in the intercalibration of the layers and in the passive material and calorimeter read-out corrections mentioned above. Additional uncertainties are assigned to cover imperfect electronics pedestal subtraction affecting the energy measurement in the cells of the calorimeter, and to the modelling of the interactions between the electrons and the detector material in GEANT4. The contribution from these sources to the relative energy-scale uncertainty is $(3-12) \times 10^{-5}$ in each $\eta$ bin, and $5.4 \times 10^{-5}$ when averaged over the full $\eta$ range after taking into account the correlation between the $\eta$ bins.

Azimuthal variations of the electron-energy response are expected from gravity-induced mechanical deformations of the EM calorimeter, and are observed especially in the endcaps, as illustrated in Fig. 8. As the $Z$-boson calibration averages over $\phi_\ell$ and the azimuthal distributions of the selected electrons differ in the two processes, a small residual effect from this modulation is expected when applying the calibration results to the $W \rightarrow ev$ sample. Related effects are discussed in Sect. 8. A dedicated correction is derived using the azimuthal dependence of the mean of the electron energy/momentum ratio, $\langle E/p \rangle$, after correcting $p$ for the momentum scale and curvature bias discussed in Sect. 7.1. The effect of this correction is a relative change of the average energy response of $3.8 \times 10^{-5}$ in $W$-boson events, with negligible uncertainty.

The $E/p$ distribution is also used to test the modelling of non-Gaussian tails in the energy response. An excess of events is observed in data at low values of $E/p$, and interpreted as the result of the mismodelling of the lateral development of EM showers in the calorimeter. Its impact is evaluated by removing the electrons with $E/p$ values in the region where the discrepancy is observed. The effect of this removal is compatible for electrons from $W$- and $Z$-boson decays within $4.9 \times 10^{-5}$, which corresponds to the statistical uncertainty of the test and is considered as an additional systematic uncertainty.

The result of the complete calibration procedure is illustrated in Fig. 9, which shows the comparison of the dielectron invariant mass distribution for $Z \rightarrow ee$ events in data and simulation. The impact of the electron-energy calibration uncertainties on the $m_W$ measurement is summarised in Table 5.
uncertainty the data-to-prediction ratio, with the error bars showing the statistical less than 0.2% of the observed distribution. The lower panel shows scale corrections are applied to the data. Background events contribute is isolation and trigger efficiencies are applied to the simulation; energy-resolution, and for reconstruction, identification, mass, the reconstruction and identification efficiency corrections have a typical uncertainty of 0.1–0.2% in the barrel, and 

\[ \eta < 0.5 \text{ MeV in the regions of } |\eta| < 1.2 \text{ and } 8 < |\eta| < 2.4. \]

\[ \eta < 0.5 \text{ MeV in the regions of } |\eta| < 1.2 \text{ and } 8 < |\eta| < 2.4. \]

7.4 Electron selection efficiency

Electron efficiency corrections are determined using samples of \( W \rightarrow e\nu \), \( Z \rightarrow ee \), and \( J/\psi \rightarrow ee \) events, and measured separately for electron reconstruction, identification and trigger efficiencies [38], as a function of electron \( \eta \) and \( p_T \). In the \( p_T \) range relevant for the measurement of the \( W \)-boson mass, the reconstruction and identification efficiency corrections have a typical uncertainty of 0.1–0.2% in the barrel, and 0.3% in the endcap. The trigger efficiency corrections have an uncertainty smaller than 0.1%, and are weakly dependent on \( p_T \).

For a data-taking period corresponding to approximately 20% of the integrated luminosity, the LAr calorimeter suffered from six front-end board failures. During this period, electrons could not be reconstructed in the region of \( 0 < \eta < 1.475 \) and \( -0.9 < \phi < -0.5 \). The data-taking conditions are reflected in the simulation for the corresponding fraction of events. However, the trigger acceptance loss is not perfectly simulated, and dedicated efficiency corrections are derived as a function of \( \eta \) and \( \phi \) to correct the mismodelling, and applied in addition to the initial corrections.

As described in Sect. 5, isolation requirements are applied to the identified electrons. Their efficiency is approximately 95% in the simulated event samples, and energy-isolation efficiency corrections are derived as for the reconstruction, identification, and trigger efficiencies. The energy-isolation efficiency corrections deviate from unity by less than 0.5%, with an uncertainty smaller than 0.2% on average.

Finally, as positively and negatively charged \( W \)-boson events have different final-state distributions, the \( W^+ \) contamination in the \( W^- \) sample, and vice versa, constitutes an additional source of uncertainty. The rate of electron charge mismeasurement in simulated events rises from about 0.2% in the barrel to 4% in the endcap. Estimates of charge mismeasurement in data confirm these predictions within better than 0.1%, apart from the high \( |\eta| \) region where differences up to 1% are observed. The electron charge mismeasurement induces a systematic uncertainty in \( m_W \) of approximately 0.5 MeV in the regions of \( |\eta| < 0.6 \) and \( 0.6 < |\eta| < 1.2 \), and of 5 MeV in the region of \( 1.8 < |\eta| < 2.4 \), separately for \( W^+ \) and \( W^- \). Since the \( W^+ \) and \( W^- \) samples contaminate each other, the effect
Table 5 Systematic uncertainties in the \( m_W \) measurement due to electron energy calibration, efficiency corrections and charge misassignment, for the different kinematic distributions and \( |\eta_\ell| \) regions, averaged over lepton charge. Combined uncertainties are evaluated as described in Sect. 2.2.

| \( |\eta_\ell| \) range | Kinematic distribution | \( \delta m_W \) [MeV] |
|----------------|------------------------|-----------------------|
|                  | \( p_T^\ell \) | \( m_T \) | \( p_T^\ell \) | \( m_T \) | \( p_T^\ell \) | \( m_T \) | \( p_T^\ell \) | \( m_T \) | Combined |
| [0.0, 0.6]       | 10.4 | 10.3 | 10.8 | 10.1 | 16.1 | 17.1 | 8.1 | 8.0 |
| [0.6, 1.2]       | 5.0 | 6.0 | 7.3 | 6.7 | 10.4 | 15.5 | 3.5 | 5.5 |
| [1.8, 2.4]       | 2.2 | 4.2 | 5.8 | 8.9 | 8.6 | 10.6 | 3.4 | 5.5 |
|                  | 2.3 | 3.3 | 2.3 | 3.3 | 2.3 | 3.3 | 2.3 | 3.3 |
|                  | 10.5 | 8.8 | 9.9 | 7.8 | 14.5 | 11.0 | 7.2 | 6.0 |
|                  | 10.4 | 7.7 | 11.7 | 8.8 | 16.7 | 12.1 | 7.3 | 5.6 |
|                  | 0.2 | 0.5 | 0.3 | 0.5 | 2.0 | 2.2 | 0.8 | 0.9 |
|                  | 0.2 | 0.2 | 0.2 | 0.2 | 1.5 | 1.5 | 0.1 | 0.1 |
| Total            | 19.0 | 17.5 | 21.1 | 19.4 | 30.7 | 30.5 | 14.2 | 14.3 |

![Fig. 10](image)

**Fig. 10** Distribution of reconstructed electrons \( \eta \) in \( Z \rightarrow e^+ e^- \) events. The data are compared to the simulation including signal and background contributions. Corrections for energy resolution, and for reconstruction, identification, isolation and trigger efficiencies are applied to the simulation; energy-scale corrections are applied to the data. Background events contribute less than 0.2% of the observed distribution. The lower panel shows the data-to-prediction ratio, with the error bars showing the statistical uncertainty in the signal.

is anti-correlated for the \( m_W \) measurements in the two different charge categories, and cancels in its combination, up to the asymmetry in the \( W^+ / W^- \) production rate. After combination, the residual uncertainty in \( m_W \) is 0.2 MeV for \( |\eta_\ell| < 1.2 \), and 1.5 MeV for \( 1.8 < |\eta_\ell| < 2.4 \), for both the \( p_T^\ell \) and \( m_T \) distributions. The uncertainties are considered as uncorrelated across pseudorapidity bins.

Figure 10 compares the \( \eta_\ell \) distribution in data and simulation for \( Z \rightarrow e^+ e^- \) events, after applying the efficiency corrections discussed above. The corresponding uncertainties in \( m_W \) due to the electron efficiency corrections are shown in Table 5.

8 Calibration of the recoil

The calibration of the recoil, \( u_T \), affects the measurement of the \( W \)-boson mass through its impact on the \( m_T \) distribution, which is used to extract \( m_W \). In addition, the recoil calibration affects the \( p_T^\ell \) and \( m_T \) distributions through the \( p_T^{\text{miss}} \), \( m_T \), and \( u_T \) event-selection requirements. The calibration procedure proceeds in two steps. First, the dominant part of the \( u_T \) resolution mismodelling is addressed by correcting the modelling of the overall event activity in simulation. These corrections are derived separately in the \( W \)- and \( Z \)-boson samples. Second, corrections for residual differences in the recoil response and resolution are derived using \( Z \)-boson events in data, and transferred to the \( W \)-boson sample.

8.1 Event activity corrections

The pile-up of multiple proton–proton interactions has a significant impact on the resolution of the recoil. As described in Sect. 4, the pile-up is modelled by overlaying the simulated hard-scattering process with additional \( pp \) interactions simulated using PYTHIA 8 with the A2 tune. The average number of interactions per bunch crossing is defined, for each event, as \( \langle \mu \rangle = L \sigma_{\text{in}} / f_{\text{BC}} \), where \( L \) is the instantaneous luminosity, \( \sigma_{\text{in}} \) is the total \( pp \) inelastic cross section and \( f_{\text{BC}} \) is the average bunch-crossing rate. The distribution of \( \langle \mu \rangle \) in the simulated event samples is reweighted to match the corresponding distribution in data. The distribution of \( \langle \mu \rangle \) is affected in particular by the uncertainty in the cross section and properties of inelastic collisions. In the simulation, \( \langle \mu \rangle \) is scaled by a factor \( \alpha \) to optimise the modelling of observed data distributions which are relevant to the modelling of \( u_T \). A value of \( \alpha = 1.10 \pm 0.04 \) is determined by minimising the \( \chi^2 \) function of the compatibility test between data and simulation for the \( \Sigma E_T^\gamma \) and \( u_T^2 \) distributions, where the uncertainty accounts for differences in the values determined using the two distributions.
After the correction applied to the average number of pileup interactions, residual data-to-prediction differences in the \( \Sigma E_T^\pi \) distribution are responsible for most of the remaining \( u_T \) resolution mismodelling. The \( \Sigma E_T^\pi \) distribution is corrected by means of a Smirnov transform, which is a mapping \( x \rightarrow x' (x) \) such that a function \( f(x) \) is transformed into another target function \( g(x) \) through the relation \( f(x) \rightarrow f (x') \equiv g(x) \) [115]. Accordingly, a mapping \( \Sigma E_T^\pi \rightarrow \Sigma E_T^\pi' \) is defined such that the distribution of \( \Sigma E_T^\pi \) in simulation, \( h_{MC}(\Sigma E_T^\pi) \), is transformed into \( h_{MC}(\Sigma E_T^{\pi'}) \) to match the \( \Sigma E_T^\pi \) distribution in data, \( h_{data}(\Sigma E_T^\pi) \). The correction is derived for \( Z \)-boson events in bins of \( p_T^{\ell \ell} \), as the observed differences in the \( \Sigma E_T^\pi \) distribution depend on the \( Z \)-boson transverse momentum. The result of this procedure is illustrated in Fig. 11a. The modified distribution is used to parameterise the recoil response corrections discussed in the next section.

