

# DRAFT

## CMS Paper

*The content of this note is intended for CMS internal use and distribution only*

2020/04/18

Archive Hash: d0ab7b7-D

Archive Date: 2020/04/18

### Measurement of differential cross sections and charge ratios for $t$ -channel single top quark production in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

#### Abstract

A measurement is presented of differential cross sections for  $t$ -channel single top quark and antiquark production in proton-proton collisions at a centre-of-mass energy of 13 TeV by the CMS experiment at the LHC. From a data set corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ , events containing one muon or electron and two or three jets are analysed. The cross section is measured as a function of the top quark transverse momentum ( $p_T$ ), rapidity, and polarisation angle, the charged lepton  $p_T$  and rapidity, and the  $p_T$  of the W boson from the top quark decay. In addition, the charge ratio is measured differentially as a function of the top quark, charged lepton, and W boson kinematic observables. The results are found to be in agreement with standard model predictions using various next-to-leading-order event generators and sets of parton distribution functions. Additionally, the spin asymmetry, sensitive to the top quark polarisation, is determined from the differential distribution of the polarisation angle at parton level to be  $0.440 \pm 0.070$ , in agreement with the standard model prediction.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	Matthias Komm, Andrea Giammanco
PDFTitle:	Measurement of differential cross sections and charge ratios for t-channel single top quark production in proton-proton collisions at $\sqrt{s} = 13$ TeV
PDFSubject:	CMS
PDFKeywords:	CMS, physics, top quark, single top

Please also verify that the abstract does not use any user defined symbols



## 1 Introduction

The three main production modes of single top quarks and antiquarks in proton-proton (pp) collisions occur via electroweak interactions and are commonly categorised through the virtuality of the exchanged W boson four-momentum. They are called  $t$  channel ( $t$  ch) when the four-momentum is space-like,  $s$  channel when it is time-like, and W-associated (tW) when the four-momentum is on shell. At the CERN LHC, the production via the  $t$  channel has the largest cross section of the three modes whose most-relevant Born-level Feynman diagrams are shown in Fig. 1. In the rest of this paper, “quark” is used to generically denote a quark or an antiquark, unless otherwise specified.

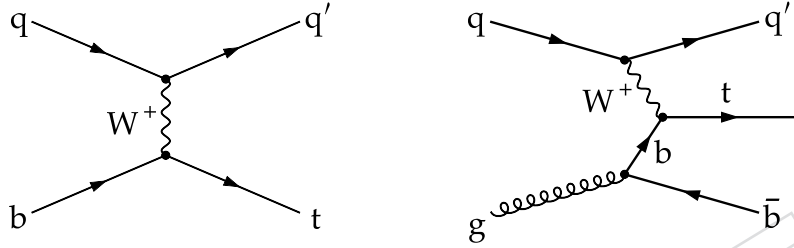


Figure 1: Born-level Feynman diagrams for single top quark production in the  $t$  channel. Corresponding diagrams also exist for single top antiquark production.

The  $t$ -channel production process was first observed by the D0 and CDF experiments at the Tevatron [1, 2]. Its inclusive cross section has been measured with high precision at the CERN LHC by the ATLAS and CMS Collaborations at  $\sqrt{s} = 7, 8$ , and 13 TeV [3–8]. Differential cross sections have been determined as well at 7 and 8 TeV [3, 5, 9].

Differential cross section measurements can contribute to constraining the effective field theory operators [10], the top quark mass, the renormalisation and factorisation scales, and the parton distribution functions (PDFs) of the proton [11]. In particular, the ratio of the  $t$ -channel top quark to antiquark production is sensitive to the ratio of the up to down quark content of the proton [12, 13]. Furthermore, differential angular distributions can be used to assess the electroweak coupling structure at the  $Wtb$  vertex. A “vector–axial-vector” (V–A) coupling is predicted in the standard model (SM), leading to the production of highly polarised top quarks [14–16]. A powerful observable to investigate the coupling structure in  $t$ -channel production is given by the top quark polarisation angle  $\theta_{\text{pol}}^*$ , defined via

$$\cos \theta_{\text{pol}}^* = \frac{\vec{p}_{q'}^* \cdot \vec{p}_\ell^*}{|\vec{p}_{q'}^*| |\vec{p}_\ell^*|}, \quad (1)$$

where the superscript signifies that the momenta of the charged lepton,  $\ell$  (muon or electron), from the top quark decay, and the spectator quark,  $q'$ , are calculated in the top quark rest frame. The normalised differential cross section as a function of  $\cos \theta_{\text{pol}}^*$  at the parton level is related to the top quark polarisation,  $P$ , as

$$\frac{1}{\sigma} \frac{d\sigma}{d \cos \theta_{\text{pol}}^*} = \frac{1}{2} \left( 1 + 2A_\ell \cos \theta_{\text{pol}}^* \right), \quad A_\ell = \frac{1}{2} P \alpha_\ell, \quad (2)$$

where  $A_\ell$  denotes the spin asymmetry and  $\alpha_\ell$  is the so-called spin-analysing power of the charged lepton [16]. The spin asymmetry and/or polarisation have been measured in pp collision data by the ATLAS and CMS Collaborations at  $\sqrt{s} = 8$  TeV using various analysis techniques [9, 17, 18].

In this paper, the differential cross section of combined single top quark and antiquark production in the  $t$  channel is measured by the CMS experiment at  $\sqrt{s} = 13$  TeV as a function of the top quark transverse momentum ( $p_T$ ), rapidity, and polarisation angle, the  $p_T$  and rapidity of the charged lepton that originates from the top quark decay, and the  $p_T$  of the W boson from the top quark decay. The spin asymmetry is further determined from the measured differential cross section with respect to the polarisation angle. Additionally, a measurement of the differential charge ratio is performed as a function of the  $p_T$  and rapidities of the top quark and charged lepton, and the  $p_T$  of the W boson. Differential cross sections are measured at both the parton and particle levels using an unfolding procedure.

The analysis strategy and the structure of the paper are outlined in the following. A brief description of the CMS detector is given in Section 2, followed by a summary of the analysed data and simulated event samples in Section 3. The reconstruction of physics objects and the event selection are detailed in Section 4. To determine the contributions from signal and backgrounds a maximum-likelihood fit (ML) is performed separately in each bin of the measurement. In the fit, shape distributions, referred to in the following as templates, are fitted to the data. For the signal and all background processes, samples of simulated events are used to determine the shape distributions, except for the templates of events containing only jets produced through the strong interaction, which are referred to as “multijet” events in this paper. The procedure to estimate the templates of multijet events based on data in a sideband region is provided in Section 5. Section 6 describes the measurement of the number of  $t$ -channel single top quark events from data through an ML fit. In the fit, statistical and experimental systematic uncertainties are profiled, where the latter encompasses uncertainties related to the reconstruction, identification, and calibration of the selected events and physics objects. The resulting distributions of the observables are validated in control and signal regions in Section 7. The fit results are input to an unfolding procedure to determine the differential cross sections and charge ratios at the parton and particle levels, as detailed in Section 8. The sources of experimental and theoretical systematic uncertainties are described in Section 9. The results are presented in Section 10 and the paper is summarised in Section 11.

## 2 The CMS detector and event reconstruction

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters (HF) extend the pseudorapidity ( $\eta$ ) coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionisation chambers embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [19].

The particle-flow (PF) algorithm [20] aims to reconstruct and identify each particle in an event with an optimised combination of information from various elements of the CMS detector. The energy of electrons is estimated from a combination of the electron momentum at the primary interaction vertex, as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of a global track estimated from reconstructed hits in the inner tracker and muon systems. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the

matching ECAL and HCAL energy deposits. Finally, the energy of neutral hadrons is obtained from the corresponding ECAL and HCAL energy deposits. In the regions  $|\eta| > 3$ , electromagnetic and hadronic shower components are identified in the HF.

Events of interest are selected using a two-tiered trigger system [21]. The first level, composed of custom hardware processors, uses information from the calorimeters and muon detectors whereas a version of the full event reconstruction software optimised for fast processing is performed at the second level, which runs on a farm of processors.

The missing transverse momentum vector,  $\vec{p}_T^{\text{miss}}$ , is defined as the projection onto the plane perpendicular to the beams of the negative vector momentum sum of all PF candidates in an event. Its magnitude is referred to as  $p_T^{\text{miss}}$ .

### 3 Data set and simulated samples

The analysed pp collision data set was recorded in 2016 by the CMS detector and corresponds to an integrated luminosity of  $35.9 \text{ fb}^{-1}$  [22]. Events were triggered by requiring at least one isolated muon candidate with  $p_T > 24 \text{ GeV}$  and  $|\eta| < 2.4$  or one electron candidate with  $p_T > 32 \text{ GeV}$  and  $|\eta| < 2.1$ , with additional requirements [23] that select genuine electrons with an efficiency of about 80%.

Various samples of simulated events are used in this measurement to evaluate the detector resolution, efficiency, and acceptance, estimate the contributions from background processes, and determine the differential cross sections at the parton and particle levels.

Single top quark events in the  $t$  channel are simulated at next-to-leading order (NLO) in the four-flavour scheme (4FS) with POWHEG v2 [24, 25] interfaced with PYTHIA v8.212 [26] for the parton shower simulation, using the CUETP8M1 [27] tune interfaced with MADSPIN [28] for simulating the top quark decay. For comparison, alternative NLO  $t$ -channel samples have been generated in the 4FS and five-flavour scheme (5FS), using MADGRAPH5\_aMC@NLO v2.2.2 [29] interfaced with PYTHIA.

The POWHEG v2 generator is also used to simulate events from top quark pair production ( $t\bar{t}$ ) at NLO. Parton showering is simulated with PYTHIA using the CUETP8M2T4 tune [30]. The production of single top quark events via the  $tW$  channel is simulated at NLO using POWHEG v1 [31] in the 5FS interfaced with PYTHIA using the CUETP8M1 tune for the parton shower simulation. The overlap with top quark pair production is removed by applying the diagram removal scheme [32]. Samples of  $W$ +jets events are generated with MADGRAPH5\_aMC@NLO v2.3.3 at NLO, and interfaced with PYTHIA using the CUETP8M1 tune. The production of leptonically decaying  $W$  bosons in association with jets is simulated with up to two additional partons at the matrix element level, and the FxFx scheme [33] is used for jet merging. Lastly,  $Z/\gamma^*$ +jets events are generated with MADGRAPH5\_aMC@NLO v2.2.2 at leading order (LO), interfaced with PYTHIA using the MLM jet matching scheme [34].

In these simulated samples, the NNPDF3.0 [35] NLO set is used as the default PDF, and a nominal top quark mass of  $172.5 \text{ GeV}$  is chosen where applicable. The simulated events are overlaid with additional collision interactions (“pileup”) according to the distribution inferred from the data. All generated events undergo a full GEANT4 [36] simulation of the detector response.

The  $t$ -channel cross section in pp collisions at  $\sqrt{s} = 13 \text{ TeV}$  is predicted to be  $\sigma_t = 136.0^{+5.4}_{-4.6} \text{ pb}$  for the top quark and  $\sigma_{\bar{t}} = 81.0^{+4.1}_{-3.6} \text{ pb}$  for the top antiquark, calculated for a top quark mass

of 172.5 GeV at NLO in quantum chromodynamics (QCD) using the HATHOR v2.1 [11, 37] program. The PDF and the strong coupling constant ( $\alpha_s$ ) uncertainties are calculated using the PDF4LHC prescription [38, 39] with the MSTW2008 NLO 68% confidence level [40, 41], CT10 [42] NLO, and NNPDF2.3 [43] NLO PDF sets, and are added in quadrature with the renormalisation and factorisation scale uncertainty. The simulated samples of single top quark and antiquark events employed in this measurement—generated with similar settings—were normalised using the predicted cross sections above. Predictions at next-to-next-to-leading order are available as well [12] and are 3% smaller than the corresponding cross sections at NLO. However, these are not utilised since they have been calculated using a different PDF set and top quark mass value.

## 4 Event selection

Proton-proton collision events containing one isolated muon or electron and two or three jets are analysed. This signature selects events where the W boson from a single top quark decays into a charged lepton and a neutrino. One of the selected jets is expected to stem from the hadronisation of a bottom quark that originates from the top quark decay. Another jet ( $j'$ ) from a light-flavoured quark (up, down, or strange) is expected from the spectator quark (labelled  $q'$  in Fig. 1) that is produced in association with the top quark. The jet from the spectator quark is characteristically found at relatively low angles with respect to the beam axis.

The reconstructed vertex with the largest value of summed physics-object  $p_T^2$  is taken to be the primary pp interaction vertex. The physics objects are the jets, clustered using the jet finding algorithm described in Refs. [44, 45] with the tracks assigned to the vertex as inputs, and the negative vector  $\vec{p}_T$  sum of those jets.

Muon candidates are accepted if they have  $p_T > 26$  GeV,  $|\eta| < 2.4$ , and pass the following identification requirements optimised for the selection of genuine muons. A global muon track must have a track fit with a  $\chi^2$  per degree of freedom  $< 10$ , have hits in the silicon tracker and muon systems, including at least six in the tracker, of which at least one must be in the pixel detector. Additionally, track segments are required in at least two muon stations to suppress signals from hadronic showers spilling into the muon system. Muon candidates are required to be isolated with a relative isolation parameter  $I_{\text{rel}}^\mu < 6\%$ , which is defined as the scalar sum of the transverse energies  $E_T$  deposited in the ECAL and HCAL within a cone of radius  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ , divided by the muon  $p_T$ . The transverse energy is defined as  $E_T = E \sin(\theta)$  with  $E$  and  $\theta$  being the energy and polar angle, respectively, of photons and charged and neutral hadrons. Here,  $\Delta\eta$  and  $\Delta\phi$  are the pseudorapidity and azimuthal angle, respectively, measured relative to the muon direction. The isolation parameter is corrected by subtracting the energy deposited by pileup, which is estimated from the energy deposited by charged hadrons within the isolation cone that are associated with pileup vertices [46].

