DRAFT
CMS Paper

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

Measurement of differential $t\bar{t}$ production cross sections for high-p_T top quarks in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$

The CMS Collaboration

Abstract

A measurement of the production cross section of high transverse momentum (p_T) top quark pairs is reported. The dataset was collected during 2016 with the CMS detector at the CERN LHC from proton-proton collisions at a center-of-mass energy of 13 TeV, and corresponds to an integrated luminosity of $35.9\,\text{fb}^{-1}$. The measurement uses events where either both top quark candidates decay hadronically and are reconstructed as large jets with $p_T > 400\,\text{GeV}$, or where one top quark decays hadronically and is identified as a single large jet with $p_T > 400\,\text{GeV}$ and the other top quark decays leptonically to a b jet, an electron or a muon, and a neutrino. The cross section is extracted differentially as a function of kinematic variables of the top quark or top quark pair system. The results are unfolded to the particle and parton levels, and are compared to various theoretical models. The measured cross section is significantly lower, by up to 40%, in the phase space of interest, compared to the theory predictions, while the normalized differential cross sections are consistent between data and theory.

1 Introduction

The top quark completes the third generation of quarks in the standard model (SM), and a precise knowledge of its properties is critical for the overall understanding of the theory. Measurements of the top-anti-top quark pair ($t\bar{t}$) production cross section confront the predictions from quantum chromodynamics (QCD), while being sensitive to physics beyond the SM. Also, the $t\bar{t}$ production process is a dominant SM background to searches for new physics phenomena and therefore its detailed characterrization is essential for new discoveries.

The large $t\bar{t}$ yield expected in proton-proton (pp) collisions at the CERN LHC allows measurements of the $t\bar{t}$ production rate in a large phase space, and, more importantly, differentially, as a function of the $t\bar{t}$ kinematic properties. Such measurements have been performed by the ATLAS [1–6] and CMS [7–15] Collaborations at 7, 8, and 13 TeV center-of-mass energies, assuming a resolved final state where the decay products of the $t\bar{t}$ pair can be reconstructed individually. This approach is possible for top quark transverse momenta, p_T, up to approximately 500 GeV. However, at higher p_T, the top quark decay products are highly collimated ("boosted") and they can no longer be reconstructed separately. In order to explore the highly boosted phase space, hadronic top quark decays are reconstructed as large-radius jets. Previous efforts in this domain by the ATLAS [16, 17] and CMS [18–20] Collaborations confirm that it is feasible to perform precise differential measurements of the $t\bar{t}$ production and have also shown interesting deviations from the theory predictions.

In this paper, a measurement of the differential boosted $t\bar{t}$ production cross section in the hadronic and lepton+jets final states is presented, using pp collisions at $\sqrt{s} = 13$ TeV recorded with the CMS detector during the 2016 LHC run and amounting to a total integrated luminosity of 35.9 fb^{-1}. In the hadronic decay channel, each W boson arising from the top quark decays into a pair of light quarks. As a result, the final state consists of at least six partons (more are possible due to initial- and final-state radiation), two of which are b quarks. Due to the high boost considered in this measurement ($p_T > 400$ GeV), the top quarks are reconstructed unambiguously as large-radius jets and the final state consists of at least two such jets. In the lepton+jets channel, one top quark decays to a hadronically-decaying W boson and a b quark (t $\to$ Wb $\to$ q$\bar{\text{q}}'$b) and is reconstructed as a single large-radius jet, while the second top quark decays to a W boson, which in turn decays to a charged lepton (e$/\mu$) and a neutrino, and a b quark (t $\to$ Wb $\to$ $\ell\nu$b). The presented measurements are performed with a much larger dataset compared to the previous CMS results and utilize improved reconstruction techniques, leading to smaller uncertainties, so that a confrontation of the theoretical predictions is possible in a wider phase space.

The paper is organized as follows: Section 2 describes the main features of the CMS detector and the triggering system. Section 3 gives the details of the Monte Carlo simulations. The object reconstruction and the event selection is discussed in Sections 4 and 5, respectively. In Section 6, the estimation of the background contributions is discussed, followed by a discussion of the signal extraction in Section 7. The systematic uncertainties are discussed in Section 8. The unfolding procedure used to obtain the particle- and parton-level cross sections and the results are presented in Section 9. Finally, Section 10 provides the summary of the paper.

2 The CMS detector

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 scintilla-

tor hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity (η) coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization 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. [21].

Events of interest are selected using a two-tiered trigger system [22]. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a time interval of less than 4 μs. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage.

3 Simulated events

Monte Carlo simulation is used to generate samples for the $t\bar{t}$ signal and to model the kinematic distributions of some of the background processes. Samples of simulated $t\bar{t}$ events have been generated at next-to-leading order (NLO) in QCD using POWHEG v2 [23–27], assuming a top quark mass of $m_t = 172.5$ GeV. Single top quark production in the t channel or in association with a W boson are simulated at NLO with POWHEG [28]. The production of W or Z bosons in association with jets (V+jets), as well as QCD multijet events, are simulated with MG5_AMC@NLO [29] at leading order (LO), with the MLM matching algorithm [30]. Samples of diboson (WW, WZ, ZZ) events are simulated using PYTHIA 8.

All simulated events are processed with PYTHIA 8.212 [31, 32] for modeling of the parton showering, hadronization, and underlying event (UE). The NNPDF 3.0 [33] Parton Distribution Functions (PDF) are used throughout, and the CUETP8M1 UE tune [34] is used for all processes except for the $t\bar{t}$, $t\bar{t}$H and single top quark processes, for which the tune CUETP8M2T4 [35] is used. The simulation of the CMS detector response is based on GEANT4 [36]. Additional pp interactions in the same or neighbouring bunch crossings (pileup) are simulated with PYTHIA and overlaid with generated events according to the pileup distribution measured in data.

The various simulated processes are normalized to the best known theoretical cross sections, namely the $t\bar{t}$, V+jets, and single top quark samples are normalized to NNLO precision in QCD [37–39].

The measured cross sections for the $t\bar{t}$ process are compared to theoretical predictions provided by the following Monte Carlo models: POWHEG combined with PYTHIA for the parton showering, as described above, or combined with HERWIG ++ [40] and the corresponding EE5C UE tune [41]. In addition, a comparison is performed with MC@NLO [29] using PYTHIA for the parton showering.

4 Object reconstruction

The global event reconstruction (also called particle-flow event reconstruction [42]) aims to reconstruct and identify each individual particle in an event, with an optimized combination of all subdetector information. In this process, the identification of the particle type (photon, electron, muon, charged hadron, neutral hadron) plays an important role in the determination of the particle direction and energy. Photons (e.g. coming from π^0 decays or from electron bremsstrahlung) are identified as ECAL energy clusters not linked to the extrapolation of any

charged particle trajectory to the ECAL. Electrons (e.g. coming from photon conversions in the tracker material or from B hadron semileptonic decays) are identified as a primary charged particle track and potentially many ECAL energy clusters corresponding to this track extrapolation to the ECAL and to possible bremsstrahlung photons emitted along the way through the tracker material. Muons (e.g. from B hadron semileptonic decays) are identified as tracks in the central tracker consistent with either a track or several hits in the muon system, and associated with calorimeter deposits compatible with the muon hypothesis. Charged hadrons are identified as charged particle tracks neither identified as electrons, nor as muons. Finally, neutral hadrons are identified as HCAL energy clusters not linked to any charged hadron trajectory, or as a combined ECAL and HCAL energy excess with respect to the expected charged hadron energy deposit.

