Measurement of jet activity produced in top-quark events with an electron, a muon and two b-tagged jets in the final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

ATLAS Collaboration
CERN, 1211 Geneva 23, Switzerland

Received: 2 November 2016 / Accepted: 20 March 2017
© CERN for the benefit of the ATLAS collaboration 2017. This article is an open access publication

Abstract Measurements of jet activity in top-quark pair events produced in proton–proton collisions are presented, using 3.2 fb$^{-1}$ of pp collision data at a centre-of-mass energy of 13 TeV collected by the ATLAS experiment at the Large Hadron Collider. Events are chosen by requiring an opposite-charge $e\mu$ pair and two $b$-tagged jets in the final state. The normalised differential cross-sections of top-quark pair production are presented as functions of additional-jet multiplicity and transverse momentum, $p_T$. The fraction of signal events that do not contain additional jet activity in a given rapidity region, the gap fraction, is measured as a function of the $p_T$ threshold for additional jets, and is also presented for different invariant mass regions of the $e\mu b\bar{b}$ system. All measurements are corrected for detector effects and presented as particle-level distributions compared to predictions with different theoretical approaches for QCD radiation. While the kinematics of the jets from top-quark decays are described well, the generators show differing levels of agreement with the measurements of observables that depend on the production of additional jets.

Contents

1 Introduction ..............................................................................................................................................
2 ATLAS detector ....................................................................................................................................
3 Data and simulation samples ..................................................................................................................
4 Object reconstruction .............................................................................................................................
5 Event selection and background estimates ............................................................................................
6 Sources of systematic uncertainty ..........................................................................................................*
7 Definition of the fiducial phase space .....................................................................................................
8 Measurement of jet multiplicities and $p_T$ spectra ..............................................................................
   8.1 Jet multiplicity results .....................................................................................................................
   8.2 Jet $p_T$ spectra results ..................................................................................................................
9 Gap fraction measurements ....................................................................................................................
   9.1 Gap fraction results in rapidity regions .............................................................................................
   9.2 Gap fraction results in $m_{e\mu b\bar{b}}$ subsamples ...........................................................................
10 Conclusions ........................................................................................................................................
References ..................................................................................................................................................

1 Introduction

Top-quark pair production final states in proton–proton ($pp$) collisions at the Large Hadron Collider (LHC) often include additional jets not directly produced in the top-quark decays. The uncertainties associated with these processes are significant in precision measurements, such as the measurement of the top-quark mass [1] and the inclusive $t\bar{t}$ production cross-section [2].

These additional jets arise mainly from hard gluon emissions from the hard-scattering interaction beyond $t\bar{t}$ production and are described by quantum chromodynamics (QCD). The higher centre-of-mass energy of the $pp$ scattering process in LHC Run 2 opens a large kinematic phase space for QCD radiation. Several theoretical approaches are available to model the production of these jets in $t\bar{t}$ processes, including next-to-leading-order (NLO) QCD calculations, parton-shower models, and methods matching fixed-order QCD with the parton shower. The aim of this analysis is to test the predictions of extra jet production in these approaches and to provide data to adjust free parameters of the models to optimise their predictions.

The jet activity is measured in events with at least two $b$-tagged jets, i.e. jets tagged as containing $b$-hadrons, and exactly one electron and exactly one muon of opposite electrical charge in the final state. Additional jets are defined as jets produced in addition to the two $b$-tagged jets required for the event selection, without requiring any matching of jets to partons. In order to probe the $p_T$ dependence of the hard-gluon emission, this analysis measures the normalised differential $t\bar{t}$ cross-sections as a function of the jet multiplicity for different transverse momentum ($p_T$) thresholds of the
additional jets. The $p_T$ of the leading additional jet is measured, as well as the $p_T$ of the leading and sub-leading jets initiated by $b$-quarks ("$b$-jets"), which are top-quark decay products in most of the events.

Furthermore, the gap fraction defined as the fraction of events with no jet activity in addition to the two $b$-tagged jets above a given $p_T$ threshold in a rapidity region in the detector, is measured as a function of the additional jets’ minimum $p_T$ threshold as defined in Refs. [3,4]. The results are presented in a fiducial phase space in which all selected final-state objects are produced within the detector acceptance following the definitions in Ref. [5].

This paper provides a measurement of additional jets in $t\bar{t}$ events in the dilepton channel for the new centre-of-mass energy of 13 TeV. Measurements similar to those presented in this paper were performed by ATLAS at 7 TeV [3,5] and have been used to tune parameters in Monte Carlo (MC) generators for LHC Run 2 [6–8]. These earlier measurements were performed in the lepton+jets channel where the inclusive jet multiplicity was measured, since it is difficult to distinguish jets originating in $W$ decays from additional jets produced by QCD radiation. Recent measurements of jet multiplicity were performed in the single lepton channel by CMS at 13 TeV [9] and in the dilepton channel, including also the gap fractions, by ATLAS and CMS at 8 TeV [4,10].

2 ATLAS detector

The ATLAS detector [11] at the LHC covers nearly the entire solid angle around the interaction point. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroid magnets. The inner-detector system is immersed in a 2T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$.

The high-granularity silicon pixel detector covers the interaction region and provides four measurements per track. The closest layer, known as the Insertable B-Layer (IBL) [12], was added in 2014 and provides high-resolution hits at small radius to improve the tracking performance. The pixel detector is followed by the silicon microstrip tracker, which provides four three-dimensional measurement points per track. These silicon detectors are complemented by the transition radiation tracker, which enables radially extended track reconstruction up to $|\eta| = 2.0$. The transition radiation tracker also provides electron identification information based on the fraction of hits (typically 30 in total) passing a higher charge threshold indicative of transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) electromagnetic calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by the steel/scintillating tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules, which are optimised for electromagnetic and hadronic measurements, respectively.

The muon spectrometer comprises separate trigger and high-precision tracking chambers, measuring the deflection of muons in a magnetic field generated by superconducting air-core toroids. The precision chamber system surrounds the region $|\eta| < 2.7$ with three layers of monitored drift tubes, complemented by cathode strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive plate chambers in the barrel, and thin-gap chambers in the endcap regions.

A two-level trigger system is used to select interesting events [13,14]. The Level-1 trigger is implemented in hardware and uses a subset of detector information to reduce the event rate to a design value of at most 100 kHz. This is followed by the high-level software-based trigger (HLT), which reduces the event rate to 1 kHz.

3 Data and simulation samples

The proton–proton ($pp$) collision data used in this analysis were collected during 2015 by the ATLAS detector and correspond to an integrated luminosity of 3.2 fb$^{-1}$ at $\sqrt{s} = 13$ TeV. The data considered in this analysis were collected under stable beam conditions, requiring that all detectors were operational. Each selected event includes interactions from an average of 14 inelastic $pp$ collisions in the same proton bunch crossing, as well as residual signals from previous bunch crossings with a 25 ns bunch spacing. These two effects are collectively referred to as “pile-up”. Events are required to pass a single-lepton trigger, either electron or muon. Multiple triggers are used to select events: either triggers with low lepton $p_T$ thresholds of 24 GeV which utilise isolation requirements to reduce the trigger rate, or triggers with higher $p_T$ thresholds but looser isolation requirements...
to increase event acceptance. The higher $p_T$ thresholds were 50 GeV for muons and 60 GeV or 120 GeV for electrons.

MC simulations are used to model background processes and to correct the data for detector acceptance and resolution effects. The nominal $t\bar{t}$ sample is simulated using the NLO Powheg-Box v2 matrix-element (ME) generator [15–17], referred to as Powheg in the following, and Pythia6 [18] (v6.427) for the parton shower (PS), hadronisation and underlying event. Powheg is interfaced to the CT10 [19] NLO parton distribution function (PDF) set, while Pythia6 uses the CTEQ6L1 PDF set [20]. Pythia simulates the underlying event and parton shower using the P2012 set of tuned parameters (tune) [21]. The “$h_{damp}$” parameter, which controls the $p_T$ of the first additional emission beyond the Born configuration, is set to the mass of the top quark ($m_t$). The main effect of this is to regulate the high-$p_T$ emission against which the $t\bar{t}$ system recoils. The choice of this $h_{damp}$ value has been found to improve the modelling of the $t\bar{t}$ system kinematics with respect to data in previous analyses [6]. In order to investigate the effects of initial- and final-state radiation, alternative Powheg+Pythia6 samples are generated with the renormalisation and factorisation scales varied by a factor of 2 (0.5) and using low (high) radiation variations of the Perugia 2012 tune and an $h_{damp}$ value of $m_t$ ($2m_t$), corresponding to less (more) parton-shower radiation [6]. These samples are called RadHi and RadLo in the following. These variations are selected to cover the uncertainties in the measurements of differential distributions in 7 TeV data [22]. Alternative samples are generated using Powheg and MadGraph5_aMC@NLO [23] (v2.2.1) with CKKW-L, referred to as MG5_aMC@NLO hereafter, both interfaced to Herwig++ [24] (v2.7.1), in order to estimate the effects of the choice of matrix-element generator. These $t\bar{t}$ samples are described in Ref. [6].

Additional $t\bar{t}$ samples are generated for comparisons with unfolded data as follows. The predictions of the ME generators Powheg and MG5_aMC@NLO are interfaced to Herwig7 [24,25] and Pythia8. In all Powheg and MG5_aMC@NLO samples mentioned above, the first emission is calculated from the leading-order real emission term, and further additional jets are simulated from parton showering, which is affected by significant theoretical uncertainties. Improved precision is expected from using Sherpa v2.2 [26], which models the inclusive and the one-additional-jet process using an NLO matrix element and up to four additional jets at leading-order (LO) accuracy using the ME + PS@NLO prescription [27]. The sample used to compare to particle-level results presented here is generated with the central scale set to $\mu^2 = m_t^2 + 0.5 \times (p_{T,T}^2 + p_{T,\tau}^2)$, where $p_{T,T}$ and $p_{T,\tau}$ refer to the $p_T$ of the top and antitop quark, respectively, and with the matching scale set to 30 GeV. Furthermore, the NNPDF 3.0 PDF [28] at next-to-next-to-leading order (NNLO) is used.

All $t\bar{t}$ samples are normalised to the cross-section calculated with the Top++2.0 program to NNLO in perturbative QCD, including soft-gluon resummation to NNLL [29], assuming a top-quark mass of 172.5 GeV.

