Measurement of lepton differential distributions and the top quark mass in $t\bar{t}$ production in $pp$ collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector

ATLAS Collaboration
CERN, 1211 Geneva 23, Switzerland

Abstract This paper presents single lepton and dilepton kinematic distributions measured in dileptonic $t\bar{t}$ events produced in 20.2 fb$^{-1}$ of $\sqrt{s} = 8$ TeV $pp$ collisions recorded by the ATLAS experiment at the LHC. Both absolute and normalised differential cross-sections are measured, using events with an opposite-charge $e\mu$ pair and one or two $b$-tagged jets. The cross-sections are measured in a fiducial region corresponding to the detector acceptance for leptons, and are compared to the predictions from a variety of Monte Carlo event generators, as well as fixed-order QCD calculations, exploring the sensitivity of the cross-sections to the gluon parton distribution function. Some of the distributions are also sensitive to the top quark pole mass; a combined fit of NLO fixed-order predictions to all the measured distributions yields a top quark mass value of $m_t^{\text{pole}} = 173.2 \pm 0.9 \pm 0.8 \pm 1.2$ GeV, where the three uncertainties arise from data statistics, experimental systematics, and theoretical sources.

1 Introduction

The top quark is the heaviest known fundamental particle, with a mass ($m_t$) that is much larger than any of the other quarks, and close to the scale of electroweak symmetry breaking. The study of its production and decay properties in proton–proton ($pp$) collisions forms an important part of the ATLAS physics program at the CERN Large Hadron Collider (LHC). Due to its large mass and production cross-section, top quark production is also a significant background to many searches for physics beyond the Standard Model, making precise predictions of absolute rates and differential distributions for top quark production a vital tool in fully exploiting the discovery potential of the LHC.

At the LHC, top quarks are primarily produced as quark-antiquark pairs ($q\bar{q}$). The inclusive $t\bar{t}$ production cross-section $\sigma_{t\bar{t}}$ has been calculated at full next-to-next-to-leading-order (NNLO) accuracy in the strong coupling constant $\alpha_S$, including the resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [1–5]. The resulting prediction at a centre-of-mass energy $\sqrt{s} = 8$ TeV is $\sigma_{t\bar{t}} = 252.9 \pm 11.7^{+6.4}_{-8.6}$ pb for a top quark mass of 172.5 GeV, calculated using the top++ 2.0 program [6]. The first uncertainty is due to parton distribution function (PDF) and $\alpha_S$ uncertainties, calculated using the PDF4LHC prescription [7] with the MSTW2008 68% [8,9], CT10 NNLO [10,11] and NNPDF 2.3 f FFN [12] PDF sets, and the second to quantum chromodynamics (QCD) scale variations. This prediction, which has a relative precision of 5.5%, agrees with measurements from ATLAS and CMS at $\sqrt{s} = 8$ TeV [13–15] which have reached a precision of 3–4%. Measurements in LHC $pp$ collisions at $\sqrt{s} = 7$ TeV [13,15] and more recently at $\sqrt{s} = 13$ TeV [16,17] are also in good agreement with the corresponding NNLO+NNLL predictions.

Going beyond the inclusive production cross-section, measurements of $t\bar{t}$ production as a function of the top quark and $t\bar{t}$ system kinematics properties allow the predictions of QCD calculations and Monte Carlo event-generator programs to be probed in more detail. These comparisons are typically more sensitive at the level of normalised differential cross-sections, i.e. shape comparisons, where both experimental and theoretical uncertainties are reduced. Measurements by ATLAS [18–21] and CMS [22–24] have generally demonstrated good agreement with the predictions of leading-order (LO) multi-leg and next-to-leading-order (NLO) event generators and calculations, though the top quark $p_T$ spectrum is measured to be softer than the predictions by both experiments; this distribution appears to be sensitive to the additional corrections contributing at NNLO [25]. Measurements of jet activity in $t\bar{t}$ events [26–29] are also sensitive to gluon radiation and hence the $t\bar{t}$ production.
produced in \( pp \) and normalised differential cross-sections for an electron (\( e^+ \mu^- \nu \bar{\nu} b\bar{b} \) final state), is particularly clean and was exploited to make the most precise ATLAS measurements of \( \sigma_{\gamma f} \) \([13,17]\). The leptons carry information about the underlying top quark kinematics, are free of the uncertainties related to the hadronic part of the final state, and are precisely measured in the detector. Measurements of the \( t\bar{t} \) differential cross-section as a function of the lepton kinematics therefore have the potential to provide a complementary view of \( t\bar{t} \) production and decay dynamics to that provided by the complete reconstruction of the \( t\bar{t} \) final state.

This paper reports such a measurement of the absolute and normalised differential cross-sections for \( t\bar{t} \rightarrow e\mu\nu \bar{\nu} b\bar{b} \) produced in \( pp \) collisions at \( \sqrt{s} = 8 \) TeV, as a function of the kinematics of the single leptons and of the dilepton system. Eight differential cross-section distributions are measured: the transverse momentum \( p_T \) and absolute pseudorapidity \( |\eta| \) of the single leptons (identical for electrons and muons), the \( p_T \), invariant mass and absolute rapidity of the dilepton system \((p_T^e, m_{ee} \) and \( |\eta^e| \)), the azimuthal angle in the transverse plane \( \phi^e\) between the two leptons, the scalar sum \( p_T^e + p_T^\mu \) of the \( p_T \) of the two leptons, and the sum \( E^e + E^\mu \) of the energies of the two leptons.\(^2\) The measurements are corrected to particle level and reported in a fiducial volume where both leptons have \( p_T > 25 \) GeV and \( \eta < 2.5 \), avoiding extrapolations into regions of leptonic phase space which are not measured. The particle-level definition includes the contribution of events where one or both \( W \) bosons decay to electrons or muons via lepton decays of \( \tau \)-leptons (\( \tau \rightarrow W \rightarrow e/\mu \)), but an alternative set of results is provided where the contributions of \( \tau \)-leptons are removed with a correction derived from simulation. The definition of the fiducial volume does not make any requirement on the presence of jets from the hadronic decay products of the \( t\bar{t} \) system. The measurements are made using events with an opposite-charge \( e\mu \) pair and one or two \( b \)-tagged jets, and extrapolated to the fiducial volume (without jet requirements), using an extension of the double-tagging technique used in the inclusive \( t\bar{t} \) cross-section measurement \([13]\). This approach minimises the systematic uncertainties due to the use of jets and \( b \)-tagging in the experimental event selection. Since the lepton kinematics are precisely measured in the ATLAS detector, a simple bin-by-bin correction technique is adequate to correct for efficiency and resolution effects, without the need for a full unfolding procedure.

The results are compared to the predictions of various NLO and LO multi-leg \( t\bar{t} \) event generators, and to fixed-order perturbative QCD predictions from the MCFM \([30]\) program, which is used to explore the sensitivity to PDFs and QCD scale uncertainties. These comparisons are complementary to previous ATLAS analyses exploring how well \( t\bar{t} \) event generators can describe the jet activity \([27]\) and production of extra heavy-flavour jets \([31]\) in the \( \sqrt{s} = 8 \) TeV \( t\bar{t} \) dilepton sample.

Some of the cross-section distributions are sensitive to the top quark mass, as suggested in Ref. \([32]\), and mass measurements are made by comparing the measured distributions to predictions from both NLO plus parton shower event generators and fixed-order QCD calculations. The former are similar to traditional measurements where the top quark mass is reconstructed from its decay products \([33–36]\), but rely only on the leptonic decay products of the \( t\bar{t} \) system and are less sensitive to experimental uncertainties related to the hadronic part of the final state. The measurements based on fixed-order QCD predictions in a well-defined renormalisation scheme correspond more directly to a measurement of the top quark pole mass \( m_{\text{pole}} \), the mass definition corresponding to that of a free particle, which may differ from that measured in direct reconstruction of the decay products by \( O(1 \) GeV \([37–39]\). Previous determinations of \( m_{\text{pole}} \) from inclusive and differential \( t\bar{t} \) cross-section measurements are compatible with the top quark mass measured from direct reconstruction, with uncertainties of 2–3 GeV \([13,15,40,41]\).

The data and Monte Carlo simulation samples used in this analysis are described in Sect. 2, followed by the event reconstruction and selection in Sect. 3, definition and determination of the fiducial differential cross-sections in Sect. 4 and systematic uncertainties in Sect. 5. Results and comparisons with predictions are given in Sect. 6. The ability of the data to constrain the gluon PDF is investigated in Sect. 7 and the determination of the top quark mass is discussed in Sect. 8. Finally, conclusions are given in Sect. 9.

\(^1\) Charge-conjugate decay modes are implied unless otherwise stated.

\(^2\) ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector, and the \( z \) axis along the beam line. Pseudorapidity is defined in terms of the polar angle \( \theta \) as \( \eta = \ln \tan \theta/2 \), and transverse momentum and energy are defined relative to the beamline as \( p_T = p \sin \theta \) and \( E_T = E \sin \theta \). The azimuthal angle around the beam line is denoted by \( \phi \), and distances in \( (\eta, \phi) \) space by \( \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \). The rapidity is defined as \( y = \frac{1}{2} \ln \left( \frac{E + p_T}{E - p_T} \right) \), where \( p_T \) is the \( z \)-component of the momentum and \( E \) is the energy of the relevant object or system.
The ATLAS detector [42] at the LHC covers nearly the entire solid angle around the collision point, and consists of an inner tracking detector surrounded by a thin superconducting solenoid magnet producing a 2 T axial magnetic field, electromagnetic and hadronic calorimeters, and an external muon spectrometer incorporating three large toroidal magnet assemblies. The analysis was performed on a sample of proton–proton collision data at √s = 8 TeV recorded by the ATLAS detector in 2012, corresponding to an integrated luminosity of 20.2 fb⁻¹. Events were required to pass a single-electron or single-muon trigger, with thresholds set to be fully efficient for leptons with p_T > 25 GeV passing offline selections. Each triggered event also includes the signals from on average 20 additional inelastic pp collisions in the same bunch crossing, referred to as pileup.

Monte Carlo simulated event samples were used to develop the analysis procedures, to compare with data, and to evaluate signal efficiencies and background contributions. An overview of the samples used for signal and background modelling is shown in Table 1, and further details are given below. Samples were processed using either the full ATLAS detector simulation [43] based on GEANT4 [44], or a faster simulation making use of parameterised showers in the calorimeters [45]. The effects of pileup were simulated by generating additional inelastic pp collisions with PYTHIA6 [46] using the A2 parameter set (tune) [47] and overlaying them on the primary simulated events. These combined events were then processed using the same reconstruction and analysis chain as the data. Small corrections were applied to the lepton trigger and selection efficiencies better to model the performance measured in data.

2 Data and simulated samples

The baseline simulated t¯t sample was produced using the NLO matrix element event generator POWHEG- BOX v1.0 (referred to hereafter as POWHEG) [48–51] using the CT10 PDFs [10], interfaced to PYTHIA6 (version 6.426) [52] with the CTEQ6L1 PDF set [53] and the Perugia 2011C (P2011C) tune [54] for parton shower, hadronisation and underlying event modelling. This setup provides an NLO QCD prediction of the t¯t production process, a leading-order prediction for the top quark decays, and an approximate treatment of the spin correlations between the quark and antiquark. The POWHEG parameter h_damp, used in the damping function that limits the resummation of higher-order effects incorporated into the Sudakov form factor, was set to m_t. This value was found to give a better modelling of the t¯t system p_T at √s = 7 TeV [55] than the setting of h_damp = ∞ used for the baseline t¯t sample in Ref. [13], which corresponds to no damping.

Alternative t¯t simulation samples used to evaluate systematic uncertainties were generated with POWHEG interfaced to HERWIG (version 6.520) [56,57] with the ATLAS AUET2 tune [58] and JIMMY (version 4.31) [59] for underlying event modelling, with MC@NLO (version 4.01) [60,61] interfaced to HERWIG + JIMMY, and with the leading-order ‘multi-leg’ event generator ALPGEN (version 2.13) [62], also interfaced to HERWIG + JIMMY. The ALPGEN samples used leading-order matrix elements for t¯t production accompanied by up to three additional light partons, and dedicated matrix elements for t¯t plus bb or c¯c production, together with the MLM parton-jet matching scheme [63] to account for double-counting of configurations generated by both the parton shower and matrix-element calculation. The effects of additional radiation in t¯t events were further studied using two additional POWHEG + PYTHIA6 samples, one using the Perugia 2012 radHi tune [54], with h_damp set to 2m_t, and
factorisation and renormalisation scales $\mu_F$ and $\mu_R$ reduced from their event generator defaults by a factor of two, giving more parton shower radiation; and one with the Perugia 2012 radLo tune [54]. $\mu_F$ and $\mu_R$ increased by a factor of two and $h_{\text{lamp}} = m_t$, giving less parton shower radiation. The parameters of these samples were chosen to span the uncertainties in jet observables measured by ATLAS in $t\bar{t}$ events at $\sqrt{s} = 7$ TeV [26,55,64]. The top quark mass was set to 172.5 GeV in all these samples, consistent with recent measurements by ATLAS [35] and CMS [36]. They were all normalised to the NNLO+NNLL cross-section prediction discussed in Sect. 1 when comparing simulation with data. Further $t\bar{t}$ simulation samples with different event generator setups were used for comparisons with the measured differential cross-sections as discussed in Sect. 6.2, and in the extraction of the top quark mass as discussed in Sect. 8.

Backgrounds to the $t\bar{t}$ event selection are classified into two types: those with two real prompt leptons from $W$ or $Z$ boson decays (including those produced via leptonic $\tau$ decays), and those where one of the reconstructed lepton candidates is misidentified, i.e. a non-prompt lepton from the decay of a bottom or charm hadron, an electron from a photon conversion, hadronic jet activity misidentified as an electron, or a muon produced from the decay in flight of a pion or kaon. The first category is dominated by the associated production of a $W$ boson and a single top quark, $Wt$, that is simulated using POWHEG + PYTHIA6 with the CT10 PDFs and the P2011C tune. The ‘diagram removal’ scheme was used to handle the interference between the $t\bar{t}$ and $Wt$ final states that occurs at NLO [65,66]. Smaller backgrounds result from $Z \rightarrow \tau \tau (\rightarrow e\mu)+$jets, modelled using ALPGEN + PYTHIA6 including leading-order matrix elements for $Zb\bar{b}$ production, and diboson ($WW$, $WZ$ and $ZZ$) production in association with jets, modelled with ALPGEN + HERWIG + JIMMY. The $Wt$ background was normalised to the approximate NNLO cross-section of 22.4 ± 1.5 pb, determined as in Ref. [67]. The inclusive $Z$ cross-section was set to the NNLO prediction from FEWZ [68], but the normalisation of the $Z \rightarrow \tau \tau$ background with $b$-tagged jets was determined with the help of data control samples as discussed in Sect. 4.2. The small diboson background was normalised to the NLO QCD inclusive cross-section predictions calculated with MCFM [69], using the ALPGEN + HERWIG prediction for the fraction of diboson events with extra jets. Production of $t\bar{t}$ in association with a $W$ or $Z$ boson, which contributes to the control sample with two same-charge leptons, was simulated with MADGRAPH [70] interfaced to PYTHIA6 with CTEQ6L1 PDFs, and normalised to NLO cross-section predictions [71,72].

Backgrounds with one real and one misidentified lepton arise from $t\bar{t}$ events with one hadronically-decaying $W$; $W+$jets production, modelled as described above for $Z+$jets; $W\gamma+$jets, modelled with SHERPA 1.4.1 [73] with CT10 PDFs; and $t$-channel single top production, modelled with ACERMC [74] with the AUET2B tune [75] and CTEQ6L1 PDFs interfaced to PYTHIA6. The normalisations of these backgrounds in the opposite-charge $e\mu$ samples were determined with the help of the corresponding same-charge $e\mu$ samples in data. Other backgrounds, including processes with two misidentified leptons, are negligible after the event selections used in this analysis.

3 Event reconstruction and selection

The analysis makes use of reconstructed electrons, muons, and $b$-tagged jets, selected exactly as described in Ref. [13]. In brief, electron candidates [76] were required to satisfy $E_T > 25$ GeV and $|\eta| < 2.47$, and to not lie within the transition region $1.37 < |\eta| < 1.52$ between the barrel and endcap electromagnetic calorimeters. Muon candidates [77] were required to satisfy $p_T > 25$ GeV and $|\eta| < 2.5$. In order to reduce background from non-prompt leptons, electrons were required to be isolated from nearby hadronic activity using both calorimeter and tracking information, and muons were required to be isolated using tracking information alone. Jets were reconstructed using the anti-$k_t$ algorithm [78,79] with radius parameter $R = 0.4$ using calorimeter energy clusters calibrated with the local cluster weighting method [80]. Jets were further calibrated using information from both simulation and data [81,82], and required to satisfy $p_T > 25$ GeV and $|\eta| < 2.5$. Jets satisfying $p_T < 50$ GeV and $|\eta| < 2.4$ were additionally required to pass pileup rejection criteria based on their associated tracks [82]. To further suppress non-isolated leptons likely to originate from heavy-flavour decays within jets, electron and muon candidates within $\Delta R < 0.4$ of selected jets were discarded. Finally, jets likely to contain $b$-hadrons were $b$-tagged using the MV1 algorithm [83], a multivariate discriminant making use of track impact parameters and reconstructed secondary vertices. A tagging working point corresponding to a 70% efficiency for tagging $b$-quark jets from top decays in $t\bar{t}$ events was used, giving a rejection factor of about 140 against light-quark and gluon jets, and about five against jets originating from charm quarks.

As in Ref. [13], events were required to have at least one reconstructed primary vertex and to have no jets with $p_T > 20$ GeV failing jet quality requirements [81]. Events having muons compatible with cosmic-ray interactions or losing substantial energy following bremsstrahlung in the calorimeter material were rejected. A preselection requiring exactly one electron and one muon selected as described above was then applied, requiring at least one selected lepton to be matched to a corresponding electron or muon.

3 The reconstructed vertex with the largest sum of $p_T^2$ for the constituent tracks was selected as the primary vertex.
Table 2. Observed numbers of opposite-sign $e\mu$ events with one and two $b$-tagged jets ($N_1$ and $N_2$) together with the estimates of backgrounds and associated total uncertainties described in Sect. 5.

<table>
<thead>
<tr>
<th>Event counts</th>
<th>$N_1$</th>
<th>$N_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>21666</td>
<td>11739</td>
</tr>
<tr>
<td>$Wt$ single top</td>
<td>2080 ± 210</td>
<td>350 ± 120</td>
</tr>
<tr>
<td>$Z(\rightarrow \tau \tau \rightarrow e\mu)$+jets</td>
<td>210 ± 40</td>
<td>7 ± 2</td>
</tr>
<tr>
<td>Diboson</td>
<td>120 ± 30</td>
<td>3 ± 1</td>
</tr>
<tr>
<td>Misidentified leptons</td>
<td>220 ± 80</td>
<td>78 ± 50</td>
</tr>
<tr>
<td>Total background</td>
<td>2630 ± 230</td>
<td>440 ± 130</td>
</tr>
</tbody>
</table>

trigger signature. Events with an opposite-charge-sign $e\mu$ pair formed the main analysis sample, with events having a same-sign pair being used to estimate the background from misidentified leptons.

A total of 66,453 data events passed the opposite-sign $e\mu$ preselection. Events were then further sub-divided according to the number of $b$-tagged jets, irrespective of the number of untagged jets, and events having one or two $b$-tagged jets were retained for further analysis. The numbers of one and two $b$-tagged jet events selected in data are shown in Table 2, compared with expected non-$t\bar{t}$ contributions from $Wt$ and dibosons evaluated from simulation, and $Z(\rightarrow \tau \tau \rightarrow e\mu)+$jets and misidentified leptons evaluated from data and simulation, as discussed in detail in Sects. 4.2 and 5 below.

In simulation, the one $b$-tagged sample is about 88% pure and the two $b$-tagged sample 96% pure in $t\bar{t}$ events, with the largest backgrounds coming from $Wt$ production in both cases. The distribution of the number of $b$-tagged jets in preselected opposite-sign $e\mu$ events is shown in Fig. 1a, compared to the predictions from simulation using POWHEG + PYTHIA6 ($PW$), MC@NLO + HERWIG (HW) and ALPGEN + HERWIG $t\bar{t}$ samples, normalising the total simulation prediction in each case using the integrated luminosity of the data sample. The distributions of the $p_T$ of $b$-tagged jets, and the reconstructed electron and muon $p_T$ and $|\eta|$ in events with at least one $b$-tagged jet are shown in Fig. 1b–f, with the total simulation prediction normalised to the same number of events as the data to facilitate shape comparisons. The distributions of the reconstructed dilepton variables $p_{\ell\ell}^e$, $m_{\ell\ell}$, $|y_{\ell\ell}|$, $|L_{\ell\ell}|$, $|\Delta\phi_{\ell\ell}|$, $p_{T\ell\ell}^e$ and $E_{T\ell\ell}^e$ and $E_{T\ell\ell}^\mu$ are shown in Fig. 2, with the simulation normalised as for Fig. 1b–f. In general the data are well described by the predictions using the different $t\bar{t}$ models, but a few differences are visible. The lepton $p_T$ spectra are softer in data than in simulation, the lepton $|\eta|$, and dilepton $|y_{\ell\ell}|$ distributions are more central than the POWHEG + PYTHIA6 and MC@NLO + HERWIG predictions.

4 The background event counts and uncertainties shown in Table 2 differ from those in Ref. [13] due to the use of different simulation samples and the estimation of the background in bins of lepton kinematic variables.

and the $\Delta\phi_{\ell\ell}^e$ distribution is slightly flatter in data than in all the predictions.

4 Fiducial cross-section determination

The cross-section measurements were made for a fiducial region, where the particle-level electron and muon were required to have opposite charge signs, to each come from $W$ decays either directly or via $W \rightarrow \tau \rightarrow e/\mu$ and to each satisfy $p_{T\ell} > 25\text{GeV}$ and $|\eta| < 2.5$. The lepton four-momenta were taken after final-state radiation, and ‘dressed’ by including the four-momenta of any photons within a cone of size $\Delta R = 0.1$ around the lepton direction, excluding photons produced from hadronic decays or interactions with the detector material. The total cross-section within this fiducial volume corresponds to the fiducial cross-section measured in Ref. [13]. According to the predictions of the baseline POWHEG+PYTHIA6 $t\bar{t}$ simulation, it is about 44% of the total $t\bar{t} \rightarrow e\mu\nu\bar{\nu}b\bar{b}$ cross-section without restrictions on the lepton acceptance and including contributions via $W \rightarrow \tau \rightarrow e/\mu$.

4.1 Cross-section extraction

The differential cross-sections were measured using an extension of the technique used in Ref. [13], counting the number of leptons or events with one ($N_1^i$) or two ($N_2^i$) $b$-tagged jets where the lepton(s) fall in bin $i$ of a differential distribution at reconstruction level. For the single-lepton distributions $p_{T\ell}$ and $|\eta|$, there are two counts per event, in the two bins corresponding to the electron and muon. For the dilepton distributions, each event contributes a single count corresponding to the bin in which the appropriate dilepton variable falls. For each measured distribution, these counts satisfy the tagging equations:

$$N_1^i = L\sigma_{1i}^e G_{e\mu}^i 2\epsilon_h(1 - C_h\epsilon_h) + N_{1,bkg}^i,$$

$$N_2^i = L\sigma_{2i}^e G_{e\mu}^i C_h(\epsilon_h^2) + N_{2,bkg}^i,$$

where $\sigma_{ii}^e$ is the absolute fiducial differential cross-section in bin $i$, and $L$ is the integrated luminosity of the sample. The reconstruction efficiency $G_{e\mu}^i$ represents the ratio of the number of reconstructed $e\mu$ events (or leptons for $p_{T\ell}$ and $|\eta|$) falling in bin $i$ at reconstruction level to the number of true $e\mu$ events (or leptons) falling in the same bin at particle level, evaluated using $t\bar{t}$ simulation without making any requirements on reconstructed or particle-level jets. It therefore corrects for both the lepton reconstruction efficiency and bin migration, where events corresponding to bin $j$ at particle level appear in a different bin $i \neq j$ at reconstruction level. The values of $G_{e\mu}^i$ in simulation are typically in the
Fig. 1 Distributions of (a) the number of $b$-tagged jets in preselected opposite-sign $e\mu$ events; and (b) the $p_T$ of $b$-tagged jets, (c) the $p_T$ of the electron, (d) the $|\eta|$ of the electron, (e) the $p_T$ of the muon and (f) the $|\eta|$ of the muon, in events with an opposite-sign $e\mu$ pair and at least one $b$-tagged jet. The reconstruction-level data are compared to the expectation from simulation, broken down into contributions from $t\bar{t}$ (Powheg+Pythia6), single top, $Z$+jets, dibosons, and events with misidentified electrons or muons. The simulation prediction is normalised to the same integrated luminosity as the data in (a) and to the same number of entries as the data in (b)–(f). The lower parts of the figure show the ratios of simulation to data, using various $t\bar{t}$ signal samples and with the cyan band indicating the data statistical uncertainty. The last bin includes the overflow in panels (b), (c) and (e).
Fig. 2 Distributions of a the dilepton $p_T^{\mu\mu}$, b invariant mass $m^{\mu\mu}$, c rapidity $|y^{\mu\mu}|$, d azimuthal angle difference $\Delta \phi^{\mu\mu}$, e lepton $p_T$ sum $p_T + p_T^e$ and f lepton energy sum $E + E^\mu$, in events with an opposite-sign $e\mu$ pair and at least one $b$-tagged jet. The reconstruction-level data are compared to the expectation from simulation, broken down into contributions from $t\bar{t}$ (POWHEG + PYTHIA6), single top, $Z$+jets, dibosons, and events with misidentified electrons or muons, normalised to the same number of entries as the data. The lower parts of the figure show the ratios of simulation to data, using various $t\bar{t}$ signal samples and with the cyan band indicating the data statistical uncertainty. The last bin includes the overflow in panels a, b, e and f.
range 0.5–0.6, with some dependence on lepton kinematics due to the varying reconstruction efficiencies with lepton $|\eta|$ and $p_T$, and the effect of isolation requirements when the leptons are close together in the detector.