Electron candidates are required to have  $p_T > 35$  GeV,  $|\eta| < 1.48$ , and fulfil a set of additional quality requirements as follows: the distance between the matched ECAL cluster position and the extrapolated electron track has to be within  $|\Delta\eta| < 3.08 \times 10^{-3}$  and  $|\Delta\phi| < 8.16 \times 10^{-2}$ ; the absolute difference between the inverse of the energy estimated from the ECAL cluster and the inverse of the electron track momentum must be less than  $12.9 \text{ MeV}^{-1}$ ; the ratio of the HCAL to the ECAL energy associated with the electron is required to be less than 4.14%; the energy-weighted lateral width of the electron shower in the ECAL along the  $\eta$  direction is restricted to  $< 9.98 \times 10^{-3}$ . Electrons from photon conversions are suppressed by requiring that the corresponding track has no missing hits in the inner layers of the tracker and that they do



not stem from a photon conversion vertex. Electron candidates have to be isolated using the so-called effective-area-corrected relative isolation parameter [47] by requiring  $I_{\text{rel}}^e < 5.88\%$ . This parameter is defined similarly to the muon isolation parameter as the sum of the charged and neutral particle energies within a cone of  $\Delta R < 0.3$  around the electron candidate, divided by the electron  $p_T$ . The relative contribution from pileup is estimated as  $A_{\text{eff}} \rho$  and subtracted from the isolation parameter, where  $A_{\text{eff}}$  denotes an  $\eta$ -dependent effective area, and  $\rho$  is the median of the  $E_T$  density in a  $\delta\eta \times \delta\phi$  region calculated using the charged particle tracks associated with the pileup vertices.

The selected muon (electron) candidate has to be within 2.0 (0.5) mm in the transverse plane and 5.0 (1.0) mm along the beam direction of the primary vertex.

Electron candidates with showers in the ECAL endcap ( $1.48 < |\eta| < 2.5$ ) are not used in the measurement because of the higher background consisting of hadrons misidentified as electrons and of electrons originating from decays of heavy-flavour hadrons, which is found to be about four times larger compared to the ECAL barrel region.

Events are rejected if additional muon or electron candidates passing looser selection criteria are present. The selection requirements for these additional muons/electrons are as follows: looser identification and isolation criteria,  $p_T > 10$  (15) GeV for muons (electrons), and  $|\eta| < 2.5$ .

The transverse W boson mass is calculated from the formula

$$m_T(W) = \sqrt{2p_T^\ell p_T^{\text{miss}} [1 - \cos(\phi^\ell - \phi^{\text{miss}})]} \quad (3)$$

using the  $p_T$  and the  $\phi$  of the charged lepton and  $\vec{p}_T^{\text{miss}}$ .

Jets are reconstructed from PF candidates and clustered by applying the anti- $k_T$  algorithm [44] with a distance parameter of 0.4 using the FASTJET package [45]. The influence of pileup is mitigated using the charged hadron subtraction technique [48]. The jet momentum is determined as the vectorial sum of all particle momenta in the jet. An offset correction is applied to the jet  $p_T$  to account for contributions from pileup. Further corrections are applied to account for the nonuniform detector response in  $\eta$  and  $p_T$  of the jets. The corrected jet momentum is found from simulation to be within 2 to 10% of the true momentum over the whole  $p_T$  spectrum and detector acceptance. The corrections are propagated to the measured  $\vec{p}_T^{\text{miss}}$ . A potential overlap of a jet with the selected lepton is removed by ignoring jets that are found within a cone of  $\Delta R < 0.4$  around a selected lepton candidate. The analysis considers jets within  $|\eta| < 4.7$  whose calibrated  $p_T$  is greater than 40 GeV, with the exception of the HCAL-HF transition region ( $2.7 < |\eta| < 3$ ) in which jets must have a  $p_T$  of at least 50 GeV to reduce the contribution from detector noise. The event is accepted for further analysis if two or three jets are present.

To reduce the large background from W+jets events, a b tagging algorithm based on a multivariate analysis (MVA) called “combined MVA” [49], which combines the results from various other b tagging algorithms, is used for identifying jets produced from the hadronisation of b quarks within the acceptance of the silicon tracker ( $|\eta| < 2.4$ ). A tight selection is applied on the discriminant of the algorithm, which gives an efficiency of  $\approx 50\%$  for jets originating from true b quarks and misidentification rates of  $\approx 0.1\%$  for light jets from u, d, or s quarks or gluons and  $\approx 3\%$  for jets from c quarks, as determined from simulation.

Corrections are applied to the simulated events to account for known differences with respect to data. Lepton trigger, reconstruction, and identification efficiencies are estimated with a “tag-and-probe” method [50] from  $Z/\gamma^*$ +jets events for data and simulation from which corrections

are derived in bins of lepton  $\eta$  and  $p_T$ . The b tagging performance in simulation is corrected to match the tagging efficiency observed in data, using scale factors that depend on the  $p_T$  and  $\eta$  of the selected jets. The scale factors are estimated by dedicated analyses performed with independent data samples [49]. In particular, the mistagging rate of non-b jets in data is determined using the “negative-tag” method [51]. A smearing of the jet momenta is applied to account for the known difference in jet energy resolution in simulation compared to data. The profile of pileup interactions is reweighted in simulation to match the one in data derived from the measured instantaneous luminosity.

To classify signal and control samples of events, different event categories are defined, denoted “ $NjMb$ ”, where  $N$  is the total number of selected jets (2 or 3) and  $M$  is the number of those jets passing the b tagging requirement (0, 1, or 2). The 2j1b category has the highest sensitivity to the signal yield, whereas the 2j0b and 3j2b categories, enriched in background processes with different compositions, are used to assess the background modelling.

One top quark candidate is reconstructed per event in the 2j1b signal category assuming  $t$ -channel single top quark production. The procedure commences by first reconstructing the W boson. The component of the neutrino candidate momentum along the beam direction  $p_z$  is found by imposing a W boson mass constraint (80.4 GeV) on the system formed by the charged lepton and  $\vec{p}_T^{\text{miss}}$ , the latter being interpreted as the projection in the transverse plane of the four-momentum of the unknown neutrino, as in Ref. [52]. The four-momentum of the top quark candidate (from which its mass,  $p_T$ , and rapidity are derived) is then calculated as the vector sum of the four-momenta of the charged lepton, the b-tagged jet, and the neutrino candidate. The other (nontagged) jet is interpreted as originating from the spectator quark, which recoils against the W boson.

## 5 Multijet background estimation

Since the probability for a simulated multijet event to mimic the final state of the signal process is very small, it becomes impractical to simulate a sufficiently large number of events for this background. Therefore, the background from multijet events in the analysis phase space region is estimated in a two-step procedure based on data in a sideband region. First, templates of the  $m_T(W)$  distribution from multijet events are obtained from data in a sideband region. Their normalisations are then estimated in a second step through a template-based ML fit to the events in the 2j1b and 3j2b categories, simultaneously with the number of signal events, as described in Section 6. In this section, a dedicated ML fit is discussed that is performed on events in the 2j0b category only for validating the procedure. The outcome of this ML fit is not used further in the measurement.

In the muon channel, the sideband region is defined by inverting the muon isolation requirement ( $I_{\text{rel}}^\mu > 20\%$ ), which results in a region dominated by multijet events. In the electron channel, the electron candidate is required to fail loose identification criteria, yielding a sideband region consisting not only of nonisolated electrons but also of electrons that fail the photon conversion criteria or are accompanied by large amounts of bremsstrahlung, thus reflecting a combination of various effects. The templates used in the ML fit are determined for this category by subtracting the contamination from other processes, estimated using simulation and which amounts to about 10 (5)% in the muon (electron) channel, from the data.

The template shapes have been validated for various observables in the 2j0b W+jets control category where the fraction of selected multijet events amounts to approximately 10 (20)% for muon (electron) events, which is comparable to those in the signal category. The  $m_T(W)$  dis-



tributions are shown in Fig. 2 for the muon (left) and electron (right) channel after the multijet templates (extracted from data) and the templates of the processes with prompt leptons (extracted from the simulated events) have been normalised to the result of a dedicated ML fit using only events in the 2j0b category. This dedicated fit encompasses only two components, which are the multijet template whose yield is unconstrained in the fit, and all other processes grouped together, with a constraint of  $\pm 30\%$  on their combined yield using a log-normal prior. The fit is performed while simultaneously profiling the impact of experimental systematic uncertainties (as discussed in Section 9) affecting the yield and shape of the templates. After the fit, the derived multijet templates and the simulated samples in both channels are found to describe the distributions of data well, thus validating the procedure for estimating the contribution of multijet events from data. For the measurement, the normalisations of the multijet templates in the 2j1b and 3j2b categories are estimated using a different procedure, as described in Section 6.

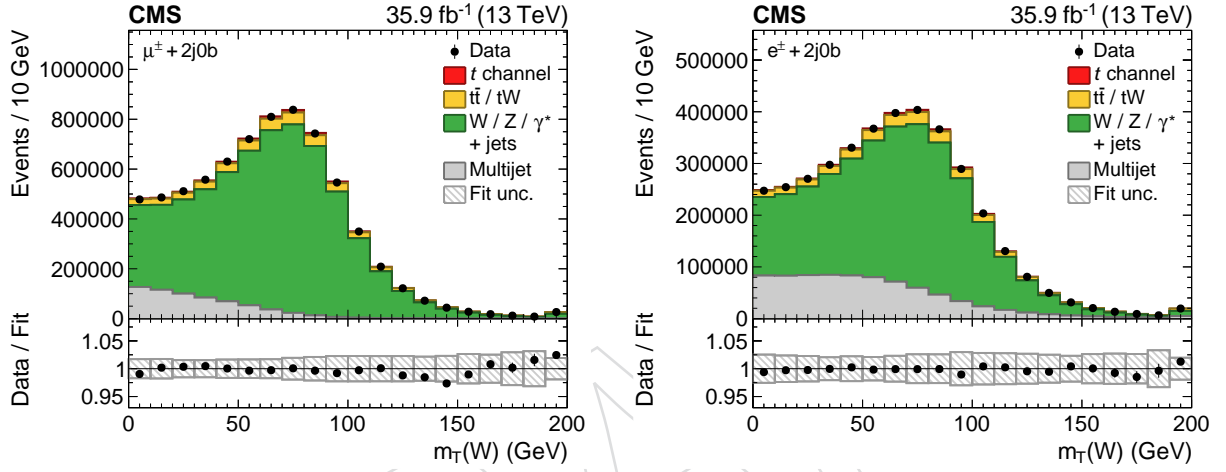


Figure 2: Distributions of the transverse W boson mass in the 2 jets, 0 b tag control category for the (left) muon and (right) electron channels after scaling the simulated and multijet templates to the result of a dedicated ML fit performed on this category of events. The hatched band displays the fit uncertainty. The lower plots give the ratio of the data to the fit results. The right-most bins include the event overflows.

## 6 Signal yield estimation

The number of  $t$ -channel single top quark events in data is determined from an ML fit using the distributions of  $m_T(W)$  and of two boosted decision tree (BDT) discriminants in the 2j1b category, and the  $m_T(W)$  distribution in the 3j2b category. Simultaneously, the background yields and the impact of the experimental systematic uncertainties, modelled using nuisance parameters that influence yield and shape, are profiled.

The first BDT, labelled  $\text{BDT}_{t\text{-ch}}$ , has been trained separately on muon and electron events to discriminate  $t$ -channel single top quark events from  $t\bar{t}$ ,  $W$ +jets, and multijet events using corresponding samples of simulated events. The following five observables have been chosen as input:

- the absolute value of the pseudorapidity of the untagged jet,  $|\eta(j')|$ ;
- the reconstructed top quark mass,  $m_{\ell\nu b}$ ;
- the transverse W boson mass,  $m_T(W)$ ;

- the distance in  $\eta$ - $\phi$  space ( $\Delta R$ ) between the b-tagged and the untagged jet,  $\Delta R(b, j')$ ;
- the absolute difference in pseudorapidity between the b-tagged jet used to reconstruct the top quark and the selected lepton,  $|\Delta\eta(b, \ell)|$ .

These have been selected based on their sensitivity for separating signal from background events, while exhibiting low correlations with the observables used to measure the differential cross sections. The resulting distribution of the  $\text{BDT}_{t\text{-ch}}$  discriminant is presented in Fig. 3 (left).