Background processes are simulated using a variety of MC generators, as described below. Details of the background estimation are described in Sect. 5. Single top-quark production in association with a W boson ($Wt$) is simulated using Powheg-Box v1+Pythia6 with the same parameters and PDF sets as those used for the nominal $t\bar{t}$ sample and is normalised to the approximate NNLO cross-section (71.7 ± 3.8 pb) described in Ref. [30]. At NLO, part of the final state of $Wt$ production is identical to the final state of $t\bar{t}$ production. The “diagram removal” (DR) generation scheme [31] is used to remove this part of the phase space from the background calculation. A sample generated using an alternative “diagram subtraction” (DS) method [31] is used to evaluate systematic uncertainties. Both samples are normalised to the generator cross-section.

The majority of backgrounds with at least one misidentified lepton in the selected sample arise from $t\bar{t}$ production in which only one of the top quarks decays semileptonically, which is simulated in the same way as the $t\bar{t}$ production in which both top quarks decay leptonically.

Sherpa v2.1, interfaced to the CT10 PDF set, is used to model Drell–Yan production, specifically $Z/\gamma^* \rightarrow \tau^+\tau^-$. For this process, Sherpa calculates matrix elements at NLO for up to two partons and at LO for up to four partons using the OpenLoops [32] and Comix [33] matrix-element generators. The matrix elements are merged with the Sherpa PS [34] using the ME + PS@NLO prescription [35]. The total cross-section is normalised to NNLO predictions calculated using the FEWZ program [36] with the MSTW2008NNLO PDF [37]. Sherpa v2.1 with the CT10 PDF set is also used to simulate electroweak diboson production [38] ($WW$, $WZ$, $ZZ$), where both bosons decay leptonically. For diboson production, Sherpa v2.1 calculates matrix elements at NLO for zero additional partons, at LO for one to three additional partons (with the exception of ZZ production, for which the one additional parton is also NLO), and using PS for all parton multiplicities of four or more.

The ATLAS detector response is simulated [39] using GEANT 4 [40]. A “fast simulation” [41], utilising parameterised showers in the calorimeter, is used in the samples chosen to estimate $t\bar{t}$ modelling uncertainties. Additional $pp$ interactions are generated using Pythia8.186 [42] with tune A2 and overlaid with signal and background processes in order to simulate the effect of pile-up. The MC simulations are reweighted to match the distribution of the average number of interactions per bunch crossing that are observed in data, referred to as “pile-up reweighting”. Corrections are applied to the MC simulation in order to improve agreement with data for the efficiencies of reconstructed objects. The same reconstruction algorithms and analysis procedures are then applied to both data and MC simulation.
4 Object reconstruction

This analysis selects reconstructed electrons, muons and jets. Electron candidates are identified by matching an inner-detector track to an isolated energy deposit in the electromagnetic calorimeter, within the fiducial region of transverse momentum $p_T > 25$ GeV and pseudorapidity $|\eta| < 2.47$. Electron candidates are excluded if the energy cluster is within the transition region between the barrel and the end-cap of the electromagnetic calorimeter, $1.37 < |\eta| < 1.52$, and if they are also reconstructed as photons. Electrons are selected using a multivariate algorithm and are required to satisfy a likelihood-based quality criterion, in order to provide high efficiency and good rejection of fake and non-prompt electrons [43,44]. Electron candidates must have tracks that pass the requirements of transverse impact parameter significance $|d_0^{\text{Sig}}| < 5$ and longitudinal impact parameter $|z_0 \sin \theta| < 0.5$ mm. Electrons must also pass isolation requirements based on inner-detector tracks and topological energy clusters varying as a function of $\eta$ and $p_T$. The track isolation cone size is given by the smaller of $\Delta R = 10 \text{ GeV}/p_T$ and $\Delta R = 0.2$, i.e. a cone which increases in size at lower $p_T$ values, up to a maximum of 0.2. These requirements result in a 95% efficiency of the isolation cuts for electrons from $Z \rightarrow e^+e^-$ decays with $p_T$ of 25 GeV and 99% for electrons with $p_T$ above 60 GeV; when estimated in simulated $t\bar{t}$ events, this efficiency is smaller by a few percent, due to the increased jet activity. Electrons that share a track with a muon are discarded. Double counting of electron energy deposits as jets is prevented by removing the closest jet with an angular distance $\Delta R < 0.2$ from a reconstructed electron. Following this, the electron is discarded if a jet exists within $\Delta R < 0.4$ of the electron, to ensure sufficient separation from nearby jet activity.

Muon candidates are identified from a track in the inner detector matching a track in the muon spectrometer; the combined track is required to have $p_T > 25$ GeV and $|\eta| < 2.5$ [45]. The tracks of muon candidates are required to have a transverse impact parameter significance $|d_0^{\text{Sig}}| < 3$ and a longitudinal impact parameter below 0.5 mm. Muons are required to meet quality criteria and the same isolation requirement as applied to electrons, to obtain the same isolation efficiency performance as for electrons. These requirements reduce the contributions from fake and non-prompt muons. Muons may leave energy deposits in the calorimeter that could be misidentified as a jet, so jets with fewer than three associated tracks are removed if they are within $\Delta R < 0.4$ of a muon. Muons are discarded if they are separated from the nearest jet by $\Delta R < 0.4$, to reduce the background from muons originating in heavy-flavour decays inside jets.

Jets are reconstructed with the anti-$k_t$ algorithm [46,47], using a radius parameter of $R = 0.4$, from topological clusters of energy deposits in the calorimeters. Jets are accepted within the range $p_T > 25$ GeV and $|\eta| < 2.5$, and are calibrated using simulation with corrections derived from data [48]. Jets likely to originate from pile-up are suppressed using a multivariate jet-vertex-tagger (JVT) [49] for candidates with $p_T < 60$ GeV and $|\eta| < 2.4$. Jets containing $b$-hadrons are $b$-tagged using a multivariate discriminant [50], which uses track impact parameters, track invariant mass, track multiplicity and secondary vertex information to discriminate those jets from light quark or gluon jets (“light jets”). The average $b$-tagging efficiency is 77% for $b$-jets in simulated dileptonic $t\bar{t}$ events with a purity of 95%. The tagging algorithm gives a rejection factor of about 130 against light jets and about 4.5 against jets originating from charm quarks (“charm jets”).

5 Event selection and background estimates

Signal events are selected by requiring exactly one electron and one muon of opposite electric charge (“opposite sign”), and at least two $b$-tagged jets. With this selection, almost all of the selected events are $t\bar{t}$ events. The other processes that pass the signal selection are events with single top quarks ($Wt$), $t\bar{t}$ events in the single-lepton decay channel with a misidentified (fake) lepton, $Z/\gamma^* \rightarrow \tau^+\tau^-$ ($\rightarrow e\mu$) and diboson events. Other backgrounds, including processes with two misidentified leptons, are negligible for the event selections used in this analysis.

Additional jets are defined as those produced in addition to the two highest-$p_T$ $b$-tagged jets. They are identified as jets above $p_T$ thresholds of 25, 40, 60 and 80 GeV, independent of the jet flavour. In very rare cases, $b$-jets may also be produced in addition to the top-quark pair, for example through splitting of a very high momentum gluon, or through the decay of a Higgs boson into a bottom–antibottom pair, leading to events with more than two $b$-tagged jets. In this case, the two selected $b$-tagged jets with the highest $p_T$ are assumed to originate from $t\bar{t}$ decay, and the others are considered as additional jets. This procedure ignores that occasionally a $b$-jet which is not the decay product of a top quark might have higher $p_T$ than those from the top-quark decays. This is a negligible effect within the uncertainties of this measurement.

The single-top background is estimated from simulation, as described in Sect. 3. The background from $t\bar{t}$ events in the lepton+jets channel with a fake lepton is estimated from a combination of data and simulation, as in Ref. [2]. This method uses the observation that samples with a same-sign
$e\mu$ pair and two $b$-tagged jets are dominated by events with a misidentified lepton, with a rate comparable to those in the opposite-sign sample. The contributions of events with misidentified leptons are therefore estimated as same-sign event counts in data, after subtraction of predicted prompt same-sign contributions multiplied by the ratio of opposite-sign to same-sign fake leptons, as predicted from the nominal $t\bar{t}$ sample.

The backgrounds from $Z/\gamma* \rightarrow \tau^+\tau^-$ and from diboson events are estimated from simulation and are below 1%. The normalisation for the $Z/\gamma* \rightarrow \tau^+\tau^-$ contribution is estimated from events with $Z/\gamma* \rightarrow e^+e^-$ or $\mu^+\mu^-$ and two $b$-jets within the acceptance of this analysis. The Monte Carlo prediction is scaled by $1.37 \pm 0.30$ to fit the observed rate.

After the event selection, only about 4.5% of the events are background, as listed in Table 1. The background is dominated by single top production (3.1%) and fake leptons (1.6%). The event yields and the relative background contributions vary with jet multiplicity and jet $p_T$ values and at low additional jet multiplicities. At high jet multiplicities ($\geq 3$ additional jets) the fake-lepton background exceeds the number of single-top events. While the number of events observed in the 0-jet bin agrees with the prediction within the uncertainties, the data exceed the predictions increasingly with jet multiplicity, reaching a 25% deviation for events with at least four additional jets above 25 GeV.

The table and figures also list the contribution of $t\bar{t}$ events with at least one additional jet identified as originating from pile-up (pile-up jets). These are signal events, but a few pile-up jets are still in the sample after object and event selection, as the background suppression of the JVT cut is very high but not 100%. Due to the presence of at least one jet that does not originate from the hard interaction, these events may appear in the wrong jet multiplicity bin. In the jet $p_T$ spectra, pile-up jets contribute at low additional-jet $p_T$ as the pile-up jets are generally softer than the jets in $t\bar{t}$ events. For the same reason, pile-up jets only contribute significantly to the jet multiplicity distributions with the 25 GeV threshold. In most of the events with remaining pile-up jets, only one of the additional jets is caused by pile-up. Any remaining pile-up jets can be identified in the simulation, but not in data. Therefore the data are corrected for pile-up jets in the unfolding procedure, as described later.

6 Sources of systematic uncertainty

The systematic uncertainties of the reconstructed objects, in the signal modelling and in the background estimates, are evaluated as described in the following.

