EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Submitted to: JHEP

CERN-EP-2020-018
2nd September 2022

Measurement of the top-quark mass using a leptonic invariant mass in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

A measurement of the top-quark mass (m_t) in the $t\bar{t} \rightarrow$ lepton + jets channel is presented, with an experimental technique which exploits semileptonic decays of b-hadrons produced in the top-quark decay chain. The distribution of the invariant mass $m_{\ell\mu}$ of the lepton, ℓ (with $\ell = e, \mu$), from the W-boson decay and the muon, μ, originating from the b-hadron decay is reconstructed, and a binned-template profile likelihood fit is performed to extract m_t. The measurement is based on data corresponding to an integrated luminosity of 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions provided by the Large Hadron Collider and recorded by the ATLAS detector. The measured value of the top-quark mass is $m_t = 174.41 \pm 0.39$ (stat.) ± 0.66 (syst.) ± 0.25 (recoil) GeV, where the third uncertainty arises from changing the Pythia8 parton shower gluon-recoil scheme, used in top-quark decays, to a recently developed setup.

Contents

1 Introduction

The large mass of the top-quark plays a role in much of the dynamics of elementary particles via loop diagrams. In the Standard Model (SM), the large top-quark mass significantly affects the radiative corrections to both the Higgs boson and W-boson masses, providing a relationship that can be used for precision tests of the consistency of the SM [1]. Furthermore, a precise measurement of the top-quark mass is required to predict the evolution of the Higgs quartic coupling at high scales [2, 3]. If performed with a precision of the order of a few hundred MeV, the direct determination of the top-quark mass from its decay products, and the indirect measurements from top-quark production cross-sections or kinematic distributions, are important not only for the constraints mentioned above, but also for the challenge of the interpretation of such measurements in the context of a strongly interacting particle theory [4, 5].

A direct measurement of the top-quark mass (m_t) is presented that uses a partial, leptonic-only, invariant mass reconstruction of the top-quark decay products. The analysis is performed from a sample of reconstructed $t\bar{t}$ events in the ℓ+jets channel, where one of the two W-bosons from the top- and antitop-quarks decays leptonically. In the top-quark decay $t \to Wb$, the invariant mass $m_{\ell\mu}$ of the lepton ℓ (with $\ell = e, \mu$) from the W-boson decay and the muon μ from a semileptonic decay of a b-hadron is constructed as the observable sensitive to the parent m_t value. The advantages of a strategy based on the invariant mass of visible leptonic decay products for the measurement of the top-quark mass rest mainly on the smaller sensitivity to the jet energy calibration and energy resolution, compared to the standard direct reconstruction methods, and on less sensitivity to top-quark production modelling (owing to the boost-invariant construction) than in methods based on the W-decay lepton alone [6]. Moreover, methods

with different types of systematic uncertainties are important when combining measurements, and for testing the consistency of the theoretical interpretation of the top-quark mass.

In this analysis the $m_{\ell\mu}$ distribution from models with different top-quark mass hypotheses is compared with data, and the optimal value of m_t is determined from a binned-template profile likelihood fit. A similar technique was first employed by the CDF Collaboration at the Tevatron collider [7], and a closely related analysis with J/ψ decays has been presented by the CMS Collaboration [8]; however, both these analyses yielded uncertainties in m_t of several GeV. Until now, the most precise measurement of the top-quark mass in the $t\bar{t} \to \ell$+jets channel by the ATLAS Collaboration was $m_t = 172.08 \pm 0.39$ (stat.) ± 0.82 (syst.) GeV, whereas combining multiple ATLAS measurements gave $m_t = 172.69 \pm 0.48$ GeV [9]. The CMS Collaboration reports its most precise combination as $m_t = 172.44 \pm 0.48$ GeV [10], and the Tevatron experiments report a combined value of $m_t = 174.30 \pm 0.65$ GeV [11]. Finally, the mass of the top-quark is indirectly determined from global electroweak fits as $m_t = 176.4 \pm 2.1$ GeV [12].

2 ATLAS experiment

The ATLAS experiment [13] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4π coverage in solid angle[1]. It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$ and consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. The innermost layer, known as the insertable B-layer [14, 15], was added in 2014 and provides high-resolution hits at small radius to improve the tracking performance. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A steel/scintillator-tile hadronic calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The endcap and forward regions are instrumented with LAr calorimeters for both the EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large air-core toroid superconducting magnets with eight coils each and a bending power of 2.0 to 7.5 Tm. It includes a system of precision tracking chambers covering the region $|\eta| < 2.7$ and fast detectors for triggering in the range $|\eta| < 2.4$. A two-level trigger system was used to select events [16]. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most 100 kHz. This is followed by the software-based high-level trigger, which reduces the event rate to around 1 kHz. An extensive software suite [17] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

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

3 Data and Simulation

3.1 Data sample and object definition

The analysis is performed with the 2015 and 2016 proton–proton collision data sample produced by the LHC at a centre-of-mass energy of $\sqrt{s} = 13$ TeV and collected by the ATLAS experiment, corresponding to an integrated luminosity of 36.1 fb^{-1} [18]. The data sample was recorded during stable beam conditions, and all relevant ATLAS detector subsystems were required to be operational. The average number of pp collisions in the same bunch crossing (referred to as pile-up) in the data sample is 24.1.

Electron candidates are reconstructed from energy deposits (clusters) in the electromagnetic calorimeter matched to reconstructed tracks in the inner detector. Candidates in the transition region $1.37 < |\eta_{\text{cluster}}| < 1.52$ between the calorimeter barrel and endcaps are excluded. Muon candidates are reconstructed from track segments in the layers of the muon spectrometer, and matched to tracks found in the inner detector. The final muon candidates are re-fitted using the complete track information from both detector systems. Jet candidates are reconstructed from three-dimensional topological EM-scale energy clusters [19] in the calorimeter using the anti-k_t jet algorithm [20, 21] with a radius parameter $R = 0.4$. The reconstructed jets are calibrated to the level of stable-particle jets by the application of a jet energy scale (JES) correction derived from simulation and *in situ* corrections based on 13 TeV data [22]. There is no dedicated energy scale correction for jets with semileptonic heavy-flavour hadron decays. The missing transverse momentum, $E_{\text{T}}^{\text{miss}}$, is defined as the magnitude of the negative vector sum of the transverse momentum, p_{T}, of all reconstructed and calibrated physics objects in the event, with an extra term added to account for soft energy in the event that is not associated with any of the reconstructed objects [23]. This soft term is calculated from inner-detector tracks matched to the primary vertex in order to make it more resilient to pile-up contamination.

3.2 Object and event selections

The event selection is designed to collect a sample of $t\bar{t}$ candidate events in the final state $\ell\nu b j j' \bar{b}$, where $\ell = e$ or μ and the jj' are the jets produced in the decay of the W-boson into quarks, and at least one of the b-initiated jets is associated with a muon from the semileptonic decay of a b-hadron. The goal is to select events where the lepton ℓ from the W-boson and the b-initiated jet with the muon from semileptonic decay come from the same top-quark.

Events are required to pass either a single-electron or single-muon trigger. Multiple trigger types were used: the lowest-threshold triggers include isolation requirements to reduce the trigger rate and had p_{T} thresholds of 20 GeV for muons and 24 GeV for electrons in 2015 data, and 26 GeV for both lepton types in 2016 data [16, 24, 25]. These triggers were complemented by others with higher p_{T} thresholds and no isolation requirements to increase event acceptance. Events must have at least one reconstructed vertex, i.e. at least two tracks with $p_{\text{T}} > 0.4$ GeV consistent with the beam-collision region in the x–y plane. If multiple vertices are reconstructed, then the primary vertex is taken to be the one with the largest sum, over the tracks assigned to it, of the transverse momentum squared of each track.

Events are further selected based on the presence of an electron or muon candidate from the decay of a W-boson, called 'primary' leptons. Primary-lepton electron candidates must satisfy a 'tight' likelihood-based identification criterion [26], be matched to the corresponding trigger, and have $p_{\text{T}} > 27$ GeV, $|\eta| < 2.47$ with the exclusion of $1.37 < |\eta| < 1.52$, longitudinal impact parameter $|z_0 \sin\theta| < 0.5$ mm

and transverse impact parameter significance $|d_0/\sigma(d_0)| < 5$, where $\sigma(d_0)$ is the uncertainty in the transverse impact parameter. Background from photon conversions, hadrons, and electrons produced away from the primary vertex ('non-prompt' electrons) is reduced by requiring the primary electron candidates to pass an isolation requirement based on the surrounding tracks and topological clusters in the calorimeter [26]. Primary-lepton muon candidates must satisfy a 'medium' quality identification criterion [27], be matched to the corresponding trigger, and have $p_{\mathrm{T}} > 27$ GeV, $|\eta| < 2.5$, longitudinal impact parameter $|z_0 \sin\theta| < 0.5$ mm and transverse impact parameter significance $|d_0/\sigma(d_0)| < 3$. Background from hadrons and from muons produced away from the primary vertex ('non-prompt' muons) is reduced by requiring primary muon candidates to pass an isolation requirement based on the surrounding tracks and topological clusters in the calorimeter, and be separated by $\Delta R > 0.4$ from the nearest selected jet. If the nearest selected jet is $\Delta R \leq 0.4$ from the muon and has less than three associated tracks (including the muon track), the muon is kept and the jet is removed from the jet list, to ensure high efficiency for muons undergoing significant energy loss in the calorimeter. Events with more than one candidate primary lepton with $p_{\mathrm{T}} > 25$ GeV are vetoed, in order to reject events from the $t\bar{t}$ dileptonic decay channel.

Jet candidates are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$, and a multivariate jet-vertex-tagger (JVT) is applied to suppress jets from pile-up [28]. During jet reconstruction, no distinction is made between identified electrons and jet energy deposits. Therefore, if any of the jets lie within $\Delta R = 0.2$ of a selected electron, the single closest jet is discarded in order to avoid double-counting electrons as jets. After this, electrons that are within $\Delta R = 0.4$ of a remaining jet are removed. Jets are identified as originating from a b-quark (b-tagged) using two techniques, one based on the reconstruction of a displaced jet (DJ tagging) and the other based on the semileptonic decay of a b-hadron into a so-called 'soft muon' (SMT tagging). For DJ tagging, multivariate techniques are used to combine information about the impact parameters of displaced tracks and the topological properties of secondary and tertiary decay vertices reconstructed within the jet [29]. The algorithm is trained on simulated $t\bar{t}$ events to discriminate b-jets from a background consisting of light-flavour jets and c-jets. A selection corresponding to an efficiency of 77% for b-jets in $t\bar{t}$ events is employed. The SMT tagging is performed by requiring the presence of a muon candidate satisfying the 'tight' quality identification criterion [27], $p_{\mathrm{T}} > 8$ GeV and $|\eta| < 2.5$, with loose requirements on the impact parameters ($|d_0| < 3$ mm, $|z_0 \sin\theta| < 3$ mm) and with a distance $\Delta R < 0.4$ from a selected jet candidate. The definition of the muon object for the SMT tagging was optimised by maximising the efficiency for muons originating from the semileptonic decays of b- and c-hadrons (selecting approximately 50% of b-jets containing a muon, which are in turn 20% of all b-jets produced in $t\bar{t}$ events), minimising the misidentification rate (about 10^{-3} per light-flavour jet, mostly due to the decays of pions and kaons), and minimising the uncertainty in the measured top-quark mass. If more than one muon satisfying the criteria above is found within a given jet, the muon with the highest p_{T} is chosen.