The  $\text{BDT}_{t\text{-ch}}$  discriminant shapes of the W+jets and  $t\bar{t}$  backgrounds are found to be very similar. To obtain sensitivity in the fit to both backgrounds individually, a second BDT, labelled  $\text{BDT}_{t\bar{t}/W}$ , has been trained separately on muon and electron events to classify events only for these two processes using the following six input observables:  $m_{\ell\nu b}$ ;  $p_T^{\text{miss}}$ ;  $\Delta R(b, j')$ ;  $|\Delta\eta(b, \ell)|$ ; the W boson helicity angle,  $\cos\theta_W^*$ , defined as the angle between the lepton momentum and the negative of the top quark momentum in the W boson rest frame [16]; and the event shape C, defined using the momentum tensor

$$S^{ab} = \frac{\sum_i^{\text{jets}, \ell, \vec{p}_T^{\text{miss}}} p_i^a p_i^b}{\sum_i^{\text{jets}, \ell, \vec{p}_T^{\text{miss}}} |\vec{p}_i|^2}, \quad (4)$$

as  $C = 3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)$ , where  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  denote the eigenvalues of the momentum tensor  $S^{ab}$  with  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ . In the two most extreme cases, the event shape C vanishes for perfectly back-to-back dijet events ( $C = 0$ ) and reaches its maximum ( $C = 1$ ) if the final-state momenta are distributed isotropically. For the measurement, the  $\text{BDT}_{t\bar{t}/W}$  discriminant is evaluated only in the phase space region defined by  $m_T(W) > 50$  GeV and  $\text{BDT}_{t\text{-ch}} < 0$ , which is found to be largely dominated by background events. Thus, the  $\text{BDT}_{t\bar{t}/W}$  input observables do not have to be selected explicitly such that they possess low correlation with the observables used to measure the differential cross sections. The resulting  $\text{BDT}_{t\bar{t}/W}$  discriminant distribution is displayed in Fig. 3 (right).

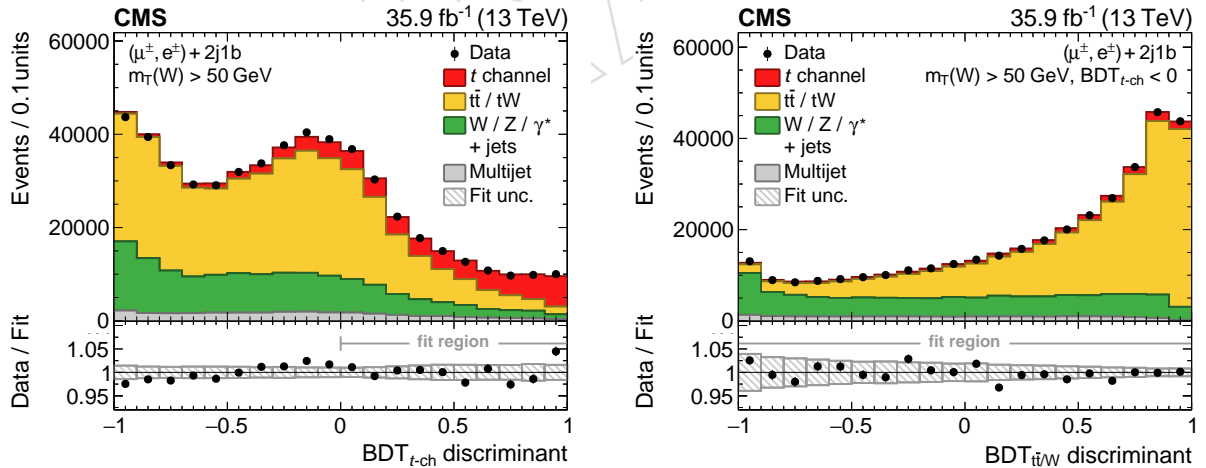


Figure 3: Distributions of the BDT discriminants in the 2 jets, 1 b tag category: (left)  $\text{BDT}_{t\text{-ch}}$  trained to separate signal from background events; (right)  $\text{BDT}_{t\bar{t}/W}$  trained to separate  $t\bar{t}$  from W+jets events in a background-dominated category. Events in the muon and electron channels have been summed. The predictions have been scaled to the result of the inclusive ML fit and the hatched band displays the fit uncertainty. The regions of the distributions used in the fits are indicated in the lower panels, which show the ratio of the data to the fit result.

The ML fit is performed using the following four distributions from events in various categories:

- the  $m_T(W)$  distribution for events with  $m_T(W) < 50$  GeV in the 2j1b category, which is particularly sensitive to the number of multijet events;
- the  $\text{BDT}_{t\bar{t}/W}$  discriminant distribution for events with  $m_T(W) > 50$  GeV and  $\text{BDT}_{t\text{-ch}} < 0$  in the 2j1b category, which defines a region enriched in  $t\bar{t}$  and  $W$ +jets but depleted of signal and multijet events;
- the  $\text{BDT}_{t\text{-ch}}$  discriminant distribution for events with  $m_T(W) > 50$  GeV and  $\text{BDT}_{t\text{-ch}} > 0$  in the 2j1b category, which is enriched in signal events;
- the  $m_T(W)$  distribution in the 3j2b category, which provides additional sensitivity to the  $t\bar{t}$  yield, and thus further reduces the correlation between the estimated yields.

The  $m_T(W)$  distributions in the 2j1b and 3j2b categories are shown in Fig. 4 on the left and right, respectively. In the fit, each distribution is split in two by separating events depending on the charge of the selected muon or electron in the event. This results in eight distributions per lepton channel and thus 16 distributions in the  $\mu/e$  combined fit. A coarser equidistant binning of the distributions, as opposed to the one shown in Figs. 3 and 4, is used in the ML fits to prevent cases where single bins are depleted of background estimates as follows: four bins are used for each of the  $m_T(W)$  and  $\text{BDT}_{t\text{-ch}}$  distributions in the 2j1b category; eight bins are used for the  $\text{BDT}_{t\bar{t}/W}$  distribution; and ten bins are used for the  $m_T(W)$  distribution in the 3j2b category.

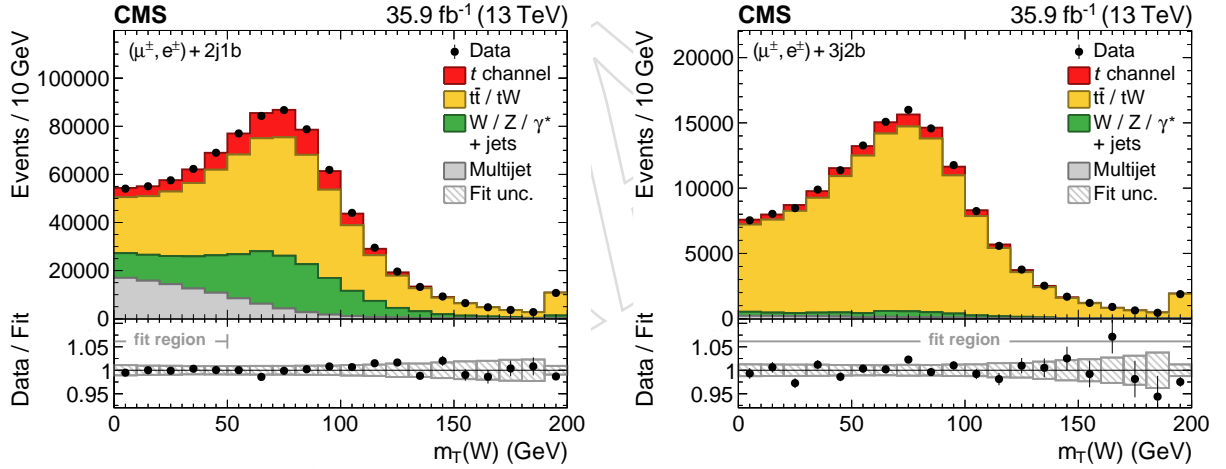


Figure 4: Distributions of the transverse W boson mass for events in the (left) 2 jets, 1 b tag and (right) 3 jets, 2 b tags categories. Events in the muon and electron channels have been summed. The predictions have been scaled to the result of the inclusive ML fit and the hatched band displays the fit uncertainty. The regions of the distributions used in the fits are indicated in the lower panels, which show the ratio of the data to the fit result. The right-most bins include the event overflows.

The yields of  $t$ -channel single top quark and antiquark events are measured independently. Background events containing top quarks ( $t\bar{t}$ ,  $tW$ ) are grouped together, and only their total yield is estimated. The top quark background yield is constrained using a log-normal prior with a width of  $\pm 10\%$  to account for the uncertainty in the theoretical  $t\bar{t}$  and  $tW$  production cross sections, and the uncertainty when two out of the four jets expected from semileptonic  $t\bar{t}$  production are not within the acceptance, as is the case in the 2j1b category. The electroweak background processes,  $W$ +jets and  $Z/\gamma^*$ +jets, are grouped together as well, and an uncertainty

of  $\pm 30\%$  in their combined yield is applied using a log-normal prior constraint. This is motivated by the theoretical uncertainty in the modelling of the  $W$  and  $Z/\gamma^*$  production rates in association with two or more (heavy-flavour) jets [53, 54]. The yields of multijet events are assumed to be independent per lepton type and event category. Their yields are constrained by a log-normal prior with a width of  $\pm 100\%$  with respect to the template normalisations obtained from data in the sideband regions. In addition, an uncertainty in the predicted lepton charge ratio per background process, accounting for charge misreconstruction and uncertainties in the charge ratio [55], is taken into account using a Gaussian prior with a width of  $\pm 1\%$  in the fit, for a total of 14 fit parameters. The impact of the finite number of simulated events on the templates is accounted for by employing the “Barlow–Beeston-lite” method [56].

Experimental systematic uncertainties, as detailed in Section 9, are profiled in the fit simultaneously with the yields and charge ratios. Each source is assigned a nuisance parameter according to which the shape and yield of the fit templates are modified.

The resulting event yields from a simultaneous fit to the data in the muon and electron channels are listed in Table 1. Overall, the distributions used in the fit, shown in Figs. 3 and 4, are found to be well modelled by the samples of simulated events and the multijet templates from data after normalising them to the fit result.

Table 1: Measured and observed event yields in the 2j1b category for each lepton channel and charge. The uncertainties in the yields are the combination of statistical and experimental systematic uncertainties.

Process	$\mu^+$	$\mu^-$	$e^+$	$e^-$
$W/Z/\gamma^* + \text{jets}$	$72\,000 \pm 6\,800$	$62\,800 \pm 5\,600$	$33\,400 \pm 3\,200$	$30\,700 \pm 2\,800$
$t\bar{t}/tW$	$142\,400 \pm 2\,400$	$143\,400 \pm 2\,500$	$84\,500 \pm 1\,400$	$84\,800 \pm 1\,500$
Multijet	$35\,150 \pm 550$	$35\,710 \pm 760$	$13\,500 \pm 1\,000$	$12\,700 \pm 1\,000$
$t$ channel (top quark)	$34\,400 \pm 1\,500$	$10 \pm 3$	$17\,720 \pm 820$	$27 \pm 2$
$t$ channel (top antiquark)	$13 \pm 2$	$21\,600 \pm 1\,600$	$25 \pm 3$	$11\,460 \pm 880$
Total	$284\,100 \pm 5\,800$	$263\,700 \pm 4\,600$	$149\,300 \pm 2\,400$	$139\,700 \pm 2\,200$
Data	283 391	260 044	148 418	138 781

For each differential cross section measurement, the observable of interest is divided into intervals, discussed in Section 8, and a fit is performed in which the signal and background yields can vary independently in each of the intervals. The likelihood  $L$  to be maximised in such fits can be expressed as

$$\ln \left( L(\vec{\beta}, \vec{v}, \vec{R}) \right) = - \sum_k^{\text{dist}} \sum_j^{\text{int}} \sum_i^{\text{bins}} \left( d_{kji} \ln p_{kji}(\vec{\beta}_j, \vec{v}, \vec{R}) - p_{kji}(\vec{\beta}_j, \vec{v}, \vec{R}) \right) + \text{constraints}, \quad (5)$$

where  $d$  denotes the number of observed events and  $p$  is the estimated yield. The summation over  $k$  denotes the 16 distributions (“dist”),  $j$  denotes the interval (“int”) in the observable (e.g. for the top quark  $p_T$ : 0–50 GeV, 50–80 GeV, 80–120 GeV, 120–180 GeV, and 180–300 GeV), and  $i$  denotes a bin in one of the 16 distributions per interval. The prediction  $\vec{p}_{kj}$ , which includes all bins  $i$  for distribution  $k$  and interval  $j$ , is given by

$$\begin{aligned} \vec{p}_{kj}(\vec{\beta}_j, \vec{v}, \vec{R}) = & \beta_{t,j} \vec{T}_{t,kj}^{t\text{-ch}}(\vec{v}) + \beta_{\bar{t},j} \vec{T}_{\bar{t},kj}^{t\text{-ch}}(\vec{v}) \\ & + \beta_{t\bar{t}/tW,j} \vec{T}_{kj}^{t\bar{t}/tW}(R_j, \vec{v}) + \beta_{W/Z/\gamma^*+\text{jets},j} \vec{T}_{kj}^{W/Z/\gamma^*+\text{jets}}(R_j, \vec{v}) \\ & + \beta_{\text{multijet},j}(\ell, r) \vec{T}_{kj}^{\text{multijet}}(R_j(\ell, r), \vec{v}), \end{aligned} \quad (6)$$

where  $\vec{v}$  are the nuisance parameters,  $R$  the charge ratios of each background process, and  $\beta$  the normalisations of the templates  $\vec{T}$ , which are independent per lepton flavour  $\ell$  and category  $r \in \{2j1b, 3j2b\}$  for the multijet templates. The profiling of systematic uncertainties leads to a correlation between the  $t$ -channel top quark and antiquark yields in the same interval of about 20–30%. These correlations are propagated to the differential cross sections for each top quark charge, and are accounted for when calculating their sum and ratio.

Since the kinematic selection of electron events is restricted to  $p_T > 35 \text{ GeV}$  and  $|\eta| < 1.48$ , which is tighter than for muon events ( $p_T > 26 \text{ GeV}$ ,  $|\eta| < 2.4$ ), the signal yields in the lowest interval of the lepton  $p_T$  and in the highest two intervals of the lepton rapidity spectra are estimated from the muon channel alone in the combined  $\mu/e$  fit.