### Table 1: Yields of data and MC events fulfilling the selection criteria

<table>
<thead>
<tr>
<th>Process</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single top ($Wt$)</td>
<td>$236 \pm 2$ (stat.) $\pm 46$ (syst.)</td>
</tr>
<tr>
<td>Fake leptons</td>
<td>$117 \pm 22$ (stat.) $\pm 120$ (syst.)</td>
</tr>
<tr>
<td>Z+jets</td>
<td>$6 \pm 3$ (stat.) $\pm 1$ (syst.)</td>
</tr>
<tr>
<td>Dibosons</td>
<td>$3.1 \pm 0.4$ (stat.) $\pm 1.5$ (syst.)</td>
</tr>
<tr>
<td>Total background</td>
<td>$362 \pm 22$ (stat.) $\pm 130$ (syst.)</td>
</tr>
<tr>
<td>$t\bar{t}$ ($\geq 1$ pile-up jet)</td>
<td>$310 \pm 2$ (stat.) $\pm 88$ (syst.)</td>
</tr>
<tr>
<td>$t\bar{t}$ (no pile-up jets)</td>
<td>$6850 \pm 11$ (stat.) $\pm 940$ (syst.)</td>
</tr>
<tr>
<td>Expected</td>
<td>$7520 \pm 25$ (stat.) $\pm 950$ (syst.)</td>
</tr>
<tr>
<td>Observed</td>
<td>$8050$</td>
</tr>
</tbody>
</table>

The jet energy scale (JES) uncertainty is evaluated by varying 19 uncertainty parameters derived from in situ analyses at $\sqrt{s} = 8$ TeV and extrapolated to data at $\sqrt{s} = 13$ TeV [48]. The JES uncertainty is 5.5% for jets with $p_T$ of 25 GeV and quickly decreases with increasing jet $p_T$, falling to below 2% for jets above 80 GeV. The uncertainty in the jet energy resolution (JER) is calculated by extrapolating the uncertainties derived at $\sqrt{s} = 8$ TeV to $\sqrt{s} = 13$ TeV [48]. The uncertainty in JER is at most 3.5% at $p_T$ of 25 GeV, quickly decreasing with increasing jet $p_T$ to below 2% for jets above 50 GeV.

Uncertainties on the efficiency for tagging $b$-jets were determined using the methods described in Ref. [51] applied to dileptonic $t\bar{t}$ events in $\sqrt{s} = 13$ TeV data. The uncertainties on mistagging of charm and light jets were determined using $\sqrt{s} = 8$ TeV data as described in Refs. [52,53]. Additional uncertainties are assigned to take into account the presence of the new IBL detector and the extrapolation to $\sqrt{s} = 13$ TeV [50].

The lepton-related uncertainties are assessed mostly using $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ decays measured in $\sqrt{s} = 13$ TeV data. The differences between the topologies of $Z$ and $t\bar{t}$ pair production events are expected not to be significant for the estimation of uncertainties.

The uncertainty associated with the amount of QCD initial- and final-state radiation is evaluated as the difference between the baseline MC sample and the corresponding RadHi and RadLo samples described in Sect. 3. The uncertainty due to the choice of parton-shower and hadronisation algorithms in the signal modelling is assessed by comparing the baseline MC sample (Powheg+Pythia6) with Powheg+Herwig++. The uncertainty due to the use of a specific NLO MC sample with its particular matching algorithm is derived from the comparison of Powheg+Herwig++ to the MG5\_aMC@NLO+Herwig++ sample.

The uncertainty due to the particular PDF used for the signal model prediction is evaluated by taking the standard deviation of variations from 100 eigenvectors of the recommended Run-2 PDF4LHC [54] set and adding them in
The multiplicity of additional jets with $p_T > 25\text{ GeV}$, $p_T > 40\text{ GeV}$, $p_T > 60\text{ GeV}$, and $p_T > 80\text{ GeV}$ for selected events at reconstruction level in data and simulation. Simulated signal events with at least one additional jet identified as pile-up are indicated in grey. The contribution of pile-up jets to the backgrounds is negligible. The lower panel shows the ratio of the total prediction to the data (solid line), the grey band represents the statistical uncertainty of the measurement, and the error bars on the solid line show the statistical uncertainty in the signal MC sample.

The uncertainty in the single top-quark background is evaluated based on the $5.3\%$ error in the approximate NNLO cross-section prediction and by comparing samples with diagram removal and diagram subtraction schemes, as described in Sect. 3. The uncertainty in the background from fake leptons is estimated to be $100\%$ from the statistical uncertainty of the same-sign event counts in data and an interpolation error using the envelope of the differences of individual subcomponents (such as photon-conversion, heavy-flavour decay leptons, for example) of misidentified lepton background between the same-sign and the opposite-sign sample.

For $Z$+jets backgrounds, the scale factor derived in the $e^+e^-$ and $\mu^+\mu^-$ channels and used to reweight the signal-region distribution is varied by $22\%$, corresponding to the difference in the scale factors derived in subsamples with and without an additional jet. This value covers the variations of the correction factor derived from subsets of events with different jet multiplicities. No theoretical uncertainty is applied to the $Z$+jets background normalisation as this is scaled to data.

The uncertainty in the amount of pile-up is estimated by changing the nominal MC reweighting factors to vary the number of interactions per bunch crossing in data up and down by $10\%$. Two methods were used to estimate the amount of interactions per bunch crossing. The first method calculated the number of interactions using the instantaneous luminosity and the inelastic proton-proton cross section [56,57]. The results of the calculation were compared to results from a data-driven method based on the number of reconstructed vertices. The difference between the correlation of the two methods in data and MC is taken as the uncertainty.

The uncertainty due to the $2–3\%$ loss of hard-scatter jets due to the JVT cut is estimated using $Z$+jet events. The uncertainty in the efficiency of the JVT cut to reduce pile-up jets is estimated by using a sideband method. The JVT cut is inverted in simulation to estimate the number of pile-up jets and derive a scale factor to describe the number of pile-up jets.
jets in data. This factor is then used to scale the predicted number of pile-up jets in the signal region (with the JVT cut applied). Scale factors are also derived using the samples with increased and decreased pile-up mentioned above, and the larger of two variations is taken as systematics.

7 Definition of the fiducial phase space

For the measurement of the jet multiplicity, the jet $p_T$ spectra and the gap fractions, the data are corrected to particle level by comparing to events from MC generators in the fiducial volume described below. The fiducial volume, i.e., the object definitions and the kinematic phase space at particle level, is designed to match the reconstruction level as closely as possible and follow closely the definitions in Refs. [4,5]. Leptons and jets are defined using particles with a mean lifetime greater than $3 \times 10^{-10}$ s, directly produced in $pp$ interactions or from subsequent decays of particles with a shorter lifetime. Leptons from $W$ boson decays ($e, \mu, \nu_e, \nu_\mu, \nu_\tau$) are identified as such by requiring that they are not hadron decay products. Electron and muon four-momenta are calculated after the addition of photon four-momenta within a cone of $\Delta R = 0.1$ around their original directions.

Jets are defined using the anti-$k_t$ algorithm with a radius parameter of 0.4. All particles are considered for jet clustering, except for leptons from $W$ boson decays as defined above (i.e., neutrinos from hadron decays are included in jets) and any photons associated with the selected electrons or muons. Jets initiated by $b$-quarks are identified as such, i.e., identified as $b$-jets if a hadron with $p_T > 5$ GeV containing a $b$-quark is associated with the jet through a ghost-matching technique as described in Ref. [58].

The cross-section is defined using events with exactly one electron and one muon with opposite-sign directly from $W$ boson decays, i.e. excluding electrons and muons from decay of the $\tau$ leptons. In addition, at least two $b$-jets each with $p_T > 25$ GeV and $|\eta| < 2.5$ are required. Following the reconstructed object selection, events with jet–electron pairs or jet–muon pairs with $\Delta R < 0.4$ are excluded. Additional jets are considered within $|\eta| < 2.5$ for $p_T$ thresholds of 25 GeV or higher, independently of their flavour.
8 Measurement of jet multiplicities and $p_T$ spectra

The multiplicities of additional reconstructed jets with different $p_T$ thresholds are corrected to particle level within the fiducial volume as defined above. Even though the kinematic range of the measurement is chosen to be the same for particle-level and reconstruction-level objects, corrections are necessary due to the efficiencies and detector resolutions that cause differences between reconstruction-level and particle-level jet distributions. Examples include events in which one or more particle-level jets do not pass the $p_T$ threshold for reconstruction-level jets and when the selection efficiency for inclusive $t\bar{t}$ events changes as a function of jet multiplicity. Furthermore, additional reconstructed jets without a corresponding particle-level jet may appear due to pile-up, or if a jet migrates into the fiducial volume due to an upward fluctuation caused by the $p_T$ resolution, or if a single particle-level jet is reconstructed as two separate jets. These effects lead to migrations between bins and are taken into account within an iterative Bayesian unfolding [59].

The reconstructed jet multiplicity measurements are corrected separately for each additional-jet $p_T$ threshold according to

$$N_{\mathrm{unfold}}^i = \frac{1}{f_{\mathrm{eff}}^i} \cdot \sum_j (M^{-1})_{\mathrm{reco},i}^{j} f_{\mathrm{accept}}^j (N_{\mathrm{data}}^j - N_{\mathrm{bg}}^j),$$

(1)

where $N_{\mathrm{unfold}}^i$ is the total number of fully corrected particle-level events with particle-level jet multiplicity $i$. The term $f_{\mathrm{eff}}^i$ represents the efficiency to reconstruct an event with $i$ additional jets, defined as the ratio of events with $i$ particle-level jets that fulfil both the fiducial volume selection at particle-level and the reconstruction-level selection, $N_{\mathrm{reco},\mathrm{part}}^i$, to the number of events that fulfil the particle-level selection, $N_{\mathrm{part}}^i$:

$$f_{\mathrm{eff}}^i = \frac{N_{\mathrm{reco},\mathrm{part}}^i}{N_{\mathrm{part}}^i}.$$  

(2)

The resulting ratio $f_{\mathrm{eff}}$ is approximately 0.33 and has very small dependence on the jet multiplicity. The analysis of different $t\bar{t}$ MC samples results in values of $f_{\mathrm{eff}}$ which vary by up to 10%. The variations of $f_{\mathrm{eff}}$ between different $p_T$ thresholds are less than 2%. The function $f_{\mathrm{accept}}^j$ is the probability of an event fulfilling the reconstruction-level selection with $j$ reconstructed jets, $N_{\mathrm{reco}}^j$, to also be within the particle-level acceptance defined in Sect. 7:

$$f_{\mathrm{accept}}^j = \frac{N_{\mathrm{reco},\mathrm{part}}^j}{N_{\mathrm{reco}}^j}.$$  

(3)

The variable $N_{\mathrm{data}}^j$ is the number of events in data with $j$ reconstructed jets and $N_{\mathrm{bg}}^j$ is the number of background events, as evaluated in Sect. 5. The resulting $f_{\mathrm{accept}}$ decreases from around 0.85 for events without additional jets to about 0.76 for the highest jet multiplicities. The MC predictions of $f_{\mathrm{accept}}$ agree within 1% for events without any additional jets and within 5% at high jet multiplicities. Only MG5_aMC@NLO+Herwig++ predicts a smaller change as a function of the number of jets.