Events must have at least one SMT-tagged jet and at least one DJ-tagged jet (which could be the same jet), among a total of at least four jet candidates with $p_{\mathrm{T}} > 30$ GeV (with the exception of the SMT-tagged jet which may have a p_{T} as low as 25 GeV). If more than one SMT-tagged jet is found in the event, only the one with the highest-p_{T} muon is considered. The SMT muon and the primary lepton must be separated by $\Delta R_{\ell,\mu} < 2$. The presence of at least one neutrino in the final state is inferred from the requirements that $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}}(W) > 60$ GeV.[2] The requirement that the SMT muon and the primary lepton must be separated by $\Delta R_{\ell,\mu} < 2$ enhances the fraction of events where both leptons come from the same top-quark, in contrast to events where the two leptons originate from different top-quarks. The selected events are categorised as same-sign (SS) events or opposite-sign (OS) events according

[2] The transverse mass is given by $m_{\mathrm{T}}(W) = \sqrt{2p_{\mathrm{T}}^{\ell}E_{\mathrm{T}}^{\mathrm{miss}}(1-\cos\Delta\phi)}$, where p_{T}^{ℓ} is the transverse momentum of the muon (electron) and $\Delta\phi$ is the azimuthal angle separation between the lepton and the direction of the missing transverse momentum.

to the charge signs of the primary lepton and the soft muon. When both leptons come from the same top-quark, opposite-sign events are enriched in direct $b \to \mu X$ decays, while same-sign events have a large contribution from sequential $b \to cX' \to \mu X''$ decays; both samples carry information about the mass of the parent top-quark though. Finally, the invariant mass of the primary lepton and the soft muon ($m_{\ell\mu}$) is required to be between 15 and 80 GeV, as this is the region most sensitive to the top-quark mass. This requirement also suppresses the Z-boson, J/ψ and Υ resonances.

3.3 Signal and background simulations

A number of Monte Carlo (MC) simulation samples are used to model the expected signal of top-quark pairs and the background. The MC samples were processed either through the full ATLAS detector simulation [30] based on GEANT4 [31] or through a faster simulation making use of parameterised showers in the calorimeters [32]. Additional simulated pp collisions generated by PYTHIA 8.186 [33] with the MSTW2008 [34, 35] set of leading-order (LO) parton distribution functions (PDFs) and a set of tuned parameters called the AUET2 tune [36] were used to model the effects of both in-time and out-of-time pile-up. They were superimposed on the MC events, matching the luminosity profile of the recorded data. All simulated samples were processed through the same reconstruction algorithms and analysis chain as the data.

Simulated MC events were corrected so that the object identification efficiencies and energy and momentum scales and resolutions matched those determined from data control samples [22, 27, 28, 37]. The modelling of SMT muons and their misidentification due to light-hadron decays and detector background ('SMT fake') was studied using control samples as well. The efficiency of muon identification in jets was calibrated using muons from J/ψ and Z decays, and checked as a function of nearby track and calorimeter activity, and of the muon's transverse impact parameter d_0. The calibration of the misidentification rate was performed using a sample of W+1-jet events simulated with SHERPA 2.2.1 [38] as in Ref. [39], and a data-to-simulation scale factor (SF) of 1.10 ± 0.14 was derived. A slight miscalibration of the p_{T} of jets that contain a soft muon was observed, and the p_{T} of these jets was corrected in the simulation with a factor of 0.967 ± 0.024. It has been measured by studying the distribution of the ratio of the p_{T} values of the SMT-tagged jet and the average non-SMT-tagged jet in $t\bar{t}$ data and simulation.

The $t\bar{t}$ sample was generated using the hvq program [40] in the POWHEG-BOX v2 generator [41, 42] with the NNPDF3.0 set of PDFs [43] and the top-quark mass set to 172.5 GeV. Additional samples with different top-quark mass hypotheses were produced in the range of m_t between 165 and 180 GeV, with steps of 0.5 GeV between 170 and 175 GeV. The samples have been produced using the appropriate top-quark decay width values predicted at next-to-next-to-leading order (NNLO) as a function of m_t [44]. The hvq program uses on-shell matrix elements for production of $t\bar{t}$ pairs at next-to-leading order (NLO) in quantum chromodynamics (QCD). Off-shell effects and top-quark decays, including spin correlations, were approximated using MADSPIN [45]. Parton showers and hadronisation were modelled by PYTHIA 8.2 [33] using a dedicated 'A14-r_b' setting of the ATLAS A14 [46] tune, as detailed in Section 3.4. The A14 tune is based on LEP and Tevatron collider data and also uses a combination of ATLAS $\sqrt{s}$=7 TeV measurements of the underlying event, jet production, Z-boson production and top-quark production in order to constrain the parameters for the showers, multiple parton interactions and colour reconnection effects. Radiation in top-quark decays was handled entirely by the parton-shower generator, which implements matrix-element corrections with an accuracy equivalent to a calculation at the NLO level. The h_{damp} parameter, which controls the p_{T} of the first additional emission beyond the Born configuration, was set to 1.5 times the top-quark mass of each sample. The main effect of the h_{damp} setting is to regulate the high-p_{T}

emission against which the $t\bar{t}$ system recoils. The EvtGen v1.2.0 [47] program was used to simulate bottom and charm hadron mixing and decays. The production fractions and the branching ratios (BR) of the decay of b-hadrons and c-hadrons into muons were rescaled to the latest values from the Particle Data Group (PDG) [1], as detailed in Section 3.4. The simulated $t\bar{t}$ event sample was normalised to the Top++2.0 [48] theoretical cross-section of 832^{+46}_{-51} pb, calculated at NNLO in QCD and including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [49–53].

The main backgrounds to candidate signal events come from the production of a single top-quark, and from a W- or Z-boson in association with jets. A small background contribution arises from diboson (WW, WZ, ZZ) production. Events not containing prompt leptons also contribute to the selected sample via the misidentification of a jet or a photon as an electron, or the presence of non-prompt electrons or muons passing the prompt isolated lepton selection. This contribution is referred to as 'multijet' background, and was estimated using data by following the matrix method described in Ref. [54].

Samples of W/Z+jets events, and diboson production in association with jets, were simulated using the Sherpa generator. In the W/Z+jets samples, matrix elements were calculated for up to two partons at NLO and four partons at LO using the Comix [55] and OpenLoops matrix-element generators and merged with the Sherpa parton shower [56], using the ME+PS@NLO prescription [57]. The CT10 PDF set was used in conjunction with dedicated parton-shower tuning developed by the Sherpa authors. The normalisation of the W+jet background and the fractions of W-bosons produced in association with heavy-flavour quarks are extracted from data, taking advantage of the intrinsic W-boson charge asymmetry in this process [58]. The Z+jets contribution is estimated from MC simulation and checked in a data control sample. The diboson+jets samples were generated following the same approach used for the W/Z+jets events but with up to one (for ZZ) or zero (for WW, WZ) additional partons at NLO, and up to three additional partons at LO. They are normalised to their respective theoretical NLO cross-sections calculated by the event generator.

Samples of Wt- and s-channel single top-quark background events were generated with Powheg-Box v1 and v2, respectively, with the CT10 PDF set. Overlaps between the $t\bar{t}$ and Wt final states were removed with the 'diagram removal' prescription [59]. Electroweak t-channel single top-quark events were generated using the Powheg-Box v1 generator which uses the four-flavour scheme for the NLO matrix-element calculations together with the fixed four-flavour PDF set CT10f4. For this process, the top-quarks were decayed using Madspin [60], preserving all spin correlations. All single top-quark samples were interfaced to Pythia 6.428 [61] with the Perugia 2012 [62] underlying-event tune. The EvtGen v1.2.0 program was used to model properties of the bottom and charm hadron decays. The single top-quark t- and s-channel samples were normalised to the approximate NNLO theoretical cross-sections [63–65].

3.4 Modelling of heavy-quark fragmentation, hadron production and decays

The modelling of the momentum transfer between the b-quark and the b-hadron is an important aspect of this analysis. The Monte Carlo event generators, such as the Pythia, Herwig [66, 67] and Sherpa programs, describe this transfer according to phenomenological models, namely the string and cluster models containing parameters which are tuned to data. The Pythia8 program uses parametric functions to describe the b-quark fragmentation function, while Herwig7 and Sherpa use a non-parametric model which handles the complete parton-shower evolution. The free parameters in those models are typically fit to measurements from e^+e^- colliders, and this analysis assumes that the b-quark fragmentation function is the same in e^+e^- and pp collisions, as supported by dedicated studies.

The Lund–Bowler parameterisation [68, 69] in Pythia8 was used. It is given by

$$f(z) = \frac{1}{z^{1+br_b m_b^2}}(1-z)^a \exp(-bm_{\mathrm{T}}^2/z),$$

where a, b and r_b are the function parameters, m_b is the b-quark mass, $m_{\mathrm{T}} = \sqrt{m_B^2 + p_{\mathrm{T}}^2}$ the b-hadron transverse mass (m_B being the b-hadron mass), and z is the fraction of the longitudinal energy carried by the b-hadron with respect to the b-quark, in the light-cone reference frame. The fragmentation function is defined at the hadronisation scale and it is evolved by the parton shower to the process scale through DGLAP evolution equations. In Pythia8, the values of a and b were fit to data sensitive to light-quark fragmentation [70], such as charged-particle multiplicities, event shapes and scaled momentum distributions. They are assumed to be universal for light- and heavy-quarks, while the r_b parameter is specific to b-quark fragmentation.

The description of the b-quark fragmentation in the ATLAS A14 tune is improved by fitting for the `StringZ:rFactB` Pythia8 parameter (corresponding to r_b) following the approach given in Refs. [71–74]. The A14 tune sets $r_b = 0.855$. The fit uses the A14 tune with e^+e^- collision data from the ALEPH, DELPHI and OPAL experiments at the LEP collider, and from the SLD experiment at the SLC collider [75–78]. The distribution of $x_B = 2p_B \cdot p_Z/m_Z^2$ from semileptonically decaying b-hadrons in $e^+e^- \to Z \to b\bar{b}$ events is used, where p_B and p_Z are the four-momenta of the b-hadron and the Z-boson, respectively. In the Z rest frame, m_Z is twice the beam energy and therefore $x_B = 2E_B/m_Z$, where E_B is the energy of the b-hadron. The fit is performed using Rivet v3.1.0 [79] to implement the measurements. The effect of the matrix-element corrections for $e^+e^- \to Z \to b\bar{b}g$ is taken into account. Eighty simulated samples of 1M $e^+e^- \to Z \to b\bar{b}$ events were produced using Pythia8 with different values of the r_b parameter in the interval [0.8-1.4] and compared to the experimental data in HEPDATA format. The extraction of the best r_b value is performed through a standard binned χ^2 test on the experimental x_B distribution where statistical and systematic uncertainties are taken into account for each of the four experiments. In addition, for the results of ALEPH, DELPHI and OPAL, bin-by-bin correlations are taken into account in the fit procedure. The SLD experiment did not provide the full covariance matrix for the total uncertainties and therefore the χ^2 fit for this experiment is performed ignoring bin-by-bin correlations. For each experiment, the χ^2 minimisation is performed and the best r_b value and its uncertainty are found. The results are summarised in Table 1. The values of χ^2/ndf for DELPHI and SLD experiments show a poor modelling of the data by the simulated templates. Therefore, before including these results into a global χ^2 combination, the uncertainties found in the r_b parameter for these two cases were rescaled by a factor of $S = \sqrt{\chi^2/\mathrm{ndf}}$ following the procedure outlined in Ref. [1]. After these uncertainties are rescaled, the four χ^2 curves are summed up to produce a single χ^2 curve taking into account the information of all four experiments. In this approach, the four experiments are considered uncorrelated since the dominant uncertainties on the measurements come from uncorrelated sources. This curve is then fitted with a parabola to find the minimum of the χ^2 and to extract the best value of the r_b parameter and its uncertainty, yielding $r_b = 1.05 \pm 0.02$. The values of the r_b parameter and its uncertainty for the four experiments, and their combination are shown in Figure 1. A similar procedure is performed, as a cross-check, exploiting the average value $\langle x_B \rangle$ of the four x_B distributions. The experimental $\langle x_B \rangle$ values are compared with the predicted value of the various simulated samples, and the best r_b values and their uncertainties are extracted. A weighted average of the single results is calculated and the result is found to be compatible with the previous method.

Table 1: Results of a fit of the r_b parameter of the Lund-Bowler b-quark fragmentation function in PYTHIA8 to different experimental data where bin-by-bin correlations for ALEPH, DELPHI and OPAL experiments were considered.