## 7 Validation of signal and background modelling

The distributions of the observables that are unfolded are validated by comparing the predictions to the data in a background-dominated as well as in a signal-enriched region before unfolding. Both regions are defined for events in the 2j1b category that also satisfy  $m_T(W) > 50 \text{ GeV}$  to suppress the contribution from multijet production. The modelling of the  $t\bar{t}/tW$  and  $W/Z/\gamma^* + \text{jets}$  backgrounds is validated in a background-dominated region obtained from events having  $\text{BDT}_{t\text{-ch}} < 0$ . To validate the modelling of the  $t$ -channel process, events are instead required to pass  $\text{BDT}_{t\text{-ch}} > 0.7$ , resulting in a sample enriched in signal events. These two regions and their selections are only defined and applied for validation purposes, and not used for measuring the differential cross sections for which the individual fit results are used in the unfolding instead.

The resulting distributions in both regions for all six observables that are unfolded are shown in Figs. 5 and 6 after the predictions have been scaled to the inclusive fit result. Overall good agreement between the data and the fit result is observed in the background-dominated region, thus validating the modelling of the  $t\bar{t}/tW$  and  $W/Z/\gamma^* + \text{jets}$  backgrounds. In the signal region, reasonable agreement is also observed.

## 8 Unfolding

The distributions from reconstructed events are affected by the detector resolution, selection efficiencies, and kinematic reconstruction, which lead to distortions with respect to the corresponding distributions at the parton or particle levels. The size of these effects varies with the event kinematics. In order to correct for these effects and determine the parton- and particle-level distributions, an unfolding method is applied to the reconstructed distributions. In this analysis, the TUNFOLD algorithm [57] is chosen, which treats unfolding as a minimisation problem of the function

$$\chi^2 = (\vec{y} - R\epsilon\vec{x})^T V_y^{-1} (\vec{y} - R\epsilon\vec{x}) + \underbrace{\tau^2 \|L(\vec{x} - \vec{x}_0)\|^2}_{\text{regularisation}} + \lambda \sum_i (\vec{y} - R\epsilon\vec{x})_i, \quad (7)$$

where  $\vec{y}$  denotes the measured yields in data,  $V_y$  is the covariance matrix of the measured yields, and  $\vec{x}$  is the corresponding differential cross section at parton or particle level. The matrices  $R$  and  $\epsilon$  denote the transition probability and selection efficiencies, respectively, both estimated from simulation. The signal yields and covariances are estimated through ML fits using the  $m_T(W)$ ,  $\text{BDT}_{t\bar{t}/W}$ , and  $\text{BDT}_{t\text{-ch}}$  distributions, as detailed in Section 6.

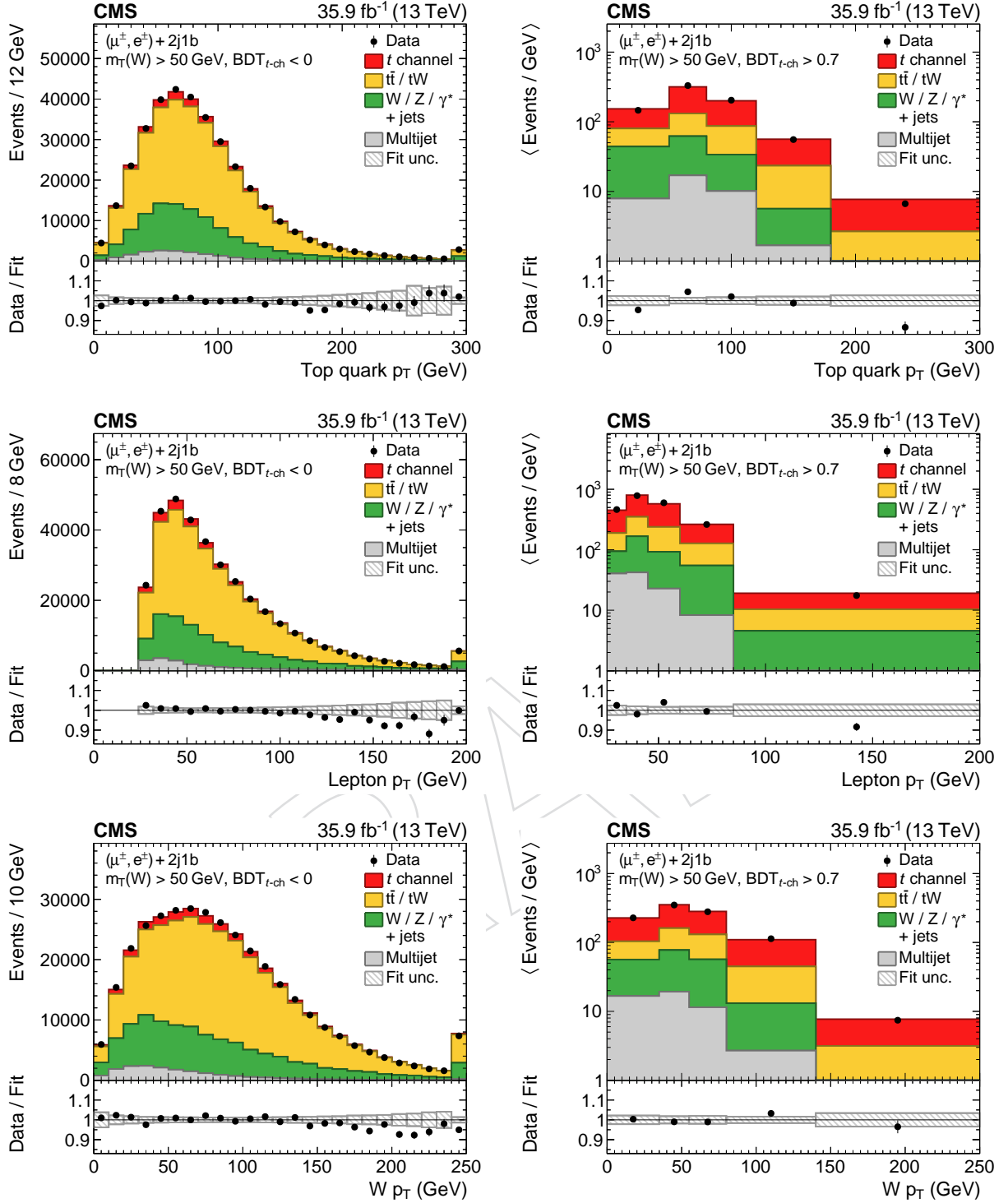


Figure 5: Distributions of the observables in a (left column) background-dominated and a (right column) signal-enriched region for events passing the 2 jets, 1 b tag selection: (upper row) top quark  $p_T$ ; (middle row) charged lepton  $p_T$ ; (lower row) W boson  $p_T$ . Events in the muon and electron channels have been summed. The predictions have been scaled to the result of the inclusive ML fit and the hatched band displays the fit uncertainty. The plots on the left give the number of events per bin, while those on the right show the number of events per bin divided by the bin width. The lower panel in each plot gives the ratio of the data to the fit results. The right-most bins include the event overflows.



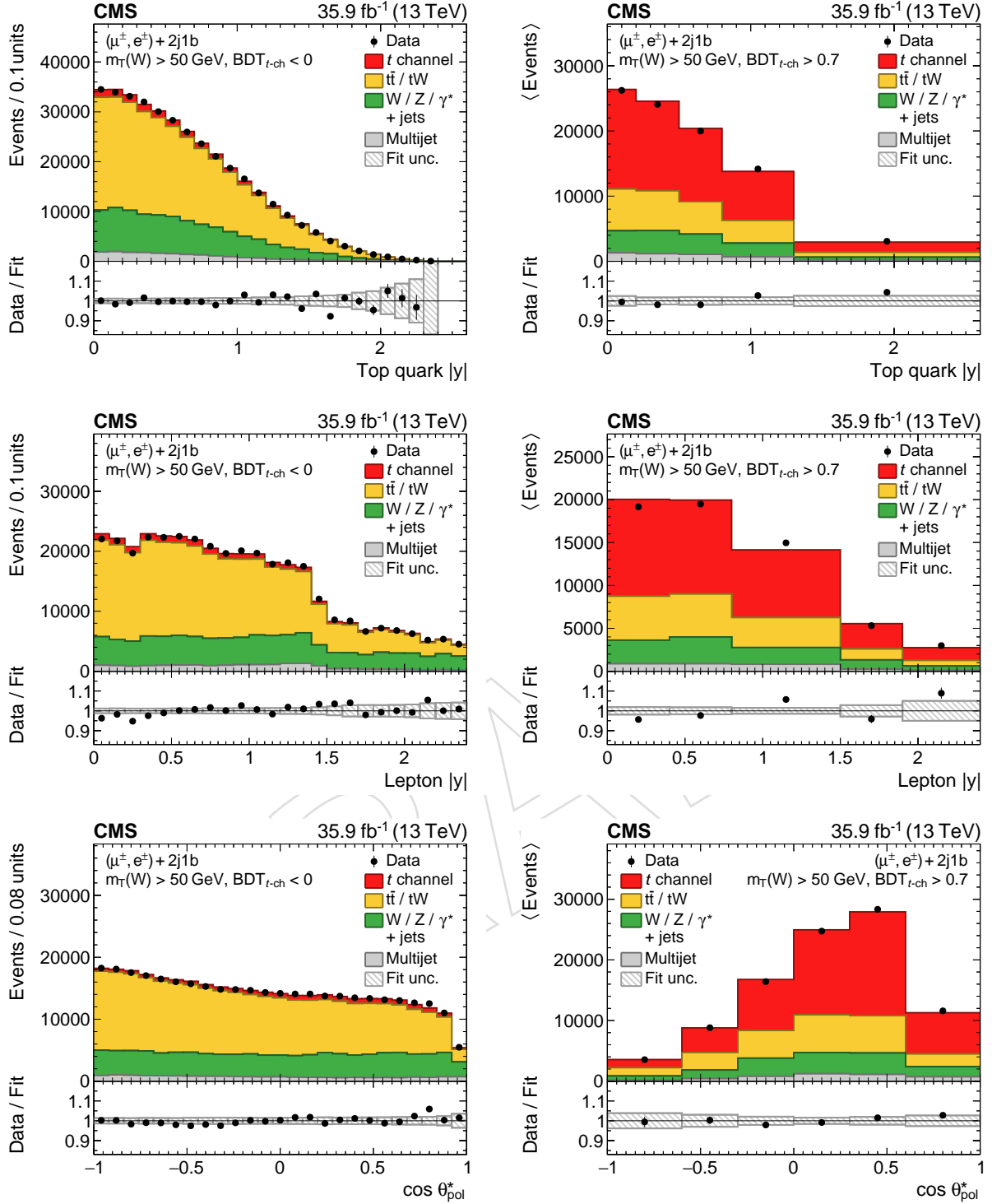


Figure 6: Distributions of the observables in a (left column) background-dominated and a (right column) signal-enriched region for events passing the 2 jets, 1 b tag selection: (upper row) top quark rapidity; (middle row) charged lepton rapidity; (lower row) cosine of the top quark polarisation angle. Events in the muon and electron channels have been summed. The predictions have been scaled to the result of the inclusive ML fit and the hatched band displays the fit uncertainty. The plots on the left give the number of events per bin, while those on the right show the number of events per bin divided by the bin width. The lower panel in each plot gives the ratio of the data to the fit results.

A penalty term, based on the curvature of the unfolded spectrum [58, 59] encoded in the matrix  $L$ , is added in the minimisation to suppress oscillating solutions originating from amplified statistical fluctuations. This “regularisation” procedure has a strength  $\tau$  that is chosen to minimise the global correlation between the unfolded bins. The “bias vector”  $\vec{x}_0$  is set to the expected spectrum from simulation. Pseudo-experiments using simulated data are performed to verify that the unfolding method estimates the uncertainties correctly, while keeping the regularisation bias at a minimum. No regularisation is applied when unfolding the lepton  $p_T$  and rapidity spectra since the migrations between bins are found to be negligible. The overall normalisation of the unfolded spectrum is determined by performing a simultaneous minimisation with respect to the Lagrange multiplier  $\lambda$ .

The parton-level top quark in simulation is defined as the generated on-shell top quark after quantum electrodynamic (QED) and QCD radiation, taking into account the intrinsic transverse momentum of initial-state partons. Events are required to contain either a muon or an electron from the top quark decay chain. This also includes muons or electrons from immediately produced  $\tau$  leptons. In such events, the  $W$  boson is chosen to be the direct daughter of the top quark. The spectator quark is selected from among the light quarks after QED and QCD radiation that are not products of the top quark decay. In case of ambiguities arising from initial-state radiation, the spectator quark that minimises the  $p_T$  of the combined spectator quark and top quark system is chosen.