The response matrix $M_{\mathrm{reco},i}^{\mathrm{part},j}$ represents the probability $P(N_{\mathrm{reco}}^i|N_{\mathrm{part}}^j)$ of finding an event with true particle-level jet multiplicity $i$ with a reconstructed jet multiplicity $j$. As shown in Fig. 3, at the higher jet $p_T$ thresholds, at least 77% of the events have the same jet multiplicity at particle level and at reconstruction level. At the 25 GeV threshold, the agreement still exceeds 64%. The worse agreement can be explained in part by the presence of pile-up jets, which leads to events with more reconstructed than particle-level jets. There are almost no events with a difference of more than one jet between particle and reconstruction-level multiplicity.

As part of the Bayesian unfolding using Eq. (1), $M_{\mathrm{reco},i}^{\mathrm{part},j}$ is calculated iteratively, i.e., the result of the first iteration is used as the reconstruction-level jet multiplicity for the following one. The corrected spectra are found to converge after four iterations of the Bayesian unfolding algorithm.

The unfolded additional-jet multiplicity distributions are normalised after the last iteration according to

$$\frac{1}{\sigma} \frac{\mathrm{d} \sigma}{\mathrm{d} N^i} = \frac{N_{\mathrm{unfold}}^i}{\sum_i N_{\mathrm{unfold}}^i},$$

(4)

where $N_{\mathrm{unfold}}^i$, as defined in Eq. (1), corresponds to the number of events with $i$ jets after full unfolding and $\sigma$ is the measured $t\bar{t}$ production cross section in the fiducial volume.

A potential bias of the unfolded results due to data statistics and the unfolding procedure is investigated using pseudo-experiments by performing Gaussian sampling of the reconstruction-level distributions with statistical power equivalent to that present in data. The size of the bias, defined as the relative difference between the unfolded and predicted particle-level distributions, is found to be within the statistical uncertainty of the data. To check the size of a potential bias of the unfolding due to the relation between reconstructed and particle level distributions, the particle-level distributions are reweighted to alternative MC samples. Pseudo-experiments are performed based on the resulting alternative spectrum at reconstruction level. The pseudo-experiments are unfolded using the original correction procedure. The relative difference between the unfolded particle-level distribution and the predicted particle-level distribution from the alternative MC sample is found to be well within the modelling uncertainty. In addition, it is ensured that differences between the nominal and alternative particle-level distributions are at least as large as the difference between data and the predicted reconstruction-level distributions.
The effect of the uncertainties listed in Sect. 6 on the unfolded multiplicity and jet spectra is evaluated as follows. The uncertainties due to detector-related effects, such as JES, JER and $b$-tagging and data statistics, are propagated through the unfolding by varying the reconstructed objects for each uncertainty component by $\pm 1 \sigma$. The modified spectrum is then used as $N_{\text{data}}^i$ in Eq. (1) for the iterative unfolding and the difference on the particle-level distribution is taken as the systematic uncertainty.

The uncertainties due to the MC modelling of the QCD initial- and final-state radiation (ISR/FSR) and the parton-shower uncertainty are evaluated by replacing the data with the corresponding alternative MC sample and using the response matrix and the correction factors from the baseline $t\bar{t}$ MC sample for unfolding. The result is compared to the particle-level distribution of the alternative MC sample and the difference is taken as a systematic uncertainty. The uncertainties due to the MC modelling of the NLO matrix element and the matching algorithm are estimated in a similar way by replacing the data with the MG5_aMC@NLO+Herwig++ sample but using the response matrix and correction factors from Powheg+Herwig++. The resulting uncertainties are symmetrised for each component.

To unfold the leading and sub-leading $b$-jet $p_T$ and the leading additional-jet $p_T$, the same ansatz is used as for the jet multiplicity measurement, but with the jet $p_T$ instead of the jet multiplicity in the matrix, the acceptance and the efficiency formula. The binning is chosen to limit the migration, such that most events have reconstruction-level jet $p_T$ in the same region as the particle-level jet $p_T$, and to limit the uncertainty due to data statistics. The efficiency correction $f_{\text{eff}}$ for the $b$-jets has a significant $p_T$ dependence: it is around 0.2 for the lowest $p_T$ bin and reaches approximately 0.35 at $p_T$ of 80 GeV. The efficiency for the additional jet varies only slightly between 0.28 and 0.31. The acceptance correction is between 0.8 and 0.9 for all jets and almost independent of $p_T$, except at very low $p_T$, at which it decreases significantly, to 0.56 for the leading additional jet. The unfolding response matrix presented in Fig. 3 shows that more than 60% of the jets are in the same $p_T$ bin at particle and reconstruction level.

The spectra are normalised after the last iteration similarly to those in the jet multiplicity measurement:

$$\frac{1}{\sigma} \frac{d\sigma}{dp_T^i} = \frac{N_{\text{PT,unfold}}^i}{\Delta p_T^i \sum_i N_{\text{PT,unfold}}^i},$$

where $N_{\text{PT,unfold}}^i$, as defined in Eq. (1), corresponds to the number of events with the jet $p_T$ in bin $i$ after full unfolding.

The measurement of the jet $p_T$ spectra is as stable as the jet multiplicity measurements and the biases are small.

### 8.1 Jet multiplicity results

The unfolded normalised cross-sections are shown in Fig. 4 and are compared to different MC predictions. Events with up to three additional jets with $p_T$ above 25 GeV are measured exclusively (four jets inclusively) and up to two additional
Fig. 4 Unfolded jet multiplicity distribution for different $p_T$ thresholds of the additional jets, for a additional jet $p_T > 25$ GeV, b additional jet $p_T > 40$ GeV, and c additional jet $p_T > 60$ GeV, and d additional jet $p_T > 80$ GeV. Comparison to different MC predictions is shown for these distribution in first panel. The middle and bottom panels show the ratios of different MC predictions of the normalised cross-section to the measurement and the ratios of Powheg+Pythia6 predictions with variation of the QCD radiation to the measurement, respectively. The shaded regions show the statistical uncertainty (dark grey) and total uncertainty (light grey).

jets exclusively (three inclusively) for the higher $p_T$ thresholds. Tables 2, 3, 4 and 5 list the detailed composition of the uncertainties for 25 to 80 GeV. The jet multiplicity distributions are measured with an uncertainty of 4–5% for one additional jet, about 10% for two additional jets, and around 20% for the highest jet multiplicity bin, except for the 80 GeV threshold where the statistical uncertainty is larger for higher jet multiplicity bins. Systematic uncertainties dominate in all the measurements. In almost all bins for all $p_T$ thresholds, the JES uncertainty dominates, followed by the modelling uncertainty.

The data are compared to Powheg and MG5_aMC@NLO matched with different shower generators, namely Pythia8, Herwig++, and Herwig7 and to Sherpa, as shown in Figs. 4 and 5. Most predictions are within uncertainties and only slight deviations are visible except for Powheg+Herwig7, which deviates significantly from the data for all $p_T$ thresholds. The MG5_aMC@NLO predictions agree within 5–10% regardless of which parton shower is used (except Herwig7), and the Powheg predictions vary slightly more. The variations are larger when using different matrix elements but the same parton shower.

The unfolded data are compared with different MC predictions using $\chi^2$ tests. Full covariance matrices are produced from the unfolding taking into account statistical and all systematic uncertainties. The correlation of the measure-
ment bins is similar for all jet $p_T$ thresholds: strong anti-correlations exist between events with no additional jet and events with any number of additional jets. Positive correlations exist between the bins with one and two additional jets. The $\chi^2$ is determined using:

$$\chi^2 = S_{n-1}^T \text{Cov}^{-1}_{n-1} S_{n-1}$$  \hspace{1cm} (6)

where $S_{n-1}$ is a column vector representing the difference between the unfolded data and the MC generator predictions of the normalised cross-section for one less than the total number of bins in the distribution, and Cov$_{n-1}$ is a matrix with $n-1$ rows and the respective $n-1$ columns of the full covariance matrix. The full covariance matrix is singular and non-invertible, as it is evaluated using normalised distributions. The $p$-values are determined using the $\chi^2$ and $n-1$ degrees of freedom. Table 6 shows the $\chi^2$ and $p$-values.

<table>
<thead>
<tr>
<th>Sources</th>
<th>Relative uncertainty in [%] in additional jets multiplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Data statistics</td>
<td>2.1</td>
</tr>
<tr>
<td>JES/JER</td>
<td>5.0</td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>0.5</td>
</tr>
<tr>
<td>ISR/FSR modelling</td>
<td>0.4</td>
</tr>
<tr>
<td>Signal modelling</td>
<td>1.9</td>
</tr>
<tr>
<td>Other</td>
<td>1.4</td>
</tr>
<tr>
<td>Total</td>
<td>6.0</td>
</tr>
</tbody>
</table>

A statistical comparison taking into account the bin correlations indicates that the agreement with data is slightly better for MG5_aMC@NLO+Herwig++, as shown in Table 6. The ratio of the data to predictions of Powheg+Pythia6 with different levels of QCD radiation both in the matrix-element calculation and in the parton shower is also shown. Powheg+Pythia6 (RadLo) does not describe the data well. The central prediction of Powheg+Pythia6 (RadHi) describes the data most consistently, which is also confirmed by high $p$-values for all $p_T$ thresholds. The Powheg+Pythia6 (RadLo) sample has $p$-values around 0.5 and the central sample mostly between 0.8 and 0.9.

<table>
<thead>
<tr>
<th>Sources</th>
<th>Relative uncertainty in [%] in additional jets multiplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Data statistics</td>
<td>1.4</td>
</tr>
<tr>
<td>JES/JER</td>
<td>0.4</td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>0.1</td>
</tr>
<tr>
<td>ISR/FSR modelling</td>
<td>0.1</td>
</tr>
<tr>
<td>Signal modelling</td>
<td>0.2</td>
</tr>
<tr>
<td>Other</td>
<td>0.8</td>
</tr>
<tr>
<td>Total</td>
<td>1.7</td>
</tr>
</tbody>
</table>

8.2 Jet $p_T$ spectra results

The particle-level normalised cross-sections differential in jet $p_T$ are shown in Fig. 6 and are compared to different MC
Fig. 5 Ratios of jet multiplicity distribution for different $p_T$ thresholds of the additional jets predicted by various MC generators to the unfolded data, for a additional jet $p_T > 25$ GeV, b additional jet $p_T > 60$ GeV. The shaded regions show the statistical uncertainty (dark grey) and total uncertainty (light grey).