Experiment	r_b	χ^2/ndf
ALEPH	1.070 ± 0.035	21/18
DELPHI	1.094 ± 0.030	73/8
OPAL	1.023 ± 0.019	18/19
SLD	1.092 ± 0.018	58/21

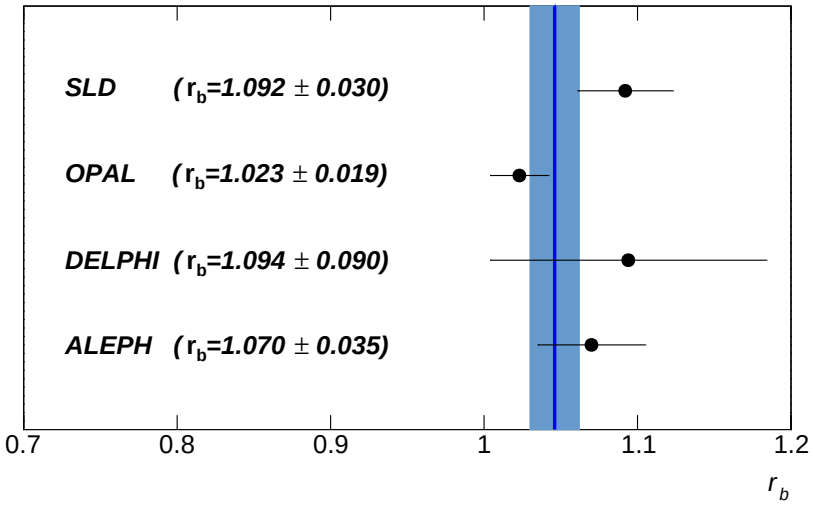


Figure 1: Fit of the r_b PYTHIA parameter to the four experiments considered: SLD, OPAL, DELPHI and ALEPH. For each experiment, the r_b fitted value is shown together with its uncertainty. These results have been obtained including bin-by-bin correlations in the experimental data (except for SLD experiment) and after rescaling the uncertainties on r_b for DELPHI and SLD experiments as explained in the text. The solid blue line represents the value of r_b obtained from the combined fit to the four experiments. The blue band represents the uncertainty from the combined fit to the four experiments.

The setting with $r_b = 1.05 \pm 0.02$ is referred to as A14-r_b and is applied in this analysis to all signal MC samples using PYTHIA8 for the simulation of the parton shower.

The production fractions of weakly decaying b- and c-hadrons observed in POWHEG+PYTHIA8 MC simulation with EVTGEN are rescaled to those from the Heavy Flavour Averaging Group (HFLAV) [80] as reported by the PDG [1] and in Ref. [81]. The production fraction values and corresponding scale factors for POWHEG+PYTHIA8 simulations are shown in Table 2. These scale factors refer only to the first weakly decaying hadron produced in the hadronisation process of b- and c-quarks. The scale factors are applied to each of these hadrons present in a MC simulated event, with the overall event weight given by the product of these scale factors. This procedure assumes that the production fractions of heavy-flavour quarks can be regarded as universal in the kinematic phase space relevant for this analysis, within the uncertainties accounted for here, as supported by recent results [82–88] which find deviations from universality only in the very low-p_{T} regime.

The branching ratios of the b- and c-hadron decays that contain a muon are also adjusted to match those measured by previous experiments [1]. Central values and relative scale factors, along with the corresponding uncertainties, are shown in Table 3; these are measured after applying the corrections for

Table 2: The production fraction values for b-hadrons and c-hadrons in the PDG and POWHEG+PYTHIA8. The relative scale factors applied to POWHEG+PYTHIA8 are also shown. The values in the PDG column are derived from Refs. [1] and [81]. The same scale factors are applied to the charge-conjugate hadrons.

Hadron	PDG	POWHEG+PYTHIA8	Scale Factor
B^0	0.404 ± 0.006	0.429	0.941 ± 0.014
B^+	0.404 ± 0.006	0.429	0.942 ± 0.014
B_s^0	0.103 ± 0.005	0.095	1.088 ± 0.052
b-baryon	0.088 ± 0.012	0.047	1.87 ± 0.26
D^+	0.226 ± 0.008	0.290	0.780 ± 0.027
D^0	0.564 ± 0.015	0.553	1.020 ± 0.027
D_s^0	0.080 ± 0.005	0.093	0.857 ± 0.054
c-baryon	0.109 ± 0.009	0.038	2.90 ± 0.24

Table 3: Hadron to muon branching ratios from the PDG and in POWHEG+PYTHIA8. The relative scale factors applied to POWHEG+PYTHIA8 are also shown. The values in the PDG column are derived from Refs. [1] and [81]. The $c \to \mu$ scale factor is applied only to the semileptonic decays of c-hadrons into muons when the c-hadrons do not come from a cascade b-hadron decay. The same scale factors are applied to the charge conjugate hadrons.

Hadronic Decay Mode	PDG	POWHEG PYTHIA8	Scale Factor
$b \to \mu$	$0.1095\ ^{+0.0029}_{-0.0025}$	0.106	$1.032\ ^{+0.0027}_{-0.0023}$
$b \to \tau$	0.0042 ± 0.0004	0.0064	0.661 ± 0.062
$b \to c \to \mu$	0.0802 ± 0.0019	0.085	0.946 ± 0.022
$b \to \bar{c} \to \mu$	0.016 ± 0.003	0.018	0.89 ± 0.17
$c \to \mu$	0.082 ± 0.005	0.084	0.976 ± 0.059

the production fractions described above. The $b \to \bar{c} \to \mu$ branching ratio was determined by averaging the direct measurement from DELPHI [89] and the predicted values computed by the LEP Electroweak Heavy Flavour Working Group [90]. The latter prediction is based on flavour-specific $B \to D$ and $B \to \Lambda_c^+$ rates measured at CLEO [91–93] in combination with the $B \to DD(X)$ rates measured in ALEPH [94] to extract the probabilities of producing the different c-hadrons from the initial b-hadron decays. The c-hadron semileptonic branching fractions were also used in the prediction. The $c \to \mu$ scale factor is applied only to the semileptonic decays of c-hadrons into muons when the c-hadrons do not come from a cascade b-hadron decay.

4 Analysis

4.1 Event yields and sample composition

The number of observed candidate events and the predicted signal and background are shown in Table 4, for both the OS and SS regions. Over 90% of the events in the sample contain a top-quark pair, including cases where the soft muon is erroneously chosen from a $t\bar{t}$ dilepton decay, whereby a muon from the prompt W decay is found near a jet or radiates a near-collinear photon mimicking a soft muon tag, and

cases where the soft muon candidate does not originate from a b decay such as in $W \to cs$ or SMT fakes. The contributions from single top-quark, W- or Z-boson in association with jets, and multijet background are not negligible. The Z+jets background gives a small contribution near the peak of the $m_{\ell\mu}$ distribution, but becomes important for $m_{\ell\mu}$ close to the Z-boson mass peak.

Of the selected $t\bar{t}$ events in the OS class, 83% are cases where the primary lepton and the soft muon belong to the decay of the same top-quark, and 10% are events where the two originate from different top-quarks. The topological requirement $\Delta R_{\ell,\mu} < 2$ is responsible for much of the purity because it is very effective in preferentially selecting the same top-quark decays. For the remaining 7% of cases, the soft muon does not originate from either of the b-quarks produced in the two top-quark decays. In the SS class, the above fractions are 57%, 41% and 2%, respectively. The rate of soft muons that are not from b-quarks is higher in OS events owing to the decay of c-quarks that come from $W \to cs$ in the top-quark decay chain.

Table 4: Event yields with $m_{\ell\mu}$ between 15 and 80 GeV, separately for OS and SS regions. Uncertainties shown include statistical and systematic contributions but not the recoil uncertainty.

Process	Yield (OS)	Yield (SS)
$t\bar{t}$ (SMT from b- or c-hadron)	55 700 ± 3400	34 800 ± 2300
$t\bar{t}$ (SMT from $W \to \mu\nu$)	2190 ± 310	4.9 ± 3.6
$t\bar{t}$ (SMT fake)	1490 ± 210	1240 ± 170
Single top t-channel	770 ± 70	490 ± 40
Single top s-channel	63 ± 6	49 ± 4
Single top Wt channel	1840 ± 140	1260 ± 100
W+jets	1600 ± 400	1080 ± 240
Z+light jets	210 ± 80	15 ± 6
Z+HF jets	550 ± 180	310 ± 100
Diboson	17.2 ± 2.9	6.3 ± 1.4
Multijet	530 ± 140	480 ± 130
Total Expected	65 000 ± 4000	39 700 ± 2500
Data	66 891	42 087

In order to better understand the nature of the sample composition in the OS and SS regions, the expected $t\bar{t}$ events shown in the first row of Table 4 are further resolved into components. Table 5 shows the components involving direct and sequential decays, and decays not belonging to the b-quark from the $t \to Wb$ transition. The direct $t \to B \to \mu$ decays are by far the dominant component in the OS sample, while the sequential $t \to B \to D \to \mu$ decays are also suppressed by the kinematic requirement on the soft muon. Decay channels involving τ leptons give a small contribution.

The data are compared with the sum of the predicted signal and backgrounds in Figures 2 and 3, for an illustrative selection of kinematic distributions of the candidate events for the OS and SS selections: $p_{\mathrm{T}}(\mu^{\mathrm{SMT}})$, $\eta(\mu^{\mathrm{SMT}})$, $p_{\mathrm{T}}(\ell^{\,\mathrm{primary}})$ and $m_{\mathrm{T}}(W)$. The uncertainties are discussed in detail in Section 5. The compatibility of the data and MC predictions is studied using a χ^2 test involving the bin-by-bin full correlation matrix, and for all distributions the level of agreement is better than 2 standard deviations. The slight excess of SMT muons in regions of high-η is associated with a potential small mismodelling of the efficiency for low-p_{T} muons in simulation. The corresponding impact on the measurement is negligible.

Table 5: Fraction of MC-simulated $t\bar{t}$ events with a soft muon originated by a b- or c-hadron split into components of direct and sequential decays, and decays not belonging to the b-quark from the $t \to Wb$ transition, separately for the opposite-sign and same-sign event selections. The letters B and D indicate b- and c-hadrons of either charge. Only MC events with two real muons are included.

	OS [%]	SS [%]
Processes involving a μ from a t or $\bar{t}$		
$t \to B \to \mu$	73.6	51.2
$t \to B \to D \to \mu$	16.7	44.2
$t \to B \to \tau \to \mu$	2.0	1.3
$t \to B \to D \to \tau \to \mu$	0.8	0.8
Processes involving a μ not from a t or $\bar{t}$		
$B \to \mu$	0.6	0.9
$D \to \mu$	5.8	1.4
Other ($\tau \to \mu$)	0.5	0.1

4.2 Extraction of the top-quark mass

The distribution of the invariant mass of the primary lepton and the soft muon, $m_{\ell\mu}$ is used to determine the mass of the parent top-quark. A binned-template profile likelihood fit is performed, with a Poisson likelihood model and systematic uncertainties included as Gaussian-constrained nuisance parameters [95]. Only the range of $m_{\ell\mu}$ between 15 and 80 GeV is considered in the fit, since the tail of the $m_{\ell\mu}$ distribution is more sensitive to $t\bar{t}$ modelling uncertainties and higher-order corrections, and to the Z+jets background. The fit is performed simultaneously for the OS and SS charge-combination samples, and Figure 4 shows the sensitivity of each of these distributions to variations of the top-quark mass, as well as the binning used by the templates. The SS sample has less sensitivity than the OS sample due to the larger incidence of sequential $b \to c \to \mu$ decays, where the SMT muon carries a smaller fraction of the parent b-quark momentum, and due to the larger fraction of events in which the leptons originate from different top-quarks.

The fit uses template histograms simulated as for the nominal $t\bar{t}$ sample but with different values for the input top-quark mass. The templates from the different mass samples are interpolated with piece-wise linear functions built bin by bin. To improve the stability of the method, the templates are smoothed assuming a linear dependence on m_t for the fraction of the total number of events in each bin. A maximum-likelihood fit is performed with three free parameters: m_t, which controls the shape of the $m_{\ell\mu}$ distribution for $t\bar{t}$ events, and the normalisation factors for $t\bar{t}$ events in the OS and SS samples. The normalisation factors ensure that the total of the $t\bar{t}$ signal and background events is always equal to the total number of selected data events, and no m_t information is extracted from the number of events.

The uncertainty due to the limited number of simulated events, and due to statistical fluctuations in the background estimates based on control samples, is evaluated by defining a new source of systematic uncertainty for each bin of the prediction, which modifies the bin content by its statistical uncertainty. Since a very large number of systematic uncertainties are considered *a priori*, a pruning procedure is applied to reduce the number of statistically insignificant systematic uncertainties affecting the prediction of each of the signal and background processes. A systematic variation of the $m_{\ell\mu}$ templates is excluded if the total predicted change is smaller than 0.05% of the nominal bin content for all bins. The impact on the total estimated uncertainty is smaller than 0.03 GeV.