The energy of photons is obtained from the ECAL measurement. The energy of electrons is determined from a combination of the track momentum at the main interaction vertex, the corresponding ECAL cluster energy, and the energy sum of all bremsstrahlung photons attached to the track. The energy of muons is obtained from the corresponding track momentum. The energy of charged hadrons is determined from a combination of the track momentum and the corresponding ECAL and HCAL energies, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

Leptons and charged hadrons are required to be compatible with originating from the primary interaction vertex. 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 with the tracks assigned to the vertex as inputs, and the associated missing transverse momentum, taken as the negative vector sum of the p_T of those jets. Charged hadrons that are associated with a pileup vertex are classified as pileup candidates and are ignored in the subsequent object reconstruction. Electron and muon objects are first identified from corresponding electron or muon particle-flow candidates. Next, jet clustering is performed on all particle-flow candidates that are not classified as pileup candidates. The jet clustering does not exclude the electron and muon particle-flow candidates, even if these have already been assigned to electron/muon objects. A dedicated overlap removal is therefore applied at analysis level to avoid double-counting.

Muons and electrons are required to have $p_T > 50(30)$ GeV and $|\eta| < 2.1$ for the l+jets (hadronic) channel. Leptons must also be isolated according to the "mini isolation" (I_{mini}) algorithm, which requires the p_T sum of tracks in a cone around the muon or electron to be less than a given fraction of the lepton p_T. The cone width varies with the lepton p_T as $R = 10/p_T^\ell$ for $p_T^\ell < 200$ GeV and $R = 0.05$ for $p_T^\ell > 200$ GeV. A cut value of $I_{mini} < 0.1$ is used. A looser selection is used to define veto leptons for the l+jets channel, requiring $p_T > 50$ GeV and $|\eta| < 2.1$, but no isolation requirement. Correction factors are applied to account for differences between data and the simulation in the modeling of the lepton identification and trigger efficiencies, determined as a function of the muon/electron $|\eta|$ and p_T.

For each event, hadronic jets are clustered from the particles reconstructed by the particle-flow algorithm using the infrared and collinear safe anti-k_T algorithm [43, 44]. Two different jet collections are considered for the analysis. Small-R jets are clustered using a distance parameter of 0.4 (used for the l+jets channel) and large-R jets using a distance parameter of 0.8 (used for the l+jets and hadronic channels). The jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance. Additional

proton-proton interactions within the same or nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, charged particles identified to be originating from pileup vertices are discarded and an offset correction is applied to correct for remaining contributions.

Jet energy corrections are derived from simulation to bring the measured response of jets to that of particle-level jets on average. In situ measurements of the momentum balance in dijet, photon + jet, Z + jet, and multijet events are used to account for any residual differences in jet energy scale between data and simulation [45]. The jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV. Additional selection criteria are applied to each jet to remove jets potentially dominated by anomalous contributions from various subdetector components or reconstruction failures.

In order to identify jets originating from hadronic decays of top quarks (t tagging) we use the 3-subjettiness, 2-subjettiness, and 1-subjettiness variables [46], τ_3, τ_2, and τ_1 computed using the jet constituents, according to

$$\tau_N = \frac{1}{\sum_k p_{T,k} R} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots \Delta R_{N,k}\}, \tag{1}$$

where N denotes the reconstructed candidate subjets and k runs over the constituent particles in the jet. The variable $\Delta R_{i,k} = \sqrt{\left(\Delta y_{i,k}\right)^2 + \left(\Delta \phi_{i,k}\right)^2}$ is the angular distance between the candidate subjet i and the particle k. The variable R is the characteristic jet radius ($R = 0.8$ in our case). The centers of hard radiation are found by performing the exclusive k_T algorithm [47, 48] on the jet constituents before the application of any grooming techniques.

In addition, a grooming technique is applied to remove soft, wide-angle radiation from the jet and to improve the mass resolution. The algorithm employed is the "modified mass drop tagger" [49, 50], also known as the *soft drop* (SD) algorithm [51], with angular exponent $\beta = 0$, soft cutoff threshold $z_{cut} < 0.1$, and characteristic radius $R_0 = 0.8$ [51]. The subjets within the large-R jets are identified by reclustering of their constituents with the Cambridge–Aachen algorithm [52, 53] and by reversing the last step of the clustering history.

Small-R jets and subjets of the large-R jets are identified as b quark candidates (b tagged) using the Combined Secondary Vertex (CSV) algorithm [54]. Data-to-simulation correction factors are used to match the b tagging efficiency observed in simulation to that measured in data. The typical efficiency for large-R (subjets) of the b tagging algorithm is 0.63 (0.58) for true b jets, while the misidentification probability for non b jets is 1%. For the subjets, the efficiency for the tagging of true b jets drops from 0.65 to 0.4 as the p_T increases from 20 GeV to 1 TeV.

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

5 Event selection

5.1 Trigger

Different triggers were employed for the collection of signal events in the l+jets and hadronic channels according to each event topology. The trigger employed in the hadronic channel required the presence of a jet with $p_T > 180$ GeV at L1. At HLT, large-R jets were reconstructed

from particle-flow candidates using the anti-k_T algorithm with a distance parameter of 0.8. The mass of the jets at HLT, after removal of soft particles, was required to be greater than 30 GeV. The selected events contained at least two such jets with $p_T > 280(200)$ GeV for the leading (trailing) one. Finally, at least one of these jets was required to be b tagged, using the CSV algorithm suitably adjusted for HLT, with an average efficiency of 90% for b jets. The aforementioned trigger ran for the entire 2016 run, collecting an integrated luminosity of 35.9 fb^{-1}. A second trigger with identical kinematic requirements but no b tagging, was employed and ran on average every 21 bunch crossings, collecting an integrated luminosity of 1.67 fb^{-1}. The events collected with the latter trigger are used for the selection of a control sample for estimating the QCD multijet background, as described later. For the l+jets channel, the data were selected using triggers requiring a single lepton, either a muon with $p_T > 40$ GeV and $|\eta| < 2.1$ or an electron with $p_T > 45$ GeV, with no isolation criteria applied, as well as two jets with $p_T > 200$ and 50 GeV, respectively.

5.2 Hadronic channel

The baseline selection is common for all regions used in the hadronic analysis and it requires at least two large-R jets in the event with $p_T > 400$ GeV and soft-drop masses in the range 50–300 GeV. Also, events with at least one signal lepton (muon or electron) are vetoed in order to suppress leptonic top decays.

In order to discriminate between events that originate from $t\bar{t}$ decays vs QCD multijet production, jet substructure variables are utilized. These variables are sensitive to the type of the jet, and in particular to whether it comes from a single-prong decay, like ordinary quark or gluon jets, or from a three-prong decay, such as the $t \to Wb \to \bar{q}q'b$ decay that is of interest here. The $\tau_{1,2,3}$ variables of the two leading, large-R jets are combined with a neural network (NN) to form a multivariate event discriminant that characterizes each event, with values close to zero indicating QCD dijet production and values close to one suggesting $t\bar{t}$ production. The variables have been chosen such that the correlation to the number of b tagged subjets is minimal, which is used to define QCD control regions. The NN consists of two hidden layers with 16 and 4 nodes, respectively, and is implemented with the TMVA toolkit [55]. More complicated architectures have not improved the discriminating capabilities of the NN. The training of the NN was performed with simulated QCD multijet (background) and $t\bar{t}$ (signal) events that satisfy the baseline selection, using the back-propagation method and a sigmoid activation function for the nodes. Furthermore, excellent agreement between data and simulation has been observed for the input variables in the phase space of the training.