The efficiency $\epsilon_i$ represents the combined probability for a jet from the quark $q$ in the $t \rightarrow Wq$ decay to fall within the detector acceptance, be reconstructed as a jet with $p_T > 25\text{ GeV}$ and be tagged as a $b$-jet. Although this quark is almost always a $b$-quark, $\epsilon_i$ also accounts for the 0.2% of top quarks that decay to $W$s or $Wd$. If the kinematics of the two $b$ quarks produced in the top quark decays are uncorrelated, the probability to tag both is given by $\epsilon_{bb} = (\epsilon_b)^2$. In practice, small correlations are present, for example due to kinematic correlations between the $b$-jets from the top quark decays, or extra $b\bar{b}$ or $c\bar{c}$ pairs produced in association with the $t\bar{t}$ system [13]. Their effects are corrected via the tagging correlation coefficient $C_b^i = \epsilon_{bb} / (\epsilon_b)^2$, whose values are taken from $t\bar{t}$ simulation. They depend slightly on the bin $i$ of the dilepton system but are always within 1–2% of unity, even for the bins at the edges of the differential distributions. The correlation $C_b^i$ also corrects for the small effects on $N_1^i$ and $N_2^i$ and $\tilde{\epsilon}_i$ of the small fraction of $t\bar{t}$ events which have additional $b$ quarks produced in association with the $t\bar{t}$ system, and the even smaller effects from mistagged light quark, charm or gluon jets in $t\bar{t}$ events. This formalism involving $\epsilon_b^i$ and $C_b^i$ allows the fraction of top quarks where the jet was not reconstructed to be inferred from the counts $N_1^i$ and $N_2^i$, minimising the exposure to systematic uncertainties from jet measurements and $b$-tagging, and allowing the fiducial cross-sections $\sigma_{f i}^i$ to be defined with no requirements on the jets in the final state.

Backgrounds from sources other than $t\bar{t} \rightarrow e\mu\nu\bar{\nu}b\bar{b}$ events also contribute to the counts $N_1^i$ and $N_2^i$, and are represented by the terms $N_1^{i,\text{bkg}}$ and $N_2^{i,\text{bkg}}$ in Eq. (1). These contributions were evaluated using a combination of simulation- and data-based methods as discussed in Sect. 4.2 below.

The tagging equations were solved numerically in each bin $i$ of each differential distribution separately. The bin ranges for each distribution were chosen according to the experimental resolution, minimising the bin-to-bin migration by keeping the bin purities (the fractions of reconstructed events in bin $i$ that originate from events which are also in bin $i$ at particle level) above about 0.9. The resolution on the reconstructed kinematic quantities is dominated by the electron energy and muon momentum measurements, and the purities for the distributions which depend mainly on angular variables are higher, around 0.96 for $|y^e\tilde{\mu}|$ and 0.99 for $|\eta^e\tilde{\mu}|$ and $\Delta\phi^e\tilde{\mu}$. For these distributions, the bin ranges were chosen so as to give about ten bins for each distribution. The bin range choices for all distributions can be seen in Tables 3, 4, 5 and 6 in Sect. 6, and the last bin of the $p_T^e$, $p_T^{\tilde{\mu}}$, $m^{e\tilde{\mu}}$, $E^e + E^{\tilde{\mu}}$ distributions includes overflow events falling above the last bin boundary, indicated by the ‘+’ sign after the upper bin limit.

The normalised fiducial differential cross-section distributions $s_{f i}^i$ were calculated from the absolute cross-sections $\sigma_{f i}^i$ determined from Eq. (1) as follows:

$$s_{f i}^i = \frac{\sigma_{f i}^i}{\Sigma_j \sigma_{f j}^i} = \frac{\sigma_{f i}^i}{\sigma_{f \text{bd}}^i},$$

where $\sigma_{f \text{bd}}^i$ is the total cross-section summed over all bins of the fiducial region. The $s_{f i}^i$ values are divided by the bin widths $W_i$, to produce the cross-sections differential in the variable $x (x = p_T^e, |\eta^e|$, etc.):

$$\frac{1}{\sigma} \left( \frac{d\sigma}{dx} \right)^{f i} = s_{f i}^i W_i.$$

The normalisation condition in Eq. (2) induces a statistical correlation between the normalised measurements in each bin. The absolute dilepton cross-section measurements are not statistically correlated between bins, but kinematic correlations between the electron and muon in each event induce small statistical correlations between bins of the absolute single lepton $p_T^e$ and $|\eta^e|$ distributions, as discussed in Sect. 4.3 below.

The measured cross-sections include contributions where one or both leptons are produced via leptonic tau decays ($t \rightarrow W \rightarrow \tau \rightarrow e/\mu$), but the fixed-order predictions discussed in Sect. 6.3 only include the direct decays $t \rightarrow W \rightarrow e/\mu$. To allow comparison with such predictions, a second set of cross-section results were derived with a bin-by-bin multiplicative correction $f^e_i$ to remove the $\tau$ contributions:

$$\sigma_{f i}^i (\text{no-}\tau) = f^e_i \sigma_{f i}^i,$$

and similarly for the normalised cross-sections $s_{f i}^i (\text{no-}\tau)$.

The corrections $f^e_i$ were evaluated from the baseline POWHEG + PYTHIA6 $t\bar{t}$ simulation and are typically close to 0.9, decreasing to 0.8–0.85 at low lepton $p_T$.

4.2 Background estimates

The $Wt$ single top and diboson backgrounds were estimated from simulation using the samples discussed in Sect. 2, whilst the $Z$+jets background (with $Z \rightarrow \tau \tau \rightarrow e\mu4\nu$) and the contribution from events with one real and one misidentified lepton were estimated using both simulation and data as discussed below. The backgrounds in both the one and two $b$-tagged samples are dominated by $Wt$ (see Table 2). The total background fraction (i.e. the predicted fraction of events in each bin which do not come from $t\bar{t}$ with two real prompt leptons) varies significantly as a function of some
of the differential variables, as shown in Fig. 3. This variation is taken into account by estimating the background contributions \( N_{j}^{1, \text{bkg}} \) and \( N_{2}^{1, \text{bkg}} \) separately in each bin of each differential distribution.

The production cross-sections for \( Z \) bosons accompanied by heavy-flavour jets are subject to large theoretical uncertainties. The background predictions from \( \text{ALPGEN} + \text{PYTHIA6} \) in each bin of each distribution were therefore normalised from data, by multiplying them by constant scale factors of \( 1.4 \pm 0.2 \) for the one \( b \)-tagged jet sample and \( 1.1 \pm 0.3 \) for the two \( b \)-tagged jet sample. These scale factors were derived from the comparison of data and simulated event yields for \( Z \rightarrow \ell\ell \) and \( Z \rightarrow \mu\mu \) plus one or two \( b \)-tagged jets, inclusively for all lepton pairs passing the kinematic selections for electrons and muons [13]. The uncertainties are dominated by the dependence of the scale factors on lepton kinematics, investigated by studying their variation with \( Z \)-boson \( p_T \), reconstructed from the \( \ell\ell \) or \( \mu\mu \) system.

The background from events with one real and one misidentified lepton was estimated using a combination of data and simulation in control regions with an electron and muon of the same charge [13]. Simulation studies showed that the samples with a same-sign electron pair and one or two \( b \)-tagged jets are dominated by events with a misidentified lepton, with rates and kinematic distributions similar to those in the opposite-sign sample. The distributions of the dilepton kinematic variables for same-sign events with at least one \( b \)-tagged jet in data are shown in Fig. 4, and compared with the predictions from simulation. The expected contributions are shown separately for events with two prompt leptons, events where the electron candidate originates from a converted photon radiated from an electron produced in a top quark decay, events with a converted photon from other sources, and events where the electron or muon originates from the decay of a bottom or charm hadron. The analogous distributions for the electron and muon \( p_T \) and \( |\eta| \) are shown in Ref. [13]. In general, the simulation models the rates and kinematic distributions of the same-sign events well. The modelling of misidentified leptons was further tested in control samples where either the electron or muon isolation requirements were relaxed in order to enhance the contributions from heavy-flavour decays, and similar levels of agreement were observed.

The contributions \( N_{j}^{\text{mis-id}} \) of events with misidentified leptons to the opposite-sign samples with \( j = 1, 2 \) \( b \)-tagged jets were estimated in each bin \( i \) of each distribution using

\[
N_{j}^{i, \text{mis-id}} = R_{j}^{i}(N_{j}^{i, \text{data, SS}} - N_{j}^{i, \text{prompt, SS}}),
\]

\[
R_{j}^{i} = \frac{N_{j}^{i, \text{mis-id, OS}}}{N_{j}^{i, \text{data, SS}}}, \tag{4}
\]

where \( N_{j}^{i, \text{data, SS}} \) is the number of observed same-sign events in bin \( i \) with \( j \) \( b \)-tagged jets, \( N_{j}^{i, \text{prompt, SS}} \) is the estimated number of events in this bin with two prompt leptons, and \( R_{j}^{i} \) is the ratio of the number of opposite- to same-sign events with misidentified leptons in bin \( i \) with \( j \) \( b \)-tagged jets. This formalism uses the observed data same-sign event rate in each bin to predict the corresponding opposite-sign contribution from misidentified leptons. It relies on simulation to predict the ratios of opposite- to same-sign rates and the prompt same-sign contribution, but not the absolute normalisation of misidentified leptons. The prompt-lepton contribution in Eq. (4) comes mainly from semileptonic \( t\bar{t} \) events with an additional \( W \) or \( Z \) boson, diboson events with two same-sign leptons, and \( t\bar{t} \rightarrow e\mu W\bar{W}b\bar{b} \) events where the electron charge was misreconstructed. These components were evaluated directly from simulation in each bin \( (i, j) \), and an uncertainty of \( \pm 50\% \) was assigned [13]. The values of \( R_{j}^{i} \) were taken from simulation, separately for each differential distribution and \( j = 1 \) and \( 2 \) \( b \)-tagged jets, and averaged over several consecutive bins \( i \) in order to reduce statistical fluctuations. The values of \( R_{1}^{i} \) range from 0.8 to 1.5, and \( R_{2}^{i} \) from 1.2 to 2.0, as the predicted background composition changes across the kinematic distributions. As in Ref. [13], uncertainties of \( \pm 0.25 \) and \( 0.5 \) were assigned to \( R_{1}^{i} \) and \( R_{2}^{i} \), based on the variation of \( R_{j}^{i} \) for different components of the misidentified lepton background, and taken to be correlated across all bins \( (i, j) \).

### 4.3 Validation of the analysis procedure

The method for the differential cross-section determination was tested on simulated events in order to check for biases and determine the expected statistical uncertainties. Pseudo-data samples corresponding to the data integrated luminosity were produced by varying the event counts \( N_{1}^{1} \) and \( N_{2}^{1} \) in each bin \( i \) independently, according to Poisson distributions with mean values predicted from a chosen \( t\bar{t} \) simulation sample plus non-\( t\bar{t} \) backgrounds. The tagging equations Eq. (1) were then solved for each pseudo-experiment using the values of \( G_{i}^{t\bar{t}}, C_{b}^{i}, N_{1}^{i, \text{bkg}}, N_{2}^{i, \text{bkg}} \) calculated with the baseline simulation samples. An initial set of 1000 pseudo-experiments was performed using the baseline simulation sample as a reference, and the mean and RMS width of the deviations of the result in each bin from the reference values were used to validate the analysis procedure. The black points in Fig. 5 show the mean deviation of the results (averaged over all pseudo-experiments) for four of the measured normalised distributions, with error bars corresponding to the uncertainty in the mean due to the finite size of the simulation samples (about 17 times the data integrated luminosity). The residual biases of the mean deviations away from the reference are compatible with zero and in all cases much smaller than...
Fig. 3 Estimated background fractions in the one and two $b$-tagged samples as functions of each lepton and dilepton differential variable, estimated from simulation alone. The error bars correspond to the statistical uncertainties of the simulation samples, and are often smaller than the marker size.
Fig. 4 Distributions of 

- a the dilepton $p_T^{e\mu}$, 
- b invariant mass $m^{e\mu}$, 
- c rapidity $|y^{e\mu}|$, 
- d azimuthal angle difference $\Delta \phi^{e\mu}$, 
- e lepton $p_T$ sum $p_T^e + p_T^\mu$, and 
- f lepton energy sum $E^e + E^\mu$, in events with a same-sign $e\mu$ pair and at least one $b$-tagged jet. The simulation prediction is normalised to the same integrated luminosity as the data, and broken down into contributions where both leptons are prompt, or one is a misidentified lepton from a photon conversion originating from a top quark decay or from background, or from heavy-flavour decay. In the $p_T^{e\mu}$, $m^{e\mu}$, $p_T^e + p_T^\mu$ and $E^e + E^\mu$ distributions, the last bin includes the overflows.
the expected statistical uncertainties in data, measured by the RMS widths and shown by the cyan bands. Similar results were obtained for the other normalised differential cross-section distributions, and for the absolute distributions. The pull distributions (i.e. the distributions of deviations divided by the estimated statistical uncertainty from each pseudo-experiment) were also found to have widths within a few percent of unity. The $\chi^2$ values for the compatibility of each measured distribution with the reference were also calculated for each pseudo-experiment and the distribution of the corresponding $p$-values across all pseudo-experiments was found to be uniform between zero and one. These tests confirm that the analysis procedure is unbiased and correctly estimates the statistical uncertainties in each bin of each distribution.

Additional pseudo-experiments were performed to test the ability of the analysis procedure to reconstruct distributions different from the reference, taking the values of $C_{e_1}^{l_{\ell}}$, $C_{b}^{l_{\ell}}$, $N_{1_{\text{bkg}}}^{l_{\ell}}$ and $N_{2_{\text{bkg}}}^{l_{\ell}}$ from the baseline samples. Tests were conducted using simulated POWHEG + PYTHIA6 and MC@NLO + HERWIG $t\bar{t}$ samples with different top mass values, a POWHEG + PYTHIA6 sample generated using the HERAPDF 1.5 [84,85] PDF set instead of CT10, and a POWHEG + PYTHIA6 sample reweighted to reproduce the top quark $p_T$ distribution calculated at NNLO from Ref. [25]. In all cases, the analysis procedure recovered the true distributions from the alternative samples within the statistical precision of the test, demonstrating the adequacy of the bin-by-bin correction procedure without the need for iteration or a more sophisticated matrix-based unfolding technique. Some examples are shown by the red points and dotted lines in Fig. 5, for an alternative sample with $m_t = 165$ GeV for $p_T$ and $p_T^{\mu\mu}$, and for HERAPDF 1.5 for $|\eta|^{e\mu}$ and $|\eta|^{\mu\mu}$, both simulation samples having about twice the statistics of the data. These figures also demonstrate the sensitivities of some of the measured distributions to $m_t$ and different PDFs.

For the single-lepton distributions $p_T^{\ell}$ and $|\eta|^\ell$, which have two entries per event, the formalism of Eq. (1) and the pseudo-experiments generated by fluctuating each bin independently

---

**Fig. 5** Results of pseudo-experiment studies on simulated events for the extraction of the normalised differential cross-section distributions for (a) $p_T^e$, (b) $p_T^\mu$, (c) $|\eta|^e$, and (d) $|\eta|^\mu$, shown as relative deviations $(\sigma - \sigma_{\text{ref}})/\sigma_{\text{ref}}$ from the reference cross-section values in the baseline POWHEG+PYTHIA6 CT10 sample with $m_t = 172.5$ GeV. The black points show the mean deviations from the reference when fitting pseudo-data samples generated with the baseline simulation sample, with error bars indicating the uncertainties due to the limited number of simulated events. The cyan bands indicate the expected statistical uncertainties for a single sample corresponding to the data integrated luminosity. The open red points show the mean deviations from the reference values when fitting pseudo-experiments generated from alternative simulation samples with $m_t = 165$ GeV (a, b) or with the HERAPDF 1.5 PDF (c, d), with error bars due to the limited size of these alternative samples. The red dotted lines show the true deviations from the reference in the alternative samples.
do not take into account correlations between the kinematics of the electron and muon in each event. This effect was checked by generating pseudo-data samples corresponding to the data integrated luminosity from individual simulated events, taken at random from a large $tt$ sample combining both full and fast simulation and corresponding to about 70 times the data integrated luminosity. The effect of neglecting the electron-muon correlations within an event was found to correspond to at most a 2% fractional overestimate of the absolute and 2% fractional underestimate of the normalised cross-section uncertainties. Hence, no corresponding corrections to the statistical uncertainties were made.

5 Systematic uncertainties

Systematic uncertainties in the measured cross-sections arise from uncertainties in the values of the input quantities $G_{ee}$, $C_b$, $N_1$, $N_2$, and $L$ used in Eq. (1). Each source of systematic uncertainty was evaluated by coherently changing the values of all relevant input quantities and re-solving Eq. (1), thus taking into account correlations of the uncertainties in e.g. $G_{ee}$ and $C_b$. The uncertainties are divided into five groups ($t\bar{t}$ modelling, leptons, jets/hadronisation, background/energy uncertainty) and are discussed in Sects. 5.1–5.5. The resulting relative uncertainties in each measured differential cross-section value are shown in the results Tables 3, 4, 5 and 6, and the grouped systematic uncertainties for the normalised differential cross-sections are shown in Fig. 6, together with the statistical and total uncertainties.

5.1 $t\bar{t}$ modelling

The uncertainties in $G_{ee}$ and $C_b$ (and $f_{\mu}$ for the $\tau$-corrected cross-sections) were evaluated using the various alternative $t\bar{t}$ simulation samples detailed in Sect. 2.

$t\bar{t}$ generator: Event generator uncertainties were evaluated by comparing the baseline POWHEG + PYTHIA6 $t\bar{t}$ sample (with $h_{\text{damp}} = m_t$) with alternative samples generated with MC@NLO interfaced to HERWIG (thus changing both the NLO hard-scattering event generator and the parton shower, hadronisation and underlying event model), and with the LO multi-leg event generator ALPGEN, also interfaced to HERWIG. The bin-by-bin shifts in $G_{ee}$ and $C_b$ were fitted with polynomial functions in order to reduce statistical fluctuations caused by the limited size of the simulated samples, and the larger of the differences between the baseline and the two alternative samples was taken in each bin to define the generator uncertainty. As also found in the inclusive cross-section analysis, a substantial part of the differences in $G_{ee}$ in the various samples arises from differences in the hadronic activity close to the leptons, which affects the efficiency of the lepton isolation requirements. These efficiencies were therefore measured in situ in $t\bar{t}$ events selected in data as discussed in Sect. 5.2 below, and the simulation uncertainties on $G_{ee}$ evaluated by considering the lepton reconstruction, identification and lepton-jet overlap requirements only. The resulting uncertainties on $G_{ee}$ are typically 0.5–1% in most regions of the phase space, varying only slightly as a function of the lepton and dilepton kinematics. The same procedure was used to evaluate uncertainties in $C_b$, and the predictions of the three simulation samples were found to agree at the 0.5–1% level, giving similar predictions for the variations of $C_b$ across the bins of the various measured distributions. Alternative $t\bar{t}$ samples generated with POWHEG + PYTHIA6, POWHEG + HERWIG (both with $h_{\text{damp}} = \infty$) were also considered, but the resulting differences in $G_{ee}$ and $C_b$ were found to be significantly less than those from the comparisons with MC@NLO + HERWIG and thus no additional uncertainty was assigned. Variations in the predictions of $f_{\mu}$ from the three $t\bar{t}$ samples were found to be at the 0.2% level, and were also taken into account for the $\tau$-corrected cross-section results.

Initial/final-state radiation: The effects on $G_{ee}$, $C_b$ and $f_{\mu}$ of uncertainties in the modelling of additional radiation in $t\bar{t}$ events were assessed as half the difference between POWHEG + PYTHIA6 samples tuned to span the uncertainties in jet activity measured in $\sqrt{s} = 7\text{ TeV}$ ATLAS data [26,55,64], as discussed in Sect. 2. The uncertainties were taken as half the difference between the upward and downward variations, and were substantially reduced by measuring the lepton isolation efficiencies from data, in the same way as for the $t\bar{t}$ generator uncertainties discussed above.

Parton distribution functions: The uncertainties in $G_{ee}$ due to limited knowledge of the proton PDFs were evaluated using the error sets of the CT10 [10], MSTW 2008 68% CL [9], and NNPDF 2.3 [12] NLO PDF sets, by reweighting the MC@NLO + HERWIG $t\bar{t}$ sample based on the $x$ and $Q^2$ values of the partons participating in the hard scattering in each event. The final uncertainty in each bin was calculated as half the envelope encompassing the predictions from all three PDF sets and their associated uncertainties, following the PDF4LHC prescription [7]. The resulting uncertainties on $G_{ee}$ are typically around 0.3% except at the high ends of the distributions, and were taken to be fully correlated across all bins.

Top quark mass: The values of $G_{ee}$ and the predicted levels of $Wt$ background depend weakly on the assumed value of $m_t$. These effects were evaluated with $t\bar{t}$ and $Wt$
Fig. 6 Relative uncertainties on the measured normalised differential cross-sections coming from data statistics, \( t\bar{t} \) modelling, leptons, jets and background, as a function of each lepton or dilepton differential variable. The total uncertainty is shown by the black lines, and also includes small contributions from the integrated luminosity and LHC beam energy uncertainties.
samples simulated with $m_t$ values of 170 and 175 GeV, and scaled to a nominal ±1 GeV mass variation. The resulting effects are at the level of 0.1–0.2% on $G^{\ell}\ell$, and are partially cancelled by the variations in the $Wt$ background, whose cross-section decreases with increasing $m_t$. The residual uncertainties are typically around 0.1% for the absolute cross-sections except at the extreme ends of the distributions, and smaller for the normalised cross-sections.

The total $t\bar{t}$ modelling uncertainties in the normalised differential cross-sections also include the small uncertainties on $G^{e\mu}$ and $C^{d}_{b}$ from the limited size of the simulated $t\bar{t}$ samples, and are shown by the green lines in Fig. 6. They are typically dominated by the $t\bar{t}$ event generator comparisons.

5.2 Lepton identification and measurement

Uncertainties in the modelling of the detector response to electrons and muons affect both $G^{e\mu}$, and the background estimates, with the largest uncertainties in the cross-section measurements coming via the former.

**Lepton identification:** The modelling of the electron and muon identification efficiencies, and the rate of electron charge misidentification, were studied using $Z \rightarrow ee/\mu\mu$, $J/\psi \rightarrow ee/\mu\mu$, and $W \rightarrow ev$ events in data and simulation [76,77], taking into account the systematic correlations across different regions of the lepton $p_T$ and $\eta$ spectrum. The uncertainties in $G^{e\mu}_\ell$ are typically below 0.5% for electron and below 0.3% for muon efficiencies, with significant cancellations in the normalised differential cross-sections.

**Lepton scales and resolution:** The electron and muon energy/momentum scales and resolutions were determined using $Z \rightarrow ee/\mu\mu$, $Z \rightarrow (ee/\mu\mu)\gamma$, $J/\psi \rightarrow ee/\mu\mu$ and $\Upsilon \rightarrow \mu\mu$ decays [77,86]. The largest uncertainty comes from the limited knowledge of the electron energy scale, which gives uncertainties varying from 0.2% to over 2% for the bins involving the highest energy electrons. The muon momentum scale uncertainties are small in comparison.

**Lepton isolation:** Building on the studies described in Ref. [13], the efficiencies of the lepton isolation requirements were measured in data, using the fractions of selected opposite-sign $e\mu$ events with at least one $b$-tagged jet where either the electron or the muon fails the isolation requirement. After correcting for the contamination from events with a misidentified lepton, these fractions give the inefficiency of the isolation requirements on signal $t\bar{t}$ events. The misidentified lepton backgrounds were measured both by using the same-sign $e\mu$ control samples discussed in Sect. 4.2 above, and by using the distributions of lepton impact parameter significance $|d_0|/\sigma_{d_0}$, where $d_0$ is the distance of closest approach of the lepton track to the event primary vertex in the transverse plane, and $\sigma_{d_0}$ its uncertainty. The isolation inefficiencies were measured as functions of lepton $p_T$ separately for the barrel ($|\eta| < 1.5$) and endcap regions of the detector. Consistent results were obtained using both misidentified lepton estimation methods, and showed that the baseline POWHEG + PYTHIA6 $t\bar{t}$ simulation sample overestimates the efficiencies of the lepton isolation requirements by up to 1% for electrons with $p_T$ in the range 40–80 GeV, and by up to 2% for muons at low $p_T$, decreasing rapidly to less than 0.5% for 40 GeV. The values of $G^{e\mu}_\ell$ from the baseline simulation were corrected for these $p_T$-dependent shifts using a reweighting technique. The corresponding uncertainties are dominated by those on the misidentified lepton subtraction (including a comparison of the same-sign and $|d_0|/\sigma_{d_0}$-based methods) and amount to typically 0.5–1% for electrons and 0.2–0.5% for muons. The effect on the normalised cross-sections is about half that on the absolute measurements, taking into account systematic correlations across lepton $p_T$ and $|\eta|$ bins.