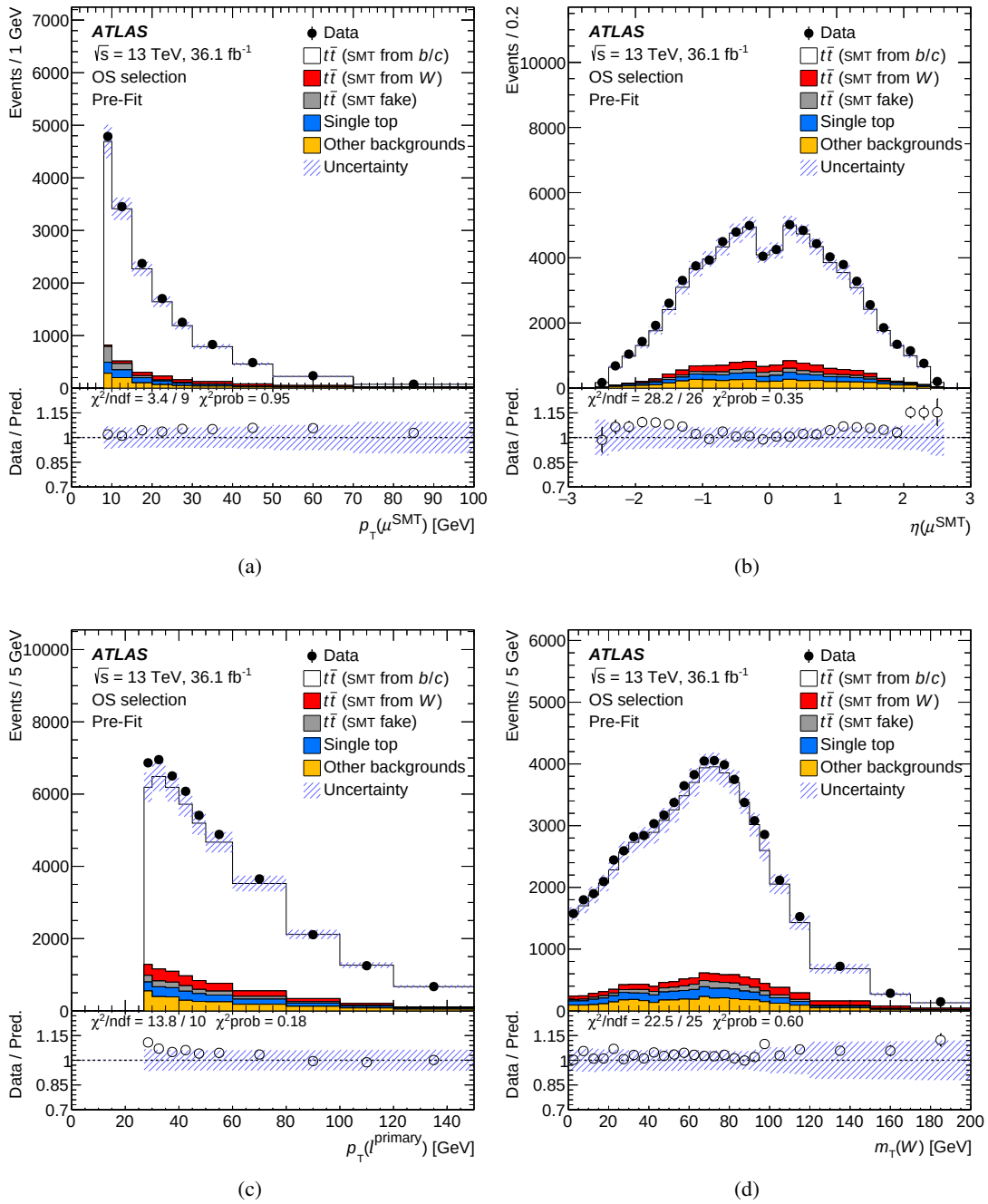


Figure 2: Comparison of data and prediction before the fit described in Section 4.2 in the OS sample, i.e. for events with primary lepton and the soft muon with opposite charges, for the (a) soft muon p_{T}, (b) soft muon η, (c) primary lepton p_{T} and (d) W-boson transverse mass. The prediction reports the expected event contribution from the signal and backgrounds. The uncertainty band includes statistical and systematic uncertainties, but does not include the recoil uncertainty.

The top-quark mass determination from the fit is found to be linear and unbiased with respect to the input top-quark mass hypothesis by means of pseudo-experiments, and its uncertainty from the likelihood ratio

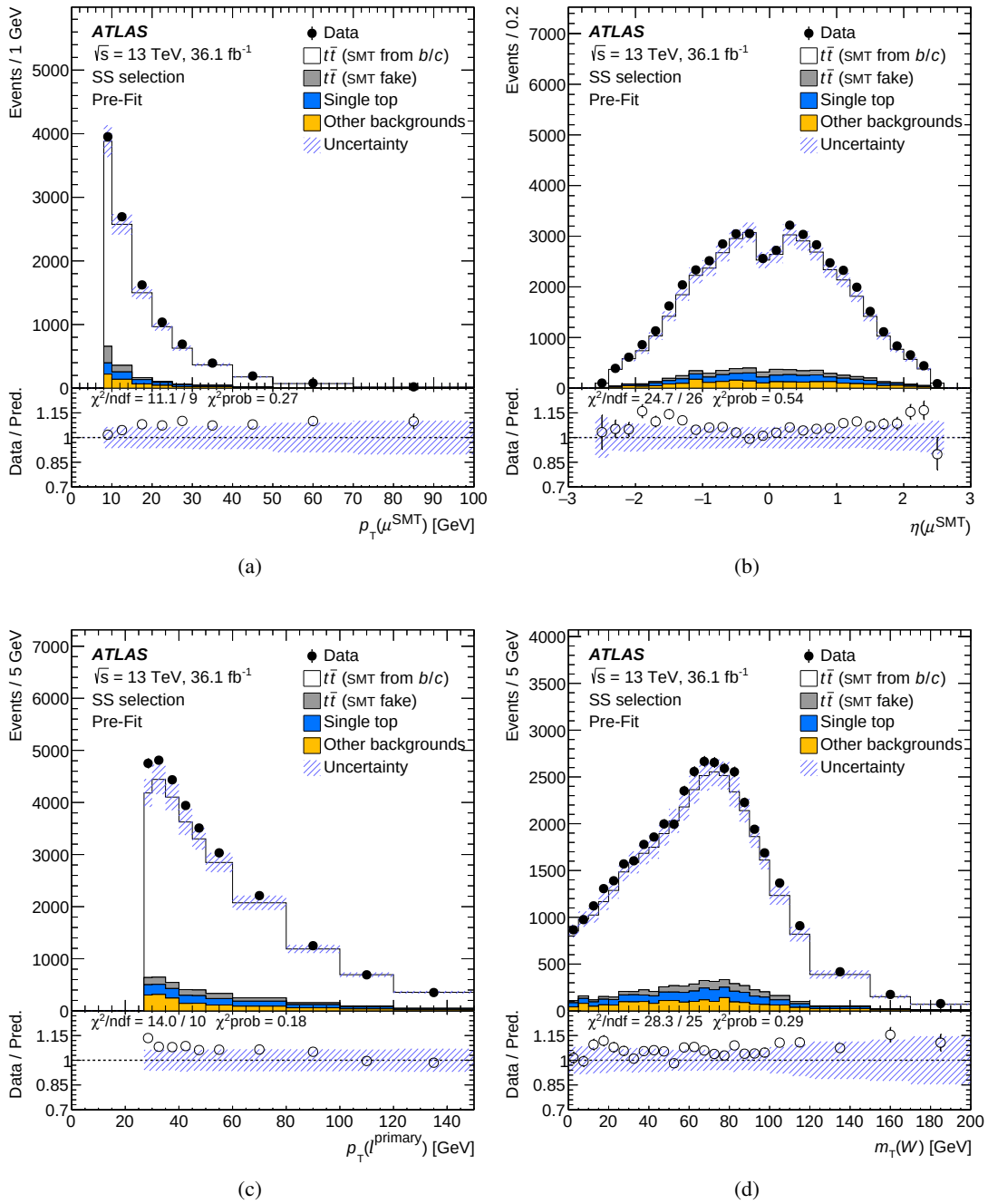
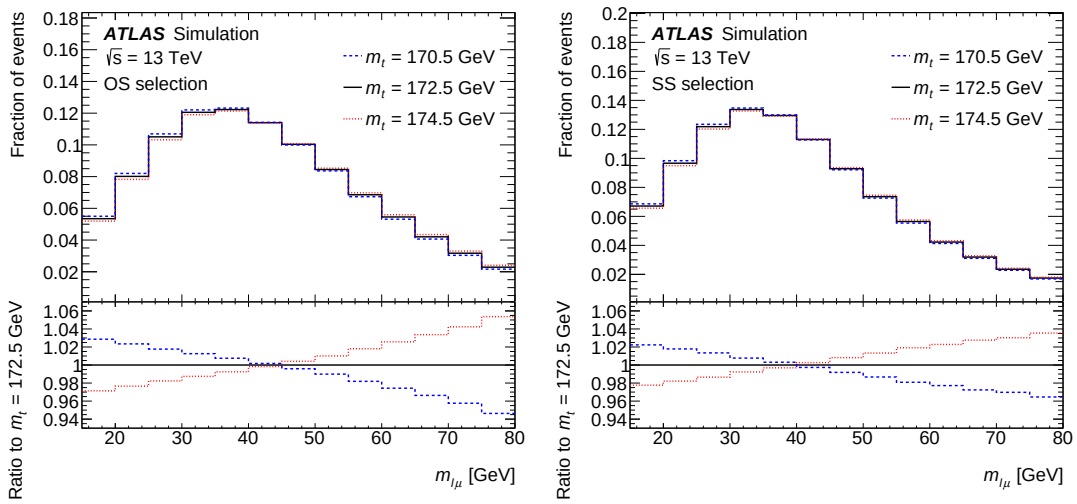


Figure 3: Comparison of data and prediction before the fit described in Section 4.2 in the SS sample, i.e. for events with primary lepton and the soft muon with same charges, for the (a) soft muon p_{T}, (b) soft muon η, (c) primary lepton p_{T} and (d) W-boson transverse mass. The prediction reports the expected event contribution from the signal and backgrounds. The uncertainty band includes statistical and systematic uncertainties, but does not include the recoil uncertainty.

is also checked to ensure it reports the correct statistical coverage. The fit method and the event selection were optimised to minimise the total uncertainty in m_t in a 'blinded' approach, i.e. using pseudo-data and



(a) (b)

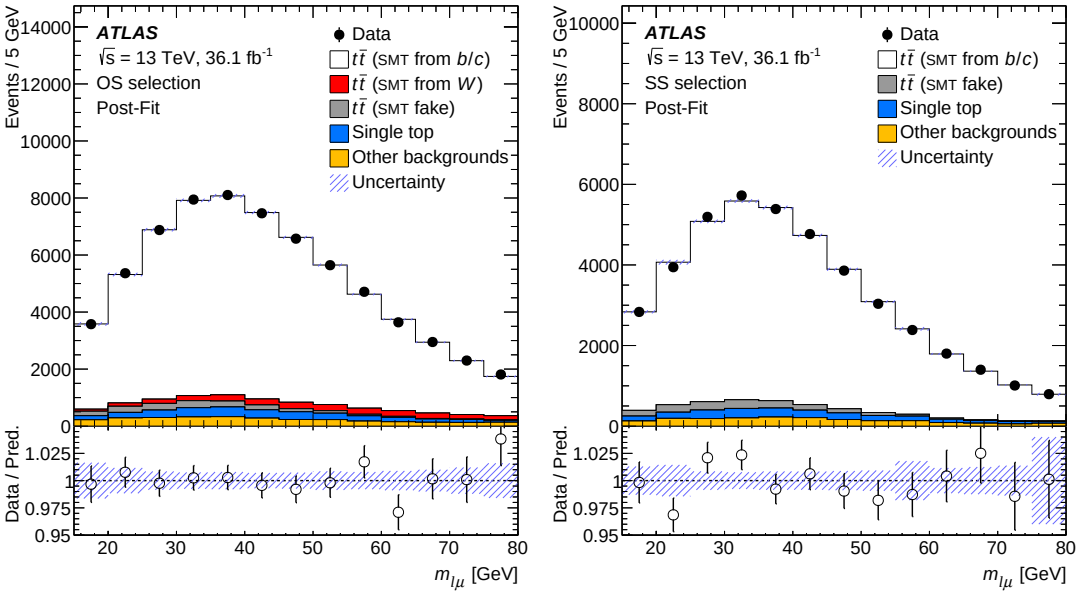
Figure 4: Sensitivity of the $m_{\ell\mu}$ distribution to different input top-quark masses from simulated events, separately for the OS and SS samples.

data without knowledge of the best-fit top-quark mass. The fit yields:

$$m_t = 174.41 \pm 0.39\ (\text{stat.}) \pm 0.66\ (\text{syst.}) \pm 0.25\ (\text{recoil})\ \text{GeV},$$

where the statistical, systematic and recoil uncertainties are described in detail in Section 5. Figure 5 shows the post-fit $m_{\ell\mu}$ distributions in the OS and SS samples; a goodness-of-fit test is performed using the saturated model technique [1, 96] and returns a probability of 56%. Figures 6 and 7 display the corresponding post-fit plots for the kinematic variables of Figures 2 and 3. The data distributions are well described by the prediction, with the primary lepton p_{T} exhibiting a slight trend which is traced to the boost of the $t\bar{t}$ system, but which has no appreciable impact on the determined top-quark mass. The likelihood scan with the best-fit top-quark mass value is shown in Figure 8.