On top of the baseline selection, sub-selection regions are defined based on the NN output, the jets' soft-drop masses, and the number of b tagged subjets in each large-R jet, which serve different analysis purposes. The signal region (SR) is where the differential measurements are performed and it contains events collected with the signal trigger where both large-R jets have a b tagged subjet and their masses lie in the range 120–220 GeV, while the NN output value is greater than 0.8. In this selection, more than 95% of the selected $t\bar{t}$ events come from hadronic top decays. The QCD control region (CR) contains events collected with the control trigger that satisfy the same requirements as the SR but with the b tagging requirement inverted (the large-R jets should not contain a b tagged subjet). Additionally, expanded regions that include SR and CR are defined for estimates of the background distributions. The signal region A (SR_A) is the same as SR without the $NN > 0.8$ requirement, while the corresponding control region (CR_A) has the b tagging condition reverted. It should be noted that the events selected in SR_A and CR_A were collected with the signal and control triggers, respectively. Finally, the signal region B (SR_B) is used to constrain some of the signal modeling uncertainties and has the same

²²² selection criteria as the *SR* but without any NN requirement.

5.3   Lepton+jets channel

²²⁴ The lepton+jets final state is identified through the presence of an electron or a muon, a small-*R*
²²⁵ jet assumed to be produced by the decay of the b quark from the leptonic top quark decay, and
²²⁶ a large-*R* jet corresponding to the hadronically decaying top quark.

²²⁷ Small-*R* jets, i.e. anti-k_t jets with a size parameter of 0.4, are required to have $p_T > 50\,\text{GeV}$ and
²²⁸ $|\eta| < 2.4$. Large-*R* jets (anti-k_t jets with a size parameter of 0.8) must fulfill $p_T > 400\,\text{GeV}$ and
²²⁹ $|\eta| < 2.4$.

²³⁰ Events are required to first pass a so-called preselection, defined as follows:

²³¹   • Exactly one signal muon or signal electron.
²³²   • Zero additional veto leptons.
²³³   • At least one leptonic-side jet (small-*R* jet in the same hemisphere as the signal lepton,
²³⁴ $0.3 < \Delta R(\ell, \text{jet}) < \pi/2$).
²³⁵   • At least one hadronic-side jet (large-*R* jet in the opposite hemisphere from the signal
²³⁶   lepton, $\Delta R(\ell, \text{jet}) > \pi/2$).
²³⁷   • $p_T^{\text{miss}} > 35\,\text{GeV}$ (muon channel) or $> 50\,\text{GeV}$ (electron channel).
²³⁸   • For events in the electron channel: 'triangular' cut, $|\Delta\phi(X, p_T^{\text{miss}})| < 1.5 \cdot \frac{p_T^{\text{miss}}}{7\,\text{GeV}}$ for
²³⁹   both X = electron and X = leading small-*R* jet, to ensure that the missing transverse
²⁴⁰   momentum vector does not point along the transverse direction of the electron or
²⁴¹   leading jet).

²⁴² The more stringent p_T^{miss} criteria and the triangular cut in the electron channel are applied to
²⁴³ further reduce background from QCD multijet production.

²⁴⁴ Events that fulfill the preselection criteria are further categorized based on whether the b(t) jet
²⁴⁵ candidate passes or fails the b(t) tagging criteria. The N-subjettiness ratio τ_3/τ_2 (abbreviated
²⁴⁶ as τ_{32}) is used as the sensitive variable to distinguish a three-prong hadronic top quark decay
²⁴⁷ from background processes, requiring $\tau_{32} < 0.81$. A data-to-simulation efficiency correction
²⁴⁸ factor is extracted simultaneously with the integrated signal yield, as described in Section 7, to
²⁴⁹ correct the t tagging efficiency in simulation to match that in data. The b(t) jet candidate is the
²⁵⁰ highest-p_T leptonic-side (hadronic-side) jet in the event. Events are divided into the following
²⁵¹ categories:

²⁵²   • 0 t tag (0t): the t jet candidate fails the t tagging requirement.
²⁵³   • 1 t tag, 0 b tag (1t0b): the t jet candidate passes the t tagging requirement, but the b
²⁵⁴   jet candidate fails the b tagging requirement.
²⁵⁵   • 1 t tag, 1 b tag (1t1b): both the t jet candidate and the b jet candidate pass their
²⁵⁶   respective tagging requirement.

²⁵⁷ These event categories are designed to have different admixtures of signal and background,
²⁵⁸ with the 0t region being the most background dominated and the 1t1b region the most signal
²⁵⁹ dominated.

6 Background estimation

The dominant background in the hadronic channel is the QCD multijet production, while the dominant background processes contributing in the l+jets channel include non-signal $t\bar{t}$, single top quark, W+jets, and QCD multijet production. Non-signal $t\bar{t}$ comprises dileptonic and hadronic final states, in addition to τ+jets events where the tau lepton does not decay to a muon or an electron.

In the hadronic channel, the QCD background is suppressed significantly with a combination of b tagging requirements on the subjets and the event NN output. The remaining contribution is estimated from a control data sample. The two elements that are determined from data are the shape $Q(x)$ of the QCD background as a function of an observable of interest x, and the absolute normalization N_{QCD}. The QCD template $Q(x)$ is taken from the control region CR_A in data. The $t\bar{t}$ contamination of the control sample, based on the simulation, is of the order of 1%. The normalization of the QCD background (N_{QCD}) is determined from a maximum likelihood fit to the data in SR_A in the soft-drop mass of the top jet candidate (this is taken as the leading large-R jet). The fit is described by the equation below:

$$D(m^t) = N_{t\bar{t}}T(m^t; k_{\text{scale}}, k_{\text{res}}) + N_{\text{QCD}}(1 + k_{\text{slope}}m^t)Q(m^t) + N_{\text{bkg}}B(m^t) \qquad (2)$$

which contains the shapes $T(m^t)$, $B(m^t)$ of the signal and the subdominant backgrounds, respectively, taken from the simulation, and the shape $Q(m^t)$ of the QCD multijet background, taken from the control sample in data. To account for a possible QCD shape difference on the m^t variable between the control and signal regions, a linear modification factor $(1 + k_{\text{slope}}m^t)$ is introduced, inspired by the simulation, but with the slope parameter k_{slope} left free in the fit. Also free in the fit are the normalization factors $N_{t\bar{t}}$, N_{QCD}, and N_{bkg}. Finally, two more nuisance parameters are introduced in the simulated $t\bar{t}$ shape, k_{scale} and k_{res}, which account for possible differences between data and simulation in the scale and resolution of the m^t parameter. The fit is performed with the ROOFIT toolkit [56] and the fit results are shown in Fig. 1 and Table 1. The fitted $t\bar{t}$ yield is significantly lower than the expectation (9885), which implies that the fiducial cross section is lower compared to the POWHEG +PYTHIA prediction with a posterior signal strength of 0.64 ± 0.03. This value is used to scale down the expected $t\bar{t}$ signal yields from the POWHEG +PYTHIA simulation in the various signal regions in the subsequent figures with data vs. simulation comparisons. Also, the nuisance parameters that control the scale and the resolution of the reconstructed mass are consistent with one, confirming the excellent agreement between data and simulation of this variable.