Fig. 6 Unfolded jet $p_T$ distribution for a leading $b$-jet, b sub-leading $b$-jet and c leading additional jet. Comparison to different MC predictions is shown for these distribution in first panel. The middle and bottom panels show the ratios of different MC predictions of the normalised cross-section to the measurement and the ratios of Powheg+Pythia6 predictions with variation of the QCD radiation to the measurement, respectively. The shaded regions show the statistical uncertainty (dark grey) and total uncertainty (light grey).
Table 6 Values of $\chi^2$/NDF and $p$-values between the unfolded normalised cross-section and the predictions for additional-jet multiplicity measurements. The number of degrees of freedom is equal to the number of bins minus one.

<table>
<thead>
<tr>
<th>Generator</th>
<th>$p_T &gt; 25$ GeV</th>
<th></th>
<th>$p_T &gt; 40$ GeV</th>
<th></th>
<th>$p_T &gt; 60$ GeV</th>
<th></th>
<th>$p_T &gt; 80$ GeV</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$/NDF</td>
<td>$p$-value</td>
<td>$\chi^2$/NDF</td>
<td>$p$-value</td>
<td>$\chi^2$/NDF</td>
<td>$p$-value</td>
<td>$\chi^2$/NDF</td>
<td>$p$-value</td>
</tr>
<tr>
<td>Powheg+Pythia6</td>
<td>0.82/4</td>
<td>0.94</td>
<td>0.83/3</td>
<td>0.84</td>
<td>1.01/3</td>
<td>0.80</td>
<td>1.82/3</td>
<td>0.61</td>
</tr>
<tr>
<td>Powheg+Pythia8</td>
<td>0.43/4</td>
<td>0.98</td>
<td>0.90/3</td>
<td>0.83</td>
<td>0.64/3</td>
<td>0.89</td>
<td>1.09/3</td>
<td>0.78</td>
</tr>
<tr>
<td>Powheg+Herwig++</td>
<td>0.51/4</td>
<td>0.97</td>
<td>0.88/3</td>
<td>0.83</td>
<td>1.46/3</td>
<td>0.69</td>
<td>2.58/3</td>
<td>0.46</td>
</tr>
<tr>
<td>Powheg+Herwig7</td>
<td>8.62/4</td>
<td>0.07</td>
<td>4.87/3</td>
<td>0.18</td>
<td>3.17/3</td>
<td>0.37</td>
<td>2.57/3</td>
<td>0.46</td>
</tr>
<tr>
<td>MG5_aMC@NLO+Pythia8</td>
<td>5.51/4</td>
<td>0.24</td>
<td>3.10/3</td>
<td>0.38</td>
<td>2.25/3</td>
<td>0.52</td>
<td>2.20/3</td>
<td>0.53</td>
</tr>
<tr>
<td>MG5_aMC@NLO+Herwig++</td>
<td>1.28/4</td>
<td>0.86</td>
<td>0.49/3</td>
<td>0.92</td>
<td>0.34/3</td>
<td>0.95</td>
<td>0.40/3</td>
<td>0.94</td>
</tr>
<tr>
<td>Sherpa v2.2</td>
<td>0.43/4</td>
<td>0.98</td>
<td>0.85/3</td>
<td>0.84</td>
<td>0.74/3</td>
<td>0.86</td>
<td>0.79/3</td>
<td>0.85</td>
</tr>
<tr>
<td>Powheg+Pythia6 (RadHi)</td>
<td>1.20/4</td>
<td>0.88</td>
<td>1.06/3</td>
<td>0.79</td>
<td>0.22/3</td>
<td>0.97</td>
<td>0.22/3</td>
<td>0.97</td>
</tr>
<tr>
<td>Powheg+Pythia6 (RadLo)</td>
<td>4.15/4</td>
<td>0.39</td>
<td>2.05/3</td>
<td>0.56</td>
<td>2.08/3</td>
<td>0.56</td>
<td>2.87/3</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Table 7 Summary of relative measurement uncertainties in [%] for the leading $b$-jet $p_T$ distribution. “Signal modelling” sources of systematic uncertainty includes the hadronisation, parton shower and NLO modelling uncertainties. “Other” sources of systematic uncertainty refers to lepton and jet selection efficiencies, background (including pile-up jets) estimations, and the PDF.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Data statistics</td>
<td>11.0</td>
<td>5.0</td>
<td>4.3</td>
<td>4.2</td>
<td>4.4</td>
<td>5.0</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>JES/JER</td>
<td>11.0</td>
<td>2.3</td>
<td>1.3</td>
<td>2.4</td>
<td>3.2</td>
<td>4.2</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>6.0</td>
<td>1.1</td>
<td>0.9</td>
<td>1.0</td>
<td>1.9</td>
<td>5.0</td>
<td>14.0</td>
<td></td>
</tr>
<tr>
<td>ISR/FSR modelling</td>
<td>6.0</td>
<td>0.9</td>
<td>1.0</td>
<td>2.1</td>
<td>3.1</td>
<td>0.9</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Signal modelling</td>
<td>9.0</td>
<td>2.0</td>
<td>5.0</td>
<td>6.0</td>
<td>2.1</td>
<td>0.4</td>
<td>15.0</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>4.4</td>
<td>3.0</td>
<td>1.4</td>
<td>1.7</td>
<td>3.0</td>
<td>2.2</td>
<td>10.0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>20.0</td>
<td>7.0</td>
<td>7.0</td>
<td>8.0</td>
<td>8.0</td>
<td>8.0</td>
<td>26.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 8 Summary of relative measurement uncertainties in [%] for the sub-leading $b$-jet $p_T$ distribution. “Signal modelling” sources of systematic uncertainty includes the hadronisation, parton shower and NLO modelling uncertainties. “Other” sources of systematic uncertainty refers to lepton and jet selection efficiencies, background (including pile-up jets) estimations, and the PDF.

<table>
<thead>
<tr>
<th>Sources</th>
<th>Relative uncertainty in sub-leading $b$-jet $p_T$ [GeV] in [%]</th>
<th>25–40</th>
<th>40–55</th>
<th>55–75</th>
<th>75–100</th>
<th>100–150</th>
<th>&gt; 150</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data statistics</td>
<td>4.0</td>
<td>4.2</td>
<td>3.9</td>
<td>6.0</td>
<td>7.0</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td>JES/JER</td>
<td>5.0</td>
<td>2.5</td>
<td>3.4</td>
<td>3.8</td>
<td>3.6</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>2.8</td>
<td>1.2</td>
<td>2.2</td>
<td>2.3</td>
<td>3.6</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td>ISR/FSR modelling</td>
<td>0.3</td>
<td>2.7</td>
<td>1.2</td>
<td>1.3</td>
<td>3.2</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Signal modelling</td>
<td>6.0</td>
<td>1.9</td>
<td>6.0</td>
<td>8.0</td>
<td>6.0</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1.4</td>
<td>1.8</td>
<td>1.9</td>
<td>3.4</td>
<td>3.1</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9.0</td>
<td>6.0</td>
<td>9.0</td>
<td>11.0</td>
<td>11.0</td>
<td>18.0</td>
<td></td>
</tr>
</tbody>
</table>

predictions. The total uncertainty in the $p_T$ measurements is 5–11%, although higher at some edges of the phase space. The uncertainty is dominated by the statistical uncertainty in almost all bins. The systematic uncertainties are listed in Tables 7, 8 and 9. JES/JER, NLO generator modelling and PS/hadronisation are all significant and one of them is always the dominant source of systematic uncertainty. JES/JER is the main source of uncertainty in the lowest $p_T$ bins of all measurements.

The predictions agree with data for all jet $p_T$ distributions as shown in Figs. 6 and 7, although the predictions of Powheg+Herwig++ and MG5_aMC@NLO+Pythia8 do not give a good description of the leading additional-jet $p_T$ distribution, which is consistent with the jet multiplicity-
Table 9  Summary of relative measurement uncertainties in [%] for the leading additional jet $p_T$ distribution. “Signal modelling” sources of systematic uncertainty includes the hadronisation, parton shower and NLO modelling uncertainties. “Other” sources of systematic uncertainty refers to lepton and jet selection efficiencies, background (including pile-up jets) estimations, and the PDF

<table>
<thead>
<tr>
<th>Sources</th>
<th>Relative uncertainty in leading additional jet $p_T$ [GeV] in [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data statistics</td>
<td>3.8</td>
</tr>
<tr>
<td>JES/JER</td>
<td>2.9</td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>0.3</td>
</tr>
<tr>
<td>ISR/FSR modelling</td>
<td>0.6</td>
</tr>
<tr>
<td>Signal modelling</td>
<td>2.5</td>
</tr>
<tr>
<td>Other</td>
<td>1.5</td>
</tr>
<tr>
<td>Total</td>
<td>6.0</td>
</tr>
</tbody>
</table>

Fig. 7  Ratios of jet $p_T$ distribution for a leading $b$-jet, b sub-leading $b$-jet and c leading additional jet predicted by various MC generators to the unfolded data. The shaded regions show the statistical uncertainty (dark grey) and total uncertainty (light grey)
Table 10 Values of $\chi^2$/NDF and $p$-values between the unfolded normalised cross-section and the predictions for the jet $p_T$ measurements. The number of degrees of freedom is equal to one less than the number of bins in the distribution.