The top quark at the particle level (called “pseudo top quark”) is defined in simulated events by performing an event reconstruction based on the set of stable simulated particles after hadronisation [60]. In the context of this study, all particles with a lifetime of more than 30 ps are considered stable. So-called “dressed” muons and electrons are constructed by accounting for the additional momenta carried by photons within a cone of  $\Delta R < 0.1$  around the corresponding prompt lepton that do not originate from hadronisation products. The  $\vec{p}_T^{\text{miss}}$  is defined as the summed momentum of all prompt neutrinos in the event. Jets at the particle level are clustered from all stable particles excluding prompt muons, prompt electrons, prompt photons, and all neutrinos using the anti- $k_T$  algorithm with a distance parameter of  $R = 0.4$ . From these objects, a pseudo top quark is reconstructed by first solving for the unknown neutrino  $p_z$  momentum, which is identical to the top quark reconstruction procedure applied to data, as described in Section 4. Events containing a single dressed muon or electron with  $p_T > 26$  GeV and  $|\eta| < 2.4$ , together with two jets with  $p_T > 40$  GeV and  $|\eta| < 4.7$ , are considered at the particle level. Jets that are closer than  $\Delta R = 0.4$  to the selected dressed muon or electron are ignored. The jet that yields a top quark mass closest to 172.5 GeV is assumed to come from the top quark decay, while the other jet is taken as the spectator jet.

The size of the binning intervals are chosen to minimise the migrations between the reconstructed bins while retaining sensitivity to the shapes of the distributions. The stability (purity) is defined as the probability that the parton- or particle-level (reconstructed) values of an observable within a certain range also have their reconstructed (parton-/particle-level) counterparts in the same range. Both quantities are found to be greater than or equal to 50% in most bins of all distributions, with the exception of a few bins at the parton level where purity and stability drop to 40%, and the first two bins of the polarisation angle distribution at the parton level where both quantities drop to about 25%. The stability and purity values are about 10% larger for the particle-level distributions than for the parton-level ones. The acceptance times efficiency for selecting  $t$ -channel single top quark events at the detector level is found to be 2–8 (20–30)% for muon events and 1–5 (10–20)% for electron events with respect to the parton (particle) level across the unfolding bins.

## 9 Systematic uncertainties

The measurements are affected by various sources of systematic uncertainty. For each systematic variation, new templates and response matrices are derived. Systematic variations can create correlations between the  $t$ -channel top quark and antiquark yields since both yields are estimated simultaneously from data through an ML fit, as described in Section 6.

The following experimental systematic uncertainties are profiled in the ML fit.

- **Background composition:** As described in Section 6, the  $Z/\gamma^*$ +jets and  $W$ +jets processes and the  $t\bar{t}$  and  $tW$  processes are separately grouped together in the ML fit. The ratios of the  $Z/\gamma^*$ +jets to the  $W$ +jets yields and the  $t\bar{t}$  to the  $tW$  yields are assigned a  $\pm 20\%$  uncertainty. This covers the uncertainty in the small  $Z/\gamma^*$ +jets and  $tW$  yields, for which the analysis has little sensitivity.
- **Multijet shape estimation:** The multijet event distributions are estimated from data by inversion of the muon isolation criterion or the electron identification criteria. The uncertainty in the shape of these distributions is estimated by varying the criteria. The requirement on the muon isolation parameter in the sideband region is modified from  $I_{\text{rel}}^\mu > 20\%$  to either  $20 < I_{\text{rel}}^\mu < 40\%$  or  $I_{\text{rel}}^\mu > 40\%$ , and the electron isolation parameter to either  $I_{\text{rel}}^e < 30\%$  or  $I_{\text{rel}}^e > 5.88\%$ , while inverting the identification criteria. Another variation is done by requiring electrons in the sideband region to explicitly pass or fail the photon conversion criterion, which is also part of the electron identification requirement.
- **Efficiency of  $b$  tagging and misidentification:** The scale factors used to reweight the  $b$  tagging and misidentification efficiencies in simulation to the ones estimated from data are varied within their uncertainties based on the true flavour of the selected jets [49].
- **Jet energy scale and resolution:** The jet energy scale and resolution corrections are varied within their uncertainties [61]. The shifts induced in the jet momenta are propagated to  $\vec{p}_T^{\text{miss}}$  as well.
- **Unclustered energy:** The contributions to  $p_T^{\text{miss}}$  of PF candidates that have not been clustered into jets are varied within their respective energy resolutions [62].
- **Pileup:** The simulated distribution of pileup interactions is modified by shifting the total inelastic pp cross section by  $\pm 5\%$  [63].
- **Lepton efficiencies:** The scale factors that account for differences in the lepton selection and reconstruction efficiencies between data and simulation are varied within their uncertainties [23, 46].

The systematic uncertainties in the theoretical modelling of the simulated samples are estimated by using new templates and response matrices in the ML fit and unfolding for each variation. For each uncertainty source, the maximum difference of the up/down variations with the result using the nominal templates and response matrix is taken as the estimated uncertainty per bin. These are added in quadrature to the experimental uncertainty per bin.

The following sources of theoretical uncertainty have been evaluated.

- **Modelling of top quark  $p_T$  in  $t\bar{t}$  events:** Differential cross section measurements of  $t\bar{t}$  production by CMS [64, 65] have shown that the  $p_T$  spectrum of top quarks in  $t\bar{t}$  events is significantly softer than predicted by NLO simulations. To correct for this effect, simulated  $t\bar{t}$  events are reweighted according to the scale factors derived from measurements at 13 TeV [65]. The difference in the predictions when using the

default  $t\bar{t}$  simulation sample is taken as an additional uncertainty.

- Top quark mass: The nominal top quark mass of 172.5 GeV is modified by  $\pm 0.5$  GeV in the simulation [66]. The difference with respect to the nominal simulation results is taken as the corresponding uncertainty.
- Parton distribution functions: The effect of the uncertainty in the PDFs is estimated by reweighting the simulated events using the recommended variations in the NNPDF3.0 NLO set, including a variation of  $\alpha_s$  [35]. The reweighting is performed using precomputed weights stored in the event record by the matrix element generator [67].
- Renormalisation/factorisation scales: A reweighting procedure similar to that used for the PDFs is carried out on simulated  $t$ -channel,  $W$ +jets, and  $t\bar{t}$  simulated events to estimate the effect of the uncertainties in the renormalisation and factorisation scales. The weights correspond to independent variations by factors of 0.5 and 2 in the scales with respect to their nominal values. The envelope of all possible combinations of up-varied/down-varied scales with the exception of the extreme up/down combinations is considered as an uncertainty. This uncertainty is evaluated independently for the  $t$ -channel,  $W$ +jets, and  $t\bar{t}$  simulated event samples.
- Parton shower: The uncertainties in the parton shower simulation are evaluated by comparing the nominal samples to dedicated samples with varied shower parameters. For  $t$ -channel single top quark production, the differences with respect to samples with a varied factorisation scale by a factor of 0.5 or 2 or with a varied POWHEG  $h_{\text{damp}}$  parameter are taken as two independent uncertainties. For the simulated  $t\bar{t}$  samples, the variation of the factorisation scale in both initial- and final-state radiation, and the  $h_{\text{damp}}$  parameter are evaluated as three independent uncertainties.
- Underlying event tune: The impact of uncertainties arising from the CUETP8M2T4 underlying event tune [30] used in the simulation of  $t\bar{t}$  events is evaluated using dedicated samples with the tune varied within its uncertainties.
- Colour reconnection: The default model of colour reconnection in PYTHIA is based on multiple-particle interactions (MPI) with early resonance decays switched off. An uncertainty in the choice of this model is taken into account by repeating the measurement using three alternative models of colour reconnection in the simulation of  $t$ -channel single top quark and  $t\bar{t}$  production: the MPI-based scheme with early resonance decays switched on, a gluon-move scheme [68], and a QCD-inspired scheme [69].
- Fragmentation model: The fragmentation of  $b$  quarks, modelled by the Bowler-Lund function [70], is varied within its uncertainties for  $t$ -channel single top quark and  $t\bar{t}$  production. Additionally, the impact when using the Peterson model [71] for  $b$  quark fragmentation instead is assessed.

In addition, an uncertainty of  $\pm 2.5\%$  in the measurement of the integrated luminosity of the data set [22] is taken into account by scaling the evaluated covariance matrix per observable accordingly.

## 10 Results

Differential cross sections of  $t$ -channel single top quark production as a function of the top quark  $p_T$ , rapidity, and polarisation angle, the  $p_T$  and rapidity of the charged lepton (muon or electron) that originates from the top quark decay, and the  $p_T$  of the  $W$  boson from the

top quark decay are presented in Figs. 7 and 8 at the parton and particle levels, respectively. The normalised differential cross sections of the same observables at the parton and particle levels are provided in Figs. 9 and 10. The total uncertainty is indicated by the vertical lines, while horizontal bars indicate the statistical and experimental uncertainties, which have been profiled in the ML fit, and thus exclude the uncertainties in the theoretical modelling and the luminosity. The differential cross sections refer to  $t$ -channel single top quark production where the top quark decays semileptonically (into either muon or electron) including events where the charged lepton stems from an intermediate  $\tau$  lepton decay. The results are compared to the predictions by the POWHEG generator interfaced with PYTHIA in the 4FS and the MADGRAPH5\_aMC@NLO generator interfaced with PYTHIA in the 4FS and 5FS.

An overall good agreement of the results with the predictions from the 4FS is observed, except for a slight deviation at low top quark  $p_T$ . The predictions from the 5FS for the top quark and W boson  $p_T$  distributions do not agree as well with the data.

Differential ratios of the top quark production rates to the sum of the top quark and antiquark rates as a function of the top quark  $p_T$  and rapidity, the  $p_T$  and rapidity of the charged lepton, and the W boson  $p_T$  are presented in Figs. 11 and 12 at the parton and particle levels, respectively. It is found that the standard definition of the charge ratio in the literature, i.e.  $\sigma_t/\sigma_{\bar{t}}$ , can yield large variances when the precision in certain intervals of the differential cross section for the top antiquark is low. Therefore, the charge ratio is defined as  $\sigma_t/\sigma_{t+\bar{t}}$  in this paper. The ratios have been calculated from the measured cross sections at the parton and particle levels, while accounting for correlations between the top quark and antiquark spectra, as detailed in Sections 6 and 9. The resulting charge ratios are compared to the predictions by the NNPDF3.0 NLO, MMHT14 NLO [72], and CT10 NLO PDF sets, which have been calculated using the POWHEG signal sample—generated in the 4FS and interfaced with PYTHIA. The uncertainty bands shown in Figs. 11 and 12 represent the total uncertainty from varying the corresponding PDF eigenvectors and  $\alpha_S$ . Within the uncertainties, the measured charge ratios are in good agreement with the predictions from all three PDF sets.

The spin asymmetry, sensitive to the top quark polarisation, is determined from the differential cross section as a function of the polarisation angle at the parton level (Fig. 7, lower right). A linear  $\chi^2$ -based fit, assuming the expected functional dependence given in Eq. (2), is used to take the correlations between the unfolded bins into account. The measured spin asymmetry in the muon and electron channel and their combination is given in Table 2.

The measured asymmetries are in good agreement with the predicted SM value of 0.436, found using POWHEG at NLO, with a negligible uncertainty. Good agreement is also found with a corresponding measurement by the ATLAS Collaboration at  $\sqrt{s} = 8$  TeV [17]. This measurement is found to be more precise than a previous analysis of the spin asymmetry at  $\sqrt{s} = 8$  TeV by the CMS Collaboration [9]. In particular, the deviation found therein, corresponding to 2.0 standard deviations, is not seen.

## 11 Summary

Differential cross sections for  $t$ -channel single top quark and antiquark production in proton-proton collisions at  $\sqrt{s} = 13$  TeV have been measured by the CMS experiment at the LHC using a sample of proton-proton collision data, corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The cross sections are determined as a function of the top quark transverse momentum ( $p_T$ ), rapidity, and polarisation angle, the charged lepton  $p_T$  and rapidity, and the  $p_T$  of the W boson from the top quark decay. In addition, the charge ratio has been measured as a