Table 1: Postfit values of the nuisance parameters of the fit on data in SR_A in the hadronic channel.

Parameter	Value
k_{res}	0.960 ± 0.026
k_{scale}	1.002 ± 0.002
k_{slope}	$(5.7 \pm 1.4) \times 10^{-3}$
N_{bkg}	400 ± 255
N_{QCD}	4539 ± 247
$N_{t\bar{t}}$	6238 ± 181

The subdominant background processes, namely single top quark production and vector bosons produced in association with jets, have a negligible contribution in the signal region (below 1% in the entire phase space) and are fixed to the simulation predictions.
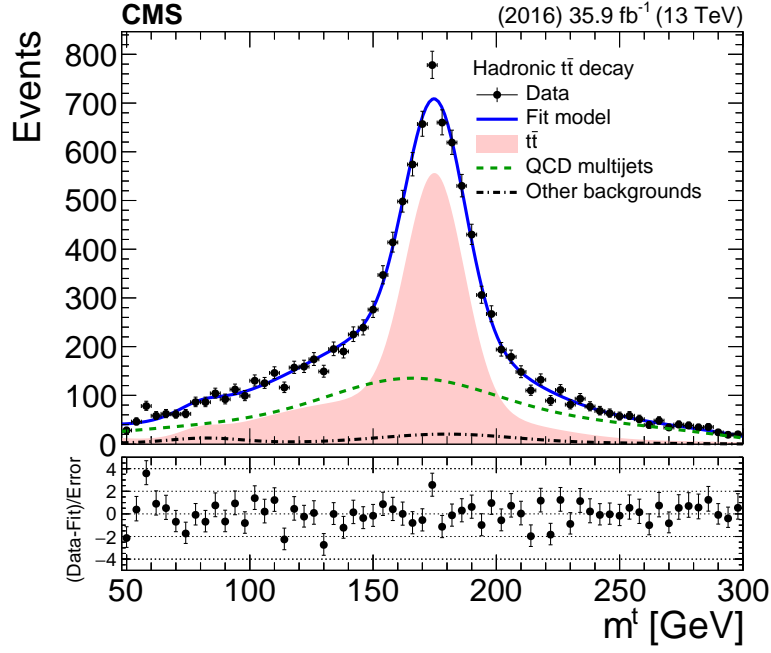
Figure 1: Result of the fit on data in SR_A in the hadronic channel. The shaded area shows the t$\bar{\text{t}}$ contribution, the dashed line shows the QCD multijet background, and the double-dashed line shows the other subdominant backgrounds. The solid line shows the combined signal plus background fit model.

Figure 2 shows the distribution of the NN output in SR_B and Figs. 3, 4 show the kinematic properties of the two top candidates and the t$\bar{\text{t}}$ system, respectively. Also, the softdrop masses of the two jets are shown in Fig. 5. In these figures the t$\bar{\text{t}}$ and QCD multijet processes are normalized according to the results of the fit in SR_A described above, while the yields of the subdominant backgrounds are taken from the simulation. Table 2 summarizes the events yields in SR.

Table 2: Observed event yields and statistical undertainties in the signal region (hadronic channel). The t$\bar{\text{t}}$ and QCD multijet yields are obtained after the fit in SR_A.

Process	Yield
t$\bar{\text{t}}$	4244 ± 127
QCD	1876 ± 102
W+jets	58 ± 29
Z+jets	12 ± 6
Single Top	83 ± 41
Data	6274
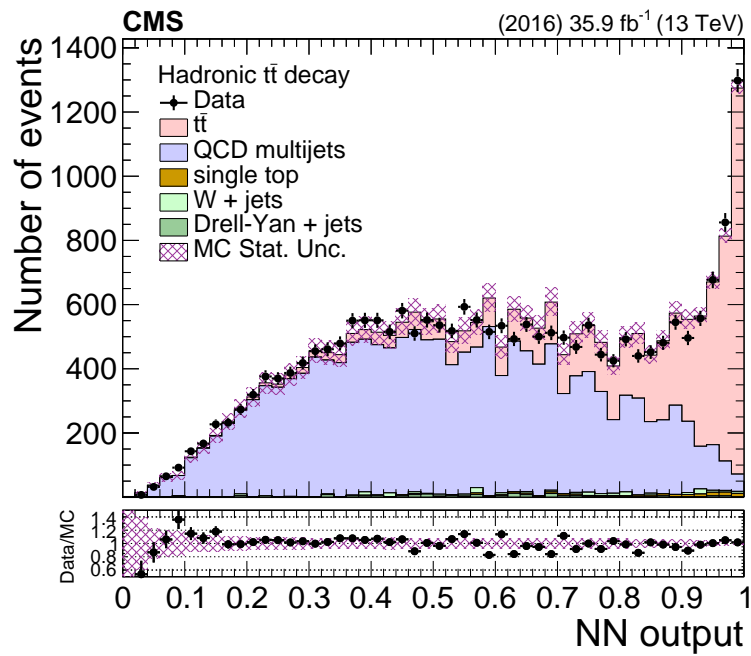
Figure 2: Data vs simulation in SR_B (same as the SR, without the NN cut) for the NN output in the hadronic channel. The $t\bar{t}$ and QCD processes are normalized according to the postfit values of the respective yields. The data points are shown with solid markers, while the shaded band represents the statistical uncertainty on the simulation.