Checks were performed by fitting the OS and SS regions separately, giving $m_t(\text{OS}) = 174.63 \pm 0.47\ (\text{stat.}) \pm 0.75\ (\text{syst.})$ GeV and $m_t(\text{SS}) = 173.88 \pm 0.74\ (\text{stat.}) \pm 1.01\ (\text{syst.})$ GeV. Checks were performed by separately fitting the electron and muon channels, different W-decay lepton charges and different configurations of b-tagging and event selection, and were found to all give consistent results. Checks also included the extraction of the top-quark mass with alternative statistical methods, namely using analytic functions for $m_{\ell\mu}$ with a parametric dependence on the top-quark mass, only using the mean value of the $m_{\ell\mu}$ distribution, and using a binned-template likelihood fit without including systematic uncertainties as nuisance parameters. In particular, the inclusion of systematic uncertainties as nuisance parameters in the fit reduces the total uncertainty by 2.6%, in line with reasonable constraints from the fit.



(a) (b)

Figure 5: Post-fit $m_{\ell\mu}$ distributions in the (a) OS sample and in the (b) SS sample. The prediction reports the event contribution for the signal and backgrounds. The uncertainty band includes statistical and systematic uncertainties, but does not include the recoil uncertainty. The SS (same sign) or OS (opposite sign) refers to the charge signs of the primary lepton and the soft muon.

5 Measurement uncertainties

The individual sources of uncertainty and the evaluation of their effect on m_t are described in the following. Many sources of systematic uncertainty are considered, corresponding to a total of 151 individual variations. Table 6 summarises the impact on m_t of the main sources of systematic uncertainty, and each systematic uncertainty is accompanied by an estimate of its statistical precision performed using the bootstrap method [97].

5.1 Statistical and datasets

The uncertainty related to the size of the data sample (data statistical uncertainty) is obtained by performing the fit while keeping constant all of the nuisance parameters associated with the systematic uncertainties. The data statistical uncertainty obtained by using both the OS and SS selections is ±0.39 GeV, while the OS sample alone yields ±0.47 GeV, highlighting the contribution to the sensitivity from the SS sample. The uncertainty due to the limited size of the simulated signal and background samples includes both the impact of the signal samples size on m_t-dependent templates used for the top mass interpolation, and the uncertainty of the backgrounds due to the limited size of the corresponding MC samples. This includes the multijet background, which is estimated with a data control sample. The uncertainty in the combined 2015–2016 integrated luminosity is 2.1% [18], obtained using the LUCID-2 detector [98] for the primary luminosity measurements. The distribution of the average number of interactions per bunch crossing (pile-up activity) in each MC sample is reweighted to match the conditions in data, and a corresponding

Table 6: Impact of main sources of uncertainty on m_t. Each row of the table corresponds to a group of individual systematic variations. For each uncertainty source the fit is repeated with the corresponding group of nuisance parameters fixed to their best-fit values. The contribution from each source is then evaluated by subtracting in quadrature the uncertainty obtained in this fit from that of the full fit. The total systematic uncertainty is different from the sum in quadrature of the different groups due to correlations among nuisance parameters in the fit. The last column shows the statistical uncertainty on each of the top-quark mass uncertainties as estimated with the bootstrap method [97].

Source	Unc. on m_t [GeV]	Stat. precision [GeV]
Statistical and datasets		
Data statistics	0.39	
Signal and background model statistics	0.17	
Luminosity	< 0.01	± 0.01
Pile-up	0.07	± 0.03
Modelling of signal processes		
Monte Carlo event generator	0.04	± 0.06
b, c-hadron production fractions	0.11	± 0.01
b, c-hadron decay BRs	0.40	± 0.01
b-quark fragmentation r_b	0.19	± 0.06
Parton shower α_S^{FSR}	0.07	± 0.04
Parton shower and hadronisation model	0.06	± 0.07
Initial-state QCD radiation	0.23	± 0.08
Colour reconnection	< 0.01	± 0.02
Choice of PDFs	0.07	± 0.01
Modelling of background processes		
Soft muon fake	0.16	± 0.03
Multijet	0.07	± 0.02
Single top	0.01	± 0.01
W/Z+jets	0.17	± 0.01
Detector response		
Leptons	0.12	± 0.01
Jet energy scale	0.13	± 0.02
Soft muon jet p_T calibration	< 0.01	± 0.01
Jet energy resolution	0.08	± 0.07
b-tagging	0.10	± 0.01
Missing transverse momentum	0.15	± 0.01
Total stat. and syst. uncertainties (excluding recoil)	0.77	± 0.03
Recoil uncertainty	0.25	
Total uncertainty	0.81	

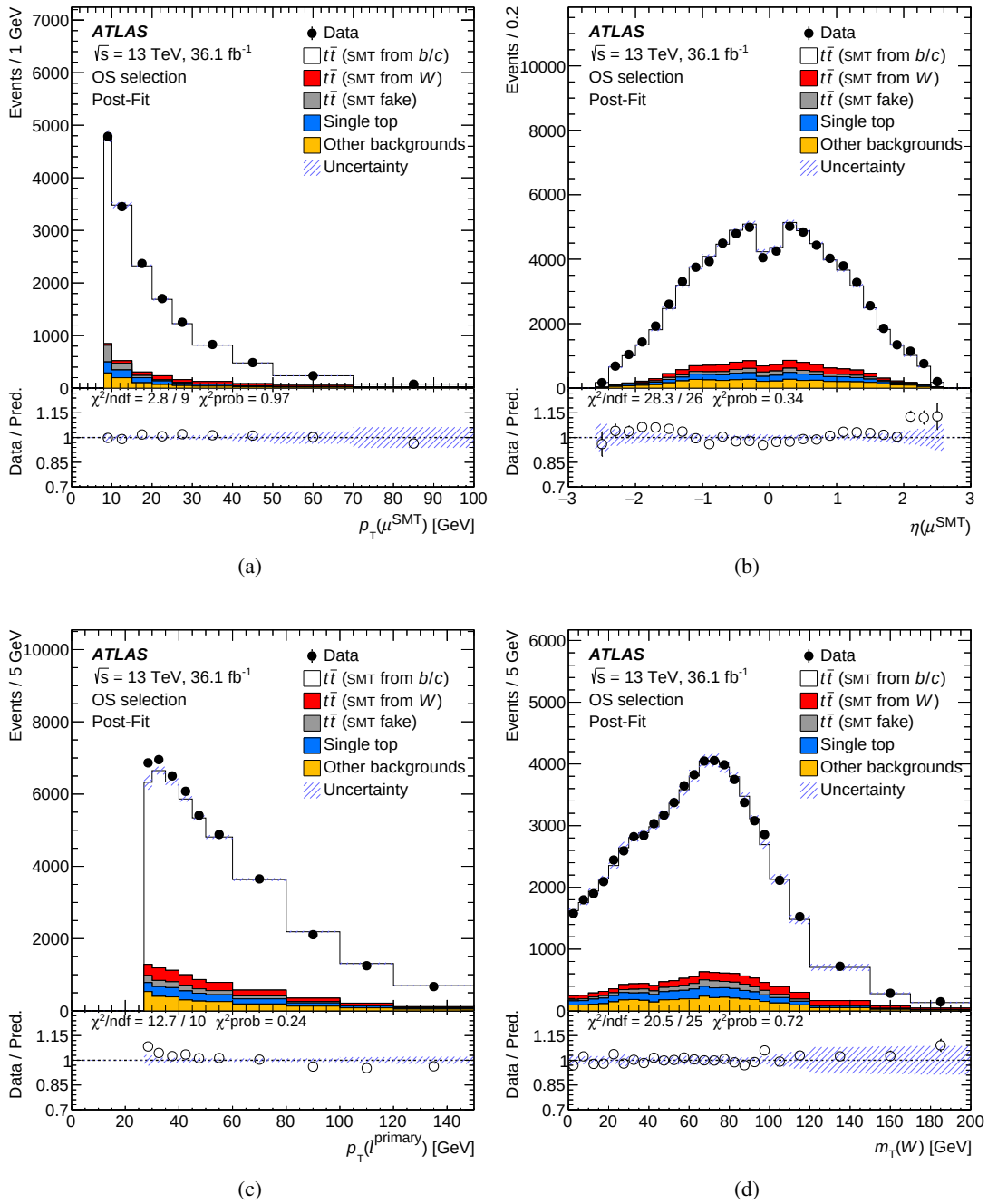


Figure 6: Post-fit comparison of data and prediction in the OS sample, i.e. for events with primary lepton and the soft muon with opposite charges, for the (a) soft muon p_{T}, (b) soft muon η, (c) primary lepton p_{T} and (d) W-boson transverse mass. The prediction reports the event contribution for the signal and backgrounds. The uncertainty band includes statistical and systematic uncertainties, but does not include the recoil uncertainty.

uncertainty is evaluated according to the uncertainty in the average number of interactions per bunch crossing.