In \( W \)-boson events, the transverse momentum of the boson can only be inferred from \( u_T \), which has worse resolution compared to \( p_T^{\ell \ell} \) in \( Z \)-boson events. To overcome this limitation, a \( p_T \)-dependent correction is defined assuming that the \( p_T \) dependence of differences between data and simulation in the \( \Sigma E_T^\pi \) distribution in \( W \)-boson events follows the corresponding differences observed in \( Z \)-boson events. The \( \Sigma E_T^\pi \) distribution to be matched by the simulation is defined as follows for \( W \)-boson events:

\[
\tilde{h}_{data}^{W}(\Sigma E_T^\pi, p_T^W) \equiv h_{data}^{Z}(\Sigma E_T^{\pi'}, p_T^{\ell \ell}) \left( \frac{h_{MC}^{W}(\Sigma E_T^\pi)}{h_{MC}^{W}(\Sigma E_T^{\pi'})} / \frac{h_{MC}^{Z}(\Sigma E_T^{\pi'})}{h_{MC}^{Z}(\Sigma E_T^{\pi})} \right),
\]

where \( p_T^W \) is the particle-level \( W \)-boson transverse momentum, and \( p_T^{\ell \ell} \) the transverse momentum measured from the decay-lepton pair, used as an approximation of the particle-level \( p_T^{Z} \). The superscripts \( W \) or \( Z \) refer to \( W \)- or \( Z \)-boson event samples, and the double ratio in the second term accounts for the differences between the inclusive distributions in \( W \)- and \( Z \)-boson events. This correction is defined separately for positively and negatively charged \( W \) bosons, so as to incorporate the dependence of the \( p_T^{W} \) distribution on the charge of the \( W \) boson. Using \( \tilde{h}_{data}^{W}(\Sigma E_T^\pi, p_T^W) \) defined in Eq. (4) as the target distribution, the \( p_T^{W} \)-dependent Smirnov transform of the \( \Sigma E_T^\pi \) distribution in \( W \)-boson events is defined as follows:

\[
h_{MC}^{W}(\Sigma E_T^\pi; p_T^W) \rightarrow h_{MC}^{W}(\Sigma E_T^{\pi'}; p_T^W) \equiv \tilde{h}_{data}^{W}(\Sigma E_T^{\pi'; p_T^W}).
\]

The validity of the approximation introduced in Eq. (4) is verified by comparing \( \tilde{h}_{data}^{W}(\Sigma E_T^\pi)/h_{MC}^{W}(\Sigma E_T^\pi) \) and \( h_{data}^{Z}(\Sigma E_T^\pi)/h_{MC}^{Z}(\Sigma E_T^\pi) \) in broad bins of \( u_T \). The associated systematic uncertainties are discussed in Sect. 8.3.

8.2 Residual response corrections

In the ideal case of beams coinciding with the z-axis, the physical transverse momentum of \( W \) and \( Z \) bosons is uniformly distributed in \( \phi \). However, an offset of the interaction point with respect to the detector centre in the transverse plane, the non-zero crossing angle between the proton beams, and \( \phi \)-dependent response of the calorimeters generate anisotropies in the reconstructed recoil distribution. Corresponding differences between data and simulation are addressed by effective corrections applied to \( u_x \) and \( u_y \) in simulation:

\[
u_x' = u_x + \left( \langle u_x \rangle_{data} - \langle u_x \rangle_{MC} \right),
\]

\[
u_y' = u_y + \left( \langle u_y \rangle_{data} - \langle u_y \rangle_{MC} \right),
\]
where \( \langle u_{x,y} \rangle_{\text{data}} \) and \( \langle u_{x,y} \rangle_{\text{MC}} \) are the mean values of these distributions in data and simulation, respectively. The corrections are evaluated in Z-boson events and parameterised as a function of \( \Sigma E_T^* \). The effect of these corrections on the recoil \( \phi \) distribution is illustrated in Fig. 11b.

The transverse momentum of Z bosons can be reconstructed from the decay-lepton pair with a resolution of 1–2 GeV, which is negligible compared to the recoil energy resolution. The recoil response can thus be calibrated from comparisons with the reconstructed \( p_T^V \) in data and simulation. Recoil energy scale and resolution corrections are derived in bins of \( \Sigma E_T^* \) and \( p_T^{\ell\ell} \) at reconstruction level, and are applied in simulation as a function of the particle-level vector-boson momentum \( p_T^V \) in both the W- and Z-boson samples. The energy scale of the recoil is calibrated by comparing the \( u_Z^\parallel + p_T^{\ell\ell} \) distribution in data and simulation, whereas resolution corrections are evaluated from the \( u_Z^\perp \) distribution. Energy-scale corrections \( b(p_T^V, \Sigma E_T^*) \) are defined as the difference between the average values of the \( u_Z^\parallel + p_T^{\ell\ell} \) distributions in data and simulation, and the energy-resolution correction factors \( r(p_T^V, \Sigma E_T^*) \) as the ratio of the standard deviations of the corresponding \( u_Z^\parallel \) distributions.

The parallel component of \( u_T \) in simulated events is corrected for energy scale and resolution, whereas the perpendicular component is corrected for energy resolution only. The corrections are defined as follows:

\[
\begin{align*}
\langle u_{\parallel}^V \rangle_{\text{corr}} &= \left[ \langle u_{\parallel}^V \rangle_{\text{MC}} - \langle u_{\parallel}^Z \rangle_{\text{data}} \right](p_T^V, \Sigma E_T^*) \cdot r(p_T^V, \Sigma E_T^*) \\
\langle u_{\perp}^V \rangle_{\text{corr}} &= \langle u_{\perp}^V \rangle_{\text{MC}} \cdot r(p_T^V, \Sigma E_T^*),
\end{align*}
\]

where \( V = W, Z \), \( u_{\parallel}^V_{\text{MC}} \) and \( u_{\parallel}^V_{\text{MC}} \) are the parallel and perpendicular components of \( u_T \) in the simulation, and \( u_{\parallel}^V_{\text{corr}} \) and \( u_{\perp}^V_{\text{corr}} \) are the corresponding corrected values. As for \( b \) and \( r \), the average \( \langle u_{\parallel}^Z \rangle_{\text{data}} \) is mapped as a function of the reconstructed \( p_T^{\ell\ell} \) in Z-boson data, and used as a function of \( p_T^V \) in both W- and Z-boson simulation. Since the resolution of \( u_T \) has a sizeable dependence on the amount of pile-up, the correction procedure is defined in three bins of \( \langle \mu \rangle \), corresponding to low, medium, and high pile-up conditions, and defined by the ranges of \( \langle \mu \rangle \in \{2.5, 6.5\} \), \( \langle \mu \rangle \in \{6.5, 9.5\} \), and \( \langle \mu \rangle \in \{9.5, 16.0\} \), respectively. Values for \( b(p_T^V, \Sigma E_T^*) \) are typically \( \mathcal{O}(100 \text{ MeV}) \), and \( r(p_T^V, \Sigma E_T^*) \) deviates from unity by 2% at most. The effect of the calibration is shown in Fig. 12 for \( Z \to \mu\mu \) events. The level of agreement obtained after corrections is satisfactory, and similar performance is observed for \( Z \to ee \) events.

A closure test of the applicability of Z-based corrections to W production is performed using W and Z samples simulated with Powheg+Herwig 6, which provide an alternative model for the description of hadronisation and the underlying event. The procedure described above is used to correct the recoil response from Powheg+Pythia 8 to Powheg+Herwig 6, where the latter is treated as pseudo-data. As shown in Fig. 13, the corrected W recoil distributions in Powheg+Pythia 8 match the corresponding distributions in Powheg+Herwig 6. For this study, the effect of the different particle-level \( p_T^W \) distributions in both samples is removed by reweighting the Powheg+Pythia 8 prediction to Powheg+Herwig 6. This study is performed applying the standard lepton selection cuts, but avoiding further kinematic selections in order to maximize the statistics available for the test.

8.3 Systematic uncertainties

The recoil calibration procedure is sensitive to the following sources of systematic uncertainty: the uncertainty of the scale factor applied to the \( \langle \mu \rangle \) distribution, uncertainties due to the Smirnov transform of the \( \Sigma E_T^* \) distribution, uncertainties in the correction of the average value of the \( u_{x,y} \) distributions, statistical uncertainties in the residual correction factors and their \( p_T \) dependence, and expected differences in the recoil response between Z- and W-boson events.

The uncertainty from the \( \langle \mu \rangle \) scale-factor \( \alpha \) is evaluated by varying it by its uncertainty and repeating all steps of the recoil calibration procedure. These variations affect the determination of \( m_W \) by less than 1 MeV.

The systematic uncertainty related to the dependence of the \( \Sigma E_T^* \) correction on \( p_T \) is estimated by comparing with the results of a \( p_T \)-inclusive correction. This source contributes, averaging over W-boson charges, an uncertainty of approximately 1 MeV for the extraction of \( m_W \) from the \( p_T^\ell \) distribution, and 11 MeV when using the \( m_{T,T} \) distribution.

The recoil energy scale and resolution corrections of Eqs. (5) and (6) are derived from the Z-boson sample and applied to W-boson events. Differences in the detector response to the recoil between W- and Z-boson processes are considered as a source of systematic uncertainty for these corrections. Differences between the \( u_{\perp}^W \) and \( u_{\perp}^Z \) distributions originating from different vector-boson kinematic properties, different ISR and FSR photon emission, and from different selection requirements are, however, discarded as they are either accurately modelled in the simulation or already incorporated in the correction procedure.

To remove the effect of such differences, the two-dimensional distribution \( h_{\Sigma E_T^*}^W(p_T, \Sigma E_T^*) \) in W-boson simulated events is corrected to match the corresponding distribution in Z-boson simulated events, treating the neutrinos in W-boson decays as charged leptons to calculate \( u_T \) as in Z-boson events. Finally, events containing a particle-level photon from final-state radiation are removed. After these corrections, the standard deviation of the \( u_{\perp} \) distribution agrees within 0.03% between simulated W- and Z-boson
events. This difference is equivalent to 6% of the size of the residual resolution correction, which increases the standard deviation of the $u_\perp$ distribution by 0.5%. Accordingly, the corresponding systematic uncertainty due to the extrapolation of the recoil calibration from $Z$- to $W$-boson events is estimated by varying the energy resolution parameter $r$ of Eqs. (5) and (6) by 6%. The impact of this uncertainty on the extraction of $m_W$ is approximately 0.2 MeV for the $p_T^\ell$ distribution, and 5.1 MeV for the $m_T$ distribution. The extrapolation uncertainty of the energy-scale correction $b$ was found to be negligible in comparison.