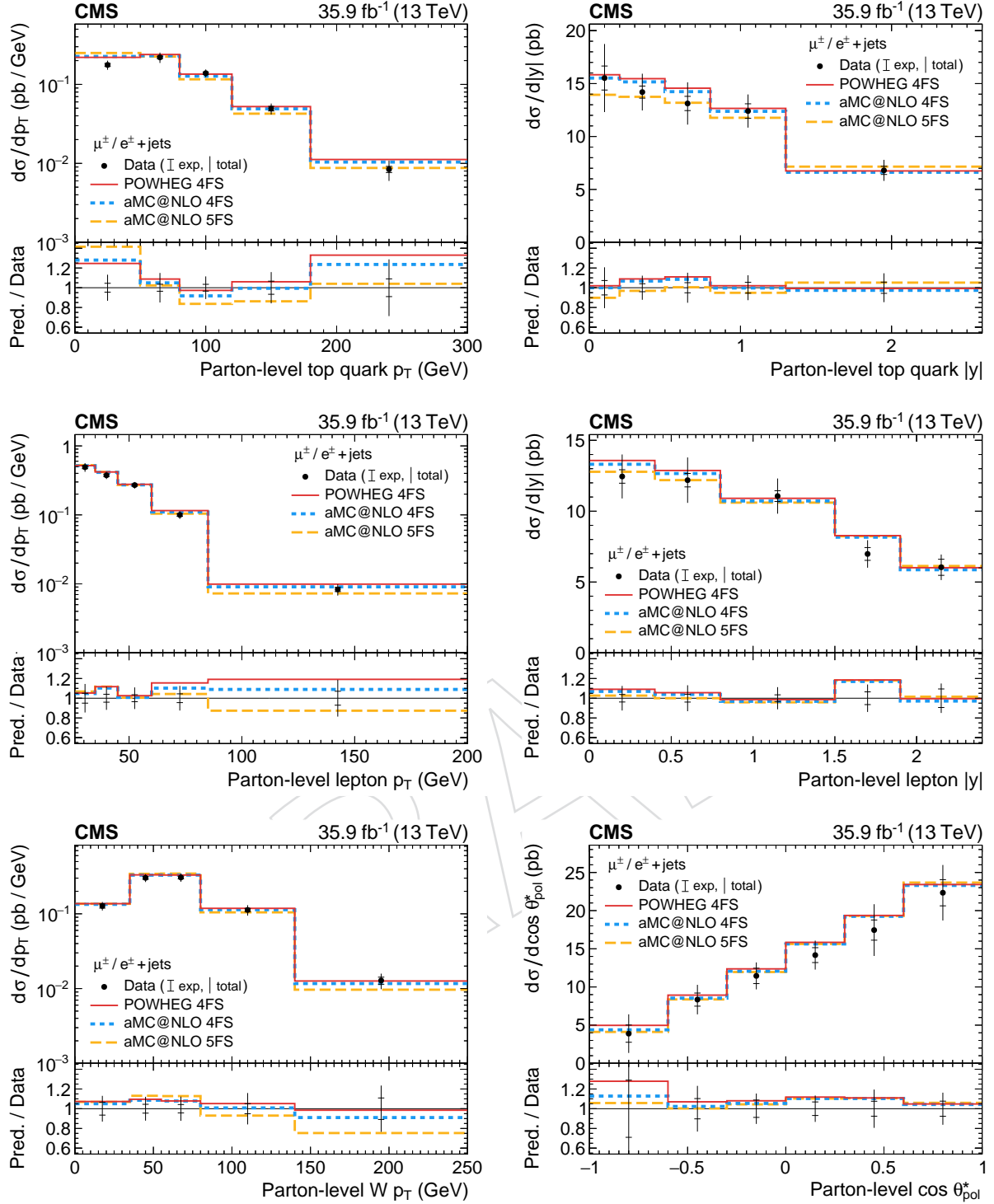


Figure 7: Differential cross sections for the sum of  $t$ -channel single top quark and antiquark production at the parton level: (upper row) top quark  $p_T$  and rapidity; (middle row) charged lepton  $p_T$  and rapidity; (lower left)  $W$  boson  $p_T$ ; (lower right) cosine of the top quark polarisation angle. The total uncertainty is indicated by the vertical lines, while horizontal bars indicate the statistical and experimental uncertainties, which have been profiled in the ML fit, and thus exclude the uncertainties in the theoretical modelling and the luminosity. Three different predictions from event generators are shown by the solid, dashed, and dotted lines. The lower panels show the ratios of the predictions to the data.



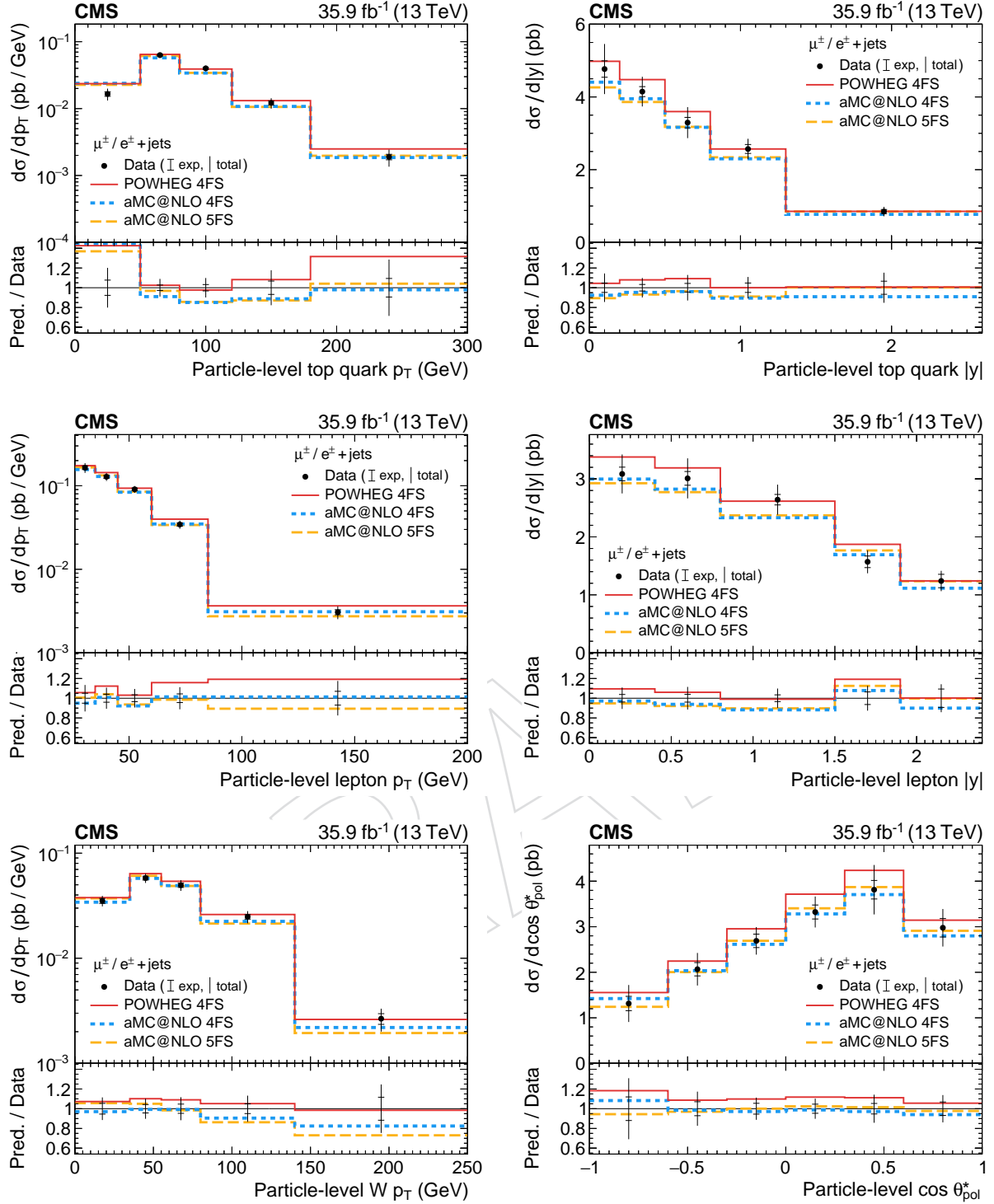


Figure 8: Differential cross sections for the sum of  $t$ -channel single top quark and antiquark production at the particle level: (upper row) top quark  $p_T$  and rapidity; (middle row) charged lepton  $p_T$  and rapidity; (lower left) W boson  $p_T$ ; (lower right) cosine of the top quark polarisation angle. The total uncertainty is indicated by the vertical lines, while horizontal bars indicate the statistical and experimental uncertainties, which have been profiled in the ML fit, and thus exclude the uncertainties in the theoretical modelling and the luminosity. Three different predictions from event generators are shown by the solid, dashed, and dotted lines. The lower panels show the ratios of the predictions to the data.

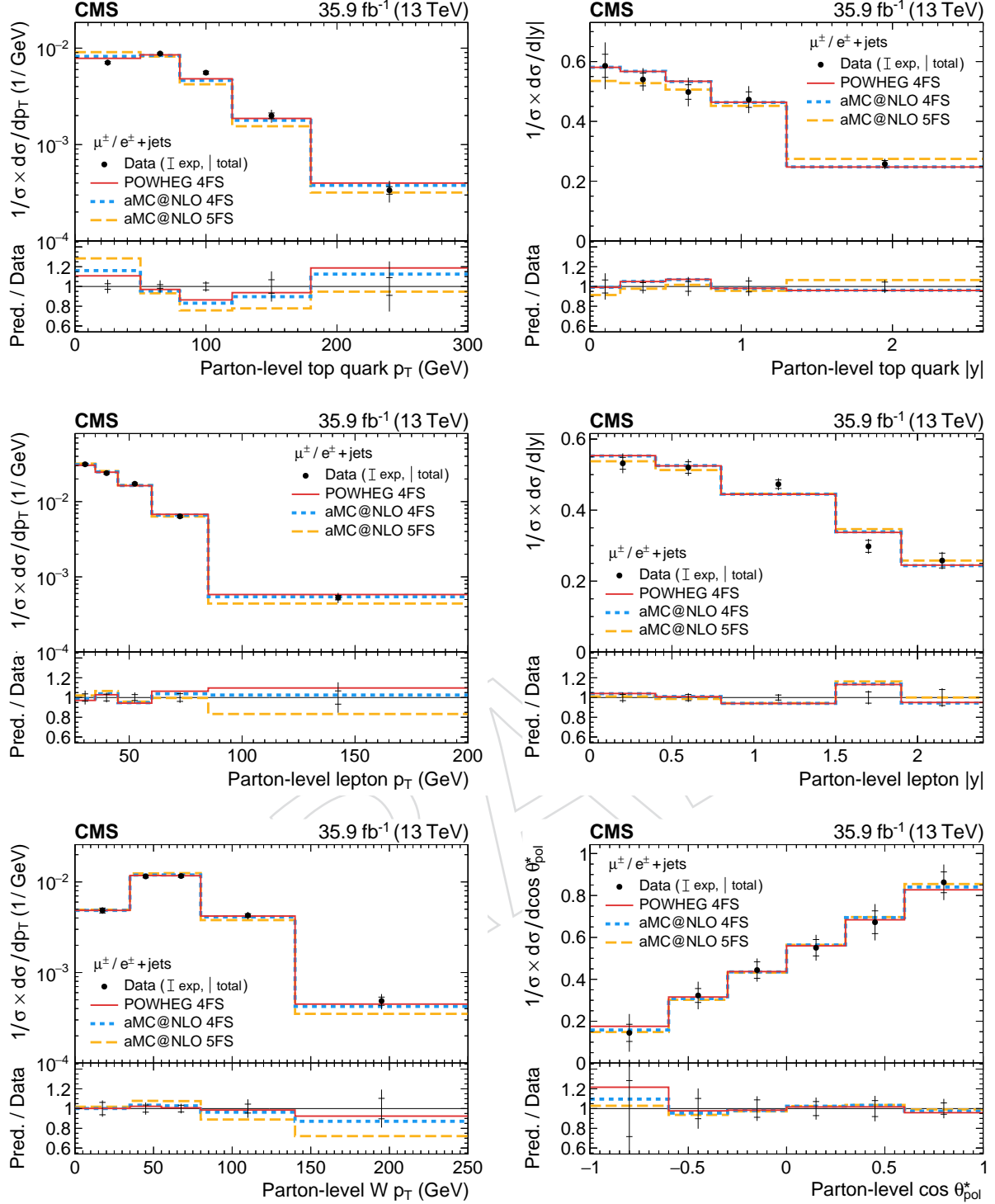


Figure 9: Normalised differential cross sections for the sum of  $t$ -channel single top quark and antiquark production at the parton level: (upper row) top quark  $p_T$  and rapidity; (middle row) charged lepton  $p_T$  and rapidity; (lower left) W boson  $p_T$ ; (lower right) cosine of the top quark polarisation angle. The total uncertainty is indicated by the vertical lines, while horizontal bars indicate the statistical and experimental uncertainties, which have been profiled in the ML fit, and thus exclude the uncertainties in the theoretical modelling. Three different predictions from event generators are shown by the solid, dashed, and dotted lines. The lower panels show the ratios of the predictions to the data.