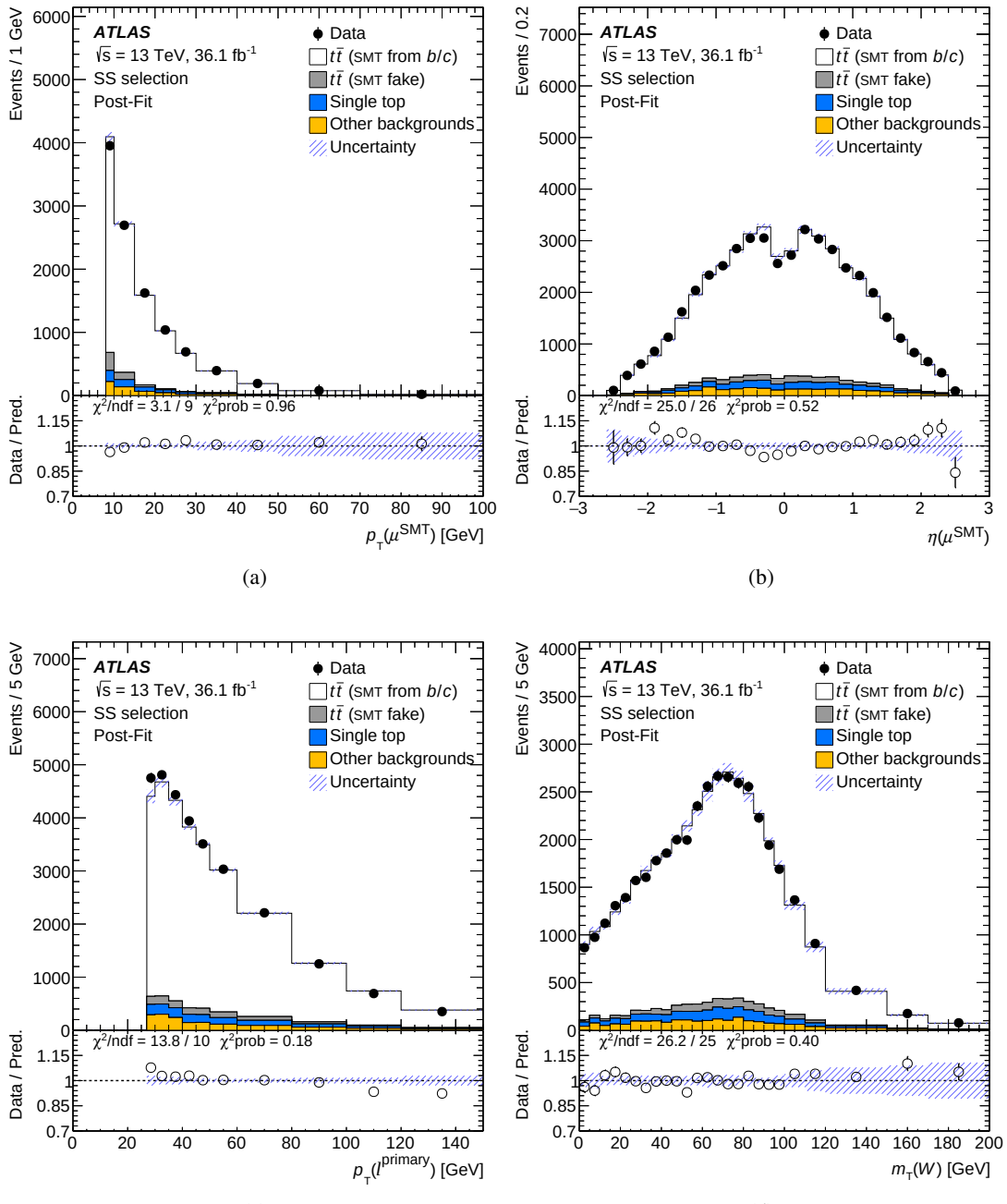


Figure 7: Post-fit comparison of data and prediction in the SS sample, i.e. for events with primary lepton and the soft muon with same charges, for the (a) soft muon p_{T}, (b) soft muon η, (c) primary lepton p_{T} and (d) W-boson transverse mass. The prediction reports the event contribution for the signal and backgrounds. The uncertainty band includes statistical and systematic uncertainties, but does not include the recoil uncertainty.

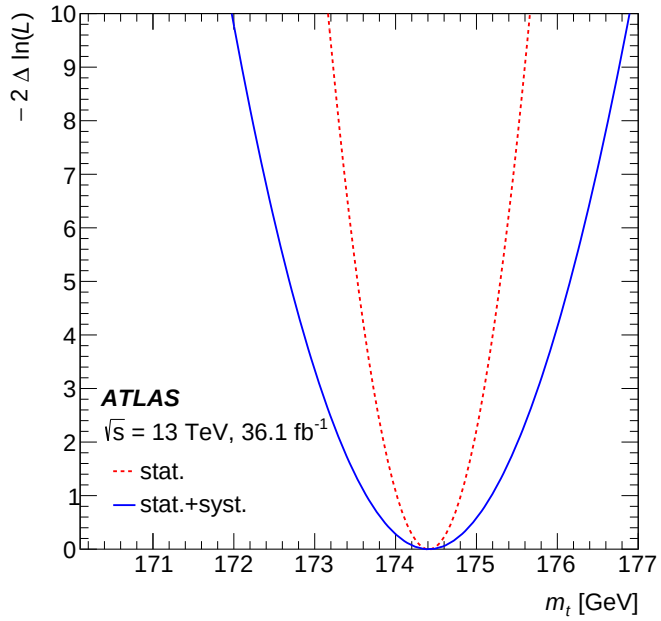


Figure 8: Likelihood scan, showing the best-fit value and the statistical and total uncertainty profiles (excluding the recoil uncertainty).

5.2 Modelling of the signal process

Uncertainties in the $t\bar{t}$ signal modelling include all sources that affect the kinematics of the lepton from the W-boson decay and the kinematics of the b-hadron giving rise to the soft muon, and also the fraction of events from different soft-muon flavour components (from b-hadrons, c-hadrons, light jets and W-bosons). The $t\bar{t}$ inclusive cross-section uncertainty does not affect the measurement, since the fit is based only on the shape of the distribution from $t\bar{t}$ events after background subtraction.

Uncertainties that depend on the choice of NLO matching scheme in the $t\bar{t}$ MC generator are estimated by comparing a sample generated with Powheg+Pythia8 with a sample generated with MadGraph5_aMC@NLO+Pythia8 [99] (referred to hereafter as aMC@NLO+Pythia8). The aMC@NLO matching requires specific settings of the Pythia8 shower to retain the NLO accuracy. The matrix-element corrections are switched off for both initial-state radiation and the global-recoil settings that are used for final-state radiation emissions. These settings are different from the nominal Powheg+Pythia8. In order to have a coherent comparison, an alternative Powheg+Pythia8 sample was generated with the same Pythia8 configuration as that used to shower aMC@NLO events. Additionally, since aMC@NLO+Pythia8 is known to describe poorly the distribution of the boost of the $t\bar{t}$ system ($p_{\mathrm{T}}^{t\bar{t}}$) [100], the $p_{\mathrm{T}}^{t\bar{t}}$ in aMC@NLO+Pythia8 is reweighted to that of the Powheg+Pythia8 sample. The full difference between the top-quark masses obtained with the two samples is considered as the positive and negative uncertainty due to the MC generator NLO matching.

Uncertainties in the b-hadron production fractions and the BRs of the inclusive decays of b- and c-hadrons into muons are derived from the uncertainties in the rescaling procedure, described in Section 3.4 and shown in Tables 2 and 3. These uncertainties are propagated through the analysis. In addition, a check was performed to verify that the impact on $m_{\ell\mu}$ due to the different admixture of D-mesons involved in

$b \to c\mu + X$ transitions was within the uncertainty assigned to $b \to \mu$ inclusive BRs. For this purpose, the exclusive decays $B^0 \to D^-\mu\nu$, $B^0 \to D^*(2010)^-\mu\nu$, $B^+ \to \overline{D^0}\mu\nu$, $B^+ \to \overline{D^*}(2007)^0\mu\nu$ and their charge conjugates were considered. For each of these decays, a kinematic selection similar to that of the main selection was applied to the events and the BR of each decay was varied within the uncertainty quoted by the PDG. The impact on $m_{\ell\mu}$ was found to be significantly smaller than the effect of varying only the BR of the inclusive $b \to \mu$ decays. The impact on $m_{\ell\mu}$ from the uncertainties in the $B^0_{(s)}$ mixing parameters is much smaller than the impact from imperfect knowledge of the b-hadron production fractions and from the heavy-quark hadrons BRs. Similarly, the impact of the modelling of the soft-muon kinematics in the exclusive semileptonic decays of b- and c-hadrons in EvtGen v1.2.0 was tested by varying the various BRs within their uncertainties. The total impact was found to be negligible with respect to the impact of the inclusive b-hadron production fractions and heavy-quark hadrons BRs uncertainties.

Uncertainties in the modelling of the parton shower and hadronisation processes include the estimation of several components. An alternative simulation of the $t\bar{t}$ sample is considered whereby the Powheg-Box generator is matched to the Herwig 7.1.3 generator for the modelling of the parton shower and hadronisation. The Herwig 7.1.3 generator release includes several improvements in the shower description for heavy-quark fragmentation, together with a new tune to e^+e^- data. The angle-ordered shower algorithm is preferred to the dipole shower for this sample, as it better describes both the shower evolution in the 7 TeV ATLAS measurement of jet shapes in $t\bar{t}$ events [101], and the x_B distribution of LEP data, although it does not describe the x_B spectrum of LEP data as well as Pythia8. The sample based on Herwig 7.1.3, when compared with the nominal $t\bar{t}$ simulation used in the fit, allows the effects of changes in the shower algorithm, and therefore in initial- and final-state emissions, to be assessed using alternative but coherent models of b-quark fragmentation and hadronisation, the underlying event and colour reconnection. The full difference between the top-quark masses obtained with the two samples of Powheg events showered with Pythia8 A14-r_b and Herwig 7.1.3 is considered as the positive and negative uncertainty from variations of the parton shower and hadronisation modelling.

To evaluate the uncertainty on the modelling of the b-quark fragmentation, additional samples were produced with the value of r_b in the fragmentation function varied by its uncertainty of ± 0.02. Additionally, the uncertainty on the DGLAP evolution of the fragmentation function from LEP to the LHC due to the final state radiation (FSR) renormalisation and factorisation scales ($\mu^{\text{FSR}}_{\text{R,F}}$, labelled α_S^{FSR} in Tab. 6), is evaluated by generating alternative event samples with the same fragmentation function at LEP, but with the scales varied up and down by factors of $\sqrt{2}$ [102].

As a check of the separate r_b and $\alpha_{\text{S}}^{\text{FSR}}$ uncertainties reported in Table 6, a sample of Powheg events showered with Pythia using the Monash tune [70] (which has different values of α_{S} and r_b than the A14-r_b tune), is used to obtain an equivalent systematic uncertainty. The change in the measured top-quark mass obtained with this sample is 0.30 ± 0.06 GeV, consistent with the uncertainties associated with $\alpha_{\text{S}}^{\text{FSR}}$, r_b and the hadronisation model, listed in Table 6.

In the modelling of the parton shower of the b-quark from $t \to Wb$ with Pythia 8.2, there is the possibility to change the default gluon recoil scheme from recoiling against the b-quark (the nominal setting, referred to here as RTB), to recoiling against the W-boson (recoilToColoured=off, referred to as RTW) [103]. Before Pythia version 8.160, the RTW was the only possibility, but it could give unphysical radiation patterns and it is now kept as an option to understand the effect this setting has in view of previous measurements. This setting changes the modelling of second and subsequent gluon emission from quarks produced by coloured resonance decays, such as the b-quark in a $t \to Wb$ process, but it has no impact for example on $Z \to b\bar{b}$ decays. A third recoil scheme has been recently made available via the UserHook

functionality of Pythia 8.2 with the top-quark itself serving as recoiler for second and subsequent gluon emission of the b-quark (referred to as RTT)[3]. The RTW and RTT setups give wider-angle gluon radiation, resulting in energy deposits that do not get clustered into the b-jet, and lower gluon-energy emission, altering the modelling of the b-quark fragmentation and hardening the b-hadron momenta. They also mildly change the W-boson p_T and the angle between the W-boson and the b-hadron resulting from the top-quark decay. The recently-developed RTT option has been considered as an additional uncertainty in the measurement, even though the implementation could only be performed based on particle-level simulation and without a dedicated tune that would normally accompany a change of setup of this nature. The change of the recoil model modifies the distribution of the momentum fraction of the b-hadron

$$x_B = \frac{1}{1 - m_W^2/m_t^2 + m_b^2/m_t^2} \frac{2p_B \cdot p_t}{m_t^2},$$

where m_w is the W-boson mass and p_t is the top four-momentum. However, the Mellin moments of this distribution derived with the RTB setup agree well with those predicted by the NLO+NLL resummation convoluted with the Kartvelishvili model tuned on ALEPH, OPAL and SLD data [71]. Therefore, the x_B distribution derived with the RTT shower is reweighted to that of the RTB simulation. The reweighting of x_B is a crude adjustment, since its distribution is controlled by the parameters of the fragmentation function and α_S^{FSR}, which are set in the A14-r_b Pythia tune, and bound by the universality of the b-quark fragmentation model. Closure tests demonstrated that the x_B reweighting provides the correct $m_{\ell\mu}$ distribution, and the correct extracted top mass, even though it does not fully address the modelling of other variables of the event. Overall, the analysis performed with the RTT shower setup as above yields a measured top-quark mass value which is larger than the one obtained with the nominal parton shower configuration by approximately 0.25 GeV. The less physically-motivated RTW setup has been also checked, and would have an effect at the same level as RTT. A one-sided shift of the measured top mass is seen with the RTT and RTW setups, but a symmetrised uncertainty of ±0.25 GeV is assigned and indicated by "recoil". This uncertainty is considered outside of the profile likelihood fit, and quoted as a separate uncertainty on the result.

The uncertainty in the modelling of initial-state radiation (ISR) is estimated by using four variations: the first two are obtained by independently changing the Pythia8 ISR renormalisation and factorisation scales up and down by a factor 2.0, the third by comparing with an alternative sample obtained doubling the h_{damp} parameter with respect to the nominal settings, the fourth one corresponds to the `Var3c` up and down variations of the Pythia8 A14 tune [104].