**Lepton trigger:** The efficiencies of the single-lepton triggers were measured in data using $Z \rightarrow ee/\mu\mu$ events [87]. Since only one lepton trigger was required to accept the $e\mu$ event, the trigger efficiency with respect to the offline event selection is about 99%, with a residual uncertainty of less than 0.2%.

The lepton-related uncertainties are shown by the blue dot-dashed lines in Fig. 6, and the largest uncertainties typically come from the electron energy scale and electron isolation uncertainties.

5.3 Jet measurement and $b$-tagging

Uncertainties in the selection and $b$-tagging of jets affect the background estimates $N^{i,bkg}_1$ and $N^{i,bkg}_2$, and to a lesser extent, the correlation $C^{d}_{b}$. The jet uncertainties also have a very small effect on $G^{e\mu}_\ell$ through the requirement that leptons be separated from selected jets by $\Delta R > 0.4$.

**Jet-related uncertainties:** The jet energy scale was varied according to the uncertainties derived from simulation and in situ calibration measurements [81], using a model with 22 orthogonal uncertainty components describing the evolution with jet $p_T$ and $|\eta|$. The effects of residual uncertainties in the modelling of the jet energy resolution [88] were assessed by smearing jet energies in simulation. The jet reconstruction efficiency was measured in data
using track-based jets, and the effect of residual uncertainties assessed in simulation by randomly discarding jets. The modelling of the pileup rejection requirement applied to jets was studied using $Z \rightarrow ee/\mu\mu+$jets events [82].

$b$-tagging uncertainties: The efficiencies for $b$-tagging jets in $t\bar{t}$ signal events were extracted from the data, but simulation was used to predict the numbers of $b$-tagged jets in the $Wt$ single top and diboson backgrounds. The corresponding uncertainties were assessed using studies of $b$-jets containing muons, charm jets containing $D^{++}$ mesons and inclusive jet events [83].

The jet- and $b$-tagging-related uncertainties are shown by the purple lines on Fig. 6, and are typically dominated by the effect of the jet energy scale on the level of $Wt$ background.

5.4 Background modelling

As well as the detector-related uncertainties discussed above, the background estimates depend on uncertainties in modelling the $Wt$ and diboson processes taken from simulation, and uncertainties in the procedures used for estimating the $Z+$jets and misidentified lepton backgrounds from data.

Single top modelling: Uncertainties in the modelling of the $Wt$ background were assessed by comparing the predictions from the baseline POWHEG + PYTHIA6 sample with those from MC@NLO + HERWIG, and from two samples generated with ACERMC + PYTHIA6 utilising different tunes to vary the amount of additional radiation, in all cases normalising the total production cross-section to the approximate NNLO prediction based on Ref. [67]. The uncertainty in this prediction was evaluated to be 6.8%. The $Wt$ background with two $b$-tagged jets is sensitive to the production of $Wt$ with an additional $b$-jet, an NLO contribution which interferes with the $t\bar{t}$ final state. The corresponding uncertainty was assessed by comparing the predictions of POWHEG + PYTHIA6 with the diagram removal and diagram subtraction schemes for handling this interference [65,66]. The latter predicts up to 25% less $Wt$ background in the one $b$-tagged and 60% less in the two $b$-tagged channels at the extreme high ends of the lepton $p_T$ and dilepton $p_T^{\mu\mu}, m^{\mu\mu}, p_T^{\mu} + p_T^\ell$ and $E_T^{\mu} + E_T^{\ell}$ distributions, but only 1–2% and 20% differences for one and two $b$-tagged $Wt$ events across the $|\eta|$ , $|y^{\mu\mu}|$ and $\Delta\phi^{\mu\mu}$ distributions, similar to the differences seen for the inclusive analysis [13]. The uncertainties due to the limited size of the $Wt$ simulation samples are negligible in comparison to the modelling uncertainties.

Diboson modelling: The uncertainties in the modelling the diboson background events (mainly $WW$) with one and two additional $b$-tagged jets were assessed by comparing the predictions from ALPGEN+HERWIG with those of SHERPA 1.4.3 [73] including the effects of massive $b$ and $c$ quarks. The resulting uncertainties in the diboson background are typically in the range 20–30%, substantially larger than the differences between recent predictions for the inclusive diboson cross-sections at NNLO in QCD [89] and the NLO predictions from MCFM used to normalise the simulated samples. The background from SM Higgs production with $H \rightarrow WW$ and $H \rightarrow \tau\tau$ is smaller than the uncertainties assigned for diboson modelling, and was neglected.

Z+jets extrapolation: The backgrounds from $Z \rightarrow \tau\tau \rightarrow e\mu$ accompanied by one or two $b$-tagged jets were extrapolated from the analogous $Z \rightarrow ee/\mu\mu$ event rates, with uncertainties of 20% for one and 30% for two additional $b$-tagged jets, as discussed in Sect. 4.2.

Misidentified leptons: Uncertainties in the numbers of events with misidentified leptons arise from the statistical uncertainties in the corresponding same-sign samples, together with systematic uncertainties in the opposite-to-same-sign ratios $R_d^i$ and the estimated contributions of prompt same-sign events. The total uncertainties in the measured cross-sections are typically 0.2–0.5%, except at the extreme ends of distributions where the same-sign data statistical uncertainties are larger.

The background uncertainties are shown by the solid red lines on Fig. 6, and are dominated by $Wt$ modelling uncertainties, in particular from the $Wt$-$t\bar{t}$ interference at the high ends of some distributions.

5.5 Luminosity and beam energy

Uncertainties in the integrated luminosity and LHC beam energy give rise to additional uncertainties in the differential cross-section results.

Luminosity: The uncertainty in the integrated luminosity is 1.9%, derived from beam-separation scans performed in November 2012 [90]. The corresponding uncertainty in the absolute cross-section measurements is slightly larger, typically about 2.1%, as the $Wt$ and diboson backgrounds were evaluated from simulation, thus becoming sensitive to the assumed integrated luminosity. The sensitivity varies with the background fractions, leaving a residual uncertainty of typically less than 0.1% in the normalised cross-section results.

Beam energy: The LHC beam energy during the 2012 $pp$ run was determined to be within 0.1% of the nominal value of 4 TeV per beam, based on the LHC magnetic model together with measurements of the revolution frequency difference of proton and lead-ion beams [91]. Following the approach used in Ref. [13] with an earlier
less precise determination of the LHC beam energy [92], an additional uncertainty corresponding to the change in cross-sections for a 0.1% change in $\sqrt{s}$ was applied to the final results, allowing them to be interpreted as measurements at exactly $\sqrt{s} = 8$ TeV. The changes in each differential cross-section bin were calculated by scaling the differences seen in POWHEG + PYTHIA6 samples generated at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 7$ TeV. The resulting values were cross-checked with an explicit NLO fixed-order calculation using SHERPA 2.1 [73], making use of the APPLGRID framework [93] to reweight an $\sqrt{s} = 8$ TeV prediction so as to change the $\sqrt{s}$ value by $\pm 0.66\%$ which was then rescaled to correspond to a $\sqrt{s}$ change of 0.1%. The changes in the absolute cross-sections are in the range 0.2–0.4%, and largely cancel in the normalised cross-sections.

These uncertainties are not shown separately in Fig. 6, but are included in the total uncertainties shown by the black lines, and given in Tables 3, 4, 5 and 6.

6 Results

The absolute differential cross-sections were determined by solving Eq. (1) separately for each bin $i$ of each lepton and dilepton differential distribution, taking the effects of systematic uncertainties into account as discussed in Sect. 5. The normalised differential cross-sections were determined from the absolute results using Eq. (2). The values of $\epsilon_L^j$, i.e. the product of jet acceptance, reconstruction and $b$-tagging probabilities in each bin, were determined to be in the range 0.5–0.6, in agreement with the simulation prediction for each bin. The results were found to be stable when changing the minimum jet $p_T$ requirement from 25 GeV up to 55 GeV, and when using $b$-tagging working points corresponding to $b$-jet efficiencies of 60–80%. The electron and muon $p_T$ and $|\eta|$ distributions were also measured separately, instead of combining them into lepton distributions with two entries per event, and found to be compatible. The bin-by-bin comparison of the electron and muon $p_T$ ($|\eta|$) distributions has a $\chi^2$ per degree of freedom of 10.9/9 (12.5/8), in both cases taking into account statistical and uncorrelated systematic uncertainties.

6.1 Fiducial cross-section measurements

The measured absolute and normalised fiducial differential cross-sections are shown in Table 3 ($p_T^e$ and $|\eta|^e$), Table 4 ($p_T^{\mu}$ and $m^{\mu}$), Table 5 ($|\eta|^\mu$ and $\Delta\phi^{\mu}$) and Table 6 ($p_T^e+p_T^\mu$ and $E^e+E^\mu$). Each table shows the measured cross-section values and uncertainties, together with a breakdown of the total uncertainties into components due to data statistics (‘Stat.’), $t\bar{t}$ modelling uncertainties (‘$t\bar{t}$ mod.’), lepton-related uncertainties (‘Lept.’), jet and $b$-tagging uncertainties (‘Jet/$b$’), background uncertainties (‘Bkg.’) and luminosity/beam energy uncertainties (‘$L/E_p$’), corresponding to the breakdown in Sects. 5.1–5.5. The rightmost columns show the cross-sections corrected to remove the contributions where one or both leptons result from $W \rightarrow \tau \rightarrow e/\mu$ decays using Eq. (3). As can also be seen from Fig. 6, the total uncertainties on the normalised differential cross-sections range from 1.2% to around 10%, typically smaller than those for the measurements as a function of the $t\bar{t}$ system kinematics in Ref. [21]. The largest uncertainties are generally statistical (from 1.1% to about 10%), with the background uncertainties also becoming large at high values of some kinematic variables. Other systematic uncertainties due to $t\bar{t}$ modelling, leptons and jets are significantly smaller than the statistical uncertainties, benefiting from cancellations between bins. The cancellations are particularly important when leptons with similar $p_T$ contribute to all bins, as is the case for $\Delta\phi^{e\mu}$ and the bulk of the $|\eta|^e$ and $|\eta|^\mu$ distributions. The uncertainties in the absolute cross-sections are substantially larger, with the systematic uncertainties due to $t\bar{t}$ modelling and leptons becoming comparable to the statistical uncertainties. The absolute cross-sections also have an uncertainty of 2.1–2.5% from the integrated luminosity measurement, depending on the background level in each bin.

The integrals of the differential cross-sections across all bins of a given distribution ($\sigma_{\text{fid}}^{\ell\ell}$ in Eq. (2)) agree in all cases within 0.4% of the integrated fiducial cross-sections of $3.455 \pm 0.025$ pb (or $3.043 \pm 0.022$ pb excluding $\tau$ contributions) measured within the same fiducial region in Ref. [13,14]. The quoted uncertainties are statistical.$^5$

The normalised differential cross-sections are shown graphically in Figs. 7 and 8; in these and later figures, the data points are plotted at the centre of each bin. The measured cross-sections are compared to the particle-level predictions from the POWHEG + PYTHIA6, MC@NLO + HERWIG and ALPGEN + HERWIG $t\bar{t}$ samples within the fiducial volume of the measurement, including the contributions from $W \rightarrow \tau \rightarrow e/\mu$ decays. Similar trends in the description of the measured distributions by the predictions can be seen as for the reconstructed distributions for events with at least one $b$-tagged jet in Figs. 1 and 2.

6.2 Comparison with event generator predictions

The measured normalised differential cross-sections are compared to a larger set of predictions from different $t\bar{t}$ Monte Carlo event generator configurations in Figs. 9, 10,

---

$^5$ The integrals of the $p_T^e$ and $|\eta|^e$ distributions correspond to twice these values, as the definitions include two leptons per event.
Table 3 Absolute and normalised differential cross-sections as functions of $p_T^\ell$ (top) and $|\eta^\ell|$ (bottom). The columns show the bin ranges, measured cross-section and total uncertainty, relative statistical uncertainty, relative systematic uncertainties in various categories (see text), total relative uncertainty, and differential cross-section corrected to remove contributions via $W \to \tau \to e/\mu$ decays. Relative uncertainties smaller than 0.05% are indicated by ‘0.0’. The last bin includes overflows where indicated by the '+' sign.

<table>
<thead>
<tr>
<th>Absolute bin (GeV)</th>
<th>$d\sigma/dp_T^\ell$ (fb/GeV)</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/b (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_b$ (%)</th>
<th>Total (%)</th>
<th>$d\sigma/dp_T^\ell$ (no $\tau$) (fb/GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25–30</td>
<td>154.8 ± 5.7</td>
<td>1.6</td>
<td>1.3</td>
<td>1.8</td>
<td>0.8</td>
<td>1.2</td>
<td>2.1</td>
<td>3.7</td>
<td>127.2 ± 4.8</td>
</tr>
<tr>
<td>30–40</td>
<td>146.1 ± 4.9</td>
<td>1.1</td>
<td>1.2</td>
<td>1.5</td>
<td>0.8</td>
<td>1.0</td>
<td>2.1</td>
<td>3.3</td>
<td>124.6 ± 4.2</td>
</tr>
<tr>
<td>40–50</td>
<td>118.8 ± 3.7</td>
<td>1.2</td>
<td>1.1</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
<td>2.1</td>
<td>3.1</td>
<td>104.3 ± 3.3</td>
</tr>
<tr>
<td>50–60</td>
<td>93.5 ± 2.9</td>
<td>1.4</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
<td>0.9</td>
<td>2.1</td>
<td>3.1</td>
<td>83.4 ± 2.6</td>
</tr>
<tr>
<td>60–80</td>
<td>60.0 ± 1.8</td>
<td>1.2</td>
<td>0.9</td>
<td>0.9</td>
<td>0.6</td>
<td>0.9</td>
<td>2.1</td>
<td>3.0</td>
<td>54.1 ± 1.6</td>
</tr>
<tr>
<td>80–100</td>
<td>32.4 ± 1.1</td>
<td>1.6</td>
<td>0.8</td>
<td>1.1</td>
<td>1.4</td>
<td>1.1</td>
<td>2.1</td>
<td>3.5</td>
<td>29.4 ± 1.0</td>
</tr>
<tr>
<td>100–120</td>
<td>16.23 ± 0.64</td>
<td>2.3</td>
<td>0.9</td>
<td>1.1</td>
<td>1.1</td>
<td>1.5</td>
<td>2.2</td>
<td>3.9</td>
<td>14.75 ± 0.58</td>
</tr>
<tr>
<td>120–150</td>
<td>7.61 ± 0.35</td>
<td>2.7</td>
<td>1.1</td>
<td>1.4</td>
<td>1.5</td>
<td>1.9</td>
<td>2.2</td>
<td>4.6</td>
<td>6.91 ± 0.32</td>
</tr>
<tr>
<td>150–200</td>
<td>2.41 ± 0.15</td>
<td>3.9</td>
<td>1.6</td>
<td>1.7</td>
<td>1.6</td>
<td>3.2</td>
<td>2.2</td>
<td>6.2</td>
<td>2.17 ± 0.13</td>
</tr>
<tr>
<td>200–300+</td>
<td>0.49 ± 0.06</td>
<td>6.7</td>
<td>3.5</td>
<td>2.3</td>
<td>2.9</td>
<td>7.5</td>
<td>2.4</td>
<td>11.5</td>
<td>0.44 ± 0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normalised bin (GeV)</th>
<th>$\frac{1}{n}d\sigma/dp_T^\ell (10^{-2}/\text{GeV})$</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/b (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_b$ (%)</th>
<th>Total (%)</th>
<th>$\frac{1}{n}d\sigma/dp_T^\ell$ (no $\tau$) (10^{-2}/\text{GeV})</th>
</tr>
</thead>
<tbody>
<tr>
<td>25–30</td>
<td>2.235 ± 0.045</td>
<td>1.5</td>
<td>0.4</td>
<td>0.9</td>
<td>0.6</td>
<td>0.8</td>
<td>0.0</td>
<td>2.0</td>
<td>2.090 ± 0.042</td>
</tr>
<tr>
<td>30–40</td>
<td>2.108 ± 0.029</td>
<td>1.0</td>
<td>0.3</td>
<td>0.7</td>
<td>0.3</td>
<td>0.5</td>
<td>0.0</td>
<td>1.4</td>
<td>2.048 ± 0.029</td>
</tr>
<tr>
<td>40–50</td>
<td>1.714 ± 0.023</td>
<td>1.1</td>
<td>0.2</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.0</td>
<td>1.3</td>
<td>1.714 ± 0.023</td>
</tr>
<tr>
<td>50–60</td>
<td>1.350 ± 0.019</td>
<td>1.3</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.0</td>
<td>1.4</td>
<td>1.370 ± 0.020</td>
</tr>
<tr>
<td>60–80</td>
<td>0.866 ± 0.011</td>
<td>1.1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.3</td>
<td>0.0</td>
<td>1.3</td>
<td>0.890 ± 0.011</td>
</tr>
<tr>
<td>80–100</td>
<td>0.4673 ± 0.0093</td>
<td>1.5</td>
<td>0.4</td>
<td>0.7</td>
<td>0.8</td>
<td>0.5</td>
<td>0.0</td>
<td>2.0</td>
<td>0.4831 ± 0.0096</td>
</tr>
<tr>
<td>100–120</td>
<td>0.2343 ± 0.0063</td>
<td>2.3</td>
<td>0.5</td>
<td>0.8</td>
<td>0.6</td>
<td>1.0</td>
<td>0.0</td>
<td>2.7</td>
<td>0.2424 ± 0.0065</td>
</tr>
<tr>
<td>120–150</td>
<td>0.1098 ± 0.0040</td>
<td>2.7</td>
<td>0.6</td>
<td>1.2</td>
<td>1.3</td>
<td>1.5</td>
<td>0.1</td>
<td>3.6</td>
<td>0.1135 ± 0.0041</td>
</tr>
<tr>
<td>150–200</td>
<td>0.0348 ± 0.0018</td>
<td>3.9</td>
<td>0.9</td>
<td>1.6</td>
<td>1.2</td>
<td>2.8</td>
<td>0.1</td>
<td>5.3</td>
<td>0.0357 ± 0.0019</td>
</tr>
<tr>
<td>200–300+</td>
<td>0.0070 ± 0.0007</td>
<td>6.7</td>
<td>2.7</td>
<td>2.3</td>
<td>2.3</td>
<td>7.2</td>
<td>0.3</td>
<td>10.7</td>
<td>0.0072 ± 0.0008</td>
</tr>
</tbody>
</table>
Table 3 continued

| Absolute bin (unit $\eta$) | $\frac{d\sigma}{d\eta}|(fb/unit\ \eta)$ | Stat. (%) | $t\bar{t}$ mod. (%) | Lept. (%) | Jet/b (%) | Bkg. (%) | $L/E_b$ (%) | Total (%) | $\frac{d\sigma}{d\eta}|(no\ \tau\ (fb/unit\ \eta))$ |
|---------------------------|------------------------------------------|-----------|---------------------|----------|-----------|---------|-------------|----------|----------------------------------|
| 0.00–0.25                 | 4590 ± 140                               | 1.2       | 1.0                 | 1.0      | 0.9       | 0.9     | 2.1         | 3.1      | 4030 ± 130                        |
| 0.25–0.50                 | 4440 ± 140                               | 1.2       | 1.0                 | 1.0      | 0.9       | 0.9     | 2.1         | 3.1      | 3900 ± 120                        |
| 0.50–0.75                 | 4230 ± 130                               | 1.2       | 1.0                 | 1.0      | 0.9       | 0.9     | 2.1         | 3.1      | 3710 ± 120                        |
| 0.75–1.00                 | 3660 ± 110                               | 1.3       | 1.0                 | 1.0      | 0.8       | 1.0     | 2.1         | 3.1      | 3210 ± 100                        |
| 1.00–1.25                 | 3100 ± 100                               | 1.5       | 1.0                 | 1.0      | 0.9       | 1.0     | 2.1         | 3.3      | 2722 ± 89                         |
| 1.25–1.50                 | 2470 ± 87                                | 1.9       | 1.1                 | 1.0      | 0.9       | 1.0     | 2.1         | 3.5      | 2173 ± 77                         |
| 1.50–1.75                 | 2035 ± 73                                | 1.9       | 1.1                 | 1.4      | 0.7       | 1.0     | 2.1         | 3.6      | 1793 ± 65                         |
| 1.75–2.00                 | 1431 ± 57                                | 2.4       | 1.2                 | 1.4      | 0.6       | 1.4     | 2.1         | 4.0      | 1263 ± 50                         |
| 2.00–2.50                 | 844 ± 34                                 | 2.2       | 1.3                 | 1.4      | 0.7       | 1.4     | 2.1         | 4.0      | 749 ± 30                          |

| Normalised bin (unit $\eta$) | $\frac{1}{\sigma}\frac{d\sigma}{d\eta}|(10^{-1}/unit\ \eta)$ | Stat. (%) | $t\bar{t}$ mod. (%) | Lept. (%) | Jet/b (%) | Bkg. (%) | $L/E_b$ (%) | Total (%) | $\frac{1}{\sigma}\frac{d\sigma}{d\eta}|(no\ \tau\ (10^{-1}/unit\ \eta))$ |
|-----------------------------|---------------------------------------------------------------|-----------|---------------------|----------|-----------|---------|-------------|----------|----------------------------------|
| 0.00–0.25                   | 6.646 ± 0.083                                                | 1.1       | 0.4                 | 0.2      | 0.1       | 0.3     | 0.0         | 1.2      | 6.632 ± 0.083                     |
| 0.25–0.50                   | 6.428 ± 0.076                                                | 1.1       | 0.3                 | 0.2      | 0.1       | 0.2     | 0.0         | 1.2      | 6.416 ± 0.076                     |
| 0.50–0.75                   | 6.117 ± 0.074                                                | 1.1       | 0.2                 | 0.2      | 0.1       | 0.2     | 0.0         | 1.2      | 6.103 ± 0.074                     |
| 0.75–1.00                   | 5.297 ± 0.070                                                | 1.3       | 0.2                 | 0.2      | 0.2       | 0.2     | 0.0         | 1.3      | 5.286 ± 0.070                     |
| 1.00–1.25                   | 4.482 ± 0.066                                                | 1.4       | 0.3                 | 0.2      | 0.2       | 0.3     | 0.0         | 1.5      | 4.484 ± 0.066                     |
| 1.25–1.50                   | 3.574 ± 0.068                                                | 1.8       | 0.4                 | 0.2      | 0.2       | 0.3     | 0.0         | 1.9      | 3.579 ± 0.068                     |
| 1.50–1.75                   | 2.944 ± 0.060                                                | 1.9       | 0.5                 | 0.6      | 0.2       | 0.4     | 0.0         | 2.0      | 2.954 ± 0.061                     |
| 1.75–2.00                   | 2.070 ± 0.054                                                | 2.3       | 0.6                 | 0.6      | 0.4       | 0.8     | 0.0         | 2.6      | 2.080 ± 0.054                     |
| 2.00–2.50                   | 1.221 ± 0.031                                                | 2.1       | 0.7                 | 0.6      | 0.4       | 1.0     | 0.0         | 2.5      | 1.233 ± 0.031                     |
Table 4  Absolute and normalised differential cross-sections as functions of $p_T^{\mu}$ (top) and $m^{\ell\mu}$ (bottom). The columns show the bin ranges, measured cross-section and total uncertainty, relative statistical uncertainty, relative systematic uncertainties in various categories (see text), total relative uncertainty, and differential cross-section corrected to remove contributions via $W \to \tau \to e/\mu$ decays. Relative uncertainties smaller than 0.05% are indicated by ‘0.0’. The last bin includes overflows where indicated by the ‘+’ sign.