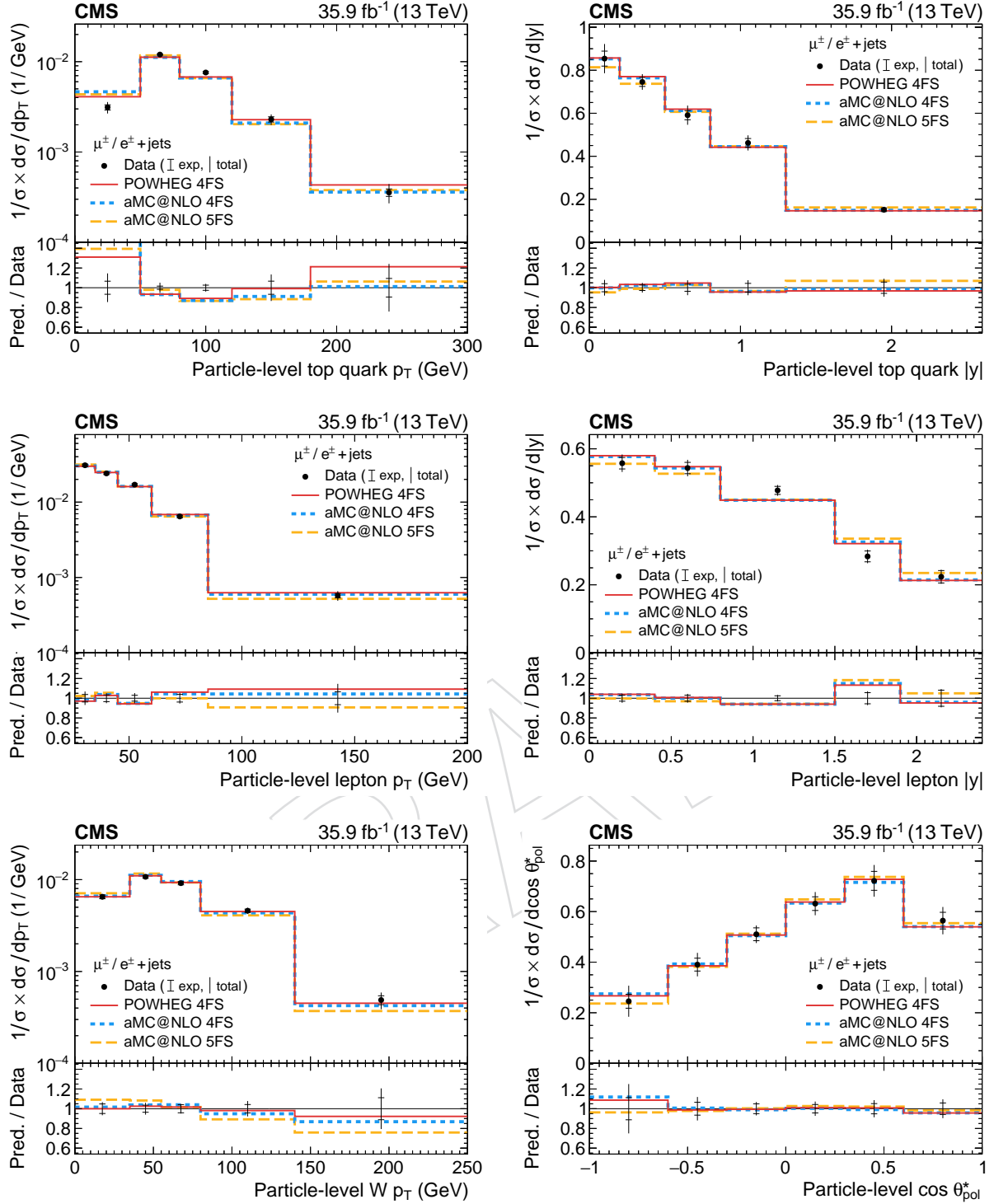


Figure 10: Normalised differential cross sections for the sum of  $t$ -channel single top quark and antiquark production at the particle level: (upper row) top quark  $p_T$  and rapidity; (middle row) charged lepton  $p_T$  and rapidity; (lower left) W boson  $p_T$ ; (lower right) cosine of the top quark polarisation angle. The total uncertainty is indicated by the vertical lines, while horizontal bars indicate the statistical and experimental uncertainties, which have been profiled in the ML fit, and thus exclude the uncertainties in the theoretical modelling. Three different predictions from event generators are shown by the solid, dashed, and dotted lines. The lower panels show the ratios of the predictions to the data.

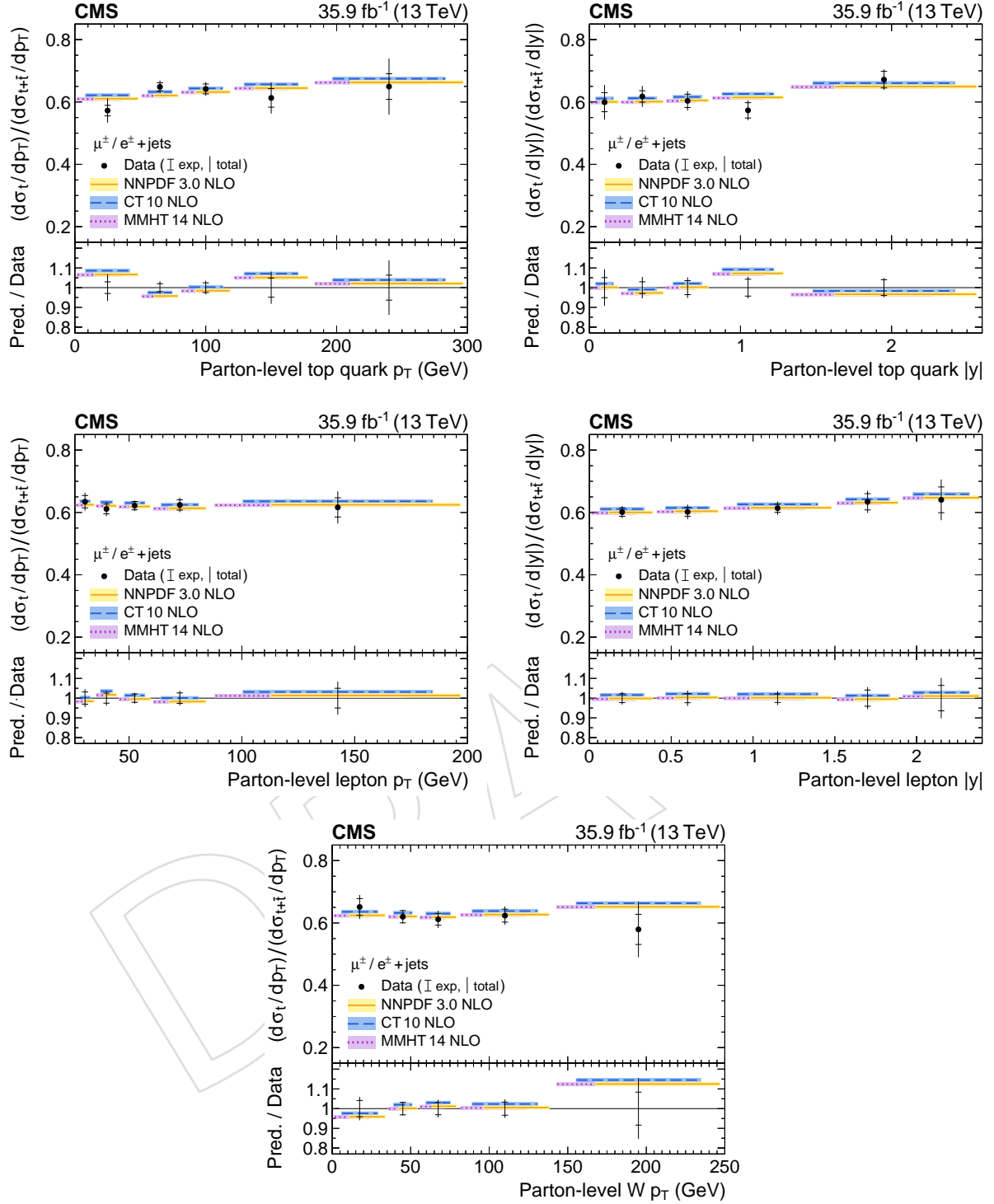


Figure 11: Ratio of the top quark to the sum of the top quark and antiquark  $t$ -channel differential cross section at the parton level: (upper row) top quark  $p_T$  and rapidity; (middle row) charged lepton  $p_T$  and rapidity; (lower row)  $W$  boson  $p_T$ . The total uncertainty is indicated by the vertical lines, while horizontal bars indicate the statistical and experimental uncertainties, which have been profiled in the ML fit, and thus exclude the uncertainties in the theoretical modelling. Predictions from three different PDF sets are shown by the solid, dashed, and dotted lines. The lower panels show the ratios of the predictions to the data.

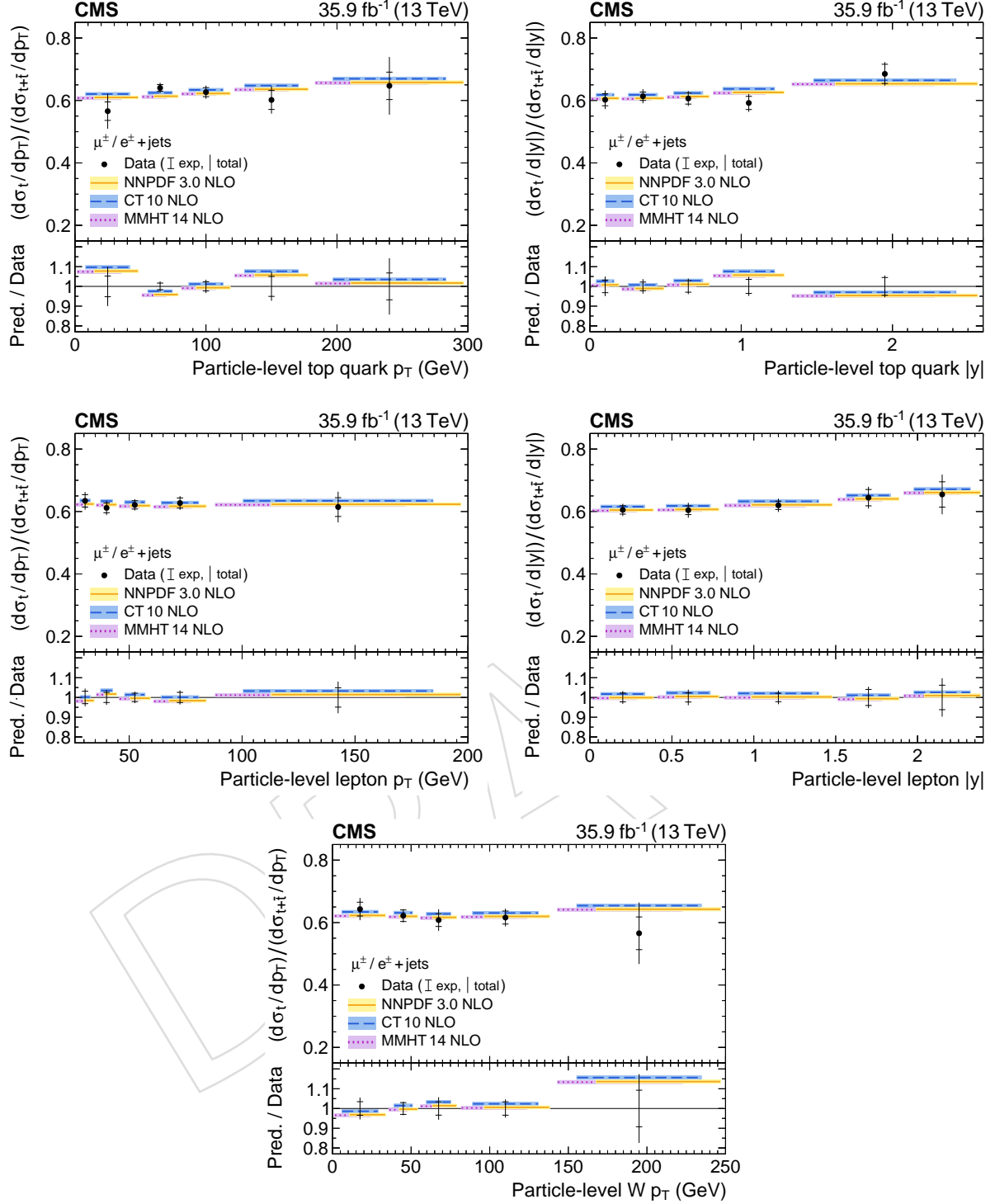


Figure 12: Ratio of the top quark to the sum of the top quark and antiquark  $t$ -channel differential cross section at the particle level: (upper row) top quark  $p_T$  and rapidity; (middle row) charged lepton  $p_T$  and rapidity; (lower row)  $W$  boson  $p_T$ . The total uncertainty is indicated by the vertical lines, while horizontal bars indicate the statistical and experimental uncertainties, which have been profiled in the ML fit, and thus exclude the uncertainties in the theoretical modelling. Predictions from three different PDF sets are shown by the solid, dashed, and dotted lines. The lower panels show the ratios of the predictions to the data.

Table 2: The measured spin asymmetry in the muon and electron channel and their combination. A breakdown of the systematic uncertainties is also provided. Minor systematic uncertainties (lepton efficiencies, pileup, and unclustered energy) have been grouped into the “Others” category.

	Central values	$A_\mu$ 0.403	$A_e$ 0.446	$A_{\mu+e}$ 0.440
Profiled uncertainties	Statistical	$\pm 0.029$	$\pm 0.038$	$\pm 0.024$
	$t\bar{t}$ /tW normalisation	$\pm 0.010$	$\pm 0.007$	$\pm 0.007$
	W/Z/ $\gamma^*$ +jets normalisation	$\pm 0.012$	$\pm 0.011$	$\pm 0.012$
	Multijet normalisation	$< 0.001$	$< 0.001$	$\pm 0.003$
	Multijet shape	$< 0.001$	$\pm 0.006$	$< 0.001$
	Jet energy scale/resolution	$\pm 0.008$	$< 0.001$	$< 0.001$
	b tagging efficiencies/misidentification	$< 0.001$	$\pm 0.009$	$\pm 0.004$
	Others	$< 0.001$	$\pm 0.003$	$\pm 0.005$
Theoretical uncertainties	Top quark mass	$\pm 0.033$	$\pm 0.063$	$\pm 0.044$
	PDF+ $\alpha_s$	$\pm 0.011$	$\pm 0.009$	$\pm 0.011$
	t channel renorm./fact. scales	$\pm 0.013$	$\pm 0.018$	$\pm 0.020$
	t channel parton shower	$\pm 0.030$	$\pm 0.008$	$\pm 0.014$
	$t\bar{t}$ renorm./fact. scales	$\pm 0.008$	$\pm 0.019$	$\pm 0.017$
	$t\bar{t}$ parton shower	$\pm 0.031$	$\pm 0.037$	$\pm 0.033$
	$t\bar{t}$ underlying event tune	$< 0.001$	$\pm 0.014$	$\pm 0.014$
	$t\bar{t}$ $p_T$ reweighting	$< 0.001$	$\pm 0.010$	$\pm 0.009$
	W+jets renorm./fact. scales	$< 0.001$	$\pm 0.019$	$\pm 0.014$
	Color reconnection	$\pm 0.036$	$\pm 0.056$	$\pm 0.031$
	Fragmentation model	$\pm 0.011$	$\pm 0.011$	$\pm 0.011$
Profiled uncertainties only (statistical+experimental)		$\pm 0.041$	$\pm 0.047$	$\pm 0.031$
Total uncertainties		$\pm 0.071$	$\pm 0.099$	$\pm 0.070$

function of the top quark, charged lepton, and W boson kinematic observables. Events containing one muon or electron and two or three jets are used. The single top quark and antiquark yields are determined through maximum-likelihood fits to the data distributions. The differential cross sections are then obtained at the parton and particle levels by unfolding the measured signal yields.