(a) p_T^1

(b) p_T^2

(c) $|y^0|$

(d) $|y^1|$
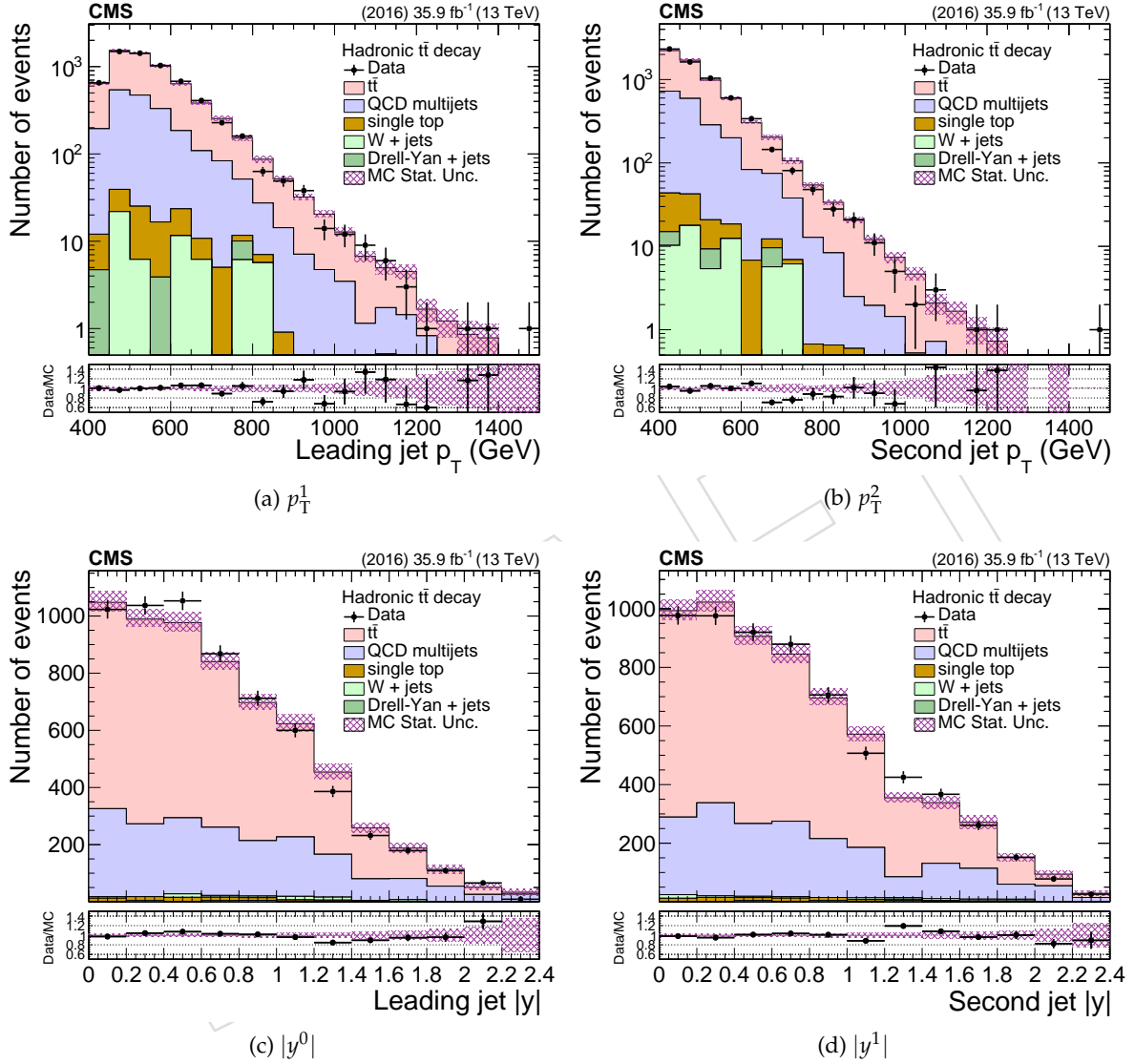
Figure 3: Data vs simulation in the signal region for the p_T and absolute rapidity of the leading and subleading large-R jets in the hadronic channel. The $t\bar{t}$ and QCD processes are normalized according to the postfit values of the respective yields. The data points are shown with solid markers, while the shaded band represents the statistical uncertainty on the simulation.

(a) $m_{t\bar{t}}$

(b) $p_{T,t\bar{t}}$

(c) $y_{t\bar{t}}$
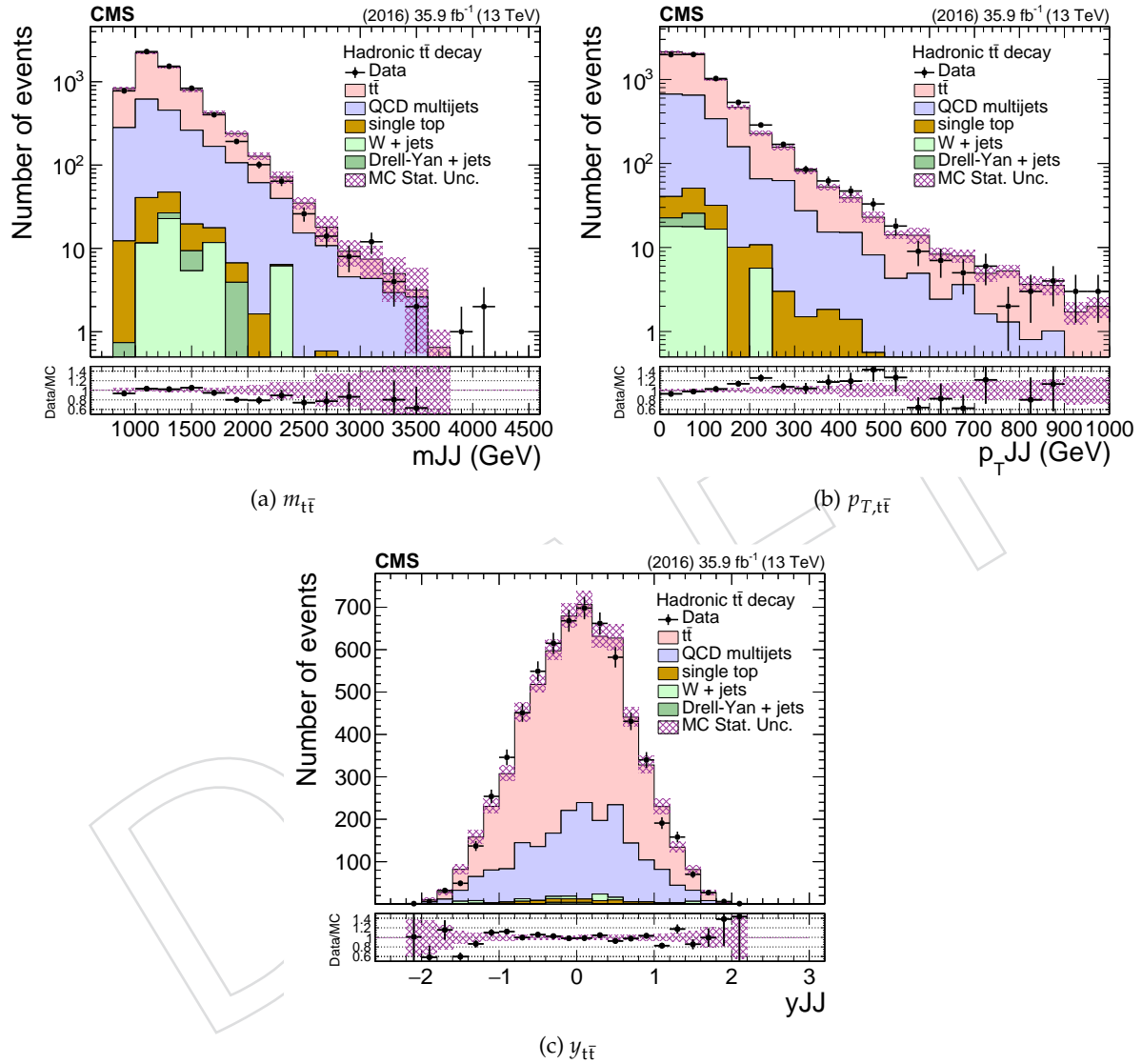
Figure 4: Data vs simulation in the signal region for the $t\bar{t}$ system variables in the hadronic channel. The $t\bar{t}$ and QCD processes are normalized according to the postfit values of the respective yields. The data points are shown with solid markers, while the shaded band represents the statistical uncertainty on the simulation.
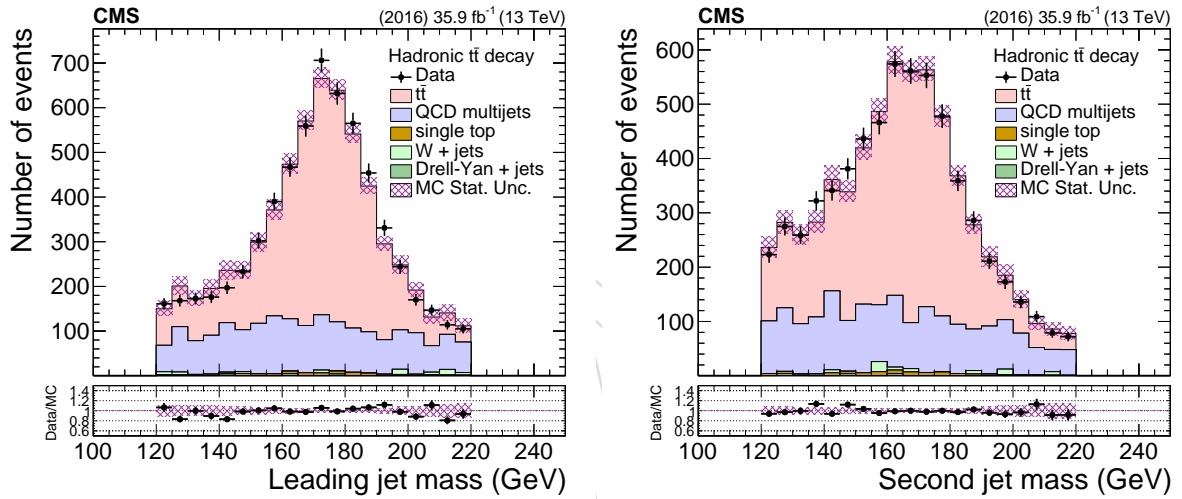
Figure 5: Data vs simulation in the signal region for the mass of the leading (left) and second (right) large-R jets in the hadronic channel. The $t\bar{t}$ and QCD processes are normalized according to the postfit values of the respective yields. The data points are shown with solid markers, while the shaded band represents the statistical uncertainty on the simulation.