<table>
<thead>
<tr>
<th>Absolute bin (GeV)</th>
<th>$d\sigma/dp_T^{\mu}$ (fb/GeV)</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/(b) (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_b$ (%)</th>
<th>Total (%)</th>
<th>$\frac{d\sigma}{dp_T^{\mu}}$ (no $\tau$) (fb/GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–20</td>
<td>11.50 ± 0.49</td>
<td>2.8</td>
<td>1.2</td>
<td>1.2</td>
<td>1.1</td>
<td>1.2</td>
<td>2.1</td>
<td>4.2</td>
<td>9.62 ± 0.41</td>
</tr>
<tr>
<td>20–40</td>
<td>26.72 ± 0.94</td>
<td>1.9</td>
<td>1.0</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0</td>
<td>2.1</td>
<td>3.5</td>
<td>22.62 ± 0.81</td>
</tr>
<tr>
<td>40–60</td>
<td>35.9 ± 1.2</td>
<td>1.6</td>
<td>0.9</td>
<td>1.2</td>
<td>1.1</td>
<td>1.0</td>
<td>2.1</td>
<td>3.4</td>
<td>30.6 ± 1.0</td>
</tr>
<tr>
<td>60–80</td>
<td>39.0 ± 1.3</td>
<td>1.5</td>
<td>0.9</td>
<td>1.1</td>
<td>0.9</td>
<td>0.9</td>
<td>2.1</td>
<td>3.2</td>
<td>34.4 ± 1.1</td>
</tr>
<tr>
<td>80–100</td>
<td>29.19 ± 0.96</td>
<td>1.7</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
<td>1.0</td>
<td>2.1</td>
<td>3.3</td>
<td>26.48 ± 0.88</td>
</tr>
<tr>
<td>100–120</td>
<td>16.38 ± 0.65</td>
<td>2.3</td>
<td>1.2</td>
<td>1.3</td>
<td>1.3</td>
<td>1.1</td>
<td>2.1</td>
<td>3.9</td>
<td>15.11 ± 0.60</td>
</tr>
<tr>
<td>120–150</td>
<td>6.53 ± 0.30</td>
<td>2.9</td>
<td>1.4</td>
<td>1.5</td>
<td>1.2</td>
<td>1.6</td>
<td>2.2</td>
<td>4.6</td>
<td>6.06 ± 0.28</td>
</tr>
<tr>
<td>150–200</td>
<td>1.39 ± 0.11</td>
<td>5.0</td>
<td>2.2</td>
<td>2.1</td>
<td>2.2</td>
<td>4.2</td>
<td>2.3</td>
<td>7.9</td>
<td>1.27 ± 0.10</td>
</tr>
<tr>
<td>200–300+</td>
<td>0.23 ± 0.04</td>
<td>9.3</td>
<td>4.2</td>
<td>2.9</td>
<td>4.1</td>
<td>13.2</td>
<td>2.5</td>
<td>17.6</td>
<td>0.20 ± 0.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normalised bin (GeV)</th>
<th>$\frac{1}{x}d\sigma/dp_T^{\mu}$ (10^{-2}/GeV)</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/(b) (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_b$ (%)</th>
<th>Total (%)</th>
<th>$\frac{1}{x}d\sigma/dp_T^{\mu}$ (no $\tau$) (10^{-2}/GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–20</td>
<td>0.332 ± 0.010</td>
<td>2.7</td>
<td>0.5</td>
<td>0.5</td>
<td>0.7</td>
<td>0.8</td>
<td>0.0</td>
<td>3.0</td>
<td>0.316 ± 0.010</td>
</tr>
<tr>
<td>20–40</td>
<td>0.772 ± 0.015</td>
<td>1.7</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>0.6</td>
<td>0.0</td>
<td>2.0</td>
<td>0.743 ± 0.015</td>
</tr>
<tr>
<td>40–60</td>
<td>1.036 ± 0.017</td>
<td>1.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>0.0</td>
<td>1.6</td>
<td>1.006 ± 0.017</td>
</tr>
<tr>
<td>60–80</td>
<td>1.127 ± 0.016</td>
<td>1.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.0</td>
<td>1.5</td>
<td>1.130 ± 0.016</td>
</tr>
<tr>
<td>80–100</td>
<td>0.843 ± 0.015</td>
<td>1.5</td>
<td>0.3</td>
<td>0.3</td>
<td>0.6</td>
<td>0.4</td>
<td>0.0</td>
<td>1.8</td>
<td>0.870 ± 0.015</td>
</tr>
<tr>
<td>100–120</td>
<td>0.473 ± 0.012</td>
<td>2.2</td>
<td>0.4</td>
<td>0.8</td>
<td>0.8</td>
<td>0.5</td>
<td>0.0</td>
<td>2.5</td>
<td>0.497 ± 0.013</td>
</tr>
<tr>
<td>120–150</td>
<td>0.1886 ± 0.0066</td>
<td>2.8</td>
<td>0.6</td>
<td>1.2</td>
<td>1.0</td>
<td>1.2</td>
<td>0.0</td>
<td>3.5</td>
<td>0.1993 ± 0.0069</td>
</tr>
<tr>
<td>150–200</td>
<td>0.0402 ± 0.0028</td>
<td>4.9</td>
<td>1.4</td>
<td>1.9</td>
<td>1.9</td>
<td>3.9</td>
<td>0.2</td>
<td>7.0</td>
<td>0.0419 ± 0.0029</td>
</tr>
<tr>
<td>200–300+</td>
<td>0.0066 ± 0.0011</td>
<td>9.2</td>
<td>3.3</td>
<td>2.7</td>
<td>3.6</td>
<td>13.0</td>
<td>0.4</td>
<td>16.9</td>
<td>0.0067 ± 0.0011</td>
</tr>
</tbody>
</table>
### Table 4 continued

<table>
<thead>
<tr>
<th>Absolute bin (GeV)</th>
<th>$\frac{1}{\beta} \frac{d\sigma}{d\eta^{\mu}}$ (fb/GeV)</th>
<th>Stat. (%)</th>
<th>$\bar{t}t$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/$b$ (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_{\eta}$ (%)</th>
<th>Total (%)</th>
<th>$\frac{1}{\beta} \frac{d\sigma}{d\eta^{\mu}}$ (no $\tau$) (fb/GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–20</td>
<td>3.37 ± 0.25</td>
<td>6.3</td>
<td>2.0</td>
<td>1.9</td>
<td>1.3</td>
<td>1.3</td>
<td>2.1</td>
<td>7.4</td>
<td>2.97 ± 0.22</td>
</tr>
<tr>
<td>20–40</td>
<td>10.94 ± 0.47</td>
<td>2.8</td>
<td>1.5</td>
<td>1.4</td>
<td>0.9</td>
<td>1.2</td>
<td>2.1</td>
<td>4.3</td>
<td>9.61 ± 0.41</td>
</tr>
<tr>
<td>40–60</td>
<td>17.66 ± 0.70</td>
<td>2.2</td>
<td>1.4</td>
<td>1.5</td>
<td>0.7</td>
<td>1.3</td>
<td>2.1</td>
<td>4.0</td>
<td>15.29 ± 0.61</td>
</tr>
<tr>
<td>60–80</td>
<td>23.98 ± 0.89</td>
<td>1.9</td>
<td>1.3</td>
<td>1.4</td>
<td>0.8</td>
<td>1.2</td>
<td>2.1</td>
<td>3.7</td>
<td>20.54 ± 0.76</td>
</tr>
<tr>
<td>80–100</td>
<td>26.00 ± 0.90</td>
<td>1.8</td>
<td>1.2</td>
<td>1.1</td>
<td>0.8</td>
<td>0.9</td>
<td>2.1</td>
<td>3.4</td>
<td>22.42 ± 0.78</td>
</tr>
<tr>
<td>100–120</td>
<td>23.03 ± 0.83</td>
<td>2.0</td>
<td>1.1</td>
<td>1.0</td>
<td>1.2</td>
<td>0.9</td>
<td>2.1</td>
<td>3.6</td>
<td>20.07 ± 0.73</td>
</tr>
<tr>
<td>120–150</td>
<td>16.71 ± 0.57</td>
<td>1.9</td>
<td>1.0</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>2.1</td>
<td>3.4</td>
<td>14.72 ± 0.51</td>
</tr>
<tr>
<td>150–200</td>
<td>9.38 ± 0.34</td>
<td>1.9</td>
<td>0.8</td>
<td>0.9</td>
<td>1.3</td>
<td>1.2</td>
<td>2.1</td>
<td>3.6</td>
<td>8.41 ± 0.30</td>
</tr>
<tr>
<td>200–250</td>
<td>4.09 ± 0.18</td>
<td>3.0</td>
<td>0.8</td>
<td>1.1</td>
<td>1.1</td>
<td>1.6</td>
<td>2.1</td>
<td>4.4</td>
<td>3.73 ± 0.16</td>
</tr>
<tr>
<td>250–300</td>
<td>1.95 ± 0.11</td>
<td>4.4</td>
<td>1.0</td>
<td>1.2</td>
<td>1.4</td>
<td>2.1</td>
<td>2.2</td>
<td>5.7</td>
<td>1.80 ± 0.10</td>
</tr>
<tr>
<td>300–400</td>
<td>0.66 ± 0.05</td>
<td>5.2</td>
<td>1.3</td>
<td>1.5</td>
<td>2.0</td>
<td>3.1</td>
<td>2.2</td>
<td>7.1</td>
<td>0.62 ± 0.04</td>
</tr>
<tr>
<td>400–500+</td>
<td>0.26 ± 0.03</td>
<td>8.0</td>
<td>2.2</td>
<td>2.1</td>
<td>2.1</td>
<td>4.8</td>
<td>2.2</td>
<td>10.3</td>
<td>0.25 ± 0.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normalised bin (GeV)</th>
<th>$\frac{1}{\beta} \frac{d\sigma}{d\eta^{\mu}}$ (10$^{-3}$/GeV)</th>
<th>Stat. (%)</th>
<th>$\bar{t}t$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/$b$ (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_{\eta}$ (%)</th>
<th>Total (%)</th>
<th>$\frac{1}{\beta} \frac{d\sigma}{d\eta^{\mu}}$ (no $\tau$) [10$^{-3}$/GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–20</td>
<td>0.973 ± 0.066</td>
<td>6.2</td>
<td>1.4</td>
<td>1.2</td>
<td>1.4</td>
<td>1.3</td>
<td>0.0</td>
<td>6.7</td>
<td>0.977 ± 0.066</td>
</tr>
<tr>
<td>20–40</td>
<td>3.157 ± 0.095</td>
<td>2.7</td>
<td>0.7</td>
<td>0.5</td>
<td>0.4</td>
<td>0.9</td>
<td>0.0</td>
<td>3.0</td>
<td>3.156 ± 0.095</td>
</tr>
<tr>
<td>40–60</td>
<td>5.10 ± 0.13</td>
<td>2.1</td>
<td>0.5</td>
<td>0.7</td>
<td>0.4</td>
<td>1.0</td>
<td>0.0</td>
<td>2.5</td>
<td>5.02 ± 0.13</td>
</tr>
<tr>
<td>60–80</td>
<td>6.92 ± 0.14</td>
<td>1.8</td>
<td>0.4</td>
<td>0.6</td>
<td>0.3</td>
<td>0.7</td>
<td>0.0</td>
<td>2.1</td>
<td>6.75 ± 0.14</td>
</tr>
<tr>
<td>80–100</td>
<td>7.51 ± 0.14</td>
<td>1.7</td>
<td>0.3</td>
<td>0.3</td>
<td>0.6</td>
<td>0.4</td>
<td>0.0</td>
<td>1.8</td>
<td>7.37 ± 0.14</td>
</tr>
<tr>
<td>100–120</td>
<td>6.65 ± 0.14</td>
<td>1.8</td>
<td>0.3</td>
<td>0.3</td>
<td>0.7</td>
<td>0.3</td>
<td>0.0</td>
<td>2.0</td>
<td>6.60 ± 0.14</td>
</tr>
<tr>
<td>120–150</td>
<td>4.823 ± 0.092</td>
<td>1.7</td>
<td>0.3</td>
<td>0.3</td>
<td>0.5</td>
<td>0.3</td>
<td>0.0</td>
<td>1.9</td>
<td>4.839 ± 0.092</td>
</tr>
<tr>
<td>150–200</td>
<td>2.707 ± 0.058</td>
<td>1.8</td>
<td>0.4</td>
<td>0.5</td>
<td>0.7</td>
<td>0.6</td>
<td>0.0</td>
<td>2.1</td>
<td>2.763 ± 0.059</td>
</tr>
<tr>
<td>200–250</td>
<td>1.180 ± 0.041</td>
<td>2.9</td>
<td>0.7</td>
<td>0.8</td>
<td>1.0</td>
<td>1.2</td>
<td>0.0</td>
<td>3.4</td>
<td>1.224 ± 0.042</td>
</tr>
<tr>
<td>250–300</td>
<td>0.563 ± 0.029</td>
<td>4.4</td>
<td>1.0</td>
<td>1.0</td>
<td>1.4</td>
<td>1.7</td>
<td>0.0</td>
<td>5.1</td>
<td>0.590 ± 0.030</td>
</tr>
<tr>
<td>300–400</td>
<td>0.191 ± 0.012</td>
<td>5.2</td>
<td>1.8</td>
<td>1.4</td>
<td>1.6</td>
<td>2.8</td>
<td>0.1</td>
<td>6.5</td>
<td>0.203 ± 0.013</td>
</tr>
<tr>
<td>400–500+</td>
<td>0.0763 ± 0.0076</td>
<td>8.0</td>
<td>2.6</td>
<td>1.9</td>
<td>2.0</td>
<td>4.5</td>
<td>0.1</td>
<td>9.9</td>
<td>0.0820 ± 0.0081</td>
</tr>
</tbody>
</table>
Table 5  Absolute and normalised differential cross-sections as functions of $|y^{e\mu}|$ (top) and $\Delta\phi^{e\mu}$ (bottom). The columns show the bin ranges, measured cross-section and total uncertainty, relative statistical uncertainty, relative systematic uncertainties in various categories (see text), total relative uncertainty, and differential cross-section corrected to remove contributions via $W \rightarrow \tau \rightarrow e/\mu$ decays. Relative uncertainties smaller than 0.05% are indicated by ‘0.0’. The bin boundaries for $\Delta\phi^{e\mu}$ correspond to exact multiples of $\pi/10$ but are quoted to two decimal places.

### Absolute bin

| Absolute bin (unit y) | $d\sigma/d|y^{e\mu}|$ (fb/unit y) | Stat. (%) | $t\bar{t}$ mod. (%) | Lept. (%) | $t\bar{t}$b (%) | Bkg. (%) | $L/E_b$ (%) | Total (%) | $d\sigma/d|y^{e\mu}|$ (no $\tau$) (fb/unit y) |
|-----------------------|----------------------------------|-----------|---------------------|-----------|----------------|---------|------------|-----------|----------------------------------|
| 0.00–0.25             | $3007 \pm 95$                   | 1.5       | 0.9                 | 1.0       | 1.0            | 0.8     | 2.1        | 3.2       | 2639 ± 84                         |
| 0.25–0.50             | $2681 \pm 86$                   | 1.5       | 0.9                 | 1.0       | 0.8            | 0.9     | 2.1        | 3.2       | 2353 ± 76                         |
| 0.50–0.75             | $2419 \pm 80$                   | 1.6       | 1.0                 | 1.0       | 1.0            | 2.1     | 3.3        | 3.3       | 2123 ± 71                         |
| 0.75–1.00             | $2026 \pm 71$                   | 1.9       | 1.1                 | 1.1       | 0.8            | 1.1     | 2.1        | 3.5       | 1780 ± 63                         |
| 1.00–1.25             | $1536 \pm 57$                   | 2.2       | 1.2                 | 1.1       | 0.8            | 1.0     | 2.1        | 3.7       | 1351 ± 50                         |
| 1.25–1.50             | $1038 \pm 43$                   | 2.7       | 1.3                 | 1.3       | 0.7            | 1.2     | 2.1        | 4.2       | 912 ± 38                          |
| 1.50–1.75             | $637 \pm 33$                    | 3.7       | 1.6                 | 1.6       | 1.3            | 1.3     | 2.2        | 5.2       | 561 ± 29                          |
| 1.75–2.00             | $321 \pm 23$                    | 5.7       | 2.0                 | 1.8       | 1.2            | 2.1     | 2.2        | 7.1       | 283 ± 20                          |
| 2.00–2.50             | $69.1 \pm 7.7$                  | 8.8       | 2.7                 | 2.3       | 2.3            | 4.7     | 2.2        | 11.1      | 61.3 ± 6.8                        |

### Normalised bin

<p>| Normalised bin (unit y) | $\frac{1}{\sigma} d\sigma/d|y^{e\mu}|$ (10^{-1}/unit y) | Stat. (%) | $t\bar{t}$ mod. (%) | Lept. (%) | $t\bar{t}$b (%) | Bkg. (%) | $L/E_b$ (%) | Total (%) | $\frac{1}{\sigma} d\sigma/d|y^{e\mu}|$ (no $\tau$) (10^{-1}/unit y) |
|------------------------|----------------------------------------------------------|-----------|---------------------|-----------|----------------|---------|------------|-----------|----------------------------------|
| 0.00–0.25              | $8.71 \pm 0.13$                                          | 1.3       | 0.3                 | 0.2       | 0.3            | 0.3     | 0.0        | 1.4       | 8.71 ± 0.13                      |
| 0.25–0.50              | $7.77 \pm 0.12$                                          | 1.4       | 0.3                 | 0.2       | 0.3            | 0.3     | 0.0        | 1.5       | 7.76 ± 0.12                      |
| 0.50–0.75              | $7.01 \pm 0.11$                                          | 1.5       | 0.3                 | 0.2       | 0.3            | 0.3     | 0.0        | 1.6       | 7.00 ± 0.11                      |
| 0.75–1.00              | $5.87 \pm 0.11$                                          | 1.7       | 0.4                 | 0.2       | 0.3            | 0.4     | 0.0        | 1.9       | 5.87 ± 0.11                      |
| 1.00–1.25              | $4.451 \pm 0.099$                                         | 2.1       | 0.5                 | 0.2       | 0.4            | 0.3     | 0.0        | 2.2       | 4.458 ± 0.099                    |
| 1.25–1.50              | $3.009 \pm 0.085$                                         | 2.6       | 0.6                 | 0.4       | 0.6            | 0.6     | 0.0        | 2.8       | 3.009 ± 0.085                    |
| 1.50–1.75              | $1.846 \pm 0.073$                                         | 3.6       | 0.8                 | 0.7       | 0.8            | 0.8     | 0.0        | 4.0       | 1.850 ± 0.073                    |
| 1.75–2.00              | $0.930 \pm 0.057$                                         | 5.6       | 1.3                 | 1.0       | 0.9            | 1.8     | 0.1        | 6.2       | 0.935 ± 0.058                    |
| 2.00–2.50              | $0.200 \pm 0.021$                                         | 8.8       | 2.1                 | 1.7       | 2.5            | 4.4     | 0.1        | 10.5      | 0.202 ± 0.021                    |</p>
<table>
<thead>
<tr>
<th>Absolute bin (rad)</th>
<th>$\frac{d\sigma}{d\Delta\phi^{\mu}}$ (fb/ rad)</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/b (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_{b}$ (%)</th>
<th>Total (%)</th>
<th>$\frac{d\sigma}{d\Delta\phi^{\mu}}$ (no $\tau$) (fb/ rad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00–0.31</td>
<td>696 ± 30</td>
<td>2.9</td>
<td>1.4</td>
<td>1.3</td>
<td>0.7</td>
<td>0.9</td>
<td>2.1</td>
<td>4.2</td>
<td>630 ± 27</td>
</tr>
<tr>
<td>0.31–0.63</td>
<td>735 ± 29</td>
<td>2.7</td>
<td>1.3</td>
<td>1.3</td>
<td>0.6</td>
<td>0.9</td>
<td>2.1</td>
<td>4.0</td>
<td>664 ± 26</td>
</tr>
<tr>
<td>0.63–0.94</td>
<td>780 ± 31</td>
<td>2.6</td>
<td>1.3</td>
<td>1.2</td>
<td>0.7</td>
<td>0.9</td>
<td>2.1</td>
<td>3.9</td>
<td>704 ± 28</td>
</tr>
<tr>
<td>0.94–1.26</td>
<td>850 ± 33</td>
<td>2.5</td>
<td>1.2</td>
<td>1.2</td>
<td>0.7</td>
<td>0.9</td>
<td>2.1</td>
<td>3.9</td>
<td>763 ± 30</td>
</tr>
<tr>
<td>1.26–1.57</td>
<td>947 ± 36</td>
<td>2.3</td>
<td>1.1</td>
<td>1.2</td>
<td>0.8</td>
<td>1.0</td>
<td>2.1</td>
<td>3.8</td>
<td>844 ± 32</td>
</tr>
<tr>
<td>1.57–1.88</td>
<td>1103 ± 41</td>
<td>2.2</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>0.9</td>
<td>2.1</td>
<td>3.7</td>
<td>977 ± 37</td>
</tr>
<tr>
<td>1.88–2.20</td>
<td>1235 ± 43</td>
<td>2.1</td>
<td>1.0</td>
<td>1.0</td>
<td>0.7</td>
<td>0.9</td>
<td>2.1</td>
<td>3.5</td>
<td>1084 ± 38</td>
</tr>
<tr>
<td>2.20–2.51</td>
<td>1410 ± 50</td>
<td>2.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.0</td>
<td>2.1</td>
<td>3.5</td>
<td>1226 ± 44</td>
</tr>
<tr>
<td>2.51–2.83</td>
<td>1575 ± 56</td>
<td>2.0</td>
<td>0.9</td>
<td>0.9</td>
<td>1.1</td>
<td>1.1</td>
<td>2.1</td>
<td>3.6</td>
<td>1353 ± 49</td>
</tr>
<tr>
<td>2.83–3.14</td>
<td>1696 ± 58</td>
<td>1.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>1.2</td>
<td>2.1</td>
<td>3.4</td>
<td>1449 ± 51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normalised bin (rad)</th>
<th>$\frac{1}{\Delta\phi} d\sigma/d\Delta\phi^{\mu} (10^{-1}/\text{rad})$</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/b (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_{b}$ (%)</th>
<th>Total (%)</th>
<th>$\frac{1}{\Delta\phi} d\sigma/d\Delta\phi^{\mu}$ (no $\tau$) (10^{-1}/\text{rad})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00–0.31</td>
<td>2.010 ± 0.060</td>
<td>2.8</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
<td>0.0</td>
<td>3.0</td>
<td>2.068 ± 0.062</td>
</tr>
<tr>
<td>0.31–0.63</td>
<td>2.121 ± 0.057</td>
<td>2.6</td>
<td>0.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.0</td>
<td>2.7</td>
<td>2.179 ± 0.058</td>
</tr>
<tr>
<td>0.63–0.94</td>
<td>2.252 ± 0.059</td>
<td>2.5</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>0.4</td>
<td>0.0</td>
<td>2.6</td>
<td>2.311 ± 0.060</td>
</tr>
<tr>
<td>0.94–1.26</td>
<td>2.454 ± 0.064</td>
<td>2.4</td>
<td>0.5</td>
<td>0.2</td>
<td>0.6</td>
<td>0.4</td>
<td>0.0</td>
<td>2.6</td>
<td>2.506 ± 0.065</td>
</tr>
<tr>
<td>1.26–1.57</td>
<td>2.732 ± 0.064</td>
<td>2.2</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.0</td>
<td>2.3</td>
<td>2.773 ± 0.065</td>
</tr>
<tr>
<td>1.57–1.88</td>
<td>3.185 ± 0.070</td>
<td>2.1</td>
<td>0.4</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3</td>
<td>0.0</td>
<td>2.2</td>
<td>3.207 ± 0.071</td>
</tr>
<tr>
<td>1.88–2.20</td>
<td>3.566 ± 0.073</td>
<td>2.0</td>
<td>0.4</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.0</td>
<td>2.0</td>
<td>3.559 ± 0.072</td>
</tr>
<tr>
<td>2.20–2.51</td>
<td>4.069 ± 0.079</td>
<td>1.9</td>
<td>0.4</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3</td>
<td>0.0</td>
<td>1.9</td>
<td>4.028 ± 0.078</td>
</tr>
<tr>
<td>2.51–2.83</td>
<td>4.546 ± 0.088</td>
<td>1.8</td>
<td>0.4</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.0</td>
<td>1.9</td>
<td>4.443 ± 0.086</td>
</tr>
<tr>
<td>2.83–3.14</td>
<td>4.897 ± 0.090</td>
<td>1.7</td>
<td>0.4</td>
<td>0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.0</td>
<td>1.8</td>
<td>4.757 ± 0.088</td>
</tr>
</tbody>
</table>
Table 6 Absolute and normalised differential cross-sections as functions of $p_T^\tau + p_T^\mu$ (top) and $E^\tau + E^\mu$ (bottom). The columns show the bin ranges, measured cross-section and total uncertainty, relative statistical uncertainty, relative systematic uncertainties in various categories (see text), total relative uncertainty, and differential cross-section corrected to remove contributions via $W \rightarrow \tau \rightarrow e/\mu$ decays. Relative uncertainties smaller than 0.05% are indicated by ‘0.0’. The last bin includes overflows where indicated by the ‘+’ sign.