In addition, the statistical uncertainty of the correction factors contributes 2.0 MeV for the $p_T^\ell$ distribution, and 2.7 MeV for the $m_T$ distribution. Finally, instead of using a binned correction, a smooth interpolation of the correction values between the bins is performed. Comparing the binned and interpolated correction parameters $b(p_T^\ell, \Sigma E_T^{\text{vis}})$ and $r(p_T^\ell, \Sigma E_T^{\text{vis}})$ leads to a systematic uncertainty in $m_W$ of 1.4 and 3.1 MeV for the $p_T^\ell$ and $m_T$ distributions, respectively. Systematic uncertainties in the $u_{x,y}$ corrections are found to be small compared to the other systematic uncertainties, and are neglected.

The impact of the uncertainties of the recoil calibration on the extraction of the $W$-boson mass from the $p_T^\ell$ and $m_T$ distributions are summarised in Table 6. The determination of $m_W$ from the $p_T^\ell$ distribution is only slightly affected by the uncertainties of the recoil calibration, whereas larger uncertainties are estimated for the $m_T$ distribution. The largest uncertainties are induced by the $\Sigma E_T^{\text{vis}}$ corrections and by the extrapolation of the recoil energy-scale and energy-resolution corrections from $Z$- to $W$-boson events. The systematic uncertainties are in general smaller for $W^-$ events.

Fig. 12 Recoil distributions for (a) $u_\parallel$, (b) $u_\perp + p_T^\ell$, (c) $u_\perp$, and (d) $u_T$ in $Z \to \mu\mu$ events. The data are compared to the simulation before and after applying the recoil corrections described in the text. The lower panels show the data-to-prediction ratios, with the vertical bars showing the statistical uncertainty.
than for $W^+$ events, as the $\Sigma E_T^\ell$ distribution in $W^-$ events is closer to the corresponding distribution in $Z$-boson events.

9 Consistency tests with $Z$-boson events

The $Z \rightarrow \ell\ell$ event sample allows several validation and consistency tests of the $W$-boson analysis to be performed. All the identification requirements of Sect. 5.1, the calibration and efficiency corrections of Sects. 7 and 8, as well as the physics-modelling corrections described in Sect. 6, are applied consistently in the $W$- and $Z$-boson samples. The $Z$-boson sample differs from the $W$-boson sample in the selection requirements, as described in Sect. 5.2. In addition to the event-selection requirements described there, the transverse momentum of the dilepton system, $p_T^{\ell\ell}$, is required to be smaller than 30 GeV.

The missing transverse momentum in $Z$-boson events is defined by treating one of the two decay leptons as a neutrino and ignoring its transverse momentum when defining the event kinematics. This procedure allows the $p_T^{\text{miss}}$ and $m_T$ variables to be defined in the $Z$-boson sample in close analogy to their definition in the $W$-boson sample. The procedure is repeated, removing the positive and negative lepton in turn.

In the $Z$-boson sample, the background contribution arising from top-quark and electroweak production is estimated using Monte Carlo samples. Each process is normalised using the corresponding theoretical cross sections, evaluated at NNLO in the perturbative expansion of the strong coupling constant. This background contributes a 0.12% fraction in each channel. In the muon channel, the background contribution from multijet events is estimated to be smaller than 0.05% using simulated event samples of $b\bar{b}$ and $c\bar{c}$ production, and neglected. In the electron channel, a data-driven estimate of the multijet background contributes about a 0.1% fraction, before applying the isolation selections, which reduce it to a negligible level.

Fig. 13 Distributions of (a) $u_T$ and (b) $u_\parallel$ in $W$ events simulated using POWHEG+PYTHIA 8 and POWHEG+HERWIG 6. The recoil response in POWHEG+PYTHIA 8 is corrected to the POWHEG+HERWIG 6 response using simulated $Z$ events following the method described in the text. The $p_T^W$ distribution in POWHEG+PYTHIA 8 is reweighted to the POWHEG+HERWIG 6 prediction. The lower panels show the ratios of POWHEG+HERWIG 6 to POWHEG+PYTHIA 8, with and without the response correction in the POWHEG+PYTHIA 8 sample.

Table 6 Systematic uncertainties in the $m_W$ measurement due to recoil corrections, for the different kinematic distributions and $W$-boson charge categories. Combined uncertainties are evaluated as described in Sect. 2.2.

<table>
<thead>
<tr>
<th>$W$-boson charge</th>
<th>Kinematic distribution</th>
<th>$W^+$</th>
<th>$W^-$</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p_T^\ell$</td>
<td>$m_T$</td>
<td>$p_T^\ell$</td>
<td>$m_T$</td>
</tr>
<tr>
<td>$\delta m_W$ [MeV]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu$ scale factor</td>
<td>0.2</td>
<td>1.0</td>
<td>0.2</td>
<td>1.0</td>
</tr>
<tr>
<td>$\Sigma E_T^\ell$ correction</td>
<td>0.9</td>
<td>12.2</td>
<td>1.1</td>
<td>10.2</td>
</tr>
<tr>
<td>Residual corrections (statistics)</td>
<td>2.0</td>
<td>2.7</td>
<td>2.0</td>
<td>2.7</td>
</tr>
<tr>
<td>Residual corrections (interpolation)</td>
<td>1.4</td>
<td>3.1</td>
<td>1.4</td>
<td>3.1</td>
</tr>
<tr>
<td>Residual corrections ($Z \rightarrow W$ extrapolation)</td>
<td>0.2</td>
<td>5.8</td>
<td>0.2</td>
<td>4.3</td>
</tr>
<tr>
<td>Total</td>
<td>2.6</td>
<td>14.2</td>
<td>2.7</td>
<td>11.8</td>
</tr>
</tbody>
</table>
Figure 14 shows the reconstructed distributions of \( p_T^{\ell\ell} \) and \( y_{\ell\ell} \) in selected Z-boson events; these distributions are not sensitive to the value of \( m_Z \). Figure 15 shows the corresponding distributions for \( p_T^{\ell} \) and \( m_T \), variables which are sensitive to \( m_Z \). Data and simulation agree at the level of 1–2\% percent in all the distributions.

The mass of the Z boson is extracted with template fits to the \( m_{\ell\ell} \), \( p_T^{\ell} \), and \( m_T \) kinematic distributions. The extraction of the Z-boson mass from the dilepton invariant mass distribution is expected to yield, by construction, the value of \( m_Z \) used as input for the muon-momentum and electron-energy calibrations, providing a closure test of the lepton calibration procedures. The \( p_T^{\ell} \) distribution is very sensitive to the physics-modelling corrections described in Sect. 6. The comparison of the value of \( m_Z \) extracted from the \( p_T^{\ell} \) distribution with the value used as input for the calibration tests the physics modelling and efficiency corrections. Finally, \( m_Z \) measurements from the \( m_T \) distribution provides a test of the recoil calibration.

Similarly to the W-boson mass, the value of \( m_Z \) is determined by minimising the \( \chi^2 \) function of the compatibility test between the templates and the measured distributions. The templates are generated with values of \( m_Z \) in steps of 4 to 25 MeV within a range of \( \pm 450 \) MeV, centred around a reference value corresponding to the LEP combined value, \( m_Z = 91187.5 \) MeV [32]. The \( \chi^2 \) function is interpolated with a second order polynomial. The minimum of the \( \chi^2 \) function yields the extracted value of \( m_Z \), and the difference between the extracted value of \( m_Z \) and the reference value is defined as \( \Delta m_Z \). The ranges used for the extraction are \([80, 100]\) GeV for the \( m_{\ell\ell} \) distributions, \([30, 55]\) GeV for the \( p_T^{\ell} \) distribution, and \([40, 120]\) GeV for the \( m_T \) distribution.
Fig. 15 The $p_T^\ell$ distribution in the a electron and b muon channels, and $m_T$ distributions in the c, e electron and d, f muon decay channels for Z events when the c, d negatively charged, or e, f positively charged lepton is removed. The data are compared to the simulation including signal and backgrounds. Detector calibration and physics-modelling corrections are applied to the simulated events. Background events contribute less than 0.2% of the observed distributions. The lower panels show the data-to-prediction ratios, with the error bars showing the statistical uncertainty.
The extraction of $m_Z$ from the $m_T$ distribution is performed separately for positively and negatively charged leptons in the event, by reconstructing $m_T$ from the kinematic properties of one of the two charged leptons and of the recoil reconstructed by treating the other as a neutrino.

Z-boson mass fits are performed using the $m_T$ and $p_T^\ell$ distributions in the electron and muon decay channels, inclusively in $\eta$ and separately for positively and negatively charged leptons. The results of the fits are summarised in Fig. 16 and Table 7. The $p_T^\ell$ fit results include all lepton reconstruction systematic uncertainties except the $Z$-based energy or momentum scale calibration uncertainties; the $m_T$ fit results include recoil calibration systematic uncertainties in addition. Physics-modelling uncertainties are neglected.

The value of $m_Z$ measured from positively charged leptons is correlated with the corresponding extraction from the negatively charged leptons. The $p_T^\ell$ distributions for positively and negatively charged leptons are statistically independent, but the $m_T$ distributions share the same reconstructed recoil event by event, and are statistically correlated. In both cases, the decay of the $Z$-boson induces a kinematical correlation between the distributions of positively and negatively charged leptons. The correlation is estimated by constructing two-dimensional $\ell^+$ and $\ell^-$ distributions, separately for $p_T^\ell$ and $m_T$, fluctuating the bin contents of these distributions within their uncertainties, and repeating the fits for each pseudodata sample. The correlation values are $-7\%$ for the $p_T^\ell$ distributions, and $-12\%$ for the $m_T$ distributions.

Accounting for the experimental uncertainties as described above, the combined extraction of $m_Z$ from the $p_T^\ell$ distribution yields a result compatible with the reference value within 0.9 standard deviations. The difference between the $m_Z$ extractions from positively and negatively charged lepton distributions is compatible with zero within 1.4 standard deviations. For the extraction from the $m_T$ distribution, the compatibility with the reference value of $m_Z$ is at the level of 1.5 standard deviations. Fits using the lepton pair invariant mass distribution agree with the reference, yielding $\Delta m_Z = 1 \pm 3$ MeV in the muon channel and $\Delta m_Z = 3 \pm 5$ MeV in the electron channel, as expected from the calibration procedure. In summary, the consistency tests based on the $Z$-boson sample agree with the expectations within the experimental uncertainties.

### 10 Backgrounds in the W-boson sample

The $W$-boson event sample, selected as described in Sect. 5.2, includes events from various background processes. Background contributions from $Z$-boson, $W \rightarrow \tau \nu$, boson pair, and top-quark production are estimated using simulation.
Contributions from multijet production are estimated with data-driven techniques.