In the l+jets channel, background events due to non-signal $t\bar{t}$, single top quark, W/Z+jets, and diboson production are estimated from simulation. The QCD multijet background is modeled with data-driven techniques, using a data sideband region dominated by multijet events. The sideband is defined by inverting the isolation requirement on the lepton and relaxing the lepton identification criteria. The predicted contributions to the sideband region from signal and other background events are subtracted from the observed data in the sideband to yield the kinematic distributions for QCD multijet events. The normalization of the multijet background is extracted from a maximum-likelihood fit, described in Section 7.2, but an initial estimate of its normalization is taken from the simulated prediction in the signal region. Also the normalization of the other background processes is constrained through the fit.

7 Signal extraction

7.1 Hadronic channel

The $t\bar{t}$ signal is extracted from data, differentially, as a function of seven variables: p_T and $|y|$ of the leading and second top jet, $t\bar{t}$ mass, p_T, and rapidity, by subtracting the contribution of the backgrounds:

$$S(x) = D(x) - R_{\text{yield}}N_{\text{QCD}}Q(x) - B(x), \tag{3}$$

where $x = p_T^{t\,1,2}, |y^{t\,1,2}|, m^{t\bar{t}}, pt^{t\bar{t}}, y^{t\bar{t}}$, $S(x)$ is the $t\bar{t}$ signal, $D(x)$ is the measured distribution in data, $Q(x)$ is the QCD shape, $B(x)$ is the subdominant backgrounds' contribution (both the shape and the normalization are taken from the simulation), and N_{QCD} is the fitted number of QCD events in SR_A. The factor R_{yield} is found (in simulation) to be independent of the b tagging requirement and therefore it is estimated from the QCD control data sample: $R_{\text{yield}} = N_{\text{QCD}}^{SR}/N_{\text{QCD}}^{SR_A} = N_{\text{QCD}}^{CR}/N_{\text{QCD}}^{CR_A} \approx 0.42$.

7.2 Lepton+jets channel

For the l+jets analysis, the signal strength, t tagging efficiency scale factor, and background normalizations are determined in a simultaneous binned maximum-likelihood fit to the data across the different analysis event categories.

The 0t, 1t0b, and 1t1b event categories are fitted simultaneously, normalizing each background component to the same cross section in all categories. The fit result is expressed in terms of a multiplicative factor, the signal strength r, applied to the input $t\bar{t}$ cross section. Different variables are used to discriminate the $t\bar{t}$ signal from the background processes. The small-R jet η distribution is used in the 0t and 1t0b categories, while the large-R jet soft-drop mass distribution is used in the 1t1b region. These distributions were chosen to have good discrimination between $t\bar{t}$, W+jets, and QCD multijet production. The $t\bar{t}$ signal and $t\bar{t}$ background distributions are merged into a single distribution for the purpose of the fit, essentially constraining the semileptonic branching ratio to be equal to that provided by the simulation.

Background normalizations and experimental sources of systematic uncertainty are treated as nuisance parameters in the fit. The uncertainties due to pileup reweighting, lepton scale factors, jet energy scale and resolution, b tagging efficiency, and t tagging efficiency are treated as uncertainties in the distribution of the input templates. Two nuisance parameters are used to describe the t tagging uncertainty: a t tag scale factor nuisance parameter is applied to the $t\bar{t}$ and single top quark samples, where we expect the t tagged jet to correspond to a true top

340 quark decay, while a t mis-tag scale factor is applied for the W+jets and QCD multijet samples.
341 The uncertainties in the luminosity and background normalizations are treated as rate uncer-
342 tainties. The event categories that are fitted are designed such that the t tagging efficiency is
343 constrained by the relative population of events in the three categories. The varying admixtures
344 of signal and background between the different regions provide power to determine the back-
345 ground normalizations. The measurement of the $t\bar{t}$ cross section is correlated with the various
346 nuisance parameters, with the strongest correlation being with the t tagging efficiency, as ex-
347 pected. For the shape uncertainties, the nuisance parameter is used to interpolate between the
348 nominal kinematic distribution and distributions corresponding to $\pm 1\sigma$ variations in the given
349 uncertainty. The uncertainties due to the theoretical modeling are evaluated independently
350 from the fit.

351 The e/μ+jets channels are fitted together, with most nuisance parameters constrained to be
352 the same in both channels. The nuisance parameters associated with the electron and muon
353 scale factors are treated separately, as are the muon and electron QCD multijet background
354 normalizations. The posterior kinematic distributions for the three fit categories are shown in
355 Fig. 6. The event counts that account for all posterior parameters are given in Table 3.

Sample	Number of events (μ+jets channel)		
	0t	1t0b	1t1b
$t\bar{t}$	18300 ± 1744	4226 ± 189	3865 ± 94
Single t	3209 ± 719	263 ± 81	144 ± 39
W+jets	22070 ± 2919	2456 ± 336	104 ± 21
Z+jets	2249 ± 597	216 ± 68	18 ± 10
Diboson	550 ± 176	34 ± 12	2 ± 1
Multijet	2761 ± 1219	166 ± 79	46 ± 25
Total	49140 ± 3735	7360 ± 408	4179 ± 107
Data	49137	7348	4187

Sample	Number of events (e+jets channel)		
	0t	1t0b	1t1b
$t\bar{t}$	9668 ± 1314	2907 ± 126	2729 ± 68
Single t	1978 ± 400	178 ± 52	98 ± 27
W+jets	15300 ± 2320	1384 ± 205	56 ± 12
Z+jets	1351 ± 378	134 ± 46	19 ± 17
Diboson	378 ± 104	20 ± 7	2 ± 1
Multijet	2879 ± 1328	208 ± 111	22 ± 23
Total	31555 ± 3030	4831 ± 274	2925 ± 80
Data	31559	4801	2953

Table 3: Posterior signal and background normalizations in the 0t, 1t0b, and 1t1b regions, together with observation in data. The uncertainties include all post-fit experimental uncertainties.