<table>
<thead>
<tr>
<th>Absolute bin (GeV)</th>
<th>$\frac{d\sigma}{dp_T^\tau + p_T^\mu}$ (fb/GeV)</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>$Jet/b$ (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_b$ (%)</th>
<th>Total (%)</th>
<th>$\frac{1}{2} \frac{d\sigma}{dp_T^\tau + p_T^\mu}$ (no $\tau$) $(10^{-2}/\text{GeV})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50–80</td>
<td>23.02 ± 0.89</td>
<td>1.7</td>
<td>1.2</td>
<td>1.8</td>
<td>0.7</td>
<td>1.5</td>
<td>2.1</td>
<td>3.8</td>
<td>18.90 ± 0.73</td>
</tr>
<tr>
<td>80–100</td>
<td>38.0 ± 1.2</td>
<td>1.5</td>
<td>1.0</td>
<td>1.2</td>
<td>0.8</td>
<td>0.9</td>
<td>2.1</td>
<td>3.3</td>
<td>33.0 ± 1.1</td>
</tr>
<tr>
<td>100–120</td>
<td>34.3 ± 1.1</td>
<td>1.6</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
<td>0.9</td>
<td>2.1</td>
<td>3.2</td>
<td>30.5 ± 1.0</td>
</tr>
<tr>
<td>120–150</td>
<td>21.00 ± 0.69</td>
<td>1.7</td>
<td>0.9</td>
<td>1.0</td>
<td>0.8</td>
<td>1.0</td>
<td>2.1</td>
<td>3.3</td>
<td>18.95 ± 0.63</td>
</tr>
<tr>
<td>150–200</td>
<td>9.11 ± 0.34</td>
<td>1.9</td>
<td>1.0</td>
<td>1.1</td>
<td>1.2</td>
<td>1.5</td>
<td>2.2</td>
<td>3.8</td>
<td>8.27 ± 0.31</td>
</tr>
<tr>
<td>200–250</td>
<td>3.03 ± 0.16</td>
<td>3.5</td>
<td>1.3</td>
<td>1.5</td>
<td>1.4</td>
<td>2.4</td>
<td>2.2</td>
<td>5.3</td>
<td>2.78 ± 0.15</td>
</tr>
<tr>
<td>250–300</td>
<td>1.08 ± 0.09</td>
<td>5.7</td>
<td>1.9</td>
<td>2.0</td>
<td>2.9</td>
<td>3.5</td>
<td>2.2</td>
<td>8.1</td>
<td>0.99 ± 0.08</td>
</tr>
<tr>
<td>300–400+</td>
<td>0.38 ± 0.04</td>
<td>7.0</td>
<td>3.3</td>
<td>2.3</td>
<td>3.3</td>
<td>6.5</td>
<td>2.3</td>
<td>11.1</td>
<td>0.35 ± 0.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normalised bin (GeV)</th>
<th>$\frac{1}{2} \frac{d\sigma}{dp_T^\tau + p_T^\mu}$ $(10^{-2}/\text{GeV})$</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>$Jet/b$ (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_b$ (%)</th>
<th>Total (%)</th>
<th>$\frac{1}{2} \frac{d\sigma}{dp_T^\tau + p_T^\mu}$ (no $\tau$) $(10^{-2}/\text{GeV})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50–80</td>
<td>0.664 ± 0.015</td>
<td>1.5</td>
<td>0.4</td>
<td>1.0</td>
<td>0.3</td>
<td>1.2</td>
<td>0.0</td>
<td>2.2</td>
<td>0.621 ± 0.014</td>
</tr>
<tr>
<td>80–100</td>
<td>1.097 ± 0.017</td>
<td>1.3</td>
<td>0.3</td>
<td>0.5</td>
<td>0.4</td>
<td>0.6</td>
<td>0.0</td>
<td>1.6</td>
<td>1.085 ± 0.018</td>
</tr>
<tr>
<td>100–120</td>
<td>0.990 ± 0.016</td>
<td>1.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.0</td>
<td>1.6</td>
<td>1.003 ± 0.016</td>
</tr>
<tr>
<td>120–150</td>
<td>0.606 ± 0.010</td>
<td>1.5</td>
<td>0.3</td>
<td>0.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.0</td>
<td>1.7</td>
<td>0.623 ± 0.010</td>
</tr>
<tr>
<td>150–200</td>
<td>0.2627 ± 0.0062</td>
<td>1.8</td>
<td>0.4</td>
<td>0.9</td>
<td>0.6</td>
<td>1.0</td>
<td>0.0</td>
<td>2.4</td>
<td>0.2716 ± 0.0063</td>
</tr>
<tr>
<td>200–250</td>
<td>0.0875 ± 0.0039</td>
<td>3.4</td>
<td>0.7</td>
<td>1.4</td>
<td>1.4</td>
<td>2.1</td>
<td>0.1</td>
<td>4.5</td>
<td>0.0912 ± 0.0041</td>
</tr>
<tr>
<td>250–300</td>
<td>0.0311 ± 0.0023</td>
<td>5.6</td>
<td>1.3</td>
<td>2.0</td>
<td>2.5</td>
<td>3.2</td>
<td>0.1</td>
<td>7.3</td>
<td>0.0326 ± 0.0024</td>
</tr>
<tr>
<td>300–400+</td>
<td>0.0110 ± 0.0011</td>
<td>7.0</td>
<td>2.5</td>
<td>2.4</td>
<td>2.9</td>
<td>6.2</td>
<td>0.2</td>
<td>10.4</td>
<td>0.0116 ± 0.0012</td>
</tr>
</tbody>
</table>
Table 6 continued

<table>
<thead>
<tr>
<th>Absolute bin (GeV)</th>
<th>$d\sigma/d(E^e + E^\mu)$ (fb/GeV)</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/b (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_b$ (%)</th>
<th>Total (%)</th>
<th>$d\sigma/d(E^e + E^\mu)$ (no $\tau$) (fb/GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50–80</td>
<td>4.05 ± 0.22</td>
<td>3.8</td>
<td>1.3</td>
<td>2.2</td>
<td>1.1</td>
<td>1.2</td>
<td>2.1</td>
<td>3.3</td>
<td>5.3 ± 0.17</td>
</tr>
<tr>
<td>80–100</td>
<td>13.68 ± 0.57</td>
<td>2.5</td>
<td>1.2</td>
<td>1.6</td>
<td>1.2</td>
<td>1.1</td>
<td>2.1</td>
<td>4.2</td>
<td>11.38 ± 0.48</td>
</tr>
<tr>
<td>100–120</td>
<td>18.36 ± 0.67</td>
<td>2.1</td>
<td>1.1</td>
<td>1.3</td>
<td>0.8</td>
<td>0.9</td>
<td>2.1</td>
<td>3.6</td>
<td>15.67 ± 0.57</td>
</tr>
<tr>
<td>120–150</td>
<td>19.10 ± 0.64</td>
<td>1.7</td>
<td>1.1</td>
<td>1.1</td>
<td>0.8</td>
<td>0.9</td>
<td>2.1</td>
<td>3.3</td>
<td>16.63 ± 0.56</td>
</tr>
<tr>
<td>150–200</td>
<td>15.79 ± 0.51</td>
<td>1.5</td>
<td>1.0</td>
<td>0.9</td>
<td>1.0</td>
<td>0.9</td>
<td>2.1</td>
<td>3.2</td>
<td>13.92 ± 0.45</td>
</tr>
<tr>
<td>200–250</td>
<td>10.04 ± 0.35</td>
<td>1.8</td>
<td>1.0</td>
<td>0.9</td>
<td>1.1</td>
<td>1.1</td>
<td>2.1</td>
<td>3.5</td>
<td>8.97 ± 0.31</td>
</tr>
<tr>
<td>250–300</td>
<td>6.24 ± 0.25</td>
<td>2.4</td>
<td>1.0</td>
<td>1.0</td>
<td>1.2</td>
<td>1.4</td>
<td>2.1</td>
<td>4.0</td>
<td>5.61 ± 0.22</td>
</tr>
<tr>
<td>300–400</td>
<td>3.04 ± 0.13</td>
<td>2.5</td>
<td>1.1</td>
<td>1.3</td>
<td>0.9</td>
<td>1.7</td>
<td>2.2</td>
<td>4.2</td>
<td>2.75 ± 0.12</td>
</tr>
<tr>
<td>400–500</td>
<td>1.20 ± 0.07</td>
<td>4.3</td>
<td>1.5</td>
<td>2.0</td>
<td>1.3</td>
<td>2.4</td>
<td>2.2</td>
<td>6.1</td>
<td>1.10 ± 0.07</td>
</tr>
<tr>
<td>500–700+</td>
<td>0.48 ± 0.04</td>
<td>4.8</td>
<td>2.3</td>
<td>2.3</td>
<td>1.6</td>
<td>3.8</td>
<td>2.2</td>
<td>7.4</td>
<td>0.44 ± 0.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normalised bin</th>
<th>Normalised $d\sigma/d(E^e + E^\mu)$ $(10^{-3}$/GeV)</th>
<th>Stat. (%)</th>
<th>$t\bar{t}$ mod. (%)</th>
<th>Lept. (%)</th>
<th>Jet/b (%)</th>
<th>Bkg. (%)</th>
<th>$L/E_b$ (%)</th>
<th>Total (%)</th>
<th>Normalised $d\sigma/d(E^e + E^\mu)$ $(10^{-3}$/GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50–80</td>
<td>1.172 ± 0.050</td>
<td>3.8</td>
<td>0.8</td>
<td>1.4</td>
<td>0.7</td>
<td>1.1</td>
<td>0.1</td>
<td>4.3</td>
<td>1.058 ± 0.046</td>
</tr>
<tr>
<td>80–100</td>
<td>3.95 ± 0.11</td>
<td>2.4</td>
<td>0.5</td>
<td>0.8</td>
<td>0.7</td>
<td>0.9</td>
<td>0.0</td>
<td>2.8</td>
<td>3.75 ± 0.11</td>
</tr>
<tr>
<td>100–120</td>
<td>5.31 ± 0.12</td>
<td>2.0</td>
<td>0.4</td>
<td>0.6</td>
<td>0.6</td>
<td>0.7</td>
<td>0.0</td>
<td>2.3</td>
<td>5.16 ± 0.12</td>
</tr>
<tr>
<td>120–150</td>
<td>5.521 ± 0.099</td>
<td>1.6</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.0</td>
<td>1.8</td>
<td>5.478 ± 0.099</td>
</tr>
<tr>
<td>150–200</td>
<td>4.564 ± 0.067</td>
<td>1.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.0</td>
<td>1.5</td>
<td>4.585 ± 0.067</td>
</tr>
<tr>
<td>200–250</td>
<td>2.904 ± 0.055</td>
<td>1.7</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
<td>0.0</td>
<td>1.9</td>
<td>2.955 ± 0.056</td>
</tr>
<tr>
<td>250–300</td>
<td>1.803 ± 0.048</td>
<td>2.3</td>
<td>0.5</td>
<td>0.5</td>
<td>0.9</td>
<td>0.7</td>
<td>0.0</td>
<td>2.6</td>
<td>1.849 ± 0.049</td>
</tr>
<tr>
<td>300–400</td>
<td>0.878 ± 0.026</td>
<td>2.4</td>
<td>0.5</td>
<td>0.8</td>
<td>0.7</td>
<td>1.1</td>
<td>0.0</td>
<td>2.9</td>
<td>0.907 ± 0.026</td>
</tr>
<tr>
<td>400–500</td>
<td>0.348 ± 0.018</td>
<td>4.2</td>
<td>1.0</td>
<td>1.7</td>
<td>1.2</td>
<td>2.0</td>
<td>0.1</td>
<td>5.2</td>
<td>0.362 ± 0.019</td>
</tr>
<tr>
<td>500–700+</td>
<td>0.1393 ± 0.0091</td>
<td>4.7</td>
<td>1.8</td>
<td>1.9</td>
<td>1.2</td>
<td>3.5</td>
<td>0.1</td>
<td>6.5</td>
<td>0.1463 ± 0.0095</td>
</tr>
</tbody>
</table>
Fig. 7 Normalised differential cross-sections as a function of a $p_T^\ell$, b $|\eta^\ell|$, c $p_T^{\mu e}$ and d $m^{\mu e}$. The measured values are shown by the black points with error bars corresponding to the data statistical uncertainties and cyan bands corresponding to the total uncertainties in each bin, and include the contributions via $W \to \tau \to e/\mu$ decays. The results are compared to the predictions from the POWHEG + PYTHIA6, MC@NLO + HERWIG and ALPGEN + HERWIG $t\bar{t}$ simulation samples. The lower plots show the ratios of predictions to data, with the error bars indicating the data statistical uncertainties and the cyan bands indicating the total uncertainties in the measurements.

The first group shows the baseline POWHEG + PYTHIA6 $t\bar{t}$ sample with $h_{damp}=m_t$ (which is also shown in Figs. 7 and 8), together with the two tunes giving more or less parton shower radiation – the Perugia 2012 radHi and radLo tunes [54] coupled to scale and $h_{damp}$ parameter variations as discussed in Sect. 2.

The second group shows a POWHEG + PYTHIA6 sample with $h_{damp} = \infty$ (i.e. no damping of the first emission), the baseline POWHEG + PYTHIA6 sample with the top quark $p_T$ spectrum reweighted to the NNLO prediction of Ref. [25], and a sample generated with POWHEG and $h_{damp} = m_t$ interfaced to PYTHIA8 (version 8.186) [46] with the
Fig. 8  Normalised differential cross-sections as a function of a $|y^{\mu e}|$, b $\Delta \phi^{\mu e}$, c $p_T^{\mu} + p_T^{e}$ and d $E^{e} + E^{\mu}$. The measured values are shown by the black points with error bars corresponding to the data statistical uncertainties and cyan bands corresponding to the total uncertainties in each bin, and include the contributions via $W \to \tau \to e/\mu$ decays. The results are compared to the predictions from the POWHEG + PYTHIA6, MC@NLO + HERWIG and ALPGEN + HERWIG tt simulation samples. The lower plots show the ratios of predictions to data, with the error bars indicating the data statistical uncertainties and the cyan bands indicating the total uncertainties in the measurements.

The third group shows a POWHEG + PYTHIA6 sample with $h_{damp} = m_t$ generated with the HERAPDF 1.5 PDF set [84,85] instead of CT10,6 and a POWHEG + PYTHIA6 sample with $h_{damp} = \infty$ and no simulation of spin correlations between the top and antitop quarks.

The fourth group shows alternative matrix-element event generators – the ALPGEN + HERWIG and MC@NLO + HERWIG samples described in Sect. 2 and shown in Figs. 7 and 8, together with a sample generated using the A14 tune [96] and the CTEQ6L1 PDF set for the parton shower, hadronisation and underlying event modelling as described in Ref. [94].
Fig. 9 Ratios of predictions of normalised differential cross-sections to data as a function of (a) $p_T^\ell$ and (b) $|\eta^\ell|$. The data statistical uncertainties are shown by the black error bars around a ratio of unity, and the total uncertainties are shown by the cyan bands. The $t\bar{t}$ predictions are shown in four groups from top to bottom, with error bars indicating the uncertainties due to the limited size of the simulated samples.

The compatibility of each prediction with each measured normalised distribution was assessed quantitatively using a $\chi^2$ test, calculated as:

$$\chi^2 = \sum_{n=1}^{n-1} (S_{(n-1)} - \Delta_{(n-1)} \Delta_{(n-1)})^2,$$

where $\Delta_{(n-1)}$ is the vector of differences between the measured and predicted normalised differential cross-section in each of the $n$ bins, excluding the last one, and $S_{(n-1)}$ is the corresponding covariance matrix, including both the experimental uncertainties in the measurement and the statistical uncertainties in the predictions. Bin-to-bin correlations in both the statistical (from the normalisation condition) and systematic uncertainties were taken into account via off-diagonal entries. The last bin of each distribution was excluded due to the normalisation condition, rendering the covariance matrix $S_{(n-1)}$ invertible. The resulting $\chi^2$ values, number of degrees of freedom $(n-1)$ and corresponding $\chi^2$ probability $p$-values are shown for each distribution and prediction in Table 8.

As can be seen from Fig. 9, in the single-lepton $p_T^\ell$ distribution, the data are softer than the predictions from POWHEG with CT10 PDFs, interfaced to HERWIG++ with the UE-EE-5 HERWIG++ author tune. The POWHEG-based predictions do not depend strongly on the choice of parton shower/hadronisation model or tune parameters controlling the amount of radiation. However, the agreement with data is improved when using HERAPDF 1.5 or reweighting to the NNLO top quark $p_T$ prediction from Ref. [25]. The predictions from the samples with alternative matrix-element event generators, i.e. MC@NLO + HERWIG and ALPGEN + HERWIG, are also harder than the data, though

$\chi^2$ value does not depend on the choice of which bin is removed.
AMC@NLO + HERWIG++ describes the data well. The $p_T^e + p_T^\mu$ and $E^e + E^\mu$ distributions (Fig. 12) show some similar features to $p_T^e$, being softer than the predictions from the POWHEG + PYTHIA6 samples with CT10, and better described with HERAPDF 1.5, and by AMC@NLO + HERWIG++.

The predictions for the single lepton $|\eta|^{e,\mu}$ and dilepton $|\eta|^{e,\mu}$ distributions (Figs. 9, 10, 11) are insensitive to the choice of parton shower/hadronisation model and tune, and are also insensitive to the top quark $p_T$ reweighting. The data distributions are more central than the predictions of all the NLO event generators (POWHEG, MC@NLO and AMC@NLO) with CT10 PDFs, but are better described by POWHEG with HERAPDF 1.5, and to a lesser extent also by ALPGEN + HERWIG, which uses the leading-order CTEQ6L1 PDF. These distributions, whose experimental measurements are limited by statistical uncertainties over the full kinematic range, are thus particularly suitable for constraining PDFs, as explored further in Sect. 7.

The dilepton $p_T^{e,\mu}$ and $m^{e,\mu}$ distributions (Fig. 10) are generally well described by all the NLO event generators, except for AMC@NLO which does not model the data well at low $m^{e,\mu}$. The $p_T^{e,\mu}$ distribution is sensitive to the amount of parton radiation, and is better described by the radHi POWHEG + PYTHIA6 sample, and by $h_{\text{damp}} = m_t$, than $h_{\text{damp}} = \infty$. Both distributions are sensitive to the modelling of $t\bar{t}$ spin correlations, and are not well-modelled by the POWHEG + PYTHIA6 sample without spin correlations.

The $\Delta\phi^{e,\mu}$ distribution (Fig. 11) is particularly sensitive to spin correlations, and has been previously used to exclude $t\bar{t}$ simulation models without spin correlation and the pair-production of supersymmetric top squarks with masses close to $m_t$, via template fits to reconstruction-level distributions [100,101]. The particle-level $\Delta\phi^{e,\mu}$ measurements shown here also exclude the prediction without spin correlations and the LO implementation of spin correlations in the ALPGEN + HERWIG sample. The $\Delta\phi^{e,\mu}$ distribution
is also sensitive to radiation, this time favouring the radHi
Powheg + Pythia6 sample.

The $\chi^2$ formalism of Eq. (5) was extended to consider
several normalised distributions simultaneously, by forming
vectors $\Delta_i$ where the index runs over the bins of several
distributions, excluding the last bin in each one to account for
the normalisation condition. The covariance matrix $S$
was extended with off-block-diagonal components encoding the
correlations between bins of different measured distributions.
The statistical correlations between distributions were evaluated
using pseudo-experiments generated by sampling from
the large simulated $t\bar{t}$ sample discussed in Sect. 4.3. The
individual sources of systematic uncertainty were assumed
to be fully correlated across the different distributions. Five
sets of combined distributions were considered: the combination
of $p_T^{\ell}$ and $p_T^{e\mu}$, combining all the information from
single and dilepton $p_T$; the combination of $p_T^{e\mu}$, $m^{\ell\mu}$
and $p_T^{e} + p_T^{\mu}$, including all the dilepton kinematic distributions
except rapidity; the combination of $|\eta^{\ell}|$ and $|\eta^{e\mu}|$, combin-
ing the single and dilepton rapidity information; the combination
of $|\eta^{\ell}|$, $|\eta^{e\mu}|$ and $E^{\ell} + E^{\mu}$, combining all the distributions
with longitudinal information; and the combination of all eight measured distributions, denoted ‘All’. The resulting $\chi^2$
values, numbers of degrees of freedom and $p$-values are
shown for each combination and prediction in Table 9.

The results for the combinations of distributions reflect the
observations for the individual distributions. The best mod-
eling of the first two combinations (involving $p_T^{\ell}$, $p_T^{e\mu}$, $m^{e\mu}$
and $p_T^{e} + p_T^{\mu}$) is given by Powheg+Pythia6 with either
HERAPDF 1.5 or with CT10 plus reweighting of the top
quark $p_T$ distribution to the NNLO prediction; the radHi vari-
ation of Powheg + Pythia6 also does well. The combinations
involving $|\eta^{\ell}|$ and $|\eta^{e\mu}|$ and the combination of all eight
evaluations are only well-described by Powheg+Pythia6
with HERAPDF 1.5, and marginally well described by the
radHi variation. All other event generator setups (in par-
Fig. 12 Ratios of predictions of normalised differential cross-sections to data as a function of (a) $p_T^{\ell^+} + p_T^{\ell^-}$ and (b) $E^\ell^+ + E^\ell^-$. The data statistical uncertainties are shown by the black error bars around a ratio of unity, and the total uncertainties are shown by the cyan bands. The $t\bar{t}$ predictions are shown in four groups from top to bottom, with error bars indicating the uncertainties due to the limited size of the simulated samples.

Table 7 Summary of particle-level simulation samples used in the comparison to the corrected data distributions in Sect. 6.2, giving the matrix-element event generator, PDF set, parton shower and associated tune parameter set. The four groups shown correspond to the four panels for each measured distribution shown in Figs. 9, 10, 11 and 12.

<table>
<thead>
<tr>
<th>Matrix-element</th>
<th>PDF</th>
<th>Parton shower</th>
<th>Tune</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>POWHEG</td>
<td>CT10</td>
<td>PYTHA6</td>
<td>P2011C</td>
</tr>
<tr>
<td></td>
<td>POWHEG</td>
<td>CT10</td>
<td>PYTHA6</td>
<td>P2012 radHi</td>
</tr>
<tr>
<td></td>
<td>POWHEG</td>
<td>CT10</td>
<td>PYTHA6</td>
<td>P2012 radLo</td>
</tr>
<tr>
<td>2</td>
<td>POWHEG</td>
<td>CT10</td>
<td>PYTHA6</td>
<td>P2011C</td>
</tr>
<tr>
<td></td>
<td>POWHEG</td>
<td>CT10</td>
<td>PYTHA6</td>
<td>P2011C</td>
</tr>
<tr>
<td></td>
<td>POWHEG</td>
<td>CT10</td>
<td>PYTHA6</td>
<td>A14</td>
</tr>
<tr>
<td>3</td>
<td>POWHEG</td>
<td>HERAPDF 1.5</td>
<td>PYTHA6</td>
<td>P2011C</td>
</tr>
<tr>
<td></td>
<td>POWHEG</td>
<td>CT10</td>
<td>PYTHA6</td>
<td>P2011C</td>
</tr>
<tr>
<td>4</td>
<td>ALPGEN</td>
<td>CTEQ6L1</td>
<td>HERWIG+JIMMY</td>
<td>AUET2</td>
</tr>
<tr>
<td></td>
<td>MC@NLO</td>
<td>CT10</td>
<td>HERWIG+JIMMY</td>
<td>AUET2</td>
</tr>
<tr>
<td></td>
<td>MG5_AMC@NLO</td>
<td>CT10</td>
<td>HERWIG++</td>
<td>UE-EE-5</td>
</tr>
</tbody>
</table>
ticular the LO multileg event generator Alpgen) fail to describe some of the distributions, but this could potentially be improved by appropriate parameter tuning and switching to a different PDF set. These results highlight the sensitivity of the differential distributions to the choice of PDF, in particular that of the gluon, as discussed further in Sect. 7. They also indicate that NNLO corrections may be important in describing the kinematics of the decay leptons, as well as for the prediction of the top quark $p_T$ spectrum as discussed in Ref. [25].

6.3 Comparison with fixed-order predictions

The comparisons described in Sect. 6.2 show that the predictions are strongly sensitive to the choice of PDF, and also to the QCD scale (whose variation approximates the effects of missing higher-order corrections) and other parameters related to the amount of radiation. In this section, these aspects are further explored using a set of predictions from the MCFM program (version 6.8) [30], combined with APPLGRID (version 1.4.73) [93] to interface to various PDF sets available in LHAPDF (version 6.1.5) [102]. Four recent NLO PDF sets were considered, namely CT14 [103], MMHT14 [104], NNPDF 3.0 [105] and HERAPDF 2.0 [97]. The data were also compared to HERAPDF 1.5 [85] for comparison with the results of Sect. 6.2; the results from these two PDF sets are similar.