<table>
<thead>
<tr>
<th>Generator</th>
<th>Leading b-jet $p_T$</th>
<th>Sub-leading b-jet $p_T$</th>
<th>Leading additional jet $p_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$/NDF $p$-value</td>
<td>$\chi^2$/NDF $p$-value</td>
<td>$\chi^2$/NDF $p$-value</td>
</tr>
<tr>
<td>Powheg+Pythia6</td>
<td>2.24/6 0.90</td>
<td>5.85/5 0.32</td>
<td>3.50/6 0.74</td>
</tr>
<tr>
<td>Powheg+Pythia8</td>
<td>1.94/6 0.93</td>
<td>6.33/5 0.28</td>
<td>2.28/6 0.89</td>
</tr>
<tr>
<td>Powheg+Herwig++</td>
<td>1.95/6 0.92</td>
<td>6.91/5 0.23</td>
<td>18.5/6 0.01</td>
</tr>
<tr>
<td>Powheg+Herwig7</td>
<td>1.26/6 0.97</td>
<td>5.44/5 0.36</td>
<td>1.95/6 0.92</td>
</tr>
<tr>
<td>MG5_aMC@NLO+Pythia8</td>
<td>1.99/6 0.92</td>
<td>6.76/5 0.24</td>
<td>10.5/6 0.10</td>
</tr>
<tr>
<td>MG5_aMC@NLO+Herwig++</td>
<td>2.03/6 0.97</td>
<td>6.94/5 0.23</td>
<td>2.97/6 0.81</td>
</tr>
<tr>
<td>Sherpa2.2</td>
<td>0.71/6 0.99</td>
<td>5.37/5 0.37</td>
<td>4.03/6 0.67</td>
</tr>
<tr>
<td>Powheg+Pythia6 (RadHi)</td>
<td>2.79/6 0.83</td>
<td>6.55/5 0.26</td>
<td>1.68/6 0.95</td>
</tr>
<tr>
<td>Powheg+Pythia6 (RadLo)</td>
<td>2.16/6 0.90</td>
<td>5.55/5 0.35</td>
<td>3.27/6 0.77</td>
</tr>
</tbody>
</table>

Fig. 8 Envelope of fractional uncertainties $\Delta f/f$ in the gap fraction $f_{gap}(Q_0)$, centred around unity, for a $|y| < 0.8$ and b $|y| < 2.1$. The statistical uncertainty is shown by the shaded area, and the total uncertainty by the solid black line. The systematic uncertainty is shown broken down into several groups, each of which includes various individual components.

Table 11 Sources of uncertainty in the gap fraction measurement as a function of $Q_0$ for the full central region $|y| < 2.1$, for a selection of $Q_0$ thresholds. “Signal modelling” sources of systematic uncertainty includes the hadronisation, parton shower and NLO modelling uncertainties. “Other” sources of systematic uncertainty refer to lepton and jet selection efficiencies, background (including pile-up jets) estimations, and the PDF.

<table>
<thead>
<tr>
<th>Sources</th>
<th>Uncertainty in $f_{gap}(Q_0)$ in [%] jet $p_T$ threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25 GeV</td>
</tr>
<tr>
<td>Data statistics</td>
<td>1.2</td>
</tr>
<tr>
<td>JES/JER</td>
<td>4.1</td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>0.5</td>
</tr>
<tr>
<td>ISR/FSR modelling</td>
<td>0.4</td>
</tr>
<tr>
<td>Signal modelling</td>
<td>4.5</td>
</tr>
<tr>
<td>Other</td>
<td>1.1</td>
</tr>
<tr>
<td>Total</td>
<td>6.3</td>
</tr>
</tbody>
</table>
Fig. 9 The measured gap fraction \( f_{\text{part}}(Q_0) \) as a function of \( Q_0 \) in different rapidity veto regions, a \( |y| < 0.8 \), b \( 0.8 < |y| < 1.5 \), c \( 1.5 < |y| < 2.1 \) and d \( |y| < 2.1 \). The data are shown by the points with error bars indicating the total uncertainty, and compared to the predictions from various \( t\bar{t} \) simulation samples shown as smooth curves.

The lower plots show the ratio of predictions to data, with the data uncertainty indicated by the shaded band, and the \( Q_0 \) thresholds corresponding to the left edges of the histogram bins, except for the first bin.
9 Gap fraction measurements

The jet activity is also studied by measuring the gap fraction \( f_{\text{gap}}(Q_0) \) defined as the fraction of events with no jet activity in addition to the two \( b \)-tagged jets above a given \( p_T \) threshold in a “veto region” defined as a rapidity region in the detector. The transverse momentum threshold is defined in two ways, and the gap fraction in two ways accordingly. First, the gap fraction is measured as the fraction of events without any additional jet in that rapidity region above a given \( p_T \) threshold \( Q_0 \): 

\[
  f_{\text{gap}}(Q_0) = \frac{n(Q_0)}{N_{\text{ft}}}, \quad (7)
\]

where \( N_{\text{ft}} \) is the total number of selected events, \( Q_0 \) is the \( p_T \) threshold for any additional jet in the veto region of these events, and \( n(Q_0) \) represents the subset of events with no additional jet with \( p_T > Q_0 \).

The second type of gap fraction is defined as the fraction of events in which the scalar \( p_T \) sum of all additional jets in the given veto region does not exceed a given threshold \( Q_{\text{sum}} \):

\[
  f_{\text{gap}}(Q_{\text{sum}}) = \frac{n(Q_{\text{sum}})}{N_{\text{ft}}}. \quad (8)
\]

Here, \( n(Q_{\text{sum}}) \) represents the subset of events in which the scalar \( p_T \) sum of all additional jets in the veto region is less than \( Q_{\text{sum}} \). The gap fraction defined using \( Q_0 \) is mainly sensitive to the leading \( p_T \) emission accompanying the \( t\bar{t} \) system, whereas the gap fraction defined using \( Q_{\text{sum}} \) is sensitive to all hard emissions accompanying the \( t\bar{t} \) system. In the following descriptions of the gap fraction measurement process, the same procedure is followed for \( Q_{\text{sum}} \) as for \( Q_0 \).

Both types of gap fraction are measured in four veto regions: \(|y| < 0.8, 0.8 < |y| < 1.5, 1.5 < |y| < 2.1\) and the full central region \(|y| < 2.1\). For each of these regions, the gap fraction is measured by counting the number of selected data events in the subset of events with no jet activity in the full central region (that had no additional jets with \( p_T > Q_0 \) within the veto region, where the sets of \( Q_0 \) and \( Q_{\text{sum}} \) threshold values correspond approximately to one standard deviation of the jet energy resolution and are the same as in the earlier publications [3,4]. The number samples are shown as ratios to data, with the data uncertainty indicated by the shaded band, and the \( Q_0 \) thresholds corresponding to the left edges of the histogram bins, except for the first bin.

Fig. 10 Ratios of prediction to data of the measured gap fraction \( f_{\text{gap}}(Q_0) \) as a function of \( Q_0 \) in different rapidity veto regions, a \(|y| < 0.8\) and b \(|y| < 2.1\). The predictions from various \( t\bar{t} \) simulation
Fig. 11 The measured gap fraction $f_{\text{part}}(Q_{\text{sum}})$ as a function of $Q_{\text{sum}}$ in different rapidity veto regions, \(a\) $|y| < 0.8$ and \(b\) $|y| < 2.1$, followed by ratios of prediction to data of the measured gap fraction $f_{\text{part}}(Q_{\text{sum}})$ as a function of $Q_{\text{sum}}$ in the same two rapidity regions. The data in \(a\) and \(b\) are shown by the points with error bars indicating the total uncertainty, and compared to the predictions from various $t\bar{t}$ simulation samples shown as smooth curves. The lower plots in \(a\) and \(b\) show the ratio of predictions to data, with the data uncertainty indicated by the shaded band, and the $Q_{\text{sum}}$ thresholds corresponding to the left edges of the histogram bins, except for the first bin.
of background events, \( N_{bg} \) and \( n_{bg}(Q_0) \), are then subtracted from these events:

\[
f_{\text{data}}^{\text{part}}(Q_0) = \frac{N_{\text{data}}(Q_0) - n_{bg}(Q_0)}{N_{\text{data}} - N_{bg}}
\]

and similarly for \( f_{\text{gap}}^{\text{part}}(Q_{\text{sum}}) \). The measured gap fraction \( f_{\text{data}}^{\text{part}}(Q_0) \) is then corrected for detector effects to particle level by multiplying it by a correction factor \( C(Q_0) \) to obtain \( f_{\text{gap}}^{\text{part}}(Q_0) \). The correction factor \( C(Q_0) \) is determined from the baseline Powheg+Pythia6 \( \bar{t}\bar{t} \) sample using the simulated gap fraction values at reconstruction level \( f_{\text{reco}}^{\text{part}}(Q_0) \), and at particle level \( f_{\text{gap}}^{\text{part}}(Q_0) \):

\[
C(Q_0) = \frac{f_{\text{part}}^{\text{part}}(Q_0)}{f_{\text{reco}}^{\text{reco}}(Q_0)}.
\]

The values of the correction factors \( C(Q_0) \) and \( C(Q_{\text{sum}}) \) deviate by less than 4% from unity at low \( Q_0 \) and \( Q_{\text{sum}} \) values in the rapidity regions (less than 8% in the \( m_{\text{subb}} \) subsamples), and approach unity at higher threshold values. The small corrections reflect the high selection efficiency and high purity of the event samples. At each threshold \( Q_0 \), the baseline simulation predicts that around 80% of the selected reconstructed events that do not have a jet with \( p_T > Q_0 \) also have no particle-level jet with \( p_T > Q_0 \). Therefore, a simple bin-by-bin correction method is considered adequate, rather than a full unfolding as used in Sect. 8.

Systematic uncertainties arise in this procedure from the uncertainties in \( C(Q_0) \) and the subtracted backgrounds. The uncertainties, as described in Sect. 6, are used to recalculate \( f_{\text{data}}^{\text{part}}(Q_0) \) and \( C(Q_0) \) to obtain the gap fraction \( f_{\text{gap}}^{\text{part}}(Q_0) \). The corresponding quantities for \( Q_{\text{sum}} \) are calculated accordingly. Figure 8 and Table 11 list the resulting relative uncertainty in \( f_{\text{gap}}^{\text{part}}(Q_0) \), \( \Delta f/f \), for the different sources of uncertainty in the full central rapidity region.