356 Figure 7 shows the p_T and y distributions for the t jet candidate in each of the three kinematic
357 regions for the combined lepton+jets channel. These distributions use the posterior t tagging
358 scale factor and background normalizations, but not the posterior values of other nuisance
359 parameters.

360 The posterior signal strength as determined in the fit is 0.77 ± 0.06, i.e. the $t\bar{t}$ simulation is
361 observed to overestimate the data by roughly 20%. Although the measured signal strength
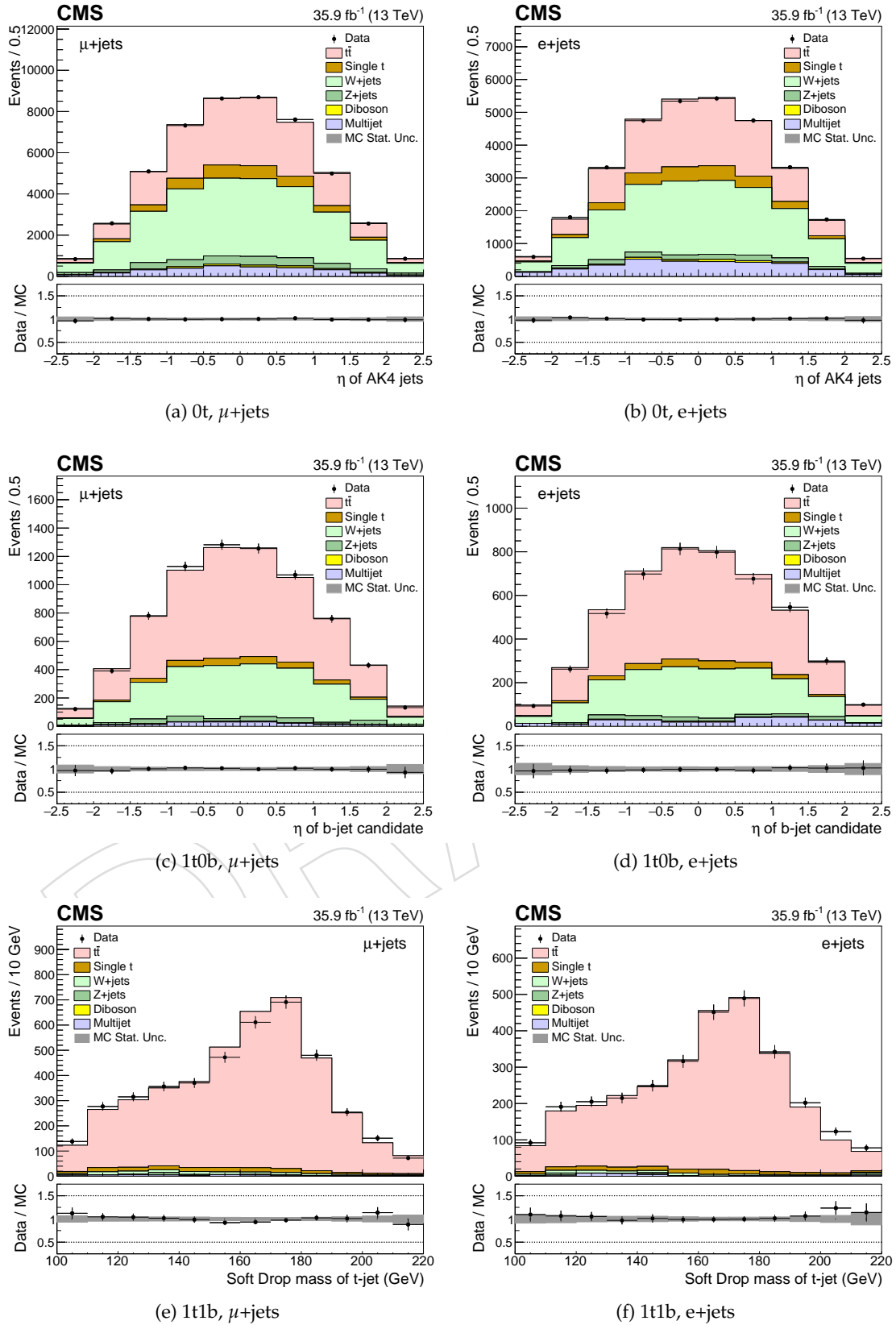
Figure 6: Posterior kinematic distributions for the maximum likelihood fit. Six different event categories are used: (top) 0t, (middle) 1t0b, and (bottom) 1t1b events, in the (top) μ+jets and (right) e+jets channel.

(a) 0t

(b) 0t

(c) 1t0b

(d) 1t0b

(e) 1t1b

(f) 1t1b
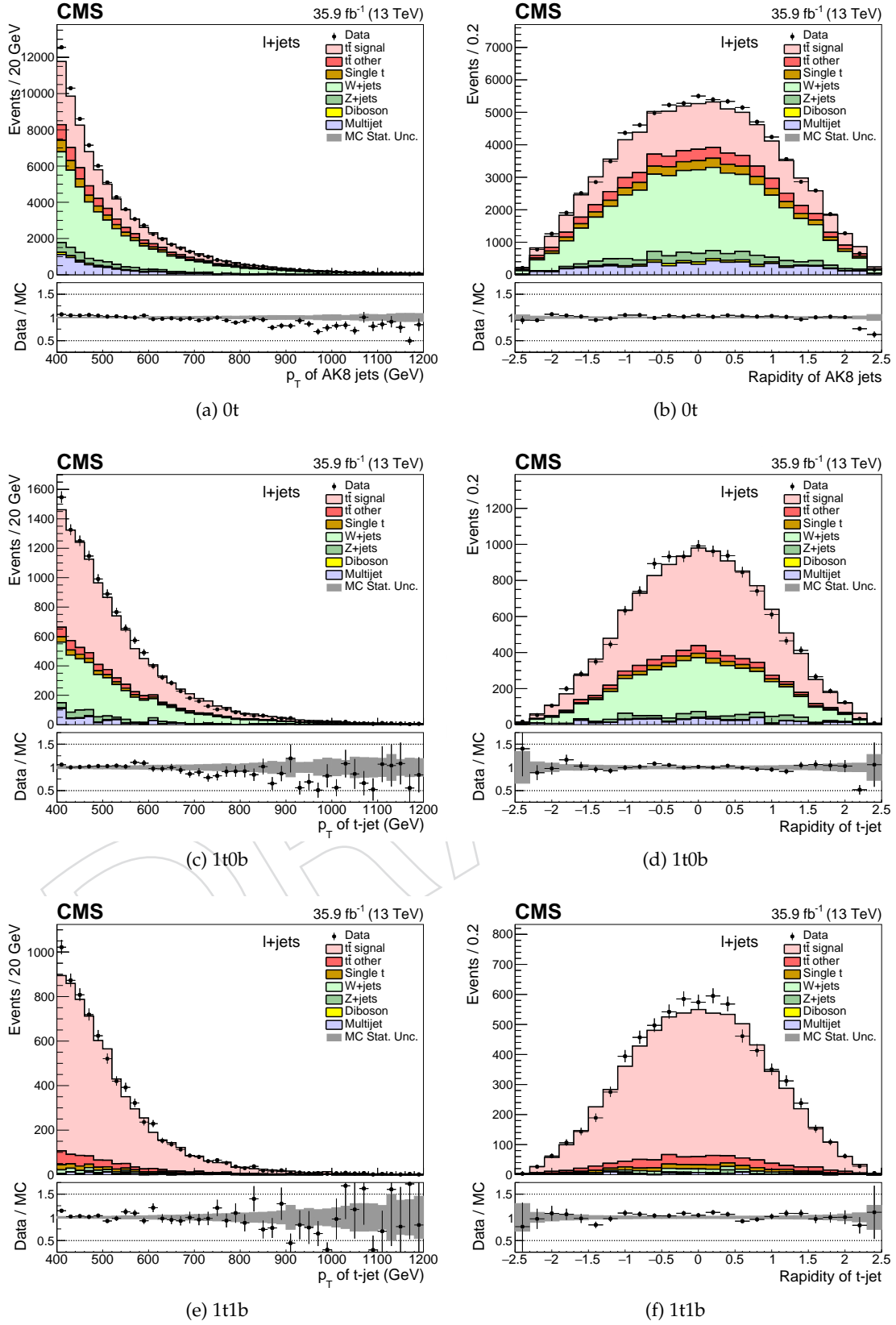
Figure 7: Distributions of the (left) p_T and (right) y of the t jet candidate for the (top) 0t, (middle) 1t0b, and (bottom) 1t1b regions, for events in the combined lepton+jets channel. Posterior t tag scale factor and background normalizations are used.