MCFM provides an NLO fixed-order prediction of the $t\bar{t}$ process in the dilepton channel, including NLO QCD corrections in both production and decay in the on-shell approximation, and full NLO spin correlations [106]. Only the direct decays of $W \rightarrow e/\mu$ are included, so these predictions were compared to the measurements corrected to remove the leptonic $\tau$ decay contributions. The top quark mass $m_t$ was set to 172.5 GeV. Informed by the discussion in Ref. [107], the central values for the QCD renormalisation and factorisation scales were set to $m_t/2$, the lower than typical ($m_t$) scale choice being intended to account for the impact of resummed soft-gluon contributions not included in the fixed-order calculations. The MCFM predictions do not include quantum electrodynamics (QED) final state photon radiation, unlike the experimental measurements where the lep-

Table 8 The $\chi^2$ values (top) and associated probabilities (bottom) for comparison of measured normalised differential fiducial cross-sections with various $t\bar{t}$ simulation samples. Probabilities smaller than $10^{-10}$ are shown as zero.

| Generator                  | $p_T^\ell$ | $|\eta^\ell|$ | $p_T^{\mu}$ | $m^{\mu}$ | $|\eta^{\mu}|$ | $\Delta\phi^{\mu}$ | $p_T^{\ell} + p_T^{\mu}$ | $E^\ell + E^{\mu}$ |
|----------------------------|------------|---------------|-------------|-----------|----------------|---------------------|------------------------|---------------------|
| POWHEG + PY6               | 13.6       | 26.3          | 7.3         | 14.6      | 46.6          | 14.0                | 11.3                   | 22.7                 |
| POWHEG + PY6 radLo         | 15.9       | 22.9          | 7.6         | 14.6      | 45.6          | 25.9                | 14.0                   | 22.0                 |
| POWHEG + PY6 radHi         | 10.0       | 28.2          | 11.0        | 12.6      | 42.0          | 4.5                 | 9.1                    | 21.4                 |
| POWHEG + PY6 $h_{damp} = \infty$ | 17.2     | 22.5          | 14.5        | 12.9      | 42.8          | 5.0                 | 15.6                   | 23.4                 |
| POWHEG + PY6 $p_T$ NNLO     | 8.3        | 28.5          | 6.3         | 12.1      | 49.2          | 7.6                 | 7.6                    | 17.4                 |
| POWHEG + PY8 $h_{damp} = \infty$ | 15.1     | 28.9          | 8.3         | 14.4      | 44.3          | 13.0                | 12.7                   | 25.8                 |
| POWHEG + PY6 HERAPDF 1.5   | 11.4       | 11.8          | 3.6         | 11.1      | 6.7           | 10.3                | 7.0                    | 1.9                  |
| POWHEG + PY6 no spin corl. | 21.8       | 23.2          | 152         | 100       | 45.3          | 279                 | 22.4                   | 27.6                 |
| ALPGEN + HW                | 31.2       | 11.6          | 15.5        | 13.7      | 15.3          | 36.0                | 27.4                   | 12.7                 |
| MC@NLO + HW               | 15.7       | 18.8          | 9.4         | 9.3       | 39.4          | 7.1                 | 11.8                   | 16.2                 |
| AMC@NLO + HW++            | 7.8        | 29.2          | 7.6         | 24.5      | 46.6          | 8.2                 | 12.0                   | 13.8                 |

| Generator                  | $p_T^\ell$ | $|\eta^\ell|$ | $p_T^{\mu}$ | $m^{\mu}$ | $|\eta^{\mu}|$ | $\Delta\phi^{\mu}$ | $p_T^{\ell} + p_T^{\mu}$ | $E^\ell + E^{\mu}$ |
|----------------------------|------------|---------------|-------------|-----------|----------------|---------------------|------------------------|---------------------|
| POWHEG + PY6               | 0.14       | $9 \times 10^{-4}$ | 0.51      | 0.20      | 2 $\times 10^{-7}$ | 0.12                | 0.13                   | $7 \times 10^{-3}$   |
| POWHEG + PY6 radLo         | 0.070      | $3 \times 10^{-3}$ | 0.48      | 0.20      | 3 $\times 10^{-7}$ | 2 $\times 10^{-3}$ | 0.052                   | $9 \times 10^{-3}$   |
| POWHEG + PY6 radHi         | 0.35       | $4 \times 10^{-4}$ | 0.20      | 0.32      | 1 $\times 10^{-6}$ | 0.87                | 0.24                    | 0.011                |
| POWHEG + PY6 $h_{damp} = \infty$ | 0.045   | $4 \times 10^{-3}$ | 0.069     | 0.30      | 1 $\times 10^{-6}$ | 0.83                | 0.029                   | $5 \times 10^{-3}$   |
| POWHEG + PY6 $p_T$ NNLO     | 0.51       | $4 \times 10^{-4}$ | 0.62      | 0.36      | 6 $\times 10^{-8}$ | 0.57                | 0.36                    | 0.043                |
| POWHEG + PY8 $h_{damp} = \infty$ | 0.089   | $3 \times 10^{-4}$ | 0.41      | 0.21      | 5 $\times 10^{-7}$ | 0.16                | 0.080                   | $2 \times 10^{-3}$   |
| POWHEG + PY6 HERAPDF 1.5   | 0.25       | 0.16          | 0.89       | 0.44      | 0.57          | 0.32                | 0.43                    | 0.99                 |
| POWHEG + PY6 no spin corl. | 0.010      | $3 \times 10^{-3}$ | 0         | 0         | 3 $\times 10^{-7}$ | 0                  | $2 \times 10^{-3}$      | $1 \times 10^{-3}$   |
| ALPGEN + HW                | $3 \times 10^{-4}$ | 0.17          | 0.051     | 0.25      | 0.054         | $4 \times 10^{-5}$ | $3 \times 10^{-4}$      | 0.17                 |
| MC@NLO + HW               | 0.073      | 0.016         | 0.31       | 0.60      | 4 $\times 10^{-6}$ | 0.62                | 0.11                    | 0.063                |
| AMC@NLO + HW++            | 0.56       | $3 \times 10^{-4}$ | 0.47      | 0.011     | 2 $\times 10^{-7}$ | 0.52                | 0.10                    | 0.13                 |
Table 9 The $\chi^2$ values (top) and associated probabilities (bottom) for comparison of combinations of measured normalised differential fiducial cross-sections with various $t\bar{t}$ simulation samples. Probabilities smaller than $10^{-10}$ are shown as zero.

| Generator              | $p_T^{\ell}, p_T^{\mu}$ | $p_T^{e\mu}$, $m^{e\mu}$, $p_T^{\ell} + p_T^{\mu}$ | $|\eta^{\ell}|, |\eta^{e\mu}|$ | $|\eta^{\ell}|, |\eta^{e\mu}|, E^e + E^\mu$ | All          |
|------------------------|--------------------------|---------------------------------------------------|--------------------------------|---------------------------------|-------------|
| POWHEG + PY6           | 20.7                     | 38.2                                              | 57.6                          | 70.0                            | 120         |
| POWHEG + PY6 radLo     | 24.6                     | 50.6                                              | 57.6                          | 70.6                            | 138         |
| POWHEG + PY6 radHi     | 16.4                     | 29.7                                              | 52.3                          | 62.8                            | 98.7        |
| POWHEG + PY6 $h_{damp} = \infty$ | 25.0                     | 40.1                                              | 54.2                          | 68.7                            | 113         |
| POWHEG + PY6 $p_T$ NNLO | 15.1                     | 30.0                                              | 60.0                          | 68.2                            | 109         |
| POWHEG + PY8 $h_{damp} = \infty$ | 23.6                     | 37.3                                              | 56.8                          | 71.3                            | 121         |
| POWHEG + PY6 HERAPDF 1.5 | 20.1                     | 29.6                                              | 22.5                          | 24.5                            | 68.6        |
| POWHEG + PY6 no spin corl. | 30.2                     | 284                                               | 58.3                          | 77.4                            | 462         |
| ALPGEN + HW            | 38.9                     | 79.3                                              | 49.3                          | 67.2                            | 154         |
| MC@NLO + HW           | 23.1                     | 35.2                                              | 54.8                          | 65.7                            | 110         |
| AMC@NLO + HW++        | 19.1                     | 45.2                                              | 63.1                          | 70.2                            | 128         |
| POWHEG + PY6           | 0.19                     | 0.058                                             | $10^{-6}$                      | $410^{-6}$                       | $10^{-4}$   |
| POWHEG + PY6 radLo     | 0.077                    | $310^{-3}$                                        | $10^{-6}$                      | $310^{-6}$                       | $210^{-6}$  |
| POWHEG + PY6 radHi     | 0.43                     | 0.28                                              | $10^{-5}$                      | $410^{-5}$                       | 0.011       |
| POWHEG + PY6 $h_{damp} = \infty$ | 0.069                    | 0.038                                             | $510^{-6}$                     | $610^{-6}$                       | $610^{-4}$  |
| POWHEG + PY6 $p_T$ NNLO | 0.51                     | 0.27                                              | $510^{-7}$                     | $710^{-6}$                       | $210^{-3}$  |
| POWHEG + PY8 $h_{damp} = \infty$ | 0.100                    | 0.071                                             | $210^{-6}$                     | $210^{-6}$                       | $110^{-4}$  |
| POWHEG + PY6 HERAPDF 1.5 | 0.21                     | 0.29                                              | 0.13                          | 0.49                            | 0.49        |
| POWHEG + PY6 no spin corl. | 0.017                    | 0                                                  | $110^{-6}$                     | $310^{-7}$                       | 0           |
| ALPGEN + HW            | $110^{-3}$               | $310^{-7}$                                        | $310^{-5}$                     | $110^{-5}$                       | $210^{-8}$  |
| MC@NLO + HW            | 0.11                     | 0.11                                              | $410^{-6}$                     | $210^{-5}$                       | $110^{-3}$  |
| AMC@NLO + HW++        | 0.26                     | 0.011                                             | $210^{-7}$                     | $410^{-6}$                       | $210^{-5}$  |

tons are dressed with nearby photons as discussed in Sect. 4. Therefore, the MCFM predictions were corrected bin-by-bin using corrections derived from two $t\bar{t}$ samples generated with PyThia8 (version 8.205) [108] and the ATTBAR tune [109] with QED final-state radiation enabled and disabled. These corrections are typically 1–2% on the absolute and always smaller than 1% on the normalised differential cross-sections. No corrections were applied to the normalised $|\eta^{\ell}|$ and $|\eta^{e\mu}|$ distributions, as the determined corrections were always smaller than 0.3% and consistent with unity within the simulation statistical uncertainties.

The ratios of the MCFM normalised differential cross-section predictions with HERAPDF 1.5 (the PDF set found to best fit the data when comparing with POWHEG + PyTHIA6 samples in Sect. 6.2) to data are shown in Fig. 13. The uncertainties in the predictions include effects from PDFs, QCD scales and the value of the strong coupling constant $\alpha_S$. For each individual component variation, the prediction was renormalised to unity before calculating the shift for each bin; the effects on the normalised cross-section predictions are typically significantly smaller than those on the absolute cross-sections. The PDF uncertainties for CT14 and MMHT were evaluated from the sum in quadrature of the symmetrised up/down variations from each individual eigenvector pair from the PDF error set. For the HERAPDF sets, each pair of eigenvector or model parameter variations was treated as an independent variation. For NNPDF 3.0, the 100 replica sets which represent the NNPDF uncertainty were used to define a full covariance matrix taking into account correlations between the bins of each distribution. The QCD scale uncertainties were evaluated by varying the renormalisation and factorisation scales $\mu_R$ and $\mu_F$ separately, and adding the variations in quadrature. Each scale was varied by factors of one-half and two from its central value $(m_t/2)$, and the resulting variations symmetrised. This procedure was used instead of taking an envelope including simultaneous variations of $\mu_F$ and $\mu_R$ in order to properly account for the correlations between bins of the normalised differential cross-section predictions. Finally, the $\alpha_S$ uncertainty was evaluated using the HERAPDF 1.5 PDF sets with $\alpha_S$ set to 0.116 and 0.120, rescaling the resulting uncertainty to
Fig. 13 Ratios of MCFM + HERAPDF 1.5 fixed-order predictions of normalised differential cross-sections to data as a function of lepton and dilepton variables. Contributions via $W \rightarrow \tau \rightarrow e/\mu$ decays are not included, and the MCFM predictions have been corrected to include QED final-state radiation effects. The total data uncertainties are shown by the error bars around unity. The separate uncertainties in the predictions from QCD scales, PDFs and the strong coupling constant $\alpha_S$ are shown by the hatched bands, and the total uncertainties in the predictions are shown by the yellow band.
\[ \Delta \alpha_S = \pm 0.0015, \text{ in line with the corresponding PDF4LHC recommendation [110].} \]

The compatibility of the predictions with the normalised cross-section data was tested quantitatively using the \( \chi^2 \) of Eq. (5), updating the covariance matrix \( S \) to also include the theoretical uncertainties discussed above, including their bin-to-bin correlations via the off-diagonal terms. The resulting \( \chi^2 \) and \( p \)-values are shown as the ‘MCFM + HERAPDF 1.5’ entries in Table 10 for individual distributions, and in Table 11 for the combinations of distributions. As can be seen from these tables and from Fig. 13, MCFM with the HERAPDF 1.5 PDF describes the data well, once all the theoretical uncertainties are taken into account. The predictions for \( p_T^e, p_T^{e\mu}, m^{e\mu}, \Delta \phi^{e\mu} \) and \( p_T^e + p_T^{e\mu} \) have large scale uncertainties, which largely cover any differences between the measurements and central predictions with scales \( \mu_R = \mu_F = m_t/2 \). The \( |\eta|^2 \) and \( |\gamma^{e\mu}| \) distributions have little scale dependence and are more sensitive to PDF variations, but are again well-described within the uncertainties of the HERAPDF 1.5 set. The \( \alpha_S \)-related uncertainties are small compared to the other two classes.

The predictions for all five PDF sets (including PDF uncertainties, scaled to 68% CL for CT14, as well as scale and \( \alpha_S \) uncertainties) are compared to the data in Fig. 14. The corresponding \( \chi^2 \) and \( p \)-values, including the PDF, scale and \( \alpha_S \) uncertainties on the predictions, are shown in Tables 10 and 11. The results for HERAPDF 1.5 and HERAPDF 2.0 are close to the data, whereas the CT14, MMHT and NNPDF 3.0 PDFs describe the data slightly less well, particularly for \( p_T^e, |\eta|^2, |\gamma^{e\mu}| \) and \( E^e + E^{e\mu} \). These conclusions are similar to those found for HERAPDF 1.5 and CT10 with the POWHEG + PYTHIA6 setup discussed in Sect. 6.2 above. However, the difference in \( \chi^2 \) between the PDF sets is smaller for the fixed-order predictions, as the explicit inclusion of PDF and scale uncertainties in the predictions renders the differences between the central predictions of each PDF less significant. The PDF comparisons would benefit from the availability of predictions including NNLO QCD effects in both the top quark production and decay, which should substantially reduce the scale uncertainties.

7 Constraints on the gluon parton distribution function

As a demonstration of the ability of the normalised differential cross-section measurements to constrain the gluon PDF, fits were performed to deep inelastic scattering (DIS) data from HERA I–II [97], with and without the addition of the constraints from \( t\bar{t} \) dilepton \( |\eta|^2, |\gamma^{e\mu}| \) and \( E^e + E^{e\mu} \) distributions. As shown in Fig. 13, these distributions are the most sensitive to PDF variations, whilst being less sensitive to QCD scale variations and the value of \( m_t \). The fits are based on the predictions from MCFM and APPLGRID discussed in Sect. 6.3, allowing predictions for arbitrary PDF variations to be obtained much faster than if a full NLO plus parton shower event generator setup were to be used. The QCD scales were set to fixed values of \( \mu_F = \mu_R = m_t/2 \). The fits were performed using the XFITTER package [111,112], which allows the PDF and other theoretical uncertainties to be included via asymmetric error propagation. In this formalism, the \( \chi^2 \) for the compatibility of the predictions with the measurement is expressed by:

\[
\chi^2 = \sum_{i,j} \left( \frac{S_\text{exp,ij}^{-1}(\zeta_i, \zeta_j)}{\zeta_i} \right) \frac{S_\text{exp,ij}(\zeta_i, \zeta_j)}{\zeta_i} \left( \zeta_{ij} - \zeta_{ij}^\text{th} \right),
\]

where \( \zeta_{ij}^\text{th} \) is the measured normalised differential cross-section in bin \( i \) (equivalent to \( \zeta_{ij}^L \) in Eq. (2)), \( \zeta_i^\text{th} \) is the corresponding theoretical prediction, \( S_\text{exp,ij} \) is the covariance matrix of experimental uncertainties including both statistical and systematic contributions, and correlations between bins, and the sums for \( i \) and \( j \) run over \( n - 1 \) bins to account for the normalisation condition. Unlike in the formulation of Eq. (5), the covariance matrix is a function of the theoretical predictions, with the statistical uncertainties being rescaled according to the difference between the measured values and the predictions using a Poisson distribution, and the systematic uncertainties being scaled in proportion to the predictions.

Following the formalism outlined in Ref. [113], the covariance matrix was decomposed into a diagonal matrix \( \mathbf{D} \) representing the uncorrelated parts of the uncertainties, and a set of coefficients \( \gamma_i^\text{exp} \) giving the one standard deviation shift in the measurement \( i \) for source \( j \), where \( j \) runs over the correlated part of the statistical uncertainties and each source of systematic uncertainty. Each source of experimental uncertainty was then associated with a ‘nuisance parameter’ \( b_{j,\text{exp}} \) parameterising the associated shift in units of standard deviation. The \( \chi^2 \) becomes a function of the set of PDF parameters \( p \) defining the theoretical prediction \( \zeta_i^\text{th} \) and the vector of experimental nuisance parameters \( b_{\text{exp}} \), and is given by:

\[
\chi^2(p, b_{\text{exp}}) = \sum_i \left( \frac{S_\text{exp}^i + \sum_j \gamma_{ij}^\text{exp} b_{j,\text{exp}} - \zeta_i^\text{th}(p)}{d_i^2} \right)^2 + \sum_j b_{j,\text{exp}}^2 + L ,
\]

where \( d_i \) are the non-zero elements of the diagonal matrix \( \mathbf{D} \), and the rescaling of the uncertainties leads to the logarithmic term \( L \), arising from the likelihood transition to \( \chi^2 \) as discussed in Refs. [113,114]. The \( \chi^2 \) was minimised as a function of the PDF parameters \( p \) and the nuisance parameters \( b_{\text{exp}} \), and the value at the minimum provides a compatibility test of the data and prediction.
Fig. 14 Ratios of MCFM fixed-order predictions of normalised differential cross-sections to data as a function of lepton and dilepton variables, using the CT14, MMHT, NNPDF 3.0, HERAPDF 1.5 and HERAPDF 2.0 PDF sets for the predictions. Contributions via $W \rightarrow \tau \rightarrow e/\mu$ decays are not included, and the MCFM predictions have been corrected to include QED final-state radiation effects. The total data uncertainties are shown by the cyan bands around unity, and the total uncertainty for each prediction (including QCD scales, PDFs, and the strong coupling constant $\alpha_S$) are shown by the vertical bars.
Table 10 The $\chi^2$ values (top) and associated probabilities (bottom) for comparison of measured normalised differential fiducial cross-sections with the predictions of MCFM with various PDF sets. Contributions via $W \to \tau \to e/\mu$ decays are not included, and the MCFM predictions have been corrected to include QED final-state radiation effects. The results take into account the uncertainties in both the measurements and predictions.

| Generator       | $p_T^\ell$ | $|\eta^\ell|$ | $p_T^{e/\mu}$ | $m^{e/\mu}$ | $|y^{e/\mu}|$ | $\Delta \phi^{e/\mu}$ | $p_T^e + p_T^{\mu}$ | $E^e + E^{\mu}$ |
|-----------------|------------|----------------|---------------|-------------|----------------|---------------------|-------------------|------------------|
| MCFM + CT14     | 11.5       | 14.1           | 7.2           | 11.2        | 13.0          | 7.2                 | 11.4              | 11.2             |
| MCFM + MMHT     | 11.3       | 12.8           | 7.2           | 11.2        | 12.6          | 7.1                 | 11.1              | 9.6              |
| MCFM + NNPDF 3.0| 11.7       | 11.3           | 7.2           | 11.4        | 9.4           | 7.3                 | 11.5              | 8.5              |
| MCFM + HERAPDF 1.5 | 9.1       | 10.9           | 6.4           | 12.1        | 8.0           | 6.9                 | 8.5               | 2.6              |
| MCFM + HERAPDF 2.0 | 8.4       | 12.0           | 6.2           | 12.4        | 8.0           | 6.8                 | 8.0               | 2.7              |
| MCFM + CT14     | 0.24       | 0.080          | 0.51          | 0.43        | 0.11          | 0.62                | 0.12              | 0.27             |
| MCFM + MMHT     | 0.26       | 0.12           | 0.51          | 0.42        | 0.13          | 0.62                | 0.13              | 0.38             |
| MCFM + NNPDF 3.0| 0.23       | 0.18           | 0.52          | 0.41        | 0.31          | 0.61                | 0.12              | 0.49             |
| MCFM + HERAPDF 1.5 | 0.43       | 0.21           | 0.61          | 0.36        | 0.44          | 0.65                | 0.29              | 0.98             |
| MCFM + HERAPDF 2.0 | 0.49       | 0.15           | 0.63          | 0.33        | 0.44          | 0.66                | 0.34              | 0.97             |

Table 11 The $\chi^2$ values (top) and associated probabilities (bottom) for comparison of combinations of measured normalised differential fiducial cross-sections with the predictions of MCFM with various PDF sets. Contributions via $W \to \tau \to e/\mu$ decays are not included, and the MCFM predictions have been corrected to include QED final-state radiation effects. The results take into account the uncertainties in both the measurements and predictions.

| Generator       | $p_T^{\ell}$, $p_T^e + p_T^{\mu}$ | $p_T^{\mu}$, $m^{e/\mu}$, $p_T^e + p_T^{\mu}$ | $|\eta^\ell|$, $|y^{e/\mu}|$ | $|\eta^\ell|$, $|y^{e/\mu}|$, $E^e + E^{\mu}$ | All |
|-----------------|----------------------------------|---------------------------------------------|-----------------------------|---------------------------------------------|-----|
| MCFM + CT14     | 19.5                             | 29.6                                        | 24.2                        | 32.4                                        | 73.0|
| MCFM + MMHT     | 19.3                             | 29.6                                        | 23.4                        | 30.7                                        | 72.0|
| MCFM + NNPDF 3.0| 19.9                             | 29.7                                        | 20.1                        | 27.4                                        | 69.3|
| MCFM + HERAPDF 1.5 | 16.1                             | 28.8                                        | 21.5                        | 26.1                                        | 68.8|
| MCFM + HERAPDF 2.0 | 15.3                             | 30.0                                        | 22.7                        | 27.4                                        | 69.0|
| MCFM + CT14     | 0.24                             | 0.28                                        | 0.086                       | 0.15                                        | 0.35|
| MCFM + MMHT     | 0.25                             | 0.28                                        | 0.10                        | 0.20                                        | 0.38|
| MCFM + NNPDF 3.0| 0.23                             | 0.28                                        | 0.22                        | 0.34                                        | 0.47|
| MCFM + HERAPDF 1.5 | 0.45                             | 0.32                                        | 0.16                        | 0.40                                        | 0.48|
| MCFM + HERAPDF 2.0 | 0.51                             | 0.27                                        | 0.12                        | 0.34                                        | 0.48|

For the PDF fits, the perturbative order of the DGLAP evolution [115–117] was set to NLO, to match the order of the MCFM predictions. The gluon PDF $g(x)$ was parameterised as a function of Bjorken-$x$ as:

$$xg(x) = Ax^B (1 - x)^C (1 + Ex^2) e^{Fx},$$

which, compared to the standard parameterisation given in Eq. (27) of Ref. [97], removes the negative $A'$ term at low $x$ and adds more flexibility at medium and high $x$ through the additional terms with the parameters $E$ and $F$. The standard parameterisations were used for the quark PDFs, giving a total of 14 free PDF parameters in the vector $p$, after imposing momentum and valance sum rules, and the constraint that the $\bar{u}$ and $\bar{d}$ contributions are equal at low $x$. Other parameters in the PDF fit were set as described in Ref. [113].

The minimised $\chi^2$ values from the fits without and with the $t\bar{t}$ data are shown in Table 12, which gives the partial $\chi^2$ for each dataset included in the fit (i.e., the contribution of that dataset to the total $\chi^2$) and the total $\chi^2$ for each fit. The partial $\chi^2$ values indicate that the $t\bar{t}$ data are well-described by the PDF derived from the combined fit, and that the description of the HERA I+II data is not degraded by the inclusion of the $t\bar{t}$ data, i.e., there is no tension between the two datasets. The ratios of the fitted gluon PDF central values with and without the $t\bar{t}$ data included are shown in Fig. 15a, together with the corresponding uncertainties. The ratio of relative uncertainties in the PDFs with and without the $t\bar{t}$ data are shown in Fig. 15b. The inclusion of the $t\bar{t}$ data reduces the uncertainty by typically 10–25% over most of the relevant $x$ range.
The gluon PDF obtained from this procedure is compared to the gluon PDFs from the CT14 [103] and NNPDF 3.0 [105] global PDF sets in Fig. 16. These PDF sets, shown by the green bands, both have a larger high-\(x\) gluon than preferred by the HERA I+II data, with or without the addition of the \(t\bar{t}\) data from this analysis. The impact of the \(t\bar{t}\) data on the global PDF sets was investigated using a profiling procedure [113,118,119], extending the \(\chi^2\) definition of Eq. (7) to incorporate a vector \(b_{th}\) of nuisance parameters \(b_{k,th}\) expressing the dependence of the theoretical prediction \(\bar{s}_i\) on the uncertainties for a particular PDF set. In this formulation, the \(\chi^2\) definition becomes:

\[
\chi^2(b_{exp}, b_{th}) = \sum_i \left( \frac{\bar{s}_i^{exp} + \sum_j \gamma_i^{exp} b_{j,exp} - \bar{s}_i^{th} - \sum_k \gamma_i^{th} b_{k,th}}{d_i^2} \right)^2 + \sum_j b_{j,exp}^2 + \sum_k b_{k,th}^2 + L,
\]

where \(b_{k,th} = \pm 1\) corresponds to the \(\pm 1\) standard deviation change of the PDF values according to the \(k\)th eigenvector of the PDF error set. The values and uncertainties of the nuisance parameters \(b_{k,th}\) after minimisation of the \(\chi^2\) of Eq. (9) give the profiled PDF with modified central values and uncertainties according to the effect of the \(t\bar{t}\) differential cross-section distributions. These profiled PDFs are shown as the orange bands in Fig. 16. Both the CT14 and NNPDF 3.0 gluon PDFs are shifted downwards at high \(x\) (corresponding to a softer gluon distribution). The effect is larger in the case of CT14, which has larger uncertainties in the gluon PDF in this region.

### 8 Extraction of the top quark mass

The normalised lepton \(p_T^\ell\) and dilepton \(p_T^{\ell\ell}, m^{\ell\ell}, p_T^e + p_T^\mu\) and \(E^e + E^\mu\) differential distributions are sensitive to the value of the top quark mass, as already shown in Fig. 5a for \(p_T^\ell\) and Fig. 5b for \(p_T^{\ell\ell}\). Provided that other theoretical uncertainties in the predictions (as discussed in Sect. 6) can be kept under control, fitting these distributions offers a complementary way to measure \(m_t\) compared to more traditional determinations from complete reconstruction of the top quark decay products [33–36]. Ref. [32] explores such an approach in detail, arguing that measurements from normalised lepton distributions are less sensitive to the modelling of perturbative and non-perturbative QCD, and are closer to the ideal of a measurement of the top quark pole mass \(m_t^{pole}\) than those employing a direct measurement of the top quark decay products. It also stresses the importance of using several different

![Fig. 15](image-url) **a** Ratio of the gluon PDF determined from the fit using HERA I+II data plus the normalised differential cross-section distributions as a function of \(|y^\ell|, |y^{\ell\ell}|\) and \(E^e + E^\mu\) in \(t\bar{t}\) events, to the gluon PDF determined from the fit using HERA I+II data alone, as a function of Bjorken-\(x\). The uncertainty bands are shown on the two PDFs as the blue and red shading. **b** Ratio of the relative uncertainty in the gluon PDF determined from the fit to HERA I+II plus \(t\bar{t}\) data to that from HERA data alone. The PDFs are shown evolved to the scale \(Q^2 = m_t^2\) in both cases.
leptonic observables to probe for inadequacies in the theoretical descriptions of the distributions which may introduce biases in the extracted $m_t$ values. Experimentally, the double-tagging technique employed here results in measurements with little uncertainty from the hadronic components of the $t\bar{t}$ system, again reducing the exposure to QCD modelling compared to the measurements based on reconstructing the top quark decay products.