10.1 Electroweak and top-quark backgrounds

The dominant sources of background contribution in the $W \rightarrow \ell \nu$ sample are $Z \rightarrow \ell \ell$ events, in which one of the two leptons escapes detection, and $W \rightarrow \tau \nu$ events, where the $\tau$ decays to an electron or muon. These background contributions are estimated using the POWHEG+PYTHIA 8 samples after applying the modelling corrections discussed in Sect. 6, which include NNLO QCD corrections to the angular coefficients and rapidity distributions, and corrections to the vector-boson transverse momentum. The $Z \rightarrow \ell \ell$ background represents 2.9% of the $W^+ \rightarrow e\nu$ sample and 4.0% of the $W^- \rightarrow e\nu$ sample. In the muon channel, the $Z \rightarrow \mu\mu$ background represents 4.8 and 6.3% of the $W^+ \rightarrow \mu\nu$ and $W^- \rightarrow \mu\nu$ samples, respectively. The $W \rightarrow \tau \nu$ background represents 1.0% of the selected sample in both channels, and the $Z \rightarrow \tau\tau$ background contributes approximately 0.12%.

The normalisation of these processes relative to the $W$-boson signal and the corresponding uncertainties are discussed in Sect. 4. A relative uncertainty of 0.2% is assigned to the normalisation of the $W \rightarrow \tau\nu$ samples with respect to the $W$-boson signal sample, to account for the uncertainty in the $\tau$-lepton branching fractions to electrons and muons. In the determination of the $W$-boson mass, the variations of $m_W$ are propagated to the $W \rightarrow \tau\nu$ background templates in the same way as for the signal.

Similarly, backgrounds involving top-quark (top-quark pairs and single top-quark) production, and boson-pair production are estimated using simulation, and normalisation uncertainties are assigned as discussed in Sect. 4. These processes represent 0.11 and 0.07% of the signal event selection, respectively.

Uncertainties in the distributions of the $W \rightarrow \tau\nu$ and $Z \rightarrow \ell\ell$ processes are described by the physics-modelling uncertainties discussed in Sect. 6, and are treated as fully correlated with the signal. Shape uncertainties for boson-pair production and top-quark production are considered negligible compared to the uncertainties in their cross sections, given the small contributions of these processes to the signal event selection.

10.2 Multijet background

Inclusive multijet production in strong-interaction processes constitutes a significant source of background. A fraction of multijet events contains semileptonic decays of bottom and charm hadrons to muons or electrons and neutrinos, and can pass the $W$-boson signal selection. In addition, inclusive jet production contributes to the background if one jet is misidentified as electron or muon, and sizeable missing transverse momentum is reconstructed in the event. In-flight decays of pions or kaons within the tracking region can mimic the $W$-boson signal in the muon channel. In the electron channel, events with photon conversions and hadrons misidentified as electrons can be selected as $W$-boson events. Due to the small selection probability for multijet events, their large production cross section, and the relatively complex modelling of the hadronisation processes, the multijet background contribution cannot be estimated precisely using simulation, and a data-driven method is used instead.

The estimation of the multijet background contribution follows similar procedures in the electron and muon decay channels, and relies on template fits to kinematic distributions in background-dominated regions. The analysis uses the distributions of $p_T^{\text{miss}}$, $m_T$, and the $p_T/ m_T$ ratio, where jet-enriched regions are obtained by relaxing a subset of the signal event-selection requirements. The first kinematic region, denoted FR1, is defined by removing the $p_T^{\text{miss}}$ and $m_T$ requirements from the event selection. A second kinematic region, FR2, is defined in the same way as FR1, but by also removing the requirement on $u_T$. Multijet background events, which tend to have smaller values of $p_T^{\text{miss}}$ and $m_T$ than the signal, are enhanced by this selection. The $p_T/ m_T$ distribution is sensitive to the angle between the $p_T^{\text{miss}}$ and $p_T$ vectors in the transverse plane. Whereas $W$-boson events are expected to peak at values of $p_T/ m_T = 0.5$, relatively large tails are observed for multijet events.

Templates of the multijet background distributions for these observables are obtained from data by inverting the lepton energy-isolation requirements. Contamination of these control regions by electroweak and top production is estimated using simulation and subtracted. In the muon channel, the anti-isolation requirements are defined from the ratio of the scalar sum of the $p_T$ of tracks in a cone of size $\Delta R < 0.2$ around the reconstructed muon to the muon $p_T$. The isolation variable $p_T^{\mu,\text{cone}}$, introduced in Sect. 5.1, is required to satisfy $c_1 < p_T^{\mu,\text{cone}} / p_T < c_2$, where the anti-isolation boundaries $c_1$ and $c_2$ are varied as discussed below. In order to avoid overlap with the signal region, the lower boundary $c_1$ is always larger than 0.1. In the electron channel, the scalar sum of the $p_T$ of tracks in a cone of size $\Delta R < 0.4$ around the reconstructed electron, defined as $p_T^{e,\text{cone}}$ in Sect. 5.1, is used to define the templates, while the requirements on the calorimeter isolation are omitted.

The multijet background normalisation is determined by fitting each of the $p_T^{\text{miss}}$, $m_T$, and $p_T/ m_T$ distributions in the two kinematic regions FR1 and FR2, using templates of these distributions based on multijet events and obtained with several ranges of the anti-isolation variables. The multijet background in the signal region is determined by correcting the multijet fraction fitted in the FR1 and FR2 for the different
efficiencies of the selection requirements of the signal region. In the electron channel, $c_1$ is varied from 4 to 9 GeV in steps of 1 GeV, and $c_2$ is set to $c_2 = c_1 + 1$ GeV. In the muon channel, $c_1$ is varied from 0.1 to 0.37 in steps of 0.03, and $c_2$ is set to $c_2 = c_1 + 0.03$. Example results of template fits in the electron and muon channels are shown in Fig. 17. The results corresponding to the various observables and to the different kinematic regions are linearly extrapolated in the isolation variables to the signal regions, denoted by $c_1 = 0$. Figure 18 illustrates the extrapolation procedure.

The systematic uncertainty in the multijet background fraction is defined as half of the largest difference between the results extrapolated from the different kinematic regions and observables. The multijet background contribution is estimated separately in all measurement categories. In the electron channel, the multijet background fraction rises from 0.58±0.08% at low $|\eta\ell|$ to 1.73±0.19% in the last measurement bin, averaging the $W^+$ and $W^-$ channels. In the muon channel, the charge-averaged multijet background fraction decreases from 0.72±0.07% to 0.49±0.03%, when going from low to high $|\eta\ell|$. The uncertainties in the multijet background fractions are sufficient to account for the observed residual discrepancies between the fitted distributions and the data (see Fig. 17). The estimated multijet background yields are consistent between $W^+$ and $W^-$, but the multijet background fraction is smaller in the $W^+$ channels due to the higher signal yield.

Corrections to the shape of the multijet background contributions and corresponding uncertainties in the distributions used to measure the $W$-boson mass are estimated with a similar procedure. The kinematic distributions in the control regions are obtained for a set of anti-isolation ranges, and parameterised with linear functions of the lower bound of the anti-isolation requirement. The distributions are extrapolated to the signal regions accordingly. Uncertainties in the extrapolated distributions are dominated by the statistical uncertainty, which is determined with a toy MC method by fluctuating within their statistical uncertainty the bin contents of the histograms in the various anti-isolation ranges. The resulting multijet background distribution is propagated to the templates, and the standard deviation of the determined values of $m_W$ yields the estimated uncertainty due to the shape of the multijet background. Uncertainties due to the choice of parameterisation are small in comparison and neglected.

Uncertainties in the normalisation of multijet, electroweak, and top-quark background processes are considered correlated across decay channels, boson charges and rapidity bins, whereas the uncertainty in the shape of multijet background is considered uncorrelated between decay channels and boson charges. The impact of the background systematic uncertainties on the determination of $m_W$ is summarised in Table 8.

11 Measurement of the $W$-boson mass

This section presents the determination of the mass of the $W$ boson from template fits to the kinematic distributions of the $W$-boson decay products. The final measured value is obtained from the combination of measurements performed using the lepton transverse momentum and transverse mass distributions in categories corresponding to the electron and muon decay channels, positively and negatively charged $W$ bosons, and absolute pseudorapidity bins of the charged lepton, as illustrated in Table 1. The number of selected events in each category is shown in Table 9.

11.1 Control distributions

The detector calibration and the physics modelling are validated by comparing data with simulated $W$-boson signal and backgrounds for several kinematic distributions that are insensitive to the $W$-boson mass. The comparison is based on a $\chi^2$ compatibility test, including statistical and systematic uncertainties, and the bin-to-bin correlations induced by the latter. The systematic uncertainty comprises all sources of experimental uncertainty related to the lepton and recoil calibration, and to the background subtraction, as well as sources of modelling uncertainty associated with electroweak corrections, or induced by the helicity fractions of vector-boson production, the vector-boson transverse-momentum distribution, and the PDFs. Comparisons of data and simulation for the $\eta\ell$, $u_T$, and $u_\parallel^\ell$ distributions, in positively and negatively charged $W$-boson events, are shown in Figs. 19 and 20 for the electron and muon decay channels, respectively.

Data and simulation agree within uncertainties for all distributions, as confirmed by the satisfactory $\chi^2$/dof values. The effect of the residual discrepancies in the $u_T$ distributions for $W^\rightarrow \ell\nu$, visible at low values in Figs. 19d and 20d, is discussed in Sect. 11.5.

11.2 Data-driven check of the uncertainty in the $p_T^W$ distribution

The uncertainty in the prediction of the $u_\parallel^\ell$ distribution is dominated by $p_T^W$ distribution uncertainties, especially at negative values of $u_\parallel^\ell$ in the kinematic region corresponding to $u_\parallel^\ell < -15$ GeV. This is illustrated in Fig. 21, which compares the recoil distributions in the POWHEG+PYTHIA 8 and POWHEG+HERWIG 6 samples, before and after the corrections described in Sect. 8.2 (the $p_T^W$ distribution predicted by POWHEG+PYTHIA 8 is not reweighted to that of POWHEG+HERWIG 6). As can be seen, the recoil corrections and the different $p_T^W$ distributions have a comparable effect on the $u_T$ distribution. In contrast, the effect of the recoil corrections is small at negative values of $u_\parallel^\ell$, whereas the
Fig. 17 Example template fits to the $a$, $b$ $p_T^{\text{miss}}$, $c$, $d$ $m_T$, and $e$, $f$ $p_T^\ell/m_T$ distributions in the FR1 kinematic region, in the $a$, $c$, $e$ electron and $b$, $d$, $f$ muon decay channels. Multijet templates are derived from the data requiring $4 \text{ GeV} < p_T^{cone} < 8 \text{ GeV}$ in the electron channel, and $0.2 < p_T^{\mu,cone}/p_T^\ell < 0.4$ in the muon channel. The data are compared to the simulation including signal and background contributions.
extrapolations are indicated by the solid lines. The thick crosses show the results of the linear extrapolation of the background estimate to the signal region, including uncertainties from the extrapolation only. The thin crosses also include the uncertainty induced by the contamination of the control regions by EW and top-quark processes.