The results are compared to various next-to-leading-order predictions, and found to be in good agreement. Furthermore, the top quark spin asymmetry, which is sensitive to the top quark polarisation, has been measured using the differential cross section as a function of the top quark polarisation angle at the parton level. The resulting value of  $0.440 \pm 0.070$  is in good agreement with the standard model prediction.

These results demonstrate a good understanding of the underlying electroweak production mechanism of single top quarks at  $\sqrt{s} = 13$  TeV and in particular of the electroweak vector–axial-vector coupling predicting highly polarized top quarks. Lastly, the differential charge ratios, sensitive to the ratio of the up to down quark content of the proton, are found to be consistent with the predictions by various sets of parton distribution functions.



## Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: BMBWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, PUT and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); NKFI (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); MES (Latvia); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MOS (Montenegro); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR, and NRC KI (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI, and FEDER (Spain); MOSTR (Sri Lanka); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie programme and the European Research Council and Horizon 2020 Grant, contract Nos. 675440, 752730, and 765710 (European Union); the Leventis Foundation; the A.P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the F.R.S.-FNRS and FWO (Belgium) under the "Excellence of Science – EOS" – be.h project n. 30820817; the Beijing Municipal Science & Technology Commission, No. Z181100004218003; the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Lendület ("Momentum") Programme and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the New National Excellence Program ÚNKP, the NKFI research grants 123842, 123959, 124845, 124850, 125105, 128713, 128786, and 129058 (Hungary); the Council of Science and Industrial Research, India; the HOMING PLUS programme of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund, the Mobility Plus programme of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Ministry of Science and Education, grant no. 3.2989.2017 (Russia); the Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia María de Maeztu, grant MDM-2015-0509 and the Programa Severo Ochoa del Principado de Asturias; the Thalís and Aristeia programmes cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

## References

- [1] D0 Collaboration, "Observation of single top quark production", *Phys. Rev. Lett.* **103**

- (2009) 092001, doi:10.1103/PhysRevLett.103.092001, arXiv:0903.0850.
- [2] CDF Collaboration, “First observation of electroweak single top quark production”, *Phys. Rev. Lett.* **103** (2009) 092002, doi:10.1103/PhysRevLett.103.092002, arXiv:0903.0885.
- [3] ATLAS Collaboration, “Comprehensive measurements of  $t$ -channel single top quark production cross sections at  $\sqrt{s} = 7$  TeV with the ATLAS detector”, *Phys. Rev. D* **90** (2014) 112006, doi:10.1103/PhysRevD.90.112006, arXiv:1406.7844.
- [4] CMS Collaboration, “Measurement of the single top quark  $t$ -channel cross section in pp collisions at  $\sqrt{s} = 7$  TeV”, *JHEP* **12** (2012) 035, doi:10.1007/JHEP12(2012)035, arXiv:1209.4533.
- [5] ATLAS Collaboration, “Fiducial, total and differential cross section measurements of  $t$ -channel single top quark production in pp collisions at 8 TeV using data collected by the ATLAS detector”, *Eur. Phys. J. C* **77** (2017) 531, doi:10.1140/epjc/s10052-017-5061-9, arXiv:1702.02859.
- [6] CMS Collaboration, “Measurement of the  $t$ -channel single top quark production cross section and of the  $|V_{tb}|$  CKM matrix element in pp collisions at  $\sqrt{s} = 8$  TeV”, *JHEP* **06** (2014) 090, doi:10.1007/JHEP06(2014)090, arXiv:1403.7366.
- [7] ATLAS Collaboration, “Measurement of the inclusive cross sections of single top quark and top antiquark  $t$ -channel production in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”, *JHEP* **04** (2017) 086, doi:10.1007/JHEP04(2017)086, arXiv:1609.03920.
- [8] CMS Collaboration, “Measurement of the single top quark and antiquark production cross sections in the  $t$  channel and their ratio in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, *Phys. Lett. B* **800** (2019) 135042, doi:10.1016/j.physletb.2019.135042, arXiv:1812.10514.
- [9] CMS Collaboration, “Measurement of top quark polarisation in  $t$ -channel single top quark production”, *JHEP* **04** (2016) 073, doi:10.1007/JHEP04(2016)073, arXiv:1511.02138.
- [10] N. P. Hartland et al., “A Monte Carlo global analysis of the standard model effective field theory: The top quark sector”, *JHEP* **04** (2019) 100, doi:10.1007/JHEP04(2019)100, arXiv:1901.05965.
- [11] P. Kant et al., “HATHOR for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions”, *Comput. Phys. Commun.* **191** (2015) 74, doi:10.1016/j.cpc.2015.02.001, arXiv:1406.4403.
- [12] E. L. Berger, J. Gao, C. P. Yuan, and H. X. Zhu, “NNLO QCD corrections to  $t$ -channel single top quark production and decay”, *Phys. Rev. D* **94** (2016) 071501, doi:10.1103/PhysRevD.94.071501, arXiv:1606.08463.
- [13] S. Alekhin, J. Blümlein, S. Moch, and R. Placakyte, “Parton distribution functions,  $\alpha_s$ , and heavy-quark masses for LHC Run II”, *Phys. Rev. D* **96** (2017) 014011, doi:10.1103/PhysRevD.96.014011, arXiv:1701.05838.

- [14] G. Mahlon and S. J. Parke, “Single top quark production at the LHC: Understanding spin”, *Phys. Lett. B* **476** (2000) 323, doi:10.1016/S0370-2693(00)00149-0, arXiv:hep-ph/9912458.
- [15] E. E. Boos and A. V. Sherstnev, “Spin effects in processes of single top quark production at hadron colliders”, *Phys. Lett. B* **534** (2002) 97, doi:10.1016/S0370-2693(02)01659-3, arXiv:hep-ph/0201271.
- [16] J. A. Aguilar-Saavedra and J. Bernabeu, “W polarisation beyond helicity fractions in top quark decays”, *Nucl. Phys. B* **840** (2010) 349, doi:10.1016/j.nuclphysb.2010.07.012, arXiv:1005.5382.
- [17] ATLAS Collaboration, “Probing the  $Wtb$  vertex structure in  $t$ -channel single-top-quark production and decay in pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector”, *JHEP* **04** (2017) 124, doi:10.1007/JHEP04(2017)124, arXiv:1702.08309.
- [18] ATLAS Collaboration, “Analysis of the  $Wtb$  vertex from the measurement of triple-differential angular decay rates of single top quarks produced in the  $t$ -channel at  $\sqrt{s} = 8$  TeV with the ATLAS detector”, *JHEP* **12** (2017) 017, doi:10.1007/JHEP12(2017)017, arXiv:1707.05393.
- [19] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [20] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [21] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.
- [22] CMS Collaboration, “CMS luminosity measurements for the 2016 data taking period”, CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- [23] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [24] 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, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [25] R. Frederix, E. Re, and P. Torrielli, “Single top  $t$ -channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO”, *JHEP* **09** (2012) 130, doi:10.1007/JHEP09(2012)130, arXiv:1207.5391.
- [26] T. Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [27] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

- [28] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”, *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.
- [29] 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, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [30] CMS Collaboration, “Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of  $t\bar{t}$  at  $\sqrt{s} = 8$  and 13 TeV”, CMS Physics Analysis Summary CMS-PAS-TOP-16-021, 2016.
- [31] E. Re, “Single top  $Wt$ -channel production matched with parton showers using the POWHEG method”, *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- [32] S. Frixione et al., “Single top hadroproduction in association with a  $W$  boson”, *JHEP* **07** (2008) 029, doi:10.1088/1126-6708/2008/07/029, arXiv:0805.3067.
- [33] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- [34] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [35] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [36] GEANT4 Collaboration, “GEANT4 — a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [37] M. Aliev et al., “HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR”, *Comput. Phys. Commun.* **182** (2011) 1034, doi:10.1016/j.cpc.2010.12.040, arXiv:1007.1327.
- [38] S. Alekhin et al., “The PDF4LHC working group interim report”, (2011). arXiv:1101.0536.
- [39] M. Botje et al., “The PDF4LHC working group interim recommendations”, (2011). arXiv:1101.0538.
- [40] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Parton distributions for the LHC”, *Eur. Phys. J. C* **63** (2009) 189, doi:10.1140/epjc/s10052-009-1072-5, arXiv:0901.0002.
- [41] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Uncertainties on  $\alpha_s$  in global PDF analyses and implications for predicted hadronic cross sections”, *Eur. Phys. J. C* **64** (2009) 653, doi:10.1140/epjc/s10052-009-1164-2, arXiv:0905.3531.
- [42] H.-L. Lai et al., “New parton distributions for collider physics”, *Phys. Rev. D* **82** (2010) 074024, doi:10.1103/PhysRevD.82.074024, arXiv:1007.2241.
- [43] NNPDF Collaboration, “Parton distributions with LHC data”, *Nucl. Phys. B* **867** (2013) 244, doi:10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303.

- [44] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_T$  jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [45] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
- [46] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV”, *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.
- [47] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Phys. Lett. B* **659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.
- [48] CMS Collaboration, “Pileup removal algorithms”, CMS Physics Analysis Summary CMS-PAS-JME-14-001, 2014.
- [49] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- [50] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV”, *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [51] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.
- [52] CMS Collaboration, “Measurement of the  $t$ -channel single top quark production cross section in pp collisions at  $\sqrt{s} = 7$  TeV”, *Phys. Rev. Lett.* **107** (2011) 091802, doi:10.1103/PhysRevLett.107.091802, arXiv:1106.3052.
- [53] S. Kallweit et al., “NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging”, *JHEP* **04** (2016) 021, doi:10.1007/JHEP04(2016)021, arXiv:1511.08692.
- [54] F. R. Anger, F. Febres Cordero, H. Ita, and V. Sotnikov, “NLO QCD predictions for  $Wb\bar{b}$  production in association with up to three light jets at the LHC”, *Phys. Rev. D* **97** (2018) 036018, doi:10.1103/PhysRevD.97.036018, arXiv:1712.05721.
- [55] C.-H. Kom and W. J. Stirling, “Charge asymmetry in W+jets production at the LHC”, *Eur. Phys. J. C* **69** (2010) 67, doi:10.1140/epjc/s10052-010-1353-z, arXiv:1004.3404.
- [56] R. Barlow and C. Beeston, “Fitting using finite Monte Carlo samples”, *Comput. Phys. Commun.* **77** (1993) 219, doi:10.1016/0010-4655(93)90005-W.
- [57] S. Schmitt, “TUnfold: An algorithm for correcting migration effects in high energy physics”, *JINST* **7** (2012) T10003, doi:10.1088/1748-0221/7/10/T10003, arXiv:1205.6201.
- [58] A. Tikhonov, “Solution of incorrectly formulated problems and the regularization method”, *Soviet Math. Dokl.* **5** (1963) 1035.
- [59] V. Blobel, “An unfolding method for high energy physics experiments”, in *Advanced statistical techniques in particle physics, Proceedings, Conference, Durham, UK, March 18–22, 2002*, p. 258. 2002. arXiv:hep-ex/0208022.



- [60] CMS Collaboration, “Object definitions for top quark analyses at the particle level”, CMS Note CMS-NOTE-2017-004, 2017.
- [61] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS”, *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- [62] CMS Collaboration, “Performance of the CMS missing transverse momentum reconstruction in pp data at  $\sqrt{s} = 8$  TeV”, *JINST* **10** (2015) P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [63] CMS Collaboration, “Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 13$  TeV”, *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161, arXiv:1802.02613.
- [64] CMS Collaboration, “Measurement of the differential cross section for top quark pair production in pp collisions at  $\sqrt{s} = 8$  TeV”, *Eur. Phys. J. C* **75** (2015) 542, doi:10.1140/epjc/s10052-015-3709-x, arXiv:1505.04480.
- [65] CMS Collaboration, “Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV”, *Phys. Rev. D* **95** (2017) 092001, doi:10.1103/PhysRevD.95.092001, arXiv:1610.04191.
- [66] 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, doi:10.1103/PhysRevD.93.072004, arXiv:1509.04044.
- [67] A. Kalogeropoulos and J. Alwall, “The SysCalc code: A tool to derive theoretical systematic uncertainties”, (2018). arXiv:1801.08401.
- [68] S. Argyropoulos and T. Sjöstrand, “Effects of color reconnection on  $t\bar{t}$  final states at the LHC”, *JHEP* **11** (2014) 043, doi:10.1007/JHEP11(2014)043, arXiv:1407.6653.
- [69] J. R. Christiansen and P. Z. Skands, “String formation beyond leading colour”, *JHEP* **08** (2015) 003, doi:10.1007/JHEP08(2015)003, arXiv:1505.01681.
- [70] M. G. Bowler, “ $e^+e^-$  production of heavy quarks in the string model”, *Z. Phys. C* **11** (1981) 169, doi:10.1007/BF01574001.
- [71] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, “Scaling violations in inclusive  $e^+e^-$  annihilation spectra”, *Phys. Rev. D* **27** (1983) 105, doi:10.1103/PhysRevD.27.105.
- [72] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, “Parton distributions in the LHC era: MMHT 2014 PDFs”, *Eur. Phys. J. C* **75** (2015) 204, doi:10.1140/epjc/s10052-015-3397-6, arXiv:1412.3989.