Several sets of top mass determinations are reported here, based either on predictions from the NLO matrix element event generator POWHEG interfaced to PYTHIA6 and the CT10 PDFs as described in Sect. 2, or on fixed-order predictions with NLO descriptions of the $t\bar{t}$ production and top quark decay from the MCFM program with various PDF sets, as described in Sect. 6.3. In the first case, $m_t$ is extracted either by using a template fit parameterising the predictions as a function of $m_t$ and finding the value which minimises the $\chi^2$ with respect to the measured data (described in Sect. 8.1), or by calculating moments of the distributions in data and comparing them to the corresponding moments of the predicted distributions for different values of $m_t$ (Sect. 8.2). In the template fit method, the comparisons between data and predictions are performed at particle level, in contrast to the template fits used for the ATLAS $m_t$ measurements based on reconstruction of the top quark decay products [120], where the comparisons are performed at detector level using the reconstructed distributions and fully-simulated Monte Carlo events. The template fit method uses the complete information from the measured distribution, taking into account the uncertainty in each bin, whereas the moments method, advocated in Ref. [32], allows different features of the distribution shapes to be emphasised via the comparisons of moments of different order. The results from these two methods are discussed and compared in Sect. 8.3.

In the mass determination from QCD fixed-order calculations, described in Sect. 8.4, $\chi^2$ values are calculated for the comparison of data with predictions at different $m_t$ values using the formalism of Eq. (9), and the best-fit $m_t$ is found by polynomial interpolation. This approach is similar to the template fit discussed above; the use of moments was not pursued as it does not exploit the full information of each distribution and does not allow the reduction of uncertainties via constrained nuisance parameters. The $m_t$ value used in the fixed-order predictions corresponds to a well-defined renormalisation scheme, which is the pole mass ($m_t^{\text{pole}}$) scheme within the MCFM implementation. Both the QCD scale uncertainties, representing the effects of missing higher-order corrections beyond NLO, and the PDF uncertainties, are included in the $\chi^2$ formalism in a natural way. This formalism also allows $m_t$ to be determined using several distributions simultaneously, giving the most precise results from any of the techniques explored here. The results from this method are discussed in Sect. 8.5 and are used to define the final measurement of the top quark mass from the distributions measured in this paper.

8.1 Mass extraction using template fits

In the template fit method, the best fit top quark mass for each measured distribution was obtained by minimising the $\chi^2$ for the comparison of that distribution with predictions at different values of $m_t$, defined analogously with Eq. (5):

$$\chi^2(m_t) = \Delta^T_{(n-1)}(m_t) S_{(n-1)}^{-1} \Delta_{(n-1)}(m_t), \quad (10)$$

$$\quad \text{with} \quad \Delta^T(m_t) = [\delta_{\xi}^{(n)}(m_t) \cdots \delta_{\xi}^{(1)}(m_t)]$$

$$\quad \text{and} \quad S_{(n-1)} = \sum_{i=1}^{n-1} S_{i}^{T} S_{i}^{-1}$$

$$\quad \text{where} \quad S_{i}^{T} = \sum_{j=1}^{n-1} \sum_{k=1}^{n-1} \delta_{\xi}^{(i)}(j) \delta_{\xi}^{(k)}(j) S_{i}^{-1}.$$
The template fit method was tested with pseudo-experiments based on fully-simulated \( t\bar{t} \) samples with \( m_t \) values in the range 165–180 GeV plus non-\( t\bar{t} \) backgrounds. The pseudo-data were processed through the complete analysis procedure starting from the observed event counts in each bin, using the methodology described in Sect. 4.3. The baseline POWHEG + PYTHIA6 \( t\bar{t} \) sample with \( m_t = 172.5 \) GeV was used as reference for the calculation of \( G_{i\mu}^t, C_\mu^t, N_{1i}^t,bkg \) and \( N_2^t,bkg \). No statistically significant biases were found for the fits based on the \( p_T^t, p_T^{\mu} \) and \( m^{\tau\mu} \) distributions, but biases of up to 0.6 GeV for \( p_T^{\tau\mu} \) and 0.9 GeV for \( E^\tau + E^{\mu} \) were found in pseudo-experiments with true \( m_t \) values 5 GeV away from the 172.5 GeV reference, still small compared to the expected statistical uncertainties using these distributions. These biases were corrected in the fit results from data discussed in Sect. 8.3 below. The pseudo-experiments were also used to check the statistical uncertainties returned by the fit via the pull distributions, which were generally found to be within 5\% of unity.

Both the data statistical uncertainty and experimental systematic uncertainties on the measurements of the differential distributions are included in the matrix \( S_{(n-1)} \) in Eq. (10). Further uncertainties in the extracted \( m_t \) value arise from the choices of PDFs and event generator setup for the predictions. The PDF uncertainties were assessed from the variations in normalised \( t\bar{t} \) differential cross-section distributions predicted by MC@NLO + HERWIG reweighted using the error sets of the CT10, MSTW and NNPDF 2.3 PDF sets as described in Sect. 5.1. The event generator setup uncertainties were assessed as the quadrature sum of a \( t\bar{t} \) generator uncertainty and a QCD radiation uncertainty. The former was obtained from the comparison of results using POWHEG + PYTHIA6 (\( h_{damp} = m_t \)) and MC@NLO + HERWIG samples (thus varying both the matrix element and parton shower generator). The latter was defined as half the variation from fits using the POWHEG + PYTHIA6 samples with radLo and radHi tunes discussed in Sect. 2. In all cases, the uncertainties were defined from the difference in \( m_t \) values obtained when fitting the two samples as pseudo-data, using the full experimental covariance matrix from the data measurement and the standard templates obtained from the POWHEG + PYTHIA6 samples as discussed above.

### 8.2 Mass extraction using moments

Top quark mass information can also be derived from a measured distribution by calculating Mellin moments of the distribution, and comparing the values observed to a calibration curve obtained from predictions with different values of \( m_t \) [32]. The \( k \)th order Mellin moment \( \mu^{(k)} \) for a distribution \( D(x) \equiv dx/dx \) as a function of a kinematic variable \( x \) is defined as:

\[
\mu^{(k)} = \frac{1}{\sigma_{\text{fid}}} \int x^k D(x) \, dx , \tag{11}
\]

where the integral is taken over the fiducial region, and the total fiducial cross-section \( \sigma_{\text{fid}} = \int D(x) \, dx \). These moments can in principle be evaluated without binning the data, since for leptonic observables, the value \( x \) for each individual event is measured with high precision. However, for the purpose of this analysis, these moments were approximated by binned moments \( \Theta^{(k)} \) evaluated as:

\[
\Theta^{(k)} = \sum_i \frac{1}{S^i_{\text{fid}}} X_i \, , \quad X_i = \langle x^k \rangle \text{ in bin } i , \tag{12}
\]

where \( S^i_{\text{fid}} \) is the fiducial weight for bin \( i \).
where $\mathcal{C}_i^i$ is the fraction of the total fiducial $t\bar{t}$ cross-section in bin $i$ (Eq. (2)) and $X_i$ is the mean value of $x$ for all the events falling in bin $i$. The values of $X_i$, which act as weights for each bin $i$ of each kinematic distribution when calculating the moment $k$, were evaluated using the baseline POWHEG + PYTHIA6 sample and kept constant when evaluating moments for the data and all simulation samples. Calibration curves for the first, second and third moments $\Theta^{(1)}$, $\Theta^{(2)}$ and $\Theta^{(3)}$ were derived using the same set of POWHEG + PYTHIA6 samples with top quark masses in the range 165–180 GeV as used for the template analysis. The dependencies of $\Theta^{(k)}$ on $m_t$ were found to be well-described by second-order polynomials $\Theta^{(k)}(m_t) = P_k(m_t)$. A constant offset in each moment was used to correct to the calibration appropriate for $h_{\text{damp}} = m_t$ samples, and the polynomial inverted to obtain the $m_t$ value corresponding to a given measured $k$th moment $\Theta^{(k)}$.

The extraction procedure was tested for bias with pseudo-experiments in the same way as for the template fit. The observed biases were of similar size to those in the template fit, and were corrected in the same way. Experimental systematic uncertainties were evaluated by calculating the moments from the normalised cross-section distribution with each bin shifted by one standard deviation of each systematic, and translating the resulting shift in $\Theta^{(k)}$ to a shift in $m_t$. Uncertainties in the predictions due to the choice of PDFs, $t\bar{t}$ generator and radiation settings were assessed in the same way, i.e. from the shifts in $\Theta^{(k)}$ predicted by each of the alternative samples.

8.3 Results from the template and moment methods

The results of applying the template and first, second and third moment methods to each of the $p_T^T$, $p_T^{e\mu}$, $p_T^{\ell}$, $E_T^e + E_T^\mu$ and $E_T^e + E_T^\mu$ distributions using predictions from POWHEG + PYTHIA6 and CT10 PDFs are shown in Table 14 and Fig. 17. The table shows the $\chi^2$ at the best fit mass for each distribution, and the breakdown of uncertainties into statistical, experimental systematic and theoretical contributions, evaluated as discussed in Sect. 8.1. For the template fits, the data statistical uncertainty was evaluated from a $\chi^2$ minimisation of Eq. (10) with only statistical uncertainties included in the covariance matrix $S$. The experimental systematic uncertainty was evaluated as the quadrature difference between the total uncertainty (when including both statistical and experimental systematic uncertainties in $S$), and the data statistical uncertainty. For the moments method, the statistical and experimental systematic uncertainties were evaluated directly on the moments $\Theta^{(k)}$ as discussed in Sect. 8.2.

The ratios of predictions to data at the best-fit top quark mass found by the application of the template fit method to each distribution are shown in Fig. 18. The data are generally well-described by these predictions, as can also be seen from the $\chi^2$ values in Table 14, except for the $E_T^e + E_T^\mu$ distribution. This distribution is quite sensitive to PDFs as well as $m_t$, and is better described by the HERAPDF PDFs than the CT10 PDFs used here to extract $m_t$, resulting in a low fitted value with a large PDF uncertainty, and a large variation between the template and moment fit results. Total uncertainties in $m_t$ of about 2 GeV are obtained from the template fits to the $p_T^T$, $p_T^{e\mu}$ and $p_T^\ell + p_T^{e\mu}$ distributions. These results have relatively small theoretical uncertainties, and the experimental uncertainties are dominated by $t\bar{t} - Wt$ interference and the electron energy scale. The $m_T^{e\mu}$ distribution is intrinsically less sensitive to $m_t$, having larger statistical, experimental and theoretical systematic uncertainties. The results from the extraction based on moments have larger uncertainties than those from the template fit, reflecting that the moments do not take into account the relative precision on the different bins of the distributions, and that the higher moments are more sensitive to the tails of the distributions, which are less precisely measured and subject to larger theoretical uncertainties. Within each distribution, the $m_t$ values from the different moments are close, though 3–4 GeV lower than the template fit results for $m_T^{e\mu}$, and up to 7 GeV lower in the case of $E_T^e + E_T^\mu$.

The central values of the template fit results from the five distributions exhibit a spread (envelope) of about 6 GeV. The results from the fits of $p_T^T$ and $p_T^\ell + p_T^{e\mu}$ lie 4–5 GeV below that from $p_T^{e\mu}$, which is close to the world-average mass value from reconstruction of top quark decay products of $173.34\pm0.76$ GeV [121]. The consistency of the fit results was assessed by combining them using the best linear unbiased estimate (BLUE) technique [122]. Correlations in the statistical uncertainties were assessed using pseudo-experiments as described in Sect. 4.3. Correlations between systematic uncertainties were determined by assuming the effects on $m_t$ from each individual experimental or theoretical component to be fully correlated between distributions. PDF uncertainties were assessed using the eigenvector pairs of the CT10 PDF only. The combination of all five distributions has a $\chi^2$ probability of 4%, indicating that the systematic uncertainties may be underestimated.

The POWHEG + PYTHIA6 $t\bar{t}$ samples used here do not provide a good modelling of the top quark $p_T$ spectrum [18–20, 22, 23], potentially biasing the results. The size of this possible bias was explored by fitting the distributions from the POWHEG + PYTHIA6 baseline sample reweighted to the top quark $p_T$ spectrum calculated at NNLO precision in Ref. [25]. The reweighted sample gives a better description of the $p_T^T$ and $p_T^\ell + p_T^{e\mu}$ distributions, as can be seen from the $\chi^2$ values for ‘POWHEG + PY6 $p_T$ NNLO’ in Table 8. The mass shifts between the baseline and reweighted samples,
Table 14  Measurements of the top quark mass from individual template fits to the lepton $p_T^\ell$ and dilepton $p_T^{\ell\ell}$, $m_\ell^{\ell\ell}$, $p_T^{\ell} + p_T^{\ell\ell}$ and $E_T + E_T^\mu$ distributions, and using the first, second and third moments of these distributions. The data are compared to predictions from POWHEG + PYTHIA6 with the CT10 PDF set. The $\chi^2$ value at the best-fit mass for each distribution (for the template fits only), the fitted mass with its total uncertainty, and the individual uncertainty contributions from data statistics, experimental systematics, and uncertainties in the predictions due to the choice of $t\bar{t}$ event generator and the modelling of QCD radiation are shown.

<table>
<thead>
<tr>
<th>Template</th>
<th>$p_T^\ell$</th>
<th>$p_T^{\ell\ell}$</th>
<th>$m_\ell^{\ell\ell}$</th>
<th>$p_T^\ell + p_T^{\ell\ell}$</th>
<th>$E_T + E_T^\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2 / N_{\text{ dof}}$</td>
<td>8.1/8</td>
<td>7.5/7</td>
<td>13.9/10</td>
<td>8.0/6</td>
<td>12.5/8</td>
</tr>
<tr>
<td>$m_t$ (GeV)</td>
<td>168.4 ± 2.3</td>
<td>173.0 ± 2.1</td>
<td>170.6 ± 4.2</td>
<td>169.4 ± 2.0</td>
<td>166.9 ± 4.0</td>
</tr>
<tr>
<td>Data statistics</td>
<td>± 1.0</td>
<td>± 0.9</td>
<td>± 2.0</td>
<td>± 0.9</td>
<td>± 1.3</td>
</tr>
<tr>
<td>Expt. systematic</td>
<td>± 1.6</td>
<td>± 1.0</td>
<td>± 3.1</td>
<td>± 1.6</td>
<td>± 1.5</td>
</tr>
<tr>
<td>PDF uncertainty</td>
<td>± 1.0</td>
<td>± 0.2</td>
<td>± 1.6</td>
<td>± 0.6</td>
<td>± 3.4</td>
</tr>
<tr>
<td>$t\bar{t}$ generator</td>
<td>± 0.4</td>
<td>± 1.4</td>
<td>± 1.4</td>
<td>± 0.4</td>
<td>± 1.1</td>
</tr>
<tr>
<td>QCD radiation</td>
<td>± 0.7</td>
<td>± 0.8</td>
<td>± 0.5</td>
<td>± 0.2</td>
<td>± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Moment 1</th>
<th>$p_T^\ell$</th>
<th>$p_T^{\ell\ell}$</th>
<th>$m_\ell^{\ell\ell}$</th>
<th>$p_T^\ell + p_T^{\ell\ell}$</th>
<th>$E_T + E_T^\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T^\ell$</td>
<td>168.2 ± 2.9</td>
<td>174.2 ± 3.8</td>
<td>166.6 ± 6.5</td>
<td>168.4 ± 2.9</td>
<td>160.8 ± 7.9</td>
</tr>
<tr>
<td>Data statistics</td>
<td>± 1.0</td>
<td>± 1.0</td>
<td>± 2.4</td>
<td>± 1.1</td>
<td>± 2.2</td>
</tr>
<tr>
<td>Expt. systematic</td>
<td>± 2.1</td>
<td>± 1.6</td>
<td>± 3.8</td>
<td>± 2.1</td>
<td>± 3.1</td>
</tr>
<tr>
<td>PDF uncertainty</td>
<td>± 1.2</td>
<td>± 0.3</td>
<td>± 2.9</td>
<td>± 1.1</td>
<td>± 6.7</td>
</tr>
<tr>
<td>$t\bar{t}$ generator</td>
<td>± 0.2</td>
<td>± 1.3</td>
<td>± 3.4</td>
<td>± 0.2</td>
<td>± 2.0</td>
</tr>
<tr>
<td>QCD radiation</td>
<td>± 1.2</td>
<td>± 3.0</td>
<td>± 1.4</td>
<td>± 1.1</td>
<td>± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Moment 2</th>
<th>$p_T^\ell$</th>
<th>$p_T^{\ell\ell}$</th>
<th>$m_\ell^{\ell\ell}$</th>
<th>$p_T^\ell + p_T^{\ell\ell}$</th>
<th>$E_T + E_T^\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T^\ell$</td>
<td>168.1 ± 3.2</td>
<td>172.2 ± 4.5</td>
<td>166.9 ± 6.9</td>
<td>167.9 ± 3.3</td>
<td>159.9 ± 9.2</td>
</tr>
<tr>
<td>Data statistics</td>
<td>± 1.2</td>
<td>± 1.1</td>
<td>± 2.8</td>
<td>± 1.3</td>
<td>± 2.6</td>
</tr>
<tr>
<td>Expt. systematic</td>
<td>± 2.3</td>
<td>± 2.0</td>
<td>± 4.3</td>
<td>± 2.4</td>
<td>± 3.4</td>
</tr>
<tr>
<td>PDF uncertainty</td>
<td>± 1.3</td>
<td>± 0.4</td>
<td>± 3.3</td>
<td>± 1.3</td>
<td>± 7.8</td>
</tr>
<tr>
<td>$t\bar{t}$ generator</td>
<td>± 0.4</td>
<td>± 1.2</td>
<td>± 3.2</td>
<td>± 0.4</td>
<td>± 2.4</td>
</tr>
<tr>
<td>QCD radiation</td>
<td>± 1.2</td>
<td>± 3.7</td>
<td>± 0.7</td>
<td>± 1.3</td>
<td>± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Moment 3</th>
<th>$p_T^\ell$</th>
<th>$p_T^{\ell\ell}$</th>
<th>$m_\ell^{\ell\ell}$</th>
<th>$p_T^\ell + p_T^{\ell\ell}$</th>
<th>$E_T + E_T^\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T^\ell$</td>
<td>168.3 ± 3.5</td>
<td>172.0 ± 5.6</td>
<td>166.4 ± 9.1</td>
<td>167.6 ± 3.8</td>
<td>160.9 ± 9.5</td>
</tr>
<tr>
<td>Data statistics</td>
<td>± 1.5</td>
<td>± 1.4</td>
<td>± 4.2</td>
<td>± 1.6</td>
<td>± 3.0</td>
</tr>
<tr>
<td>Expt. systematic</td>
<td>± 2.5</td>
<td>± 2.6</td>
<td>± 6.0</td>
<td>± 2.7</td>
<td>± 3.7</td>
</tr>
<tr>
<td>PDF uncertainty</td>
<td>± 1.5</td>
<td>± 0.6</td>
<td>± 4.1</td>
<td>± 1.4</td>
<td>± 7.8</td>
</tr>
<tr>
<td>$t\bar{t}$ generator</td>
<td>± 0.6</td>
<td>± 1.1</td>
<td>± 3.5</td>
<td>± 0.7</td>
<td>± 2.4</td>
</tr>
<tr>
<td>QCD radiation</td>
<td>± 1.1</td>
<td>± 4.6</td>
<td>± 0.2</td>
<td>± 1.4</td>
<td>± 0.2</td>
</tr>
</tbody>
</table>

Table 14  Measurements of the top quark mass using templates derived from POWHEG + PYTHIA6 with the CT10 PDF set. The results from fitting templates of the single lepton $p_T^\ell$ and dilepton $p_T^{\ell\ell}$, $m_\ell^{\ell\ell}$, $p_T^\ell + p_T^{\ell\ell}$ and $E_T + E_T^\mu$ distributions, and using the first, second and third moments of these distributions, are shown. For comparison, the world-average of mass measurements from reconstruction of the top quark decay products and its uncertainty [121] is shown by the cyan band.
Fig. 18 Ratios of predictions of normalised differential cross-sections to data as a function of (a) $p_T\ell$, (b) $p_T^{\mu\ell}$, (c) $m_{\mu\ell}$, (d) $p_T^{\mu\ell} + p_T^{\ell\mu}$ and (e) $E^\ell + E^{\mu}$, with the prediction taken from Powheg + Pythia6 with the CT10 PDF at the best-fit top quark mass $m_t$, for each distribution. The data statistical uncertainties are shown by the black error bars around a ratio of unity, and the total experimental uncertainties by the cyan band representing the amount that the top quark mass measured in data would be shifted upwards if the templates were based on reweighted samples, are shown in Table 13. These shifts are larger (1.3–1.8 GeV) for $p_T^{\mu\ell} + p_T^{\ell\mu}$ and $p_T^{\ell\mu}$ than for $p_T^{\mu\ell}$ (0.3 GeV), and would bring the results shown in Fig. 17 into closer agreement with each other. However, given that this reweighting is relatively crude, and does not take into account the potential NNLO effects on other distributions important for modelling the lepton and dilepton kinematics (e.g. the invariant mass and rapidity of the $t\bar{t}$ system), the shifts are taken to be purely indicative, and no attempt has been made to correct the quoted central values for these effects. The predictions for the $p_T^{\ell\mu}$ and $p_T^{\mu\ell} + p_T^{\ell\mu}$ distributions are also sensitive to the choice of PDF. The PDF uncertainties shown for $p_T^{\ell\mu}$ and $p_T^{\mu\ell} + p_T^{\ell\mu}$ in Table 14 are significantly larger than those for $p_T^{\ell\mu}$, and as shown in Sect. 6.2, the
QCD corrections are parameterised by $\mu$ simultaneously to extract distributions, as discussed below, where several distributions are fitted simultaneously corresponding to the precision of the experimental measurements. These limitations are addressed by the approach discussed below, where several distributions are fitted simultaneously to extract $m_t$ whilst constraining the uncertainties in the theoretical predictions.

8.4 Mass extraction using fixed-order predictions

The NLO fixed-order predictions for each distribution were generated using MCFM as discussed in Sect. 6.3, for top quark masses in the range 161–180 GeV in steps of 0.5 GeV, with various PDF choices. The $\chi^2$ for the consistency of each prediction with the data was calculated using Eq. (9), incorporating both PDF and QCD scale uncertainties into the theoretical uncertainties represented by the nuisance parameters $b_{th}$. The central scales were again chosen to be $\mu_F = \mu_R = m_t/2$, with the values varying with $m_t$ in the mass scan, and independent variations of $\mu_F$ and $\mu_R$ by factors of two and one-half defining the one standard deviation up and down scale variations. The $\chi^2$ was evaluated at each mass point, and interpolated using a fourth-order polynomial. The asymmetric uncertainty in the fitted value of $m_t$ was defined as the points at which the $\chi^2$ increases by one unit either side of the minimum point. This uncertainty naturally includes both experimental statistical and systematic uncertainties in the measurements, and theoretical uncertainties due to PDFs and QCD scale choices.

In this method, the top quark mass can be extracted from each measured distribution individually, or from the combination of several distributions, where the sum $i$ in Eq. (9) runs over the bins of all considered distributions, and the experimental covariance matrix includes both statistical and systematic correlations between bins of the same and different distributions, evaluated as discussed in Sect. 6.2. When fitting several distributions simultaneously, the system is over-constrained, profiling the various sources of theoretical uncertainty. For example, when including all eight measured distributions, the $|\eta^T|$ and $|\eta^{\mu \ell}|$ distributions have little sensitivity to $m_t$, but constrain the PDF parameters. The $\Delta \phi^{e\mu}$ distribution constrains the QCD scale parameters $\mu_F$ and $\mu_R$, under the assumption that uncertainties in higher-order QCD corrections are parameterised by $\mu_F$ and $\mu_R$ in a way that can be transported from one distribution to another. Two alternative dynamical scale choices were also tested in order to probe this assumption, as discussed in Sect. 8.5 below.

Potential biases in the method were checked by using predictions with $m_t = 172.5$ GeV as pseudo-data, and considering both experimental and theoretical uncertainties in the $\chi^2$ definition. The resulting fitted values of $m_t$ were within $0.1$ GeV of the input value for all five fitted individual distributions ($p_T^T, p_T^{\mu \ell}, m^{e\mu}, p_T^{e\mu}$ and $E^e + E^{\mu}$), and $0.01$ GeV from the input value for a combined fit of all eight distributions, also including $|\eta^T|, |\eta^{e\mu}|$ and $\Delta \phi^{e\mu}$. The widths of the pull distributions were found to be compatible with unity, confirming the validity of the uncertainty estimates from the fits.