Table 8: Systematic uncertainties in the $m_W$ measurement due to electroweak, top-quark, and multijet background estimation, for fits to the $p_T^\ell$ and $m_T$ distributions, in the electron and muon decay channels, with positively and negatively charged $W$ bosons.

<table>
<thead>
<tr>
<th>Kinematic distribution</th>
<th>Decay channel</th>
<th>$p_T^\ell$</th>
<th>$m_T$</th>
<th>$W^\pm$</th>
<th>$W^\mp$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W \rightarrow \tau \nu$ (fraction, shape)</td>
<td>$W^+$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>$Z \rightarrow ee$ (fraction, shape)</td>
<td>$W^+$</td>
<td>3.3</td>
<td>4.8</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$Z \rightarrow \mu \mu$ (fraction, shape)</td>
<td>$W^+$</td>
<td>–</td>
<td>–</td>
<td>3.5</td>
<td>4.5</td>
</tr>
<tr>
<td>$Z \rightarrow \tau \tau$ (fraction, shape)</td>
<td>$W^+$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>$W W, W Z, Z Z$ (fraction)</td>
<td>$W^+$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Top (fraction)</td>
<td>$W^+$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Multijet (fraction)</td>
<td>$W^+$</td>
<td>3.2</td>
<td>3.6</td>
<td>1.8</td>
<td>2.4</td>
</tr>
<tr>
<td>Multijet (shape)</td>
<td>$W^+$</td>
<td>3.8</td>
<td>3.1</td>
<td>1.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Total</td>
<td>$W^+$</td>
<td>6.0</td>
<td>6.8</td>
<td>4.3</td>
<td>5.3</td>
</tr>
</tbody>
</table>

Table 9: Numbers of selected $W^+$ and $W^-$ events in the different decay channels in data, inclusively and for the various $|\eta_\ell|$ categories.

| $|\eta_\ell|$ range | $W^+ \rightarrow \mu^+ \nu$ | $W^+ \rightarrow \mu^\mp \nu$ | $W^- \rightarrow \mu^- \bar{\nu}$ | $W^- \rightarrow \mu^- \bar{\nu}$ | $W^+ \rightarrow e^+ \nu$ | $W^+ \rightarrow e^- \nu$ | $W^- \rightarrow e^- \bar{\nu}$ | $W^- \rightarrow e^- \bar{\nu}$ |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0–0.8           | 1283332         | 1063131         | 1377773         | 885582          | 460918          | 3234960         | 956620          | 3397716         |
| 0.8–1.4         |                 |                 |                 |                 |                 |                 |                 |                 |
| 1.4–2.0         |                 |                 |                 |                 |                 |                 |                 |                 |
| 2.0–2.4         |                 |                 |                 |                 |                 |                 |                 |                 |
| Inclusive       |                 |                 |                 |                 |                 |                 |                 |                 |

A difference in the $p_T^W$ distributions has a large impact in this region.

The sensitivity of the $u_T^W$ distribution is exploited to validate the modelling of the $p_T^W$ distribution by Pythia 8 AZ, and its theory-driven uncertainty, described in Sect. 6.5.2, with a data-driven procedure. The parton-shower factorisation scale $\mu_F$ associated with the $c \bar{c} \rightarrow W$ processes constitutes the main source of uncertainty in the modelling of the $p_T^W$ distribution. Variations of the $u_T^W$ distribution induced by changes in the factorisation scale of the $c \bar{c} \rightarrow W$ processes are parameterised and fitted to the data. The $u_T^W$ distribution is predicted for the two boundary values of $\mu_F$, and assumed to vary linearly as a function of $\mu_F$. Variations induced by changes in $\mu_F$ are parameterised using a variable...
Fig. 19 The a, b $\eta_{\ell}$, (c,d) $u_T$, and e, f $u_L^\parallel$ distributions for a, c, e $W^+$ events and b, d, f $W^-$ events in the electron decay channel. The data are compared to the simulation including signal and background contributions. Detector calibration and physics-modelling corrections are applied to the simulated events. The lower panels show the data-to-prediction ratios, the error bars show the statistical uncertainty, and the band shows the systematic uncertainty of the prediction. The $\chi^2$ values displayed in each figure account for all sources of uncertainty and include the effects of bin-to-bin correlations induced by the systematic uncertainties.
Fig. 20 The a, b $\eta_{\ell}$, (c,d) $u_T$, and e, f $u_T^\parallel$ distributions for a, c, e $W^+$ events and b, d, f $W^-$ events in the muon decay channel. The data are compared to the simulation including signal and background contributions. Detector calibration and physics-modelling corrections are applied to the simulated events. The lower panels show the data-to-prediction ratios, the error bars show the statistical uncertainty, and the band shows the systematic uncertainty of the prediction. The $\chi^2$ values displayed in each figure account for all sources of uncertainty and include the effects of bin-to-bin correlations induced by the systematic uncertainties.
s defined in units of the initially allowed range, i.e. values of $s = -1, 0, +1$ correspond to half the effect of changing from $\mu_F = m_V$ to $\mu_F = m_V/2, 2m_V$, respectively. The optimal value of $s$ is determined by fitting the fraction of events in the kinematic region $-30 < u_T < -15$ GeV. The fit accounts for all experimental and modelling uncertainties affecting the $u_T$ distribution, and gives a value of $s = -0.22 \pm 1.06$. The best-fit value of $s$ confirms the good agreement between the the PYTHIA 8 AZ prediction and the data; its uncertainty is dominated by PDF and recoil-calibration uncertainties, and matches the variation range of $\mu_F$ used for the initial estimation of the $p_T^{W}$ distribution uncertainty.

This validation test supports the PYTHIA 8 AZ prediction of the $p_T^{W}$ distribution and the theory-driven associated uncertainty estimate. On the other hand, as shown in Fig. 22, the data disagree with the DYRes and POWHEG MiNLO+PYTHIA 8 predictions. The latter are obtained reweighting the initial $p_T^{W}$ distribution in POWHEG+PYTHIA 8 according to the product of the $p_T^{Z}$ distribution of PYTHIA 8 AZ, which matches the measurement of Ref. [44], and $R_{W/Z}(p_T)$ as predicted by DYRes and POWHEG MiNLO+PYTHIA 8. The uncertainty bands in the DYRes prediction are calculated using variations of the factorisation, renormalisation and resummation scales $\mu_F$, $\mu_R$ and $\mu_{ES}$ following the procedure described in Ref. [116,117]. The uncertainty obtained applying correlated scale variations in $W$ and $Z$ production does not cover the observed difference with the data. The potential effect of using $R_{W/Z}(p_T)$ as predicted by DYRes instead of PYTHIA 8 AZ for the determination of $m_W$ is discussed in Sect. 11.5.

11.3 Results for $m_W$ in the measurement categories

Measurements of $m_W$ are performed using the $p_T^{Z}$ and $m_T$ distributions, separately for positively and negatively charged $W$ bosons, in three bins of $|\eta_\ell|$ in the electron decay channel, and in four bins of $|\eta_\ell|$ in the muon decay channel, leading to a total of 28 $m_W$ determinations. In each category, the value of $m_W$ is determined by a $\chi^2$ minimisation, comparing the $p_T^{Z}$ and $m_T$ distributions in data and simulation for different values of $m_W$. The templates are generated with values of $m_W$ in steps of 1 to 10 MeV within a range of $\pm 400$ MeV, centred around the reference value used in the Monte Carlo signal samples. The statistical uncertainty is estimated from the half width of the $\chi^2$ function at the value corresponding to one unit above the minimum. Systematic uncertainties due to physics-modelling corrections, detector-calibration corrections, and background subtraction, are discussed in Sects. 6–8 and 10, respectively.

The lower and upper bounds of the range of the $p_T^{Z}$ distribution used in the fit are varied from 30 to 35 GeV, and from 45 to 50 GeV respectively, in steps of 1 GeV. For the $m_T$ distribution, the boundaries are varied from 65 to 70 GeV, and from 90 to 100 GeV. The total measurement uncertainty is evaluated for each range, after combining the measurement categories as described in Sect. 11.4 below. The

---

3 Half the effect is used because only one of the two quarks in the initial state is heavy, as discussed in Sect. 6.5.2.
smallest total uncertainty in $m_W$ is found for the fit ranges $32 < p_T^W < 45$ GeV and $66 < m_T < 99$ GeV. The optimisation is performed before the unblinding of the $m_W$ value and the optimised range is used for all the results described below.

The final measurement uncertainty is dominated by modelling uncertainties, with typical values in the range 25–35 MeV for the various charge and $|\eta|$ categories. Lepton-calibration uncertainties are the dominant sources of experimental systematic uncertainty for the extraction of $m_W$ from the $p_T^W$ distribution. These uncertainties vary from about 15 MeV to about 35 MeV for most measurement categories, except the highest $|\eta|$ bin in the muon channel where the total uncertainty of about 120 MeV is dominated by the muon momentum linearity uncertainty. The uncertainty in the calibration of the recoil is the largest source of experimental systematic uncertainty for the $m_T$ distribution, with a typical contribution of about 15 MeV for all categories. The determination of $m_W$ from the $p_T^W$ and $m_T$ distributions in the various categories is summarised in Table 10, including an overview of statistical and systematic uncertainties. The results are also shown in Fig. 23. No significant differences in the values of $m_W$ corresponding to the different decay channels and to the various charge and $|\eta|$ categories are observed.

The comparison of data and simulation for kinematic distributions sensitive to the value of $m_W$ provides further validation of the detector calibration and physics modelling. The comparison is performed in all measurement categories. The $\eta$-inclusive $p_T^W$, $m_T$ and $p_T^{\text{miss}}$ distributions for positively and negatively charged $W$ bosons are shown in Figs. 24 and 25 for the electron and muon decay channels, respectively. The value of $m_W$ used in the predictions is set to the overall measurement result presented in the next section. The $\chi^2$ values quantifying the comparison between data and prediction are calculated over the full histogram range and account for all sources of uncertainty. The bin-to-bin correlations induced by the experimental and physics-modelling systematic uncertainties are also accounted for. Overall, satisfactory agreement is observed. The deficit of data visible for $p_T^W \sim 40–42$ GeV in the $W^+ \rightarrow e\nu$ channel does not strongly affect the mass measurement, as the observed effect differs from that expected from $m_W$ variations. Cross-checks of possible sources of this effect were performed, and its impact on the mass determination was shown to be within the corresponding systematic uncertainties.

11.4 Combination and final results

The measurements of $m_W$ in the various categories are combined accounting for statistical and systematic uncertainties and their correlations. The statistical correlation of the $m_W$ values determined from the $p_T^W$ and $m_T$ distributions is evaluated with the bootstrap method [118], and is approximately 50% for all measurement categories.