The modelling of the underlying event and of colour reconnection (CR) can affect the amount of radiation emitted from the b-quark, as well as modify the kinematic distribution of the b-hadron. An underlying-event uncertainty is estimated using the corresponding eigentunes of the A14 Pythia8 tune. Variations of colour reconnection parameters are also provided by the A14 eigentunes, determined from measurements of the underlying event in jet production. Samples are generated where the colour reconnection strength in the Pythia8 default model is set to its maximum value (all hadrons are reconnected) and are compared with a setting with the colour reconnection switched off. To account for the possibility of colour reconnection also affecting the top-quark decay products, a comparison with the 'Early Resonance Decay' (ERD) model is performed [105]. In this model, the top-quarks and W-bosons are allowed to decay before CR takes place, so the top-quark decay products directly participate in CR. The impact on the measured top-quark mass is found to be negligible.

[3] The code used for the RTT UserHook was provided by the Pythia authors directly.

The systematic uncertainty due to the choice of PDFs is evaluated using the PDF4LHC15 error set [106] applied to the nominal $t\bar{t}$ MC settings, and is obtained by means of event reweighting for 30 PDF replicas. The m_t value is extracted for each of the 30 cases, and the total systematic uncertainty due to this effect is computed as the sum in quadrature of the single variations.

5.3 Modelling of background processes

Several sources of uncertainty are considered for the normalisation and shape of the background contributions. The relative uncertainty in the normalisation of the soft-muon component of $t\bar{t}$ events which arises from light-hadron decays and detector background ('Soft muon fake') is derived from the uncertainty in the calibration of the misidentification rate. An additional uncertainty is derived from the difference in shape and normalisation of the $m_{\ell\mu}$ distribution of SMT mistags between $t\bar{t}$ samples generated with Powheg+Pythia8 and Sherpa. For the $t\bar{t}$ dilepton component, uncertainties in the modelling of the ISR, in the NLO matching and in the parton shower and hadronisation model are estimated in the same way as for the nominal $t\bar{t}$ sample.

An uncertainty of $+5\%/-4\%$ is applied to the total cross-section for single top-quark production [63–65]. An additional uncertainty in the amount of initial- and final-state radiation is evaluated in a manner similar to that used for $t\bar{t}$. The uncertainty in the interference between Wt and $t\bar{t}$ production at NLO is assessed by comparing the default 'diagram removal' scheme with an alternative 'diagram subtraction' scheme [59]. The uncertainty in the event generator for the t-channel is evaluated by comparing it with a sample simulated with aMC@NLO+Pythia8. The uncertainty in the parton shower and hadronisation models for the t- and Wt-channels is derived from a comparison with samples showered with Herwig++. The impact of m_t variations on the single top-quark background has been found to yield negligible effects.

An uncertainty of 30% is applied to the Z+jets background normalisation, for both its light-flavour-jet and heavy-flavour-jet ($Z+c\bar{c}$ and $Z+b\bar{b}$) components. It was validated in a control region around the Z-boson mass peak, where the normalisations of the Z+light-jet and Z+heavy-flavour-jet are simultaneously extracted with a combined fit and are found to be in agreement with the theoretical expectation for Z+jets. The uncertainty in the normalisation and in the flavour composition of W+jets is assessed using data control regions. The total normalisation and flavour fraction uncertainty is about 22% for $Wb(b)$ and Wcc, approximately 45% for Wc, and about 23% for W+light-jets.

For the multijet background, a 30% systematic uncertainty is assigned to the predicted yields, based on comparisons with data yields in control regions similar to the signal region but enriched in events from the multijet background; the e+jet and μ+jet events are treated as uncorrelated. For the small diboson background, a 50% normalisation uncertainty is assigned and includes uncertainties in the inclusive cross-section and additional jet production [107].

5.4 Detector response

Uncertainties associated with leptons arise from the trigger, reconstruction, identification, and isolation requirements, as well as the lepton momentum scale and resolution. The reconstruction, identification and isolation efficiency for electrons and muons, as well as the efficiency of the trigger, differ slightly between data and simulation and are compensated for by dedicated SFs. Efficiency SFs are derived using data and simulated samples of $Z \to \ell^+\ell^-$ ($\ell = e, \mu$), and are applied to the simulation to correct for differences. The effect of uncertainties in these SFs is propagated through the analysis. The total uncertainty in efficiency

SFs, for the high-p_{T} leptons, is $< 0.5\%$ for muons across the entire p_{T} spectrum [27] and for electrons with $p_{\mathrm{T}} > 30$ GeV, while it exceeds 1% for lower-p_{T} electrons [26]. Additional sources of uncertainty originate from the uncertainty in the corrections applied to adjust the lepton momentum scale and resolution in the simulation to match those in data. They are measured using reconstructed $Z \to \ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$ dilepton invariant mass distributions [26, 27]. To evaluate the effect of momentum scale and resolution uncertainties, the analysis is repeated with the lepton momentum varied by $\pm 1\sigma$ and with the lepton momentum smeared, respectively. A systematic uncertainty due to the electron charge misidentification is taken from Ref. [26]. Scale factors correcting for the differences in electron charge misidentification rates between data and simulation are computed using $Z \to e^+e^-$ events.

Uncertainties associated with jets arise from the efficiency of jet reconstruction and identification based on the JVT variable, as well as the JES and the jet energy resolution (JER). Although the observable $m_{\ell\mu}$ does not involve jets, the various jet uncertainties impact the analysis through the event selection. The JES and its uncertainty were derived by combining information from test-beam data, LHC collision data and simulation [22]. The JES uncertainty is about 5.5% for jets with $p_{\mathrm{T}} = 25$ GeV and decreases quickly with increasing jet p_{T}. It is below 1.5% for central jets with p_{T} in the range of approximately 100 GeV to 1.5 TeV. The highest-p_{T} jet in this analysis has an average p_{T} of around 130 GeV, with a typical range between 50 GeV and 450 GeV. The uncertainty from the soft muon jet p_{T} calibration affects the measured top-quark mass marginally, through the event selection. The magnitude of the JER uncertainty variation is parameterised in jet p_{T} and η [108], and the uncertainty is propagated by smearing the jet p_{T} in the simulation. The uncertainty in the efficiency to pass the JVT requirement is evaluated by varying the scale factors within their uncertainties [28].

The efficiencies of DJ tagging in simulated samples are corrected to match efficiencies in data. Correction scale factors are derived for jets originating from b-, c- and light-quarks separately in dedicated calibration analyses [29, 109, 110]. For jets originating from b- and c-quarks, SFs are derived as a function of p_{T}, whereas the light-jet efficiency is scaled by p_{T}- and η-dependent factors. Uncertainties in the correction scale factors are estimated by varying each source of uncertainty up and down by one standard deviation and are taken as uncorrelated between b-jets, c-jets, and light-jets. The same SFs are verified to also be applicable to jets containing soft muons by repeating the calibration procedure on a dedicated sample of events. An additional set of MC-to-MC correction factors are introduced to account for the different parton shower and hadronisation model used in the calibration samples compared to those used in this analysis. Furthermore, the efficiency of tagging the hadronic decays of τ-leptons in simulation is treated in the same way as the efficiency of tagging c-jets, and a specific uncertainty is considered for this simplified approach. An additional check is performed by changing the event selection such that there is always a DJ tagged jet other than the SMT tagged jet in the event, and the measured top-quark mass is consistent with the value measured using the nominal event selection.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is affected by uncertainties associated with lepton and jet energy scales and resolutions, which are appropriately propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Additional small uncertainties associated with the modelling of the underlying event, in particular its impact on the p_{T} scale and resolution of unclustered energy, are also taken into account [23].

The total systematic uncertainty is computed from the variance (σ^2) difference between the total uncertainty returned by the fit and the data statistics uncertainty. The plot in Figure 9 shows the ranking of the main systematic uncertainties, including the pulls and the impact of constraining the systematic uncertainties. Uncertainties in Figure 9 are from the individual nuisance parameters. They may be included within a grouped category in Table 6. The leading uncertainties are due to the modelling of the b-quark fragmentation b-hadron and decay. In particular, the BRs for direct and sequential decays are important

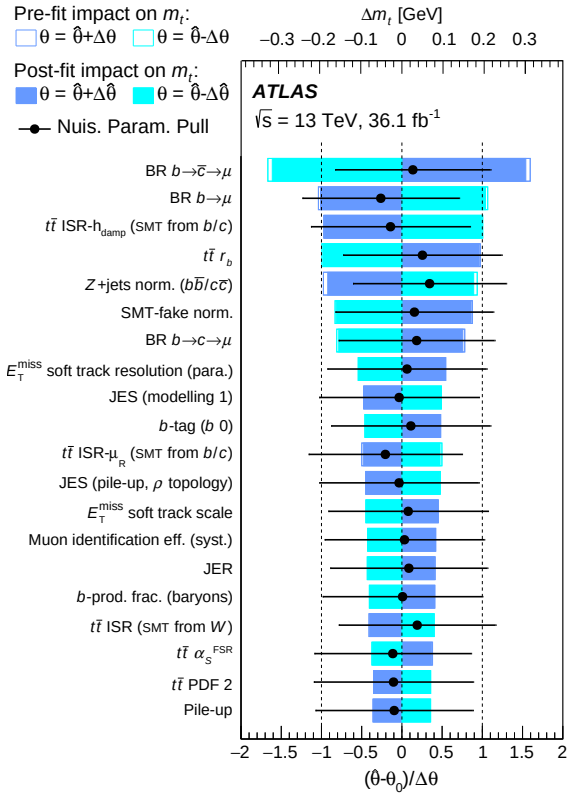


Figure 9: Ranking, from top to bottom, of the main systematic uncertainties (excluding recoil) showing the pulls and the impact of the systematic uncertainties on m_t, from the combined OS and SS binned-template profile likelihood fit to data. The θ's represent each systematic variation. The upper scale shows the uncertainty contribution to the measured top-quark mass. The PDF 2 is the set number 2 of the PDF4LHC15 error set [106].

because the SMT muon p_{T} is softer when it is produced from c-hadrons in a cascade b-hadron decay, than when it comes directly from a semileptonic b-hadron decay. The BRs also impact the charge-signs combination of the primary lepton and the soft muon. Nearly all of the main systematic uncertainties are largely uncorrelated with those dominant in previous top-quark direct reconstruction measurements [9], and the uncertainty from jet energy calibration is sub-dominant with a value of ±0.13 GeV.

6 Conclusions

A direct measurement of the top-quark mass has been performed using a technique that exploits a partial, leptonic-only, invariant mass reconstruction of the top-quark decay products. The analysis uses data corresponding to an integrated luminosity of 36.1 fb^{-1} of $\sqrt{s}$ = 13 TeV pp collisions provided by the Large Hadron Collider and recorded by the ATLAS detector, and is based on the invariant mass, $m_{\ell\mu}$, of the lepton ℓ (with $\ell = e, \mu$) from the W-boson decay and the muon μ from a semileptonic decay of a b-hadron. A binned-template profile likelihood fit to the $m_{\ell\mu}$ distribution is performed to determine the most probable top-quark mass value. The result, $m_t = 174.41 \pm 0.39$ (stat.) ± 0.66 (syst.) ± 0.25 (recoil) GeV, corresponds to the most precise single measurement to date of the top-quark mass from the direct reconstruction of its decay products by the ATLAS Collaboration, and more precise than those performed previously with

similar techniques [7, 8]. The third uncertainty arises from using a recently developed setup of the Pythia8 parton shower gluon-recoil scheme in top quark decays.

Taking into account the correlation between uncertainties, the result is consistent at the level of approximately two standard deviations with the current ATLAS combination of top-quark mass measurements from the reconstruction of the top-quark decay [9]. A similar level of consistency is found with the equivalent combination at CMS [10], while agreement with the latest Tevatron combination [11] is good. Agreement within one standard deviation is also found with the indirect prediction of the top-quark mass from global electroweak fits [12]. The main sources of systematic uncertainty are due to the modelling of top-quark pair production and b-quark fragmentation and decay, with uncertainties from backgrounds also being significant. On the other hand, the uncertainty due to the calibration of the jet energies is sub-dominant, which is advantageous in future combinations of this result with those from the standard reconstruction of the top-quark decay products.

References

[1] Particle Data Group, *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001.

[2] G. Degrassi et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, JHEP **08** (2012) 98, arXiv: 1205.6497 [hep-ph].