8.5 Mass results from fixed-order predictions

The results of the fits to NLO QCD fixed-order predictions with MCFM and the CT14 PDF set are shown for the individual distributions in Table 15, and the results using the CT14, MMHT, NNPDF 3.0, HERAPDF 2.0, ABM 11 [123] and NNPDF 3.0_nojet [105] PDF sets are shown in Fig. 19. As shown in Sect. 7, the constraint on the gluon PDF from the leptonic $t\bar{t}$ measurements is consistent with the PDF determination from DIS data. The use of the NNPDF 3.0_nojet PDF set, which does not include Tevatron and LHC jet production data, allows the effects on $m_t$ of any possible tension between DIS and jet data in the determination of the gluon PDF to be tested. The results from combined fits to all eight distributions, using predictions from all six PDF sets, are shown in Table 16 and Fig. 19. In Tables 15 and 16, the decomposition of the total uncertainty from each mass fit into statistical, experimental and theoretical (PDF and QCD scales) uncertainties was obtained in analogy to the numerical procedure outlined in Ref. [124]. For each individual source of statistical or systematic uncertainty (corresponding to a nuisance parameter $b_{j,exp}$ or $b_{k,th}$ in Eq. (9)), the data were shifted by plus or minus one standard deviation, and a new $m_t$ value obtained by re-minimising the $\chi^2$ function. The resulting shifts in $m_t$ were added in quadrature to obtain the decomposition into the various categories. The quadrature sum of the decomposed uncertainties agrees with the total to within $10\%$ in all cases, the residual differences being due to non-linearity between the uncertainty sources and the extracted values of $m_t$.

The MCFM fixed-order results for individual distributions shown in Table 15 and Fig. 19 show some similar patterns to those from the POWHEG + PYTHIA6-based template fits shown in Table 14 and Fig. 17. The results from $p_T^T$ and $p_T^{e\mu}$ are close, the largest $m_t$ values come from $p_T^{e\mu}$, the smallest from $E^e + E^{\mu}$ and the least precise determination is obtained from $m^{e\mu}$. The envelope of the central values
Table 15: Measurements of the top quark mass from individual fits to the lepton $p_T^T$ and dilepton $p_T^{\mu\mu}$, $m_T^{\mu\mu}$, $p_T^T + p_T^{\mu\mu}$ and $E^\tau + E^{\mu}$ distributions, using fixed-order predictions from MCFM with the CT14 PDF set. The $\chi^2$ value at the best-fit mass for each distribution, the fitted mass with its total uncertainty, and the individual uncertainty contributions from data statistics, experimental systematics, and uncertainties in the predictions from PDF and QCD scale effects are shown.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\chi^2$</th>
<th>$N_{\text{dof}}$</th>
<th>$m_T^{\mu\mu}$ (GeV)</th>
<th>$m_T^{\mu\mu}$</th>
<th>$p_T^T$</th>
<th>$p_T^{\mu\mu}$</th>
<th>$E^\tau + E^{\mu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data statistics</td>
<td>$\pm 2.0$</td>
<td>$\pm 1.4$</td>
<td>$+3.8$</td>
<td>$-4.0$</td>
<td>$\pm 1.4$</td>
<td>$+2.0$</td>
<td></td>
</tr>
<tr>
<td>Expt. systematic</td>
<td>$+2.5$</td>
<td>$-2.3$</td>
<td>$+2.9$</td>
<td>$-3.3$</td>
<td>$+1.5$</td>
<td>$-1.6$</td>
<td></td>
</tr>
<tr>
<td>PDF uncertainty</td>
<td>$\pm 0.5$</td>
<td>$\pm 0.1$</td>
<td>$+1.1$</td>
<td>$\pm 0.5$</td>
<td>$+1.4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCD scales</td>
<td>$\pm 1.1$</td>
<td>$+0.7$</td>
<td>$-0.8$</td>
<td>$+2.6$</td>
<td>$+0.4$</td>
<td>$-0.5$</td>
<td></td>
</tr>
</tbody>
</table>

Table 16: Measurements of the top quark mass from combined fits to all eight lepton and dilepton distributions, using fixed-order predictions from MCFM with the CT14, MMHT, NNPDF 3.0, HERAPDF 2.0, ABM 11 and NNPDF 3.0 nojet PDF sets, and various choices for the central QCD factorisation and renormalisation scales $\mu_F$ and $\mu_R$. The upper section of the table gives the results for $\mu_F = \mu_R = m_T^{\mu\mu}/2$, showing the $\chi^2$ values at the best-fit mass for each PDF set, the fitted mass with its total uncertainty, and the breakdown of individual uncertainty contributions from data statistics, experimental systematics, and uncertainties in the predictions from PDF and QCD scale effects. Uncertainties given as ‘0.0’ are smaller than 0.05 GeV. The lower parts of the table give the $\chi^2$ values, fitted mass and total uncertainty for alternative scale choices of $\mu_F = \mu_R = H_T/4$ and $E^\tau/2$.

<table>
<thead>
<tr>
<th>CT14</th>
<th>MMHT</th>
<th>NNPDF 3.0</th>
<th>HERAPDF 2.0</th>
<th>ABM 11</th>
<th>NNPDF nojet</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_F = \mu_R = m_T^{\mu\mu}/2$</td>
<td>$\chi^2/N_{\text{dof}}$</td>
<td>$m_T^{\mu\mu}$ (GeV)</td>
<td>$m_T^{\mu\mu}$</td>
<td>$m_T^{\mu\mu}$</td>
<td>$m_T^{\mu\mu}$</td>
</tr>
<tr>
<td></td>
<td>71/68</td>
<td>67/68</td>
<td>67/68</td>
<td>67/68</td>
<td>67/68</td>
</tr>
<tr>
<td></td>
<td>70/68</td>
<td>173.5 ± 1.2</td>
<td>173.4 ± 1.2</td>
<td>173.2 ± 1.2</td>
<td>172.9 ± 1.2</td>
</tr>
<tr>
<td>Data statistics</td>
<td>$\pm 0.9$</td>
<td>$\pm 0.9$</td>
<td>$\pm 0.9$</td>
<td>$\pm 0.9$</td>
<td>$\pm 0.9$</td>
</tr>
<tr>
<td>Expt. systematic</td>
<td>$+0.7$</td>
<td>$+0.8$</td>
<td>$+0.8$</td>
<td>$+0.8$</td>
<td>$+0.8$</td>
</tr>
<tr>
<td>PDF uncertainty</td>
<td>$\pm 0.1$</td>
<td>$\pm 0.1$</td>
<td>$\pm 0.1$</td>
<td>$\pm 0.1$</td>
<td>$\pm 0.1$</td>
</tr>
<tr>
<td>QCD scales</td>
<td>$\pm 0.1$</td>
<td>$\pm 0.1$</td>
<td>$\pm 0.1$</td>
<td>$\pm 0.1$</td>
<td>$\pm 0.1$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\mu_F = \mu_R = H_T/4$</th>
<th>$\chi^2/N_{\text{dof}}$</th>
<th>$m_T^{\mu\mu}$ (GeV)</th>
<th>$m_T^{\mu\mu}$</th>
<th>$m_T^{\mu\mu}$</th>
<th>$m_T^{\mu\mu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>69/68</td>
<td>67/68</td>
<td>64/64</td>
<td>61/68</td>
<td>66/68</td>
<td>60/68</td>
</tr>
<tr>
<td>$\mu_F = \mu_R = E^\tau/2$</td>
<td>$\chi^2/N_{\text{dof}}$</td>
<td>$m_T^{\mu\mu}$ (GeV)</td>
<td>$m_T^{\mu\mu}$</td>
<td>$m_T^{\mu\mu}$</td>
<td>$m_T^{\mu\mu}$</td>
</tr>
<tr>
<td>71/68</td>
<td>70/68</td>
<td>66/68</td>
<td>64/68</td>
<td>68/68</td>
<td>64/68</td>
</tr>
</tbody>
</table>

The $\chi^2$ values are reasonable, indicating a satisfactory description of the data by the predictions at the best-fit $m_T$ values. The various distributions show different relative sensitivities to the PDF and QCD scale uncertainties.

As shown in Table 16, the combination of all eight measured distributions (including $|\eta^T|$, $|y^{\mu\mu}|$ and $\Delta\phi^{\mu\mu}$ which are not sensitive to $m_T$) significantly reduces the theoretical uncertainties due to both PDF and QCD scale effects. The $\chi^2$ values for the combined description of all eight distributions are reasonable for all PDFs, implying that there is no significant tension between the mass fit results from the individual distributions, once the correlations between the distributions are taken into account. Several additional tests using the predictions based on NNPDF 3.0 were performed to probe the compatibility of the top quark mass values extracted from the different distributions, and the accuracy of the physics modelling used to perform the extraction. The combined fit was repeated removing one distribution at a time. The largest shift of $-1.4 \pm 1.1$ GeV was observed when removing the $p_T^{\mu\mu}$ distribution, where the uncertainty corresponds to the quadrature difference of the fit uncertainties with and without the $p_T^{\mu\mu}$ distribution included. The removal of any other single distribution changed the result by less than 0.3 GeV, and a fit to only the five distributions directly sensitive to $m_T$ (excluding $|\eta^T|$, $|y^{\mu\mu}|$ and $\Delta\phi^{\mu\mu}$) gave a result of
173.1 ± 1.2 GeV, corresponding to a shift of −0.1 GeV with respect to the eight-distribution result. Finally, the individual measurements from the five directly-sensitive distributions were combined using the HAVERAGER program [125, 126]. Correlated statistical and systematic uncertainties were taken into account using nuisance parameters, but post-fit correlations between these nuisance parameters were neglected, unlike in the simultaneous fit approach with XFITTER. The average of the five measurements is 173.4 ± 1.6 GeV with a χ² of 6.4/4, in reasonable agreement with the result from the simultaneous fit of the five distributions. No additional uncertainty was included as a result of these tests.

The combined-fit χ² values in Table 16 are smallest for the HERAPDF 2.0 and NNPDF 3.0_nojet PDF sets, which do not include the constraints on the gluon PDF from LHC and Tevatron jet data in the region relevant for t¯t production. However, the m_t values resulting from the NNPDF 3.0 and NNPDF 3.0_nojet PDFs are close, indicating that the results are not sensitive to whether the jet data are included or not. Amongst the ‘global fit’ PDF sets incorporating a larger set of experimental data, the smallest χ² values result from the fit with NNPDF 3.0, though the values from the other PDFs are also reasonable. The results using NNPDF 3.0 were therefore used to define the central m_t value from the combined fit to all eight distributions, and an additional uncertainty of 0.3 GeV, corresponding to half the difference of the envelope encompassing all the other PDFs, was added in quadrature to the PDF uncertainty from NNPDF 3.0 alone. The effect of the uncertainty in the value of α_S was found to be 0.01 GeV. The residual dependence of the measured differential cross-sections on the top quark mass assumed in the simulation (see Sect. 5.1) is very small. A ±5 GeV variation around the baseline value of m_t = 172.5 GeV was assumed, giving a 0.1 GeV change on the result of the combined fit.

The choice of a fixed central scale, μ_F = μ_R = m_t/2 is expected to provide a good description of the inclusive t¯t cross-section and differential distributions in the kinematic regions dominated by top quarks with relatively low p_T. However, dynamical scales, which vary as a function of the top quark kinematics, are expected to be more appropriate for modelling the regions with high p_T [107]. Two alternative dynamical central scale choices for the t¯t production process were explored to test the sensitivity of the results to this choice:

- μ_F = μ_R = H_T/4 where H_T is defined as \( \sqrt{m_t^2 + p_T(t)^2} \) and \( p_T(t) \) are the transverse momentum of the top quark and antiquark, corresponding to one of the dynamical scales suggested in Ref. [107].
- μ_F = μ_R = E_T/2 where E_T is defined as \( \sqrt{m_t^2 + p_T(t)^2} \) and \( p_T(t) \) is the p_T of the t¯t system, analogously to a scale \( \sqrt{m_W^2 + p_T(W)^2} \) used in the description of jet production in association with W bosons [127, 128].

In both cases, the central scale for the top quark decay process \( t \to bℓν + X \) was fixed at m_t/2. The corresponding predictions for the top quark p_T spectrum from MCFM with NNPDF 3.0 and these scale choices are shown in Fig. 20, and compared to the ATLAS \( 8 \) TeV measurement using t¯t events with a lepton and at least four jets [20]. Unlike the predictions of POWHEG + PYTHIA6 used in Sect. 8.3, the MCFM
predictions with central scale choices of $\mu_F = \mu_R = m_t/2$, $H_T/4$ and $E_T/2$ provide good descriptions of the measured top quark $p_T$ spectrum, whereas $\mu_F = \mu_R = m_t$ is too hard. The results from the combined fit to all eight distributions with these scale choices and all six PDF sets are shown in the lower part of Table 16, and displayed graphically in Fig. 21. In the same way as for the fixed central scale, the actual factorisation and normalisation scales used in the predictions were allowed to vary independently around the dynamical central scales, with one standard deviation variations corresponding to factors of two and one-half. The $\chi^2$ values for the fits with a central scale of $H_T/4$ are all improved compared to those for $m_t/2$, reflecting a generally better description of the high-$p_T$ tails of the distributions. The $\chi^2$ values from the $E_T/2$ fits lie between the other two choices. The largest difference in the $m_t$ values from a dynamical scale and the fixed scale with any PDF (1.1 GeV for $E_T/2$ vs. $m_t/2$ with the CT14 PDF) was used to define an additional theoretical uncertainty due to the choice of the functional form of the QCD scales.

The final top quark mass value from the combination of all distributions is:

$$m_t^{\text{pole}} = 173.2 \pm 0.9 \pm 0.8 \pm 1.2 \text{ GeV},$$

where the three uncertainties arise from data statistics, experimental systematic effects, and uncertainties in the theoretical predictions, giving a total uncertainty of 1.6 GeV. The theoretical uncertainty is dominated by the comparison of results with different QCD central scale choices. Figure 22 shows a comparison with previous determinations of the top quark pole mass from the inclusive $t\bar{t}$ production cross-section [13,15,40] and from the invariant mass distribution of the $t\bar{t}$ plus one jet system [41]. The present result is in agreement with these other results, all of which have larger uncertainties. It is also in agreement with the Tevatron and LHC average measurement of 173.34 ± 0.76 GeV from reconstruction of the top quark decay products [121], as well as with more precise recent results using similar techniques [35,36,129]. However, the precision of the present pole mass result is not sufficient to probe potential differences between it and the other techniques at the 1 GeV level.
The theoretical uncertainty of 1.2 GeV on the final result using fixed-order predictions is significantly smaller than the uncertainties due to $\bar{t}t$ modelling and potential NNLO effects in the top quark $p_T$ spectrum for the fits based on POWHEG + PYTHIA6 templates. In the fixed-order approach, the potential missing NNLO corrections are absorbed into the variations of the QCD scales $\mu_F$ and $\mu_R$, which are significantly constrained by the fit to the complete set of distributions, including those with little sensitivity to $m_t$. However, there remains a significant uncertainty of about 1 GeV due to the choice of the functional form of the QCD scales, limiting the gain from the combined fit. This approach would therefore benefit significantly from the availability of fixed-order calculations including NNLO effects in the top quark production and decay [130], which should reduce the uncertainties due to scale choices. Off-shell and interference effects in the $pp \rightarrow WWb\bar{b} \rightarrow e\mu\nu\bar{\nu}b\bar{b} + X$ process (including both $\bar{t}t$ and single top $Wt$ contributions) [131–137], as well as NLO electroweak corrections [138,139], were not considered in this analysis. They are expected to be small compared to the theoretical uncertainties of the current result, but likely cannot be neglected in a determination of $m_t$ based on NNLO QCD predictions. These theoretical advances would allow the power of the full set of distributions to be utilised more effectively, especially in view of the likely reduction in the experimental statistical and systematic uncertainties from the larger $\bar{t}t$ samples now becoming available from LHC running at $\sqrt{s} = 13$ TeV.

9 Conclusions

Lepton and dilepton differential cross-section distributions have been measured in $\bar{t}t \rightarrow e\mu\nu\bar{\nu}b\bar{b}$ events selected from 20.2 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 8$ TeV recorded by the ATLAS detector at the LHC. The absolute and normalised cross-sections were measured using opposite-charge $e\mu$ events with one or two $b$-tagged jets, and corrected to a fiducial volume corresponding to the experimental acceptance of the leptons and no requirements on jets. Eight single lepton and dilepton differential distributions were measured, with relative uncertainties varying in the range 1–10%, and presented with and without the contribution from leptonic decays of $\tau$-leptons produced in the $W$ decays.

The results were compared to the predictions of various $\bar{t}t$ NLO and LO multileg matrix element event generators interfaced to several parton shower and hadronisation models. These generally give a good description of the distributions, though some distributions are modelled poorly by certain event generators. Those involving rapidity information are better described by the HERAPDF PDF sets than the CT10 set used as default. The distributions also show some sensitivity to NNLO corrections in the description of the top quark $p_T$ spectrum. The data are sensitive to the gluon PDF around $x \approx 0.1$ and have the potential to reduce PDF uncertainties in this region.

Several of the measured distributions are sensitive to the top quark mass, in a way which is complementary to traditional measurements of $m_t$ using the invariant mass of the reconstructed top quark decay products. Various techniques for extracting the top quark mass from the measured distributions were explored, including fits using templates from POWHEG + PYTHIA6 simulated samples, mass determinations based on moments of the distributions, and fits to fixed-order NLO QCD predictions, giving access to the top quark pole mass in a well-defined renormalisation scheme as implemented in MCFM. The most precise result was obtained from a fit of fixed-order predictions to all eight measured distributions simultaneously, extracting $m_t^{\text{pole}}$ whilst simultaneously constraining uncertainties due to PDFs and QCD scales. The final result is:

$$m_t^{\text{pole}} = 173.2 \pm 0.9 \pm 0.8 \pm 1.2 \text{ GeV},$$

where the three uncertainties arise from data statistics, experimental systematic effects, and uncertainties in the theoretical predictions. This result is in excellent agreement with other determinations of $m_t^{\text{pole}}$ from inclusive and differential cross-sections, and traditional measurements based on reconstruction of the top quark decay products.

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 acknowl-
edge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CF, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNNRF and DNNRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; SRNSGF, 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; MNE-IJaF, Romania; MES of Russia and NRC KI, Russian Federation; JINR, MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallaceberg 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; BSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhultme 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 large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [140].

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP$^3$.

References

15. CMS Collaboration, Measurement of the $t\bar{t}$ production cross section in the $e\mu$ channel in proton–proton collisions at $\sqrt{s} = 7$ and 8 TeV. JHEP 08, 029 (2016). https://doi.org/10.1007/JHEP08(2016)029. arXiv:1603.02303 [hep-ex]
25 Department of Physics, Brandeis University, Waltham, MA, USA
26 (a) Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil; (b) Electrical Circuits Department, Federal University of Juiz de Fora (UFJF), Juiz de Fora, Brazil; (c) Federal University of Sao Joao del Rei (UFSJ), Sao Joao del Rei, Brazil; (d) Instituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil
27 Physics Department, Brookhaven National Laboratory, Upton, NY, USA
28 (a) Transilvania University of Brasov, Brasov, Romania; (b) Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania; (c) Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi, Romania; (d) Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj-Napoca, Romania; (e) University Politehnica Bucharest, Bucharest, Romania; (f) West University in Timisoara, Timisoara, Romania
29 Departamento de Física, Universidade de Buenos Aires, Buenos Aires, Argentina
30 Cavendish Laboratory, University of Cambridge, Cambridge, UK
31 Department of Physics, Carleton University, Ottawa, ON, Canada
32 CERN, Geneva, Switzerland
33 Enrico Fermi Institute, University of Chicago, Chicago, IL, USA
34 (a) Departamento de Física, Pontificia Universidad Católica de Chile, Santiago, Chile; (b) Departamento de Física, Universidad Técnica Federico Santa María, Valparaíso, Chile
35 (a) Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; (b) Department of Physics, Nanjing University, Nanjing, Jiangsu, China; (c) Physics Department, Tsinghua University, Beijing 100084, China; (d) University of Chinese Academy of Science (UCAS), Beijing, China
36 (a) Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei, Anhui, China; (b) School of Physics, Shandong University, Jinan, Shandong, China; (c) Department of Physics and Astronomy, Key Laboratory for Particle Physics, Astrophysics and Cosmology, Ministry of Education, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University, Shanghai (also at PKU-CHEP), Shanghai, China
37 Université Clermont Auvergne, CNRS/IN2P3, LPC, Clermont-Ferrand, France
38 Nevis Laboratory, Columbia University, Irvington, NY, USA
39 Niels Bohr Institute, University of Copenhagen, Kobenhavn, Denmark
40 (a) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Frascati, Italy; (b) Dipartimento di Fisica, Università della Calabria, Rende, Italy
41 (a) Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Krakow, Poland; (b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow, Poland
42 Institute of Nuclear Physics, Polish Academy of Sciences, Krakow, Poland
43 Physics Department, Southern Methodist University, Dallas, TX, USA
44 Physics Department, University of Texas at Dallas, Richardson, TX, USA
45 DESY, Hamburg and Zeuthen, Germany
46 Lehrstuhl für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany
47 Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany
48 Department of Physics, Duke University, Durham, NC, USA
49 SUPA-School of Physics and Astronomy, University of Edinburgh, Edinburgh, UK
50 INFN e Laboratori Nazionali di Frascati, Frascati, Italy
51 Fakultät für Mathematik und Physik, Albert-Ludwigs-Universität, Freiburg, Germany
52 Département de Physique Nucléaire et Corpusculaire, Université de Genève, Geneva, Switzerland
53 (a) INFN Sezione di Genova, Genoa, Italy; (b) Dipartimento di Fisica, Università di Genova, Genoa, Italy
54 (a) E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi, Georgia; (b) High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia
55 II Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany
56 SUPA-School of Physics and Astronomy, University of Glasgow, Glasgow, UK
57 II Physikalisches Institut, Georg-August-Universität, Göttingen, Germany
58 Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, Grenoble, France
INFN Sezione di Napoli, Naples, Italy; Dipartimento di Fisica, Università di Napoli, Naples, Italy
Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, USA
Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands
Nikhef National Institute for Subatomic Physics, University of Amsterdam, Amsterdam, The Netherlands
Department of Physics, Northern Illinois University, DeKalb, IL, USA
Budker Institute of Nuclear Physics, SB RAS, Novosibirsk, Russia
Department of Physics, New York University, New York, NY, USA
Ohio State University, Columbus, OH, USA
Department of Physics, Oklahoma State University, Stillwater, OK, USA
Palacky University, RCPTM, Olomouc, Czech Republic
Center for High Energy Physics, University of Oregon, Eugene, OR, USA
LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
Department of Physics, University of Pittsburgh, Pittsburgh, PA, USA
Department of Physics and Astronomy, University of Washington, Seattle, WA, USA
Department of Physics and Astronomy, University of Sheffield, Sheffield, UK
Department of Physics, Shinshu University, Nagano, Japan
Department Physik, Universität Siegen, Siegen, Germany
Department of Physics, Simon Fraser University, Burnaby, BC, Canada
SLAC National Accelerator Laboratory, Stanford, CA, USA
(a) Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovak Republic; (b) Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
(a) Department of Physics, University of Cape Town, Cape Town, South Africa; (b) Department of Physics, University of Johannesburg, Johannesburg, South Africa; (c) School of Physics, University of the Witwatersrand, Johannesburg, South Africa
(a) Department of Physics, Stockholm University, Stockholm, Sweden; (b) The Oskar Klein Centre, Stockholm, Sweden
Physics Department, Royal Institute of Technology, Stockholm, Sweden
Departments of Physics and Astronomy and Chemistry, Stony Brook University, Stony Brook, NY, USA
Department of Physics and Astronomy, University of Sussex, Brighton, UK
School of Physics, University of Sydney, Sydney, Australia
Institute of Physics, Academia Sinica, Taipei, Taiwan
Department of Physics, Technion: Israel Institute of Technology, Haifa, Israel
Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece
International Center for Elementary Particle Physics and Department of Physics, The University of Tokyo, Tokyo, Japan
Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo, Japan
Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
Tomsk State University, Tomsk, Russia
Department of Physics, University of Toronto, Toronto, ON, Canada
(a) INFN-TIFPA, Trento, Italy; (b) University of Trento, Trento, Italy
(a) TRIUMF, Vancouver, BC, Canada; (b) Department of Physics and Astronomy, York University, Toronto, ON, Canada
Faculty of Pure and Applied Sciences, and Center for Integrated Research in Fundamental Science and Engineering, University of Tsukuba, Tsukuba, Japan
Department of Physics and Astronomy, Tufts University, Medford, MA, USA
Department of Physics and Astronomy, University of California Irvine, Irvine, CA, USA
(a) INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine, Italy; (b) ICTP, Trieste, Italy; (c) Dipartimento di Chimica, Fisica e Ambiente, Università di Udine, Udine, Italy
Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
Department of Physics, University of Illinois, Urbana, IL, USA
Instituto de Fisica Corpuscular (IFIC), Centro Mixto Universidad de Valencia-CSIC, Valencia, Spain
Department of Physics, University of British Columbia, Vancouver, BC, Canada
Department of Physics and Astronomy, University of Victoria, Victoria, BC, Canada
Department of Physics, University of Warwick, Coventry, UK
Waseda University, Tokyo, Japan
Department of Particle Physics, The Weizmann Institute of Science, Rehovot, Israel
Department of Physics, University of Wisconsin, Madison, WI, USA
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität, Würzburg, Germany
Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal, Germany
Department of Physics, Yale University, New Haven, CT, USA
Yerevan Physics Institute, Yerevan, Armenia
Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France
Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan

a Also at Department of Physics, King’s College London, London, UK
b Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
c Also at Novosibirsk State University, Novosibirsk, Russia