The systematic uncertainties have specific correlation patterns across the $m_W$ measurement categories. Muon-momentum and electron-energy calibration uncertainties are uncorrelated between the different decay channels, but largely correlated between the $p_T^W$ and $m_T$ distributions. Recoil-calibration uncertainties are correlated between electron and muon decay channels, and they are small for $p_T^W$ distributions. The PDF-induced uncertainties are largely cor-
Table 10  Results of the $m_W$ measurements in the electron and muon decay channels, for positively and negatively charged $W$ bosons, in different lepton-$|\eta|$ ranges, using the $m_T$ and $p_T^2$ distributions in the optimised fitting range. The table shows the statistical uncertainties, together with all experimental uncertainties, divided into muon-, electron-, recoil- and background-related uncertainties, and all modelling uncertainties, separately for QCD modelling including scale variations, parton shower and angular coefficients, electroweak corrections, and PDFs. All uncertainties are given in MeV.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$W^+ \rightarrow \mu \nu$, $</td>
<td>\eta</td>
<td>&lt; 0.8$</td>
<td>80371.3</td>
<td>29.2</td>
<td>12.4</td>
<td>0.0</td>
<td>15.2</td>
<td>8.1</td>
<td>9.9</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^+ \rightarrow \mu \nu$, $0.8 &lt;</td>
<td>\eta</td>
<td>&lt; 1.4$</td>
<td>80354.1</td>
<td>32.1</td>
<td>19.3</td>
<td>0.0</td>
<td>13.0</td>
<td>6.8</td>
<td>9.6</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^+ \rightarrow \mu \nu$, $1.4 &lt;</td>
<td>\eta</td>
<td>&lt; 2.0$</td>
<td>80426.3</td>
<td>30.2</td>
<td>35.1</td>
<td>0.0</td>
<td>14.3</td>
<td>7.2</td>
<td>9.3</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^+ \rightarrow \mu \nu$, $2.0 &lt;</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>80334.6</td>
<td>40.9</td>
<td>112.4</td>
<td>0.0</td>
<td>14.4</td>
<td>9.0</td>
<td>8.4</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^- \rightarrow \mu \nu$, $</td>
<td>\eta</td>
<td>&lt; 0.8$</td>
<td>80375.5</td>
<td>30.6</td>
<td>11.6</td>
<td>0.0</td>
<td>13.1</td>
<td>8.5</td>
<td>9.5</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^- \rightarrow \mu \nu$, $0.8 &lt;</td>
<td>\eta</td>
<td>&lt; 1.4$</td>
<td>80417.5</td>
<td>36.4</td>
<td>18.5</td>
<td>0.0</td>
<td>12.2</td>
<td>7.7</td>
<td>9.7</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^- \rightarrow \mu \nu$, $1.4 &lt;</td>
<td>\eta</td>
<td>&lt; 2.0$</td>
<td>80379.4</td>
<td>35.6</td>
<td>33.9</td>
<td>0.0</td>
<td>10.5</td>
<td>8.1</td>
<td>9.7</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^- \rightarrow \mu \nu$, $2.0 &lt;</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>80334.2</td>
<td>52.4</td>
<td>123.7</td>
<td>0.0</td>
<td>11.6</td>
<td>10.2</td>
<td>9.9</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^+ \rightarrow e \nu$, $</td>
<td>\eta</td>
<td>&lt; 0.6$</td>
<td>80352.9</td>
<td>29.4</td>
<td>0.0</td>
<td>19.5</td>
<td>13.1</td>
<td>15.3</td>
<td>9.9</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^+ \rightarrow e \nu$, $0.6 &lt;</td>
<td>\eta</td>
<td>&lt; 1.2$</td>
<td>80381.5</td>
<td>30.4</td>
<td>0.0</td>
<td>21.4</td>
<td>15.1</td>
<td>13.2</td>
<td>9.6</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^+ \rightarrow e \nu$, $1.8 &lt;</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>80352.4</td>
<td>32.4</td>
<td>0.0</td>
<td>26.6</td>
<td>16.4</td>
<td>32.8</td>
<td>8.4</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^- \rightarrow e \nu$, $</td>
<td>\eta</td>
<td>&lt; 0.6$</td>
<td>80415.8</td>
<td>31.3</td>
<td>0.0</td>
<td>16.4</td>
<td>11.8</td>
<td>15.5</td>
<td>9.5</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^- \rightarrow e \nu$, $0.6 &lt;</td>
<td>\eta</td>
<td>&lt; 1.2$</td>
<td>80297.5</td>
<td>33.0</td>
<td>0.0</td>
<td>18.7</td>
<td>11.2</td>
<td>12.8</td>
<td>9.7</td>
<td>3.4</td>
</tr>
<tr>
<td>$W^- \rightarrow e \nu$, $1.8 &lt;</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>80423.8</td>
<td>42.8</td>
<td>0.0</td>
<td>33.2</td>
<td>12.8</td>
<td>35.1</td>
<td>9.9</td>
<td>3.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$W^+ \rightarrow \mu \nu$, $</td>
<td>\eta</td>
<td>&lt; 0.8$</td>
<td>80327.7</td>
<td>22.1</td>
<td>12.2</td>
<td>0.0</td>
<td>2.6</td>
<td>5.1</td>
<td>9.0</td>
<td>6.0</td>
</tr>
<tr>
<td>$W^+ \rightarrow \mu \nu$, $0.8 &lt;</td>
<td>\eta</td>
<td>&lt; 1.4$</td>
<td>80357.3</td>
<td>25.1</td>
<td>19.1</td>
<td>0.0</td>
<td>2.5</td>
<td>4.7</td>
<td>8.9</td>
<td>6.0</td>
</tr>
<tr>
<td>$W^+ \rightarrow \mu \nu$, $1.4 &lt;</td>
<td>\eta</td>
<td>&lt; 2.0$</td>
<td>80446.9</td>
<td>23.9</td>
<td>33.1</td>
<td>0.0</td>
<td>2.5</td>
<td>4.9</td>
<td>8.2</td>
<td>6.0</td>
</tr>
<tr>
<td>$W^+ \rightarrow \mu \nu$, $2.0 &lt;</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>80334.1</td>
<td>34.5</td>
<td>110.1</td>
<td>0.0</td>
<td>2.5</td>
<td>6.4</td>
<td>6.7</td>
<td>6.0</td>
</tr>
<tr>
<td>$W^- \rightarrow \mu \nu$, $</td>
<td>\eta</td>
<td>&lt; 0.8$</td>
<td>80427.8</td>
<td>23.3</td>
<td>11.6</td>
<td>0.0</td>
<td>2.6</td>
<td>5.8</td>
<td>8.1</td>
<td>6.0</td>
</tr>
<tr>
<td>$W^- \rightarrow \mu \nu$, $0.8 &lt;</td>
<td>\eta</td>
<td>&lt; 1.4$</td>
<td>80395.6</td>
<td>27.9</td>
<td>18.3</td>
<td>0.0</td>
<td>2.5</td>
<td>5.6</td>
<td>8.0</td>
<td>6.0</td>
</tr>
<tr>
<td>$W^- \rightarrow \mu \nu$, $1.4 &lt;</td>
<td>\eta</td>
<td>&lt; 2.0$</td>
<td>80380.6</td>
<td>28.1</td>
<td>35.2</td>
<td>0.0</td>
<td>2.6</td>
<td>5.6</td>
<td>8.0</td>
<td>6.0</td>
</tr>
<tr>
<td>$W^- \rightarrow \mu \nu$, $2.0 &lt;</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>80315.2</td>
<td>45.5</td>
<td>116.1</td>
<td>0.0</td>
<td>2.6</td>
<td>7.6</td>
<td>8.3</td>
<td>6.0</td>
</tr>
<tr>
<td>$W^+ \rightarrow e \nu$, $</td>
<td>\eta</td>
<td>&lt; 0.6$</td>
<td>80336.5</td>
<td>22.2</td>
<td>0.0</td>
<td>20.1</td>
<td>2.5</td>
<td>6.4</td>
<td>9.0</td>
<td>5.3</td>
</tr>
<tr>
<td>$W^+ \rightarrow e \nu$, $0.6 &lt;</td>
<td>\eta</td>
<td>&lt; 1.2$</td>
<td>80345.8</td>
<td>22.8</td>
<td>0.0</td>
<td>21.4</td>
<td>2.6</td>
<td>6.7</td>
<td>8.9</td>
<td>5.3</td>
</tr>
<tr>
<td>$W^+ \rightarrow e \nu$, $1.8 &lt;</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>80344.7</td>
<td>24.0</td>
<td>0.0</td>
<td>30.8</td>
<td>2.6</td>
<td>11.9</td>
<td>6.7</td>
<td>5.3</td>
</tr>
<tr>
<td>$W^- \rightarrow e \nu$, $</td>
<td>\eta</td>
<td>&lt; 0.6$</td>
<td>80351.0</td>
<td>23.1</td>
<td>0.0</td>
<td>19.8</td>
<td>2.6</td>
<td>7.2</td>
<td>8.1</td>
<td>5.3</td>
</tr>
<tr>
<td>$W^- \rightarrow e \nu$, $0.6 &lt;</td>
<td>\eta</td>
<td>&lt; 1.2$</td>
<td>80309.8</td>
<td>24.9</td>
<td>0.0</td>
<td>19.7</td>
<td>2.7</td>
<td>7.3</td>
<td>8.0</td>
<td>5.3</td>
</tr>
<tr>
<td>$W^- \rightarrow e \nu$, $1.8 &lt;</td>
<td>\eta</td>
<td>&lt; 2.4$</td>
<td>80413.4</td>
<td>30.1</td>
<td>0.0</td>
<td>30.7</td>
<td>2.7</td>
<td>11.5</td>
<td>8.3</td>
<td>5.3</td>
</tr>
</tbody>
</table>
related between electron and muon decay channels, but significantly anti-correlated between positively and negatively charged \( W \) bosons, as discussed in Sect. 6. Due to the different balance of systematic uncertainties and to the variety of correlation patterns, a significant reduction of the uncertainties in the measurement of \( m_W \) is achieved by combining the different decay channels and the charge and \(|\eta|\) categories.

As discussed in Sect. 2, the comparison of the results from the \( p_T^e \) and \( m_T \) distributions, from the different decay channels, and in the various charge and \(|\eta|\) categories, provides a test of the experimental and physics modelling corrections. Discrepancies between the positively and negatively charged lepton categories, or in the various \(|\eta|\) bins would primarily indicate an insufficient understanding of physics-modelling effects, such as the PDFs and the \( p_T^W \) distribution. Inconsistencies between the electron and muon channels could indicate problems in the calibration of the muon-momentum and electron-energy responses. Significant differences between results from the \( p_T^e \) and \( m_T \) distributions would point to either problems in the calibration of the recoil, or to an incorrect modelling of the transverse-momentum distribution of the \( W \) boson. Several measurement combinations are performed, using the best linear unbiased estimate (BLUE) method \cite{119,120}. The results of the combinations are verified with the HERAverager program \cite{121}, which gives very close results.