### 9.1 Gap fraction results in rapidity regions

Figure 9 shows the measured gap fractions \( f_{\text{gap}}^{\text{part}}(Q_0) \) in data, corrected to the particle level. The gap fraction \( f_{\text{gap}}^{\text{part}}(Q_0) \) is compared to various MC generator predictions in Fig. 10, and Fig. 11 shows the measured gap fractions \( f_{\text{gap}}^{\text{part}}(Q_{\text{sum}}) \) compared to various MC generators, corrected to the particle level. The predictions of Sherpa and MG5_aMC@NLO+Herwig++ agree well with each other and are within the uncertainties of the data, while Powheg+Pythia8 has slightly higher gap fractions, i.e., predicts too little radiation. Similarly to the jet multiplicity measurements, Powheg+Pythia6 (RadHi) agrees well with data, while the nominal and the Powheg+Pythia6 (RadLo) samples give similar but too high predictions compared to data. The results in Fig. 9d

### Table 12 Values of \( \chi^2 \) for the comparison of the measured gap fraction distributions with the predictions from various \( t\bar{t} \) generator configurations, for the four rapidity regions as a function of \( Q_0 \). The \( \chi^2 \) and \( p \)-values correspond to 18 degrees of freedom.

| \( Q_0 \) | \( |y| < 0.8 \) | \( 0.8 < |y| < 1.5 \) | \( 1.5 < |y| < 2.1 \) | \( |y| < 2.1 \) |
|---|---|---|---|---|
| Generator | \( \chi^2 \) | \( p \)-value | \( \chi^2 \) | \( p \)-value | \( \chi^2 \) | \( p \)-value | \( \chi^2 \) | \( p \)-value |
| Powheg+Pythia6 | 18.5 | 0.42 | 8.0 | 0.98 | 14.8 | 0.67 | 17.4 | 0.50 |
| Powheg+Pythia8 | 13.3 | 0.77 | 8.7 | 0.97 | 11.8 | 0.86 | 15.0 | 0.66 |
| Powheg+Herwig++ | 24.4 | 0.14 | 10.2 | 0.93 | 19.6 | 0.36 | 30.8 | 0.03 |
| Powheg+Herwig7 | 18.5 | 0.42 | 14.1 | 0.72 | 14.6 | 0.69 | 18.7 | 0.41 |
| MG5_aMC@NLO+Pythia8 | 33.7 | 0.01 | 26.2 | 0.09 | 18.0 | 0.45 | 58.6 | 3.4 \times 10^{-6} |
| MG5_aMC@NLO+Herwig++ | 14.1 | 0.72 | 8.5 | 0.97 | 18.9 | 0.40 | 8.4 | 0.97 |
| MG5_aMC@NLO+Herwig7 | 22.2 | 0.22 | 25.7 | 0.11 | 20.3 | 0.32 | 44.6 | 4.7 \times 10^{-4} |
| Sherpa v2.2 | 12.1 | 0.84 | 11.6 | 0.87 | 14.5 | 0.70 | 14.2 | 0.71 |
| Powheg+Pythia6 RadHi | 10.7 | 0.91 | 6.8 | 0.99 | 13.0 | 0.79 | 11.1 | 0.89 |
| Powheg+Pythia6 RadLo | 23.1 | 0.19 | 12.6 | 0.82 | 17.4 | 0.50 | 24.6 | 0.14 |
Table 13 Values of $\chi^2$ for the comparison of the measured gap fraction distributions with the predictions from various $t\bar{t}$ generator configurations, for the four rapidity regions as a function of $Q_{\text{sum}}$. The $\chi^2$ and $p$-values correspond to 22 degrees of freedom.

| Generator                          | $|y| < 0.8$ | $0.8 < |y| < 1.5$ | $1.5 < |y| < 2.1$ | $|y| < 2.1$ |
|-----------------------------------|------------|------------------|------------------|------------|
|                                   | $\chi^2$   | $p$-value        | $\chi^2$ | $p$-value | $\chi^2$ | $p$-value | $\chi^2$ | $p$-value |
| Powheg+Pythia6                   | 17.5       | 0.74             | 8.6     | 1.00      | 19.0      | 0.64      | 29.0     | 0.15      |
| Powheg+Pythia8                   | 12.4       | 0.95             | 9.7     | 0.99      | 17.7      | 0.72      | 30.8     | 0.10      |
| Powheg+Herwig++                  | 17.4       | 0.74             | 11.5    | 0.97      | 21.9      | 0.46      | 34.6     | 0.04      |
| Powheg+Herwig7                   | 15.3       | 0.85             | 14.0    | 0.90      | 16.4      | 0.79      | 32.8     | 0.06      |
| MG5_aMC@NLO+Pythia8              | 30.5       | 0.11             | 22.1    | 0.45      | 20.7      | 0.54      | 55.7     | $9.4 \times 10^{-5}$ |
| MG5_aMC@NLO+Herwig++             | 17.8       | 0.72             | 10.4    | 0.98      | 18.4      | 0.68      | 23.6     | 0.37      |
| MG5_aMC@NLO+Herwig7              | 21.2       | 0.51             | 27.3    | 0.20      | 24.4      | 0.32      | 54.7     | $1.3 \times 10^{-4}$ | |
| Sherpa v2.2                      | 6.6        | 1.00             | 9.5     | 0.99      | 14.4      | 0.89      | 19.1     | 0.64      |
| Powheg+Pythia6 RadHi             | 10.3       | 0.98             | 8.8     | 0.99      | 15.4      | 0.85      | 26.5     | 0.23      |
| Powheg+Pythia6 RadLo             | 23.0       | 0.40             | 12.6    | 0.94      | 21.6      | 0.49      | 40.7     | $8.9 \times 10^{-3}$ |

Fig. 13 Distribution of the reconstructed invariant mass of the $e\mu + 2$ $b$-jets system $m_{e\mu bb}$ in data, compared to simulation. The shaded band represents the statistical uncertainty in data. The lower plot shows the ratio of the distribution of invariant mass in simulation compared to data.

can directly be compared with the jet multiplicity results in Figs. 4 and 5 in the one additional jet bin. Here the Powheg+Pythia8 predictions are below data for all distributions which proves the consistency of the measurements. The $p_T$ distribution of the first additional jet shown in Fig. 6 contains only events with at least one additional jet and differs in this respect from the gap fraction distribution which includes events with no additional jet. However, the results are also consistent as Powheg+Pythia8 predicts a slightly softer $p_T$ spectrum for the additional jet which leads to the observed effect that less jets above the 25 GeV threshold are observed.

The matrix of statistical and systematic correlations is shown in Fig. 12 for the gap fraction measurement at different values of $Q_0$ for the full central $|y| < 2.1$ rapidity region. Nearby points in $Q_0$ are highly correlated, while well-separated $Q_0$ points are less correlated. The full covariance matrix, including correlations, is used to calculate a $\chi^2$ value for the compatibility of each of the NLO generator predictions with the data in each veto region. The results are given in Tables 12 and 13. An analysis of the $p$-values confirms that Powheg+Herwig++, MG5_aMC@NLO+Herwig7, MG5_aMC@NLO+Pythia8 and Powheg+Pythia6 (RadLo) are not consistent with the data. Powheg+Pythia6 (RadHi) has the best $p$-values among the QCD shower variations of Powheg+Pythia6.

9.2 Gap fraction results in $m_{e\mu bb}$ subsamples

The gap fraction is also measured over the full central veto region $|y| < 2.1$ after dividing the data sample into four regions of $m_{e\mu bb}$. The distribution of reconstructed $m_{e\mu bb}$ in the selected $e\mu + 2$ $b$-tagged jets events is reasonably well-reproduced by the nominal $t\bar{t}$ simulation sample, as shown in Fig. 13. The distribution is divided into four regions at both reconstruction and particle level: $m_{e\mu bb} < 300$ GeV, $300$ GeV $< m_{e\mu bb} < 425$ GeV, $425$ GeV $< m_{e\mu bb} < 600$ GeV and $m_{e\mu bb} > 600$ GeV. These boundaries are chosen to minimise migration between the regions. In the baseline simulation, around 85% of the reconstructed events in each $m_{e\mu bb}$ region belong to the corresponding region at particle level. The corresponding correction factors $C_m(Q_0)$ which translate the measured gap fraction in the reconstruction-level $m_{e\mu bb}$ region to the corresponding particle-level gap fractions $f_m(Q_0)$ are of similar
Fig. 14 The measured gap fraction $f_m(Q_0)$ as a function of $Q_0$ in the full central veto region $|y| < 2.1$ for the invariant mass regions (a) $m_{\mu\mu} < 300 \text{ GeV}$, (b) $300 \text{ GeV} < m_{\mu\mu} < 425 \text{ GeV}$, (c) $425 \text{ GeV} < m_{\mu\mu} < 600 \text{ GeV}$ and (d) $m_{\mu\mu} > 600 \text{ GeV}$. The data are shown by the points with error bars indicating the total uncertainty, and compared to the predictions from various $t\bar{t}$ simulation samples shown as smooth curves. The lower plots show the ratio of predictions to data, with the data uncertainty indicated by the shaded band, and the $Q_0$ thresholds corresponding to the left edges of the histogram bins, except for the first bin.
Fig. 15 Ratios of prediction to data of the measured gap fraction $f^\text{part}_{\text{gap}}(Q_0)$ as a function of $Q_0$ in the full central veto region $|y| < 2.1$ for the invariant mass regions (a) $m_{\mu\ell\mu\ell} < 300$ GeV and (b) $425 < m_{\mu\ell\mu\ell} < 600$ GeV. The predictions from various $t\bar{t}$ simulation samples are shown as ratios to data, with the data uncertainty indicated by the shaded band, and the $Q_0$ thresholds corresponding to the left edges of the histogram bins, except for the first bin.

Table 14 Measurements of $\chi^2$ comparing the measured gap fraction distributions with predictions from various $t\bar{t}$ generator configurations, for the four invariant mass $m_{\mu\ell\mu\ell}$ regions as a function of $Q_0$. The $\chi^2$ and $p$-values correspond to 18 degrees of freedom.