[3] S. Alekhin, A. Djouadi and S. Moch, *The top quark and Higgs boson masses and the stability of the electroweak vacuum*, Phys. Lett. B **716** (2012) 214, arXiv: 1207.0980 [hep-ph].

[4] P. Nason, *The Top Mass in Hadronic Collisions*, 2019, arXiv: 1712.02796 [hep-ph].

[5] G. Corcella, *The top-quark mass: challenges in definition and determination*, Front. in Phys. **7** (2019) 54, arXiv: 1903.06574 [hep-ph].

[6] ATLAS Collaboration, *Measurement of lepton differential distributions and the top quark mass in $t\bar{t}$ production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 804, arXiv: 1709.09407 [hep-ex].

[7] CDF Collaboration, *Measurement of the Top Quark Mass Using the Invariant Mass of Lepton Pairs in Soft Muon b-tagged Events*, Phys. Rev. D **80** (2009) 051104, arXiv: 0906.5371 [hep-ex].

[8] CMS Collaboration, *Measurement of the mass of the top quark in decays with a J/ψ meson in pp collisions at 8 TeV*, JHEP **12** (2016) 123, arXiv: 1608.03560 [hep-ex].

[9] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results*, Eur. Phys. J. C **79** (2019) 290, arXiv: 1810.01772 [hep-ex].

[10] CMS Collaboration, *Measurement of the top quark mass using proton–proton data at $\sqrt{s} = 7$ and 8 TeV*, Phys. Rev. D **93** (2016) 072004, arXiv: 1509.04044 [hep-ex].

[11] The Tevatron Electroweak Working Group, *Combination of CDF and D0 results on the mass of the top quark using up to 9.7 fb^{-1} at the Tevatron*, (2016), arXiv: 1608.01881.

[12] J. Haller et al., *Update of the global electroweak fit and constraints on two-Higgs-doublet models*, Eur. Phys. J. C **78** (2018) 675, arXiv: 1803.01853 [hep-ph].

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

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

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

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

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

[18] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: https://cds.cern.ch/record/2677054.

[19] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex].

[20] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[21] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph].

[22] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex].

[23] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex].

[24] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex].

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

[26] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [hep-ex].

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

[28] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex].

[29] ATLAS Collaboration,
Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV,
JHEP **08** (2018) 089, arXiv: 1805.01845 [hep-ex].

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

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

[32] ATLAS Collaboration,
The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim,
ATL-PHYS-PUB-2010-013, 2010, URL: https://cds.cern.ch/record/1300517.

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

[34] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*,
Eur. Phys. J. C **63** (2009) 189, arXiv: 0901.0002 [hep-ph].

[35] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt,
Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections,
Eur. Phys. J. C **64** (2009) 653, arXiv: 0905.3531 [hep-ph].

[36] ATLAS Collaboration, *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*,
ATL-PHYS-PUB-2011-009, 2011, URL: https://cds.cern.ch/record/1363300.

[37] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, JINST **14** (2019) P03017,
arXiv: 1812.03848 [hep-ex].

[38] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007,
arXiv: 0811.4622 [hep-ph].

[39] ATLAS Collaboration, *Measurements of charge and CP asymmetries in b-hadron decays using top-quark events collected by the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV*,
JHEP **02** (2017) 071, arXiv: 1610.07869 [hep-ex].

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

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

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

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

[44] J. Gao, C. S. Li and H. X. Zhu, *Top-Quark Decay at Next-to-Next-to-Leading Order in QCD*,
Phys. Rev. Lett. **110** (2013) 042001, arXiv: 1011.3540 [hep-ph].

[45] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081,
arXiv: hep-ph/0702198.

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

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

[48] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph].

[49] M. Cacciari, M. Czakon, M. Mangano, A. Mitov and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. B **710** (2012) 612, arXiv: 1111.5869 [hep-ph].

[50] P. Bärnreuther, M. Czakon and A. Mitov, *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t}$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv: 1204.5201 [hep-ph].

[51] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **12** (2012) 54, arXiv: 1207.0236 [hep-ph].

[52] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the quark-gluon reaction*, JHEP **01** (2013) 80, arXiv: 1210.6832 [hep-ph].

[53] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph].

[54] ATLAS Collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, ATLAS-CONF-2014-058, 2014, URL: https://cds.cern.ch/record/1951336.

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

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

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

[58] ATLAS Collaboration, *Measurements of differential cross sections of top quark pair production in association with jets in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, JHEP **10** (2018) 159, arXiv: 1802.06572 [hep-ex].

[59] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph].

[60] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph].

[61] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175.

[62] P. Z. Skands, *Tuning Monte Carlo generators: The Perugia tunes*, Phys. Rev. D **82** (2010) 074018, arXiv: 1005.3457 [hep-ph].

[63] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, Phys. Rev. D **82** (2010) 054018, arXiv: 1005.4451 [hep-ph].

[64] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, arXiv: 1001.5034 [hep-ph].

[65] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503, arXiv: 1103.2792 [hep-ph].

[66] M. Bähr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: 0803.0883 [hep-ph].

[67] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph].

[68] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand, *Parton fragmentation and string dynamics*, Phys. Rept. **97** (1983) 31.

[69] M.G. Bowler, *e^+e^- Production of heavy quarks in the string model*, Z. Phys. C **11** (1981) 169.

[70] P. Skands, S. Carrazza and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, Eur. Phys. J. C **74** (2014) 3024, arXiv: 1404.5630 [hep-ph].

[71] G. Corcella and V. Drollinger, *Bottom-quark fragmentation: Comparing results from tuned event generators and resummed calculations*, Nucl. Phys. B **730** (2005) 82, arXiv: hep-ph/0508013 [hep-ph].

[72] G. Corcella and F. Mescia, *A phenomenological study of bottom-quark fragmentation in top-quark decay*, Eur. Phys. J. C **65** (2010) 171, arXiv: 0907.5158 [hep-ph].

[73] G. Corcella and F. Mescia, *Erratum to: A phenomenological study of bottom-quark fragmentation in top-quark decay*, Eur. Phys. J. C **68** (2010) 687.

[74] G. Corcella, R. Franceschini and D. Kim, *Fragmentation uncertainties in hadronic observables for top-quark mass measurements*, Nucl. Phys. B **929** (2018) 485, arXiv: 1712.05801 [hep-ph].

[75] ALEPH Collaboration, *Study of the fragmentation of b quarks into B mesons at the Z peak*, Phys. Lett. B **512** (2001) 30, arXiv: hep-ex/0106051.

[76] DELPHI Collaboration, *A study of the b-quark fragmentation function with the DELPHI detector at LEP I and an averaged distribution obtained at the Z Pole*, Eur. Phys. J. C **71** (2011) 1557, arXiv: 1102.4748 [hep-ex].

[77] OPAL Collaboration, *Inclusive analysis of the b quark fragmentation function in Z decays at LEP*, Eur. Phys. J. C **29** (2003) 463, arXiv: hep-ex/0210031 [hep-ex].

[78] SLD Collaboration, *Precise Measurement of the b-Quark Fragmentation Function in Z^0 Boson Decays*, Phys. Rev. Lett. **84** (2000) 4300, arXiv: hep-ex/9912058 [hep-ex].

[79] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, SciPost Phys. **8** (2020) 026, arXiv: 1912.05451 [hep-ph].

[80] Heavy Flavour Averaging Group (HFAG), *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, (2014), arXiv: 1412.7515 [hep-ex].

[81] E. Lohrmann, *A Summary of Charm Hadron Production Fractions*, 2011, arXiv: 1112.3757 [hep-ex].

[82] LHCb Collaboration, *Evidence for modification of b quark hadronization in high-multiplicity pp collisions at $\sqrt{s}$ = 13 TeV*, 2022, arXiv: 2204.13042 [hep-ex].

[83] LHCb collaboration, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, J. High Energ. Phys. **04** (2013) 001, arXiv: 1301.5286 [hep-ex].

[84] LHCb Collaboration, *Measurement of f_s/f_u Variation with Proton-Proton Collision Energy and B-Meson Kinematics*, Phys. Rev. Lett. **124** (2020) 122002, arXiv: 1910.09934 [hep-ex].

[85] LHCb Collaboration, *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B^0_s decay branching fractions*, Phys. Rev. D **104** (2021) 032005, arXiv: 2103.06810 [hep-ex].

[86] LHCb Collaboration, *Measurement of b hadron production fractions in 7 TeV pp collisions*, Phys. Rev. D **85** (2012) 032008, arXiv: 1111.2357 [hep-ex].

[87] LHCb Collaboration, *Measurement of b hadron fractions in 13 TeV pp collisions*, Phys. Rev. D **100** (2019) 031102, arXiv: 1902.06794 [hep-ex].

[88] ALICE Collaboration, *Charm-quark fragmentation fractions and production cross section at midrapidity in pp collisions at the LHC*, Phys. Rev. D **105** (2022) L011103, arXiv: 2105.06335 [hep-ex].

[89] DELPHI Collaboration, *Measurement of the semileptonic b branching fractions and average b mixing parameter in Z decays*, Eur. Phys. J. C **20** (2001) 455, arXiv: hep-ex/0105080.

[90] LEP Electroweak Working Group and the SLD Heavy Flavour and Electroweak Groups, *Final input parameters for the LEP/SLD heavy flavour analyses*, LEPHF 2001-01, 2001, URL: https://www.cern.ch/LEPEWWG/heavy/.

[91] CLEO Collaboration, *Flavor-Specific Inclusive B Decays to Charm*, Phys. Rev. Lett. **80** (1998) 1150, arXiv: hep-ex/9710028 [hep-ex].

[92] CLEO Collaboration, *Measurements of $B \to D_s^+ X$ decays*, Phys. Rev. D **53** (1996) 4734.

[93] CLEO Collaboration, *Study of flavor-tagged baryon production in B decay*, Phys. Rev. D **55** (1997) 13.

[94] ALEPH Collaboration, *Observation of doubly-charmed B decays at LEP*, Eur. Phys. J. C **4** (1998) 387.

[95] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN-OPEN-2012-016, 2012, URL: https://cds.cern.ch/record/1456844.

[96] S. Baker and R. Cousins, *Clarification of the use of chi-square and likelihood functions in fits to histograms*, Nucl. Instrum. Meth. **221** (1984) 437.

[97] ATLAS Collaboration,
Evaluating statistical uncertainties and correlations using the bootstrap method,
ATL-PHYS-PUB-2021-011, 2021, URL: https://cds.cern.ch/record/2759945.

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

[99] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph].

[100] ATLAS Collaboration, *Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector*, Eur. Phys. J. C **79** (2019) 1028, arXiv: 1908.07305 [hep-ex], Erratum: Eur. Phys. J. C **80** (2020) 1092.

[101] ATLAS Collaboration,
Measurement of jet shapes in top-quark pair events at $\sqrt{s}$ = 7 TeV using the ATLAS detector, Eur. Phys. J. C **73** (2013) 2676, arXiv: 1307.5749 [hep-ex].

[102] S. Mrenna and P. Skands, *Automated parton-shower variations in Pythia 8*, Phys. Rev. D **94** (2016) 074005, arXiv: 1605.08352 [hep-ph].

[103] H. Brooks and P. Skands, *Coherent showers in decays of colored resonances*, Phys. Rev. D **100** (7 2019) 076006, arXiv: 1907.08980 [hep-ph].

[104] ATLAS Collaboration,
Simulation of top-quark production for the ATLAS experiment at $\sqrt{s}$ = 13 TeV,
ATL-PHYS-PUB-2016-004, 2016, URL: https://cds.cern.ch/record/2120417.

[105] S. Argyropoulos and T. Sjöstrand, *Effects of color reconnection on $t\bar{t}$ final states at the LHC*, JHEP **11** (2014) 043, arXiv: 1407.6653 [hep-ph].

[106] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: 1510.03865 [hep-ph].

[107] ATLAS Collaboration, *Multi-boson simulation for 13 TeV ATLAS analyses*,
ATL-PHYS-PUB-2016-002, 2016, URL: https://cds.cern.ch/record/2119986.

[108] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-015, 2015,
URL: https://cds.cern.ch/record/2037613.

[109] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018,
URL: https://cds.cern.ch/record/2306649.

[110] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s}$ = 13 TeV*, ATLAS-CONF-2018-006, 2018,
URL: https://cds.cern.ch/record/2314418.