Table 11 shows an overview of partial \( m_W \) measurement combinations. In the first step, determinations of \( m_W \) in the electron and muon decay channels from the \( m_T \) distribution are combined separately for the positive- and negative-charge categories, and together for both \( W \)-boson charges. The results are compatible, and the positively charged, negatively charged, and charge-inclusive combinations yield values of \( \chi^2/\text{dof} \) corresponding to 2/6, 7/6, and 11/13, respectively. Compatibility of the results is also observed for the corresponding combinations from the \( p_T^e \) distribution, with values of \( \chi^2/\text{dof} \) of 5/6, 10/6, and 19/13, for positively charged, negatively charged, and charge-inclusive combinations, respectively. The \( \chi^2 \) compatibility test validates the consistency of the results in the \( W \rightarrow e\nu \) and \( W \rightarrow \mu\nu \) decay channels. The precision of the determination of \( m_W \) from the \( m_T \) distribution is slightly worse than the result obtained from the \( p_T^e \) distribution, due to the larger uncertainty induced by the recoil calibration. In addition, the impact of PDF- and \( p_T^W \)-related uncertainties on the \( p_T^e \) fits is limited by the optimisation of the fitting range. In the second step, determinations of \( m_W \) from the \( p_T^e \) and \( m_T \) distributions are combined separately for the electron and the muon decay channels. The results are compatible, with values of \( \chi^2/\text{dof} \) of 4/5 and 8/5 in the electron channel for the \( p_T^e \) and \( m_T \) distributions, respectively, and values of 7/7 and 3/7 in the muon channel for the \( p_T^e \) and \( m_T \) distributions, respectively. The \( m_W \) determinations in the electron and in the muon channels agree, further validating the consistency of the electron and muon calibrations. Agreement between the \( m_W \) determinations from the \( p_T^e \) and \( m_T \) distributions supports the calibration of the recoil, and the modelling of the transverse momentum of the \( W \) boson.

The results are summarised in Fig. 26. The combination of all the determinations of \( m_W \) reported in Table 10 has a value of \( \chi^2/\text{dof} \) of 29/27, and yields a final result of
\[
m_W = 80369.5 \pm 6.8(\text{stat.}) \pm 10.6(\text{exp. syst.}) \pm 13.6(\text{mod. syst.}) \text{ MeV},
\]
where the first uncertainty is statistical, the second corresponds to the experimental systematic uncertainty, and the
Fig. 24 The a, b $p_T^{ℓ}$, c, d $m_T$, and e, f $p_T^{miss}$ distributions for a, c, e $W^+$ events and b, d, f $W^-$ events in the electron decay channel. The data are compared to the simulation including signal and background contributions. Detector calibration and physics-modelling corrections are applied to the simulated events. For all simulated distributions, $m_W$ is set according to the overall measurement result. The lower panels show the data-to-prediction ratios, the error bars show the statistical uncertainty, and the band shows the systematic uncertainty of the prediction. The $\chi^2$ values displayed in each figure account for all sources of uncertainty and include the effects of bin-to-bin correlations induced by the systematic uncertainties.
Fig. 25 The a, b \( \ell_T \), c, d \( m_T \), and e, f \( p_T^{\text{miss}} \) distributions for a, c, e \( W^+ \) events and b, d, f \( W^- \) events in the muon decay channel. The data are compared to the simulation including signal and background contributions. Detector calibration and physics-modelling corrections are applied to the simulated events. For all simulated distributions, \( m_W \) is set according to the overall measurement result. The lower panels show the data-to-prediction ratios, the error bars show the statistical uncertainty, and the band shows the systematic uncertainty of the prediction. The \( \chi^2 \) values displayed in each figure account for all sources of uncertainty and include the effects of bin-to-bin correlations induced by the systematic uncertainties.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_T, W^+, e-\mu$</td>
<td>80370.0</td>
<td>12.3</td>
<td>8.3</td>
<td>6.7</td>
<td>14.5</td>
<td>9.7</td>
<td>9.4</td>
<td>3.4</td>
<td>16.9</td>
<td>30.9</td>
<td>2/6</td>
</tr>
<tr>
<td>$m_T, W^-, e-\mu$</td>
<td>80381.1</td>
<td>13.9</td>
<td>8.8</td>
<td>6.6</td>
<td>11.8</td>
<td>10.2</td>
<td>9.7</td>
<td>3.4</td>
<td>16.2</td>
<td>30.5</td>
<td>7/6</td>
</tr>
<tr>
<td>$m_T, W^\pm, e-\mu$</td>
<td>80375.7</td>
<td>9.6</td>
<td>7.8</td>
<td>5.5</td>
<td>13.0</td>
<td>8.3</td>
<td>9.6</td>
<td>3.4</td>
<td>10.2</td>
<td>25.1</td>
<td>11/13</td>
</tr>
<tr>
<td>$p_T^{\ell}, W^+, e-\mu$</td>
<td>80352.0</td>
<td>9.6</td>
<td>6.5</td>
<td>8.4</td>
<td>2.5</td>
<td>5.2</td>
<td>8.3</td>
<td>5.7</td>
<td>14.5</td>
<td>23.5</td>
<td>5/6</td>
</tr>
<tr>
<td>$p_T^{\ell}, W^-, e-\mu$</td>
<td>80383.4</td>
<td>10.8</td>
<td>7.0</td>
<td>8.1</td>
<td>2.5</td>
<td>6.1</td>
<td>8.1</td>
<td>5.7</td>
<td>13.5</td>
<td>23.6</td>
<td>10/6</td>
</tr>
<tr>
<td>$p_T^{\ell}, W^\pm, e-\mu$</td>
<td>80369.4</td>
<td>7.2</td>
<td>6.3</td>
<td>6.7</td>
<td>2.5</td>
<td>4.6</td>
<td>8.3</td>
<td>5.7</td>
<td>9.0</td>
<td>18.7</td>
<td>19/13</td>
</tr>
<tr>
<td>$p_T^{\ell}, W^\pm, e$</td>
<td>80347.2</td>
<td>9.9</td>
<td>0.0</td>
<td>14.8</td>
<td>2.6</td>
<td>5.7</td>
<td>8.2</td>
<td>5.3</td>
<td>8.9</td>
<td>23.1</td>
<td>4/5</td>
</tr>
<tr>
<td>$m_T, W^\pm, e$</td>
<td>80364.6</td>
<td>13.5</td>
<td>0.0</td>
<td>14.4</td>
<td>13.2</td>
<td>12.8</td>
<td>9.5</td>
<td>3.4</td>
<td>10.2</td>
<td>30.8</td>
<td>8/5</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^+, e$</td>
<td>80345.4</td>
<td>11.7</td>
<td>0.0</td>
<td>16.0</td>
<td>3.8</td>
<td>7.4</td>
<td>8.3</td>
<td>5.0</td>
<td>13.7</td>
<td>27.4</td>
<td>1/5</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^-, e$</td>
<td>80359.4</td>
<td>12.9</td>
<td>0.0</td>
<td>15.1</td>
<td>3.9</td>
<td>8.5</td>
<td>8.4</td>
<td>4.9</td>
<td>13.4</td>
<td>27.6</td>
<td>8/5</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^\pm, e$</td>
<td>80349.8</td>
<td>9.0</td>
<td>0.0</td>
<td>14.7</td>
<td>3.3</td>
<td>6.1</td>
<td>8.3</td>
<td>5.1</td>
<td>9.0</td>
<td>22.9</td>
<td>12/11</td>
</tr>
<tr>
<td>$p_T^{\ell}, W^\pm, \mu$</td>
<td>80382.3</td>
<td>10.1</td>
<td>10.7</td>
<td>0.0</td>
<td>2.5</td>
<td>3.9</td>
<td>8.4</td>
<td>6.0</td>
<td>10.7</td>
<td>21.4</td>
<td>7/7</td>
</tr>
<tr>
<td>$m_T, W^\pm, \mu$</td>
<td>80381.5</td>
<td>13.0</td>
<td>11.6</td>
<td>0.0</td>
<td>13.0</td>
<td>6.0</td>
<td>9.6</td>
<td>3.4</td>
<td>11.2</td>
<td>27.2</td>
<td>3/7</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^+, \mu$</td>
<td>80364.1</td>
<td>11.4</td>
<td>12.4</td>
<td>0.0</td>
<td>4.0</td>
<td>4.7</td>
<td>8.8</td>
<td>5.4</td>
<td>17.6</td>
<td>27.2</td>
<td>5/7</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^-, \mu$</td>
<td>80398.6</td>
<td>12.0</td>
<td>13.0</td>
<td>0.0</td>
<td>4.1</td>
<td>5.7</td>
<td>8.4</td>
<td>5.3</td>
<td>16.8</td>
<td>27.4</td>
<td>3/7</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^\pm, \mu$</td>
<td>80382.0</td>
<td>8.6</td>
<td>10.7</td>
<td>0.0</td>
<td>3.7</td>
<td>4.3</td>
<td>8.6</td>
<td>5.4</td>
<td>10.9</td>
<td>21.0</td>
<td>10/15</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^+, e-\mu$</td>
<td>80352.7</td>
<td>8.9</td>
<td>6.6</td>
<td>8.2</td>
<td>3.1</td>
<td>5.5</td>
<td>8.4</td>
<td>5.4</td>
<td>14.6</td>
<td>23.4</td>
<td>7/13</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^-, e-\mu$</td>
<td>80383.6</td>
<td>9.7</td>
<td>7.2</td>
<td>7.8</td>
<td>3.3</td>
<td>6.6</td>
<td>8.3</td>
<td>5.3</td>
<td>13.6</td>
<td>23.4</td>
<td>15/13</td>
</tr>
<tr>
<td>$m_T-p_T^{\ell}, W^\pm, e-\mu$</td>
<td>80369.5</td>
<td>6.8</td>
<td>6.6</td>
<td>6.4</td>
<td>2.9</td>
<td>4.5</td>
<td>8.3</td>
<td>5.5</td>
<td>9.2</td>
<td>18.5</td>
<td>29/27</td>
</tr>
</tbody>
</table>
11.5 Additional validation tests

The final combination of $m_W$, presented above, depends only on template fits to the $p_T^\ell$ and $m_T$ distributions. As a validation test, the value of $m_W$ is determined from the $p_T^\ell$ distribution, performing a fit in the range $30 < p_T^\ell < 60$ GeV. Consistent results are observed in all measurement categories, leading to combined results of $80364 \pm 26$ (stat) MeV and $80367 \pm 23$ (stat) MeV for the electron and muon channels, respectively.

Several additional studies are performed to validate the stability of the $m_W$ measurement. The stability of the result with respect to different pile-up conditions is tested by dividing the event sample into three bins of $(u_T)$, namely $[2.5, 6.5]$, $[6.5, 9.5]$, and $[9.5, 16]$. In each bin, $m_W$ measurements are performed independently using the $p_T^\ell$ and $m_T$ distributions. This categorisation also tests the stability of $m_W$ with respect to data-taking periods, as the later data-taking periods have on average more pile-up due to the increasing LHC luminosity.