<table>
<thead>
<tr>
<th>$Q_0$</th>
<th>$m &lt; 300$ GeV</th>
<th>$300 &lt; m &lt; 425$ GeV</th>
<th>$425 &lt; m &lt; 600$ GeV</th>
<th>$m &gt; 600$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>$p$-value</td>
<td>$\chi^2$</td>
<td>$p$-value</td>
</tr>
<tr>
<td>Powheg+Pythia6</td>
<td>5.1</td>
<td>1.00</td>
<td>21.1</td>
<td>0.28</td>
</tr>
<tr>
<td>Powheg+Pythia8</td>
<td>4.4</td>
<td>1.00</td>
<td>16.7</td>
<td>0.55</td>
</tr>
<tr>
<td>Powheg+Herwig++</td>
<td>14.6</td>
<td>0.69</td>
<td>19.8</td>
<td>0.35</td>
</tr>
<tr>
<td>Powheg+Herwig7</td>
<td>9.1</td>
<td>0.96</td>
<td>16.5</td>
<td>0.56</td>
</tr>
<tr>
<td>MG5_aMC@NLO+Pythia8</td>
<td>27.5</td>
<td>0.07</td>
<td>27.6</td>
<td>0.07</td>
</tr>
<tr>
<td>MG5_aMC@NLO+Herwig++</td>
<td>4.6</td>
<td>1.00</td>
<td>18.2</td>
<td>0.44</td>
</tr>
<tr>
<td>MG5_aMC@NLO+Herwig7</td>
<td>20.9</td>
<td>0.29</td>
<td>23.6</td>
<td>0.17</td>
</tr>
<tr>
<td>Sherpa v2.2</td>
<td>7.8</td>
<td>0.98</td>
<td>11.3</td>
<td>0.88</td>
</tr>
<tr>
<td>Powheg+Pythia6 RadHi</td>
<td>4.1</td>
<td>1.00</td>
<td>15.5</td>
<td>0.63</td>
</tr>
<tr>
<td>Powheg+Pythia6 RadLo</td>
<td>7.4</td>
<td>0.99</td>
<td>24.9</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Figures 14 and 15 show the measured gap fractions as a function of $Q_0$ in four $m_{\mu\ell\mu\ell}$ regions in data, compared to the same set of predictions as shown in Figs. 9 and 10. Tables 14 and 15 give the $\chi^2$ and $p$-values taking into account bin-by-bin correlations of the gap fractions compared to the predictions from the different generators. Figure 16 gives an alternative presentation of the gap fraction $f_m(Q_0)$ as a function of $m_{\mu\ell\mu\ell}$ for four different $Q_0$ values. The level of agreement between the data and the various predictions is consistent with the results of the gap fraction in rapidity bins. Only in the lowest mass region the Powheg+Pythia8 prediction agrees very well, while MG5_aMC@NLO+Herwig++ and Sherpa are at the lower edge of the uncertainties.
The different measurements are compared to various Monte Carlo predictions and give consistent results. Even though many predictions are within the uncertainty band of the measurements, the proper evaluation of the compatibility of the models, taking into account the bin-by-bin correlations within each measurement, revealed that Powheg+Pythia6 (RadHi), MG5_aMC@NLO+Herwig++ and Sherpa describe the data best for all observables. Powheg+Pythia6 (RadLo), MG5_aMC@NLO+Pythia8 and all predictions involving Herwig7 do not describe the data well.

All studied combinations of the matrix element generators MG5_aMC@NLO and Powheg with the shower generators Herwig++, Pythia6 and Pythia8 provided no systematic trend indicating that one of the matrix element generators describes the data better for all parton shower generators. We also have no indication that one of the parton shower generators describes the data systematically better for both matrix element generators. This observation suggests that the matching between the parton shower and matrix element calculation plays an important role, and motivates further study in this area. The predictions of Sherpa which use NLO matrix elements consistently matched with up to four additional jets at LO show similar good agreement with data as the best of the MG5_aMC@NLO and Powheg predictions.

**Acknowledgements** We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSRMC, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, 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; FOM and NWO, Netherlands; RCN, Norway; MINISW and NCN, Poland; FCT, Portugal;
MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSL, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [60].

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 SCOAP3.

References


4. ATLAS Collaboration, Measurement of jet energy in top quark events using the $e\mu$ final state with two $b$-tagged jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. JHEP 9, 74 (2016) arXiv:1606.09490 [hep-ex]


22. ATLAS Collaboration, Measurements of normalized differential cross sections for $t \bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. Phys. Rev. D 90, 072004 (2014)


(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) National Institute of Physics and Nuclear Engineering, Bucharest, Romania; (c) Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj-Napoca, Romania; (d) University Politehnica Bucharest, Bucharest, Romania; (e) West University in Timisoara, Timisoara, Romania

29 Departamento de Física, Universidad 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

36 (a) Department of Modern Physics, 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, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University (also affiliated with PKU-CHEP), Shanghai, China

37 Laboratoire de Physique Corpusculaire, Université Clermont Auvergne, Université Blaise Pascal, CNRS/IN2P3, Clermont-Ferrand, France

38 Nevis Laboratory, Columbia University, Irvington, NY, USA

39 Niels Bohr Institute, University of Copenhagen, Copenhagen, 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, Kraków, Poland; (b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Kraków, Poland

42 Institute of Nuclear Physics, Polish Academy of Sciences, Kraków, 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 Laboratori Nazionali di Frascati, Frascati, Italy

51 Fakultät für Mathematik und Physik, Albert-Ludwigs-Universität, Freiburg, Germany

52 Département de Physique Nucleaire 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

59 Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge, MA, USA

60 (a) Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; (b) Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; (c) ZITI Institut für technische Informatik, Ruprecht-Karls-Universität Heidelberg, Mannheim, Germany

61 Faculty of Applied Information Science, Hiroshima Institute of Technology, Hiroshima, Japan
62 (a) Department of Physics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong; (b) Department of Physics, The University of Hong Kong, Hong Kong, China; (c) Department of Physics and Institute for Advanced Study, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
63 Department of Physics, National Tsing Hua University, Hsinchu, Taiwan
64 Department of Physics, Indiana University, Bloomington, IN, USA
65 Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck, Austria
66 University of Iowa, Iowa City, IA, USA
67 Department of Physics and Astronomy, Iowa State University, Ames, IA, USA
68 Joint Institute for Nuclear Research, JINR Dubna, Dubna, Russia
69 KEK, High Energy Accelerator Research Organization, Tsukuba, Japan
70 Graduate School of Science, Kobe University, Kobe, Japan
71 Faculty of Science, Kyoto University, Kyoto, Japan
72 Kyoto University of Education, Kyoto, Japan
73 Department of Physics, Kyushu University, Fukuoka, Japan
74 Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina
75 Physics Department, Lancaster University, Lancaster, UK
76 (a) INFN Sezione di Lecce, Lecce, Italy; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy
77 Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK
78 Department of Physics, Jožef Stefan Institute and University of Ljubljana, Ljubljana, Slovenia
79 School of Physics and Astronomy, Queen Mary University of London, London, UK
80 Department of Physics, Royal Holloway University of London, Surrey, UK
81 Department of Physics and Astronomy, University College London, London, UK
82 Louisiana Tech University, Ruston, LA, USA
83 Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université Paris-Diderot and CNRS/IN2P3, Paris, France
84 Fysiska institutionen, Lunds universitet, Lund, Sweden
85 Departamento de Física Teorica C-15, Universidad Autonoma de Madrid, Madrid, Spain
86 Institut für Physik, Universität Mainz, Mainz, Germany
87 School of Physics and Astronomy, University of Manchester, Manchester, UK
88 CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
89 Department of Physics, University of Massachusetts, Amherst, MA, USA
90 Department of Physics, McGill University, Montreal, QC, Canada
91 School of Physics, University of Melbourne, Melbourne, Victoria, Australia
92 Department of Physics, The University of Michigan, Ann Arbor, MI, USA
93 Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA
94 (a) INFN Sezione di Milano, Milan, Italy; (b) Dipartimento di Fisica, Università di Milano, Milan, Italy
95 B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Republic of Belarus
96 Research Institute for Nuclear Problems of Byelorussian State University, Minsk, Republic of Belarus
97 Group of Particle Physics, University of Montreal, Montreal, QC, Canada
98 P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, Russia
99 Institute for Theoretical and Experimental Physics (ITEP), Moscow, Russia
100 National Research Nuclear University MEPhI, Moscow, Russia
101 D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
102 Fakultät für Physik, Ludwig-Maximilians-Universität München, Munich, Germany
103 Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Munich, Germany
104 Nagasaki Institute of Applied Science, Nagasaki, Japan
105 Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan
106 (a) INFN Sezione di Napoli, Naples, Italy; (b) Dipartimento di Fisica, Università di Napoli, Naples, Italy
107 Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, USA
108 Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands
109 Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, The Netherlands
110 Department of Physics, Northern Illinois University, DeKalb, IL, USA
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 Física Corpuscular (IFIC) and Departamento de Física Atomica, Molecular y Nuclear and Departamento de Ingeniería Electrónica and Instituto de Microelectrónica de Barcelona (IMB-CNM), University of Valencia and 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

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
d Also at TRIUMF, Vancouver, BC, Canada
e Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY, USA
f Also at Physics Department, An-Najah National University, Nablus, Palestine
g Also at Department of Physics, California State University, Fresno, CA, USA
h Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
i Also at Departament de Fisica de la Universitat Autonoma de Barcelona, Barcelona, Spain
j Also at Departamento de Física e Astronomia, Faculdade de Ciencias, Universidade do Porto, Porto, Portugal
k Also at Tomsk State University, Tomsk, Russia
l Also at Universita di Napoli Parthenope, Napoli, Italy
m Also at Institute of Particle Physics (IPP), Canada
n Also at National Institute of Physics and Nuclear Engineering, Bucharest, Romania
Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia

p Also at Department of Physics, The University of Michigan, Ann Arbor, MI, USA

q Also at Centre for High Performance Computing, CSIR Campus, Rosebank, Cape Town, South Africa

r Also at Louisiana Tech University, Ruston, LA, USA

s Also at Institut Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain

t Also at Graduate School of Science, Osaka University, Osaka, Japan

u Also at Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, Netherlands

v Also at Department of Physics, The University of Texas at Austin, Austin, TX, USA

w Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia

x Also at CERN, Geneva, Switzerland

y Also at Georgian Technical University (GTU), Tbilisi, Georgia

z Also at Ochadai Academic Production, Ochanomizu University, Tokyo, Japan

aa Also at Manhattan College, New York, NY, USA

ab Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan

ac Also at School of Physics, Shandong University, Shandong, China

ad Also at Department of Physics, California State University, Sacramento CA, USA

ae Also at Moscow Institute of Physics and Technology, State University, Dolgoprudny, Russia

af Also at Departement de Physique Nucleaire et Corpusculaire, Université de Genève, Geneva, Switzerland

ag Also at Eotvos Lorand University, Budapest, Hungary

ah Also at Departments of Physics and Astronomy and Chemistry, Stony Brook University, Stony Brook NY, USA

ai Also at International School for Advanced Studies (SISSA), Trieste, Italy

aj Also at Department of Physics and Astronomy, University of South Carolina, Columbia, SC, USA

ak Also at Institut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Barcelona, Spain

al Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China

am Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria

an Also at Faculty of Physics, M.V.Lomonosov Moscow State University, Moscow, Russia

ao Also at Institute of Physics, Academia Sinica, Taipei, Taiwan

ap Also at National Research Nuclear University MEPhI, Moscow, Russia

aq Also at Department of Physics, Stanford University, Stanford, CA, USA

ar Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary

as Also at Giresun University, Faculty of Engineering, Turkey

at Also at Flensburg University of Applied Sciences, Flensburg, Germany

au Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia

av Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France

Deceased