The calibration of the recoil and the modelling of the $p_T^W$ distribution are tested by performing $m_W$ fits in two bins of the recoil corresponding to $[0, 15]$ GeV and $[15, 30]$ GeV, and in two regions corresponding to positive and negative values of $u_T^\ell$. The analysis is also repeated with the $p_T^{\ell\text{miss}}$ requirement removed from the signal selection, leading to a lower recoil modelling uncertainty but a higher multijet background contribution. The stability of the $m_W$ measurements upon removal of this requirement is studied, and consistent results are obtained. All $m_W$ determinations are consistent with the nominal result. An overview of the validation tests is shown in Table 12, where only statistical uncertainties are given. Fitting ranges of $30 < p_T^W < 50$ GeV and $65 < m_T < 100$ GeV are used for all these validation tests, to minimise the statistical uncertainty.

The lower and upper bounds of the range of the $p_T^W$ and $m_T$ distributions are varied as in the optimisation procedure described in Sect. 11.3. The statistical and systematic uncertainties are evaluated for each range, and are only partially correlated between different ranges. Figure 27 shows measured values of $m_W$ for selected ranges of the $p_T^\ell$ and $m_T$ distributions, where only the uncorrelated statistical and systematic uncertainties with respect to the optimal range are shown. The observed variations are all within two standard deviations of the uncorrelated uncertainties, and small compared to the overall uncertainty of the measurement, which is illustrated by the band on Fig. 27. The largest dependence on the kinematic ranges used for the fits is observed for variations of the upper bound of the $p_T^\ell$ distribution in the $W^+ \rightarrow e\nu$ channel, and is related to the shape of the data-to-prediction ratio for this distribution in the region $40 < p_T^\ell < 42$ GeV, as discussed in Sect. 11.3.

The effect of the residual discrepancies in the $u_T$ distributions for $W^- \rightarrow \ell\nu$, visible at low values in Figs. 19-(d) and 20-(d), is estimated by adjusting, in turn, the particle-level $p_T^W$ distribution and the recoil calibration...
Table 12  Summary of consistency tests for the determination of \( m_W \) in several additional measurement categories. The \( \Delta m_W \) values correspond to the difference between the result for each category and the inclusive result for the corresponding observable (\( p_T^W \) or \( m_T \)). The uncertainties correspond to the statistical uncertainty of the fit to the data of each category alone. Fitting ranges of \( 30 < p_T^Z < 50 \text{ GeV} \) and \( 65 < m_T < 100 \text{ GeV} \) are used.

<table>
<thead>
<tr>
<th>Decay channel</th>
<th>Kinematic distribution</th>
<th>( W \rightarrow e\nu )</th>
<th>( W \rightarrow \mu\nu )</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_T^W ) in MeV</td>
<td>( m_T )</td>
<td>( p_T^W )</td>
<td>( m_T )</td>
<td>( p_T^W )</td>
</tr>
<tr>
<td>( (\mu) ) in [2.5, 6.5]</td>
<td>8 ± 14</td>
<td>14 ± 18</td>
<td>-21 ± 12</td>
<td>0 ± 16</td>
</tr>
<tr>
<td>( (\mu) ) in [6.5, 9.5]</td>
<td>-6 ± 16</td>
<td>6 ± 23</td>
<td>12 ± 15</td>
<td>-8 ± 22</td>
</tr>
<tr>
<td>( (\mu) ) in [9.5, 16]</td>
<td>-1 ± 16</td>
<td>3 ± 27</td>
<td>25 ± 16</td>
<td>35 ± 26</td>
</tr>
<tr>
<td>( \Delta p_T^W ) ( m_T ) in [0, 15] GeV</td>
<td>0 ± 11</td>
<td>-8 ± 13</td>
<td>5 ± 10</td>
<td>8 ± 12</td>
</tr>
<tr>
<td>( \Delta p_T^W ) ( m_T ) in [15, 30] GeV</td>
<td>10 ± 15</td>
<td>0 ± 24</td>
<td>-4 ± 14</td>
<td>-18 ± 22</td>
</tr>
<tr>
<td>( p_T^W ) ( m_T ) ( &lt; 0 \text{ GeV} )</td>
<td>8 ± 15</td>
<td>20 ± 17</td>
<td>3 ± 13</td>
<td>-1 ± 16</td>
</tr>
<tr>
<td>( p_T^W ) ( m_T ) ( &gt; 0 \text{ GeV} )</td>
<td>-9 ± 10</td>
<td>1 ± 14</td>
<td>-12 ± 10</td>
<td>10 ± 13</td>
</tr>
<tr>
<td>No ( p_T^\text{miss-cut} )</td>
<td>14 ± 9</td>
<td>-1 ± 13</td>
<td>10 ± 8</td>
<td>-6 ± 12</td>
</tr>
</tbody>
</table>

Fig. 27  Stability of the combined measurement of \( m_W \) with respect to variations of the kinematic ranges of \( a \) \( p_T^W \) and \( b \) \( m_T \) used for the template fits. The optimal \( m_T \) range is used for the \( p_T^W \) variations, and the optimal \( p_T^W \) range is used for the \( m_T \) variations. The effect on the result of symmetric variations of the fitting range boundaries, and its dependence on variations of the lower (upper) boundary for two values of the upper (lower) boundary for \( p_T^W \) (\( m_T \)) are shown. The bands and solid lines respectively show the statistical and total uncertainty on the difference with the optimal result.

11.6 Measurement of \( m_{W^+} - m_{W^-} \)

The results presented in the previous sections can be used to derive a measurement of the mass difference between the positively and negatively charged W bosons, \( m_{W^+} - m_{W^-} \). Starting from the \( m_W \) measurement results in the 28 categories described above, 14 measurements of \( m_{W^+} - m_{W^-} \) can be constructed by subtraction of the results obtained from the \( W^+ \) and \( W^- \) samples in the same decay channel and \( |\eta| \) category. In practice, the \( m_W \) values measured in \( W^+ \) and \( W^- \) events are subtracted linearly, as are the effects of systematic uncertainties on these measurements, while the uncertainty contributions of a statistical nature are added in quadrature. Contrarily to the \( m_W \) measurement discussed above, no blinding procedure was applied for the measurement of \( m_{W^+} - m_{W^-} \).
In this process, uncertainties that are anti-correlated between $W^+$ and $W^-$ largely cancel for the $m_W$ measurement. On the physics-modelling side, the fixed-order PDF uncertainty and the parton shower PDF uncertainty give the largest contributions, while other sources of uncertainty only weakly depend on charge and tend to cancel. Among the sources of uncertainty related to lepton calibration, the track sagitta correction dominates in the muon channel, whereas several residual uncertainties contribute in the electron channel. Most lepton and recoil calibration uncertainties tend to cancel. Background systematic uncertainties contribute as the $Z$ and multijet background fractions differ in the $W^+$ and $W^-$ channels. The dominant statistical uncertainties arise from the size of the data and Monte Carlo signal samples, and of the control samples used to derive the multijet background.

The $m_{W^+} - m_{W^-}$ measurement results are shown in Table 13 for the electron and muon decay channels, and for the combination. The electron channel measurement combines six categories ($p_T^\ell$ and $m_T$ fits in three $|\eta_\ell|$ bins), while the muon channel has four $|\eta_\mu|$ bins and eight categories in total. The fully combined result is

$$m_{W^+} - m_{W^-} = -29.2 \pm 12.8 \text{ (stat.)} \pm 7.0 \text{ (exp. syst.)} \pm 23.9 \text{ (mod. syst.)} \text{ MeV}$$

where the first uncertainty is statistical, the second corresponds to the experimental systematic uncertainty, and the third to the physics-modelling systematic uncertainty.

### 12 Discussion and conclusions

This paper reports a measurement of the $W$-boson mass with the ATLAS detector, obtained through template fits to the kinematic properties of decay leptons in the electron and muon decay channels. The measurement is based on proton–proton collision data recorded in 2011 at a centre-of-mass energy of $\sqrt{s} = 7$ TeV at the LHC, and corresponding to an integrated luminosity of 4.6 fb$^{-1}$. The measurement relies
on a thorough detector calibration based on the study of Z-boson events, leading to a precise modelling of the detector response to electrons, muons and the recoil. Templates for the W-boson kinematic distributions are obtained from the NLO MC generator POWHEG, interfaced to PYTHIA8 for the parton shower. The signal samples are supplemented with several additional physics-modelling corrections allowing for the inclusion of higher-order QCD and electroweak corrections, and by fits to measured distributions, so that agreement between the data and the model in the kinematic distributions is improved. The W-boson mass is obtained from the transverse-momentum distribution of charged leptons and from the transverse-mass distributions, for positively and negatively charged W bosons, in the electron and muon decay channels, and in several kinematic categories. The individual measurements of $m_W$ are found to be consistent and their combination yields a value of

$$m_W = 80370 \pm 7 \text{ (stat.)} \pm 11 \text{ (exp. syst.)}$$

$$= 80370 \pm 19 \text{ MeV},$$

where the first uncertainty is statistical, the second corresponds to the experimental systematic uncertainty, and the third to the physics-modelling systematic uncertainty. A measurement of the $W^+$ and $W^-$ mass difference yields $m_{W^+} - m_{W^-} = -29 \pm 28 \text{ MeV}$.

The W-boson mass measurement is compatible with the current world average of $m_W = 80385 \pm 15 \text{ MeV}$ [29], and similar in precision to the currently leading measurements performed by the CDF and D0 collaborations [22, 23]. An overview of the different $m_W$ measurements is shown in Fig. 28. The compatibility of the measured value of $m_W$ in the context of the global electroweak fit is illustrated in Figs. 29 and 30. Figure 29 compares the present measurement with earlier results, and with the SM prediction updated with regard to Ref. [16] using recent measurements of the top-quark and Higgs boson masses, $m_t = 172.84 \pm 0.70 \text{ GeV}$ [122] and $m_H = 125.09 \pm 0.24 \text{ GeV}$ [123]. This update gives a numerical value for the SM prediction of $m_W = 80356 \pm 8 \text{ MeV}$. The corresponding two-dimensional 68 and 95% confidence limits for $m_W$ and $m_t$ are shown in Fig. 30, and compared to the present measurement of $m_W$ and the average of the top-quark mass determinations performed by ATLAS [122].

The determination of the W-boson mass from the global fit of the electroweak parameters has an uncertainty of 8 MeV, which sets a natural target for the precision of the experimental measurement of the mass of the W boson. The modelling uncertainties, which currently dominate the overall uncertainty of the $m_W$ measurement presented in this paper, need to be reduced in order to fully exploit the larger data samples available at centre-of-mass energies of 8 and 13 TeV. Better knowledge of the PDFs, as achievable with the inclusion in PDF fits of recent precise measurements of $W$- and $Z$-boson rapidity cross sections with the ATLAS detector [41], and improved QCD and electroweak predictions for Drell–Yan production, are therefore crucial for future measurements of the W-boson mass at the LHC.

Acknowledgements We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; SRNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNISW and NCN, Poland; FCT, Portugal; MINE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BRF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and of large non-WLCG
References