extrapolated from the fit is not directly used to compute an integrated cross section, it serves as an indicator of the level of agreement between the measured inclusive $t\bar{t}$ cross section and the prediction from simulation in the boosted regime.

8 Systematic uncertainties

The systematic uncertainties considered in this analysis originate from both experimental and theoretical sources. The former includes all uncertainties related to the differences in the particle reconstruction and identification performance between data and simulation, as well as the background modeling. The latter are related to the simulation itself and affect primarily the unfolded results through the acceptance, efficiency, and migration matrices. For each systematic variation, described in detail below, the differential cross sections are re-measured and the difference with respect to the nominal result is taken as the effect of this variation on the measurement.

A number of experimental uncertainties affect the measurement. The dominant uncertainties in the hadronic channel are due to the jet energy scale and the subjet b tagging efficiency. In the l+jets channel, the t tagging and b tagging efficiency uncertainties result in the largest contributions. The different sources are described in detail in the list below.

- *QCD multijet background (hadronic):* The fitted QCD yield uncertainty is used. The impact of the shape uncertainties and the different pileup profiles in the control and signal regions are of the order of 1%. In order to estimate the uncertainties of the shape of the QCD multijet background, a test was performed with simulated events, comparing the distributions of each variable between the signal region *SR* and the control region *CR* (see Section 5).

- *Subdominant backgrounds (hadronic):* The expected yield of the subdominant background processes estimated from the simulation (single top quark production and vector bosons produced in association with jets) is varied by 50%, leading to a negligible uncertainty (smaller than 1%).

- *Background estimate (l+jets):* An a priori uncertainty of 30% is applied to the single top quark and W+jets background normalizations. An additional uncertainty in the flavor composition of the W+jets is considered, allowing the light- and heavy-flavor components to vary separately with a 30% normalization uncertainty. For the QCD multijet normalization, a 50% uncertainty is used to reflect the combined uncertainty in the normalization and the extraction of the kinematic distributions from the data sideband region. These are all constrained in the maximum likelihood fit.

- *Jet energy scale:* The uncertainty in the energy scale of each reconstructed large-*R* jet, the jet energy scale (JES) uncertainty, is a leading experimental uncertainty for the hadronic channel. It is divided in 24 independent sources [45] and treated as follows: for each variation a new jet collection is created and the event interpretation is repeated. This results not only in variations of the p_T scale itself, but may also lead to different top jet candidates. The JES uncertainty, per jet, is of the order $1-2\%$, p_T and η dependent. The effect on the measured cross section is typically of the order 10% but it can be much larger at high top quark p_T. For the l+jets channel, the jet energy scale uncertainty is estimated for both small-*R* and large-*R* jets by shifting the jet energy in simulation up and down by their p_T and η dependent uncertainties, with a resulting impact on the differential cross section measurements of 1–10%.

- *Jet energy resolution:* The impact on the measurement due to the jet energy resolution

(JER) is determined by smearing the jets according to the JER uncertainty [45]. The effect on the cross section is relatively small, at the level of 2%.

- *t tagging efficiency (l+jets):* The t tagging efficiency and its associated uncertainty is simultaneously extracted with the signal strength and background normalizations in the likelihood fit for the l+jets analysis. This is described in Section 7.

- *Subjet b tagging efficiency (hadronic):* The uncertainty in the identification of b-subjets within the large-R jets (estimated in [54]) is a leading experimental uncertainty in the hadronic channel. The effect on the cross sections is of the order of 10%, relatively flat in all the observables. Unlike the uncertainty associated with the JES, the b-subjet tagging uncertainty therefore largely cancels in the normalized cross sections.

- *b tagging efficiency (l+jets):* For the l+jets channel, the small-R jet b tagging efficiency observed in the simulation is corrected to match that in data using p_{T} and η dependent scale factors [54]. The resulting uncertainty in the differential cross-sections varies between about 1–2%. The b tagging efficiency and light jet misidentification rates are treated as fully correlated.

- *Pileup:* The uncertainty related to the modeling of additional pileup interactions is a subdominant uncertainty. The impact on the measurement is estimated by varying the total inelastic cross section used to weight the simulated events by $\pm 4.6\%$ [57]. The effect on the cross sections is negligible (below 1%).

- *Trigger (hadronic):* The uncertainty associated with the trigger, accounting for the difference between the simulated and observed trigger efficiency, is well below 1% in the phase space for the hadronic channel.

- *Lepton identification and trigger (l+jets):* The performance of the lepton identification, reconstruction, trigger, and isolation provides a small source of uncertainty. Correction factors used to modify the simulation to match the efficiencies observed in data are estimated with a tag-and-probe method using $Z \to \ell\ell$ decays. The corresponding uncertainty is determined by varying the correction factors up or down within their uncertainties. The resulting uncertainties depend on p_{T} and η and range between 1–5% and 1–7% for muon and electrons, respectively.

- *Luminosity:* The uncertainty in the measurement of the integrated luminosity is 2.5% [58].

The theoretical uncertainties are divided into two sub-categories: the ones related to the matrix element of the hard scatter process and the ones related to the modeling of the parton shower and the underlying event. The first category (consisting of the first three sources below) is evaluated by variations of the simulated event weights, while the second category is evaluated with dedicated, alternative Monte Carlo samples with modified parameters.

- *Parton distribution functions:* The uncertainty due to parton distribution functions (PDFs) is estimated by applying event weights corresponding to the 100 replicas of the NNPDF set [33]. For each observable we compute its standard deviation from the 100 variations.

- *Renormalization and factorization scales:* This source of systematic uncertainty is estimated by applying event weights corresponding to different factorization and renormalization scale options. Both scales are varied independently by a factor of two up or down in the event generation, omitting the two cases where the scales are varied in opposite directions, and taking the envelope of the six resulting variations.

- *Strong coupling constant (α_S):* The uncertainty associated with the α_S is estimated

by applying event weights corresponding to higher and lower values of α_S for the matrix element using the variations of the NNPDF set [33].

- *Final and initial state radiation:* The uncertainty in the initial state radiation (ISR) and the final state radiation (FSR) is estimated from alternative Monte Carlo samples with reduced and increased value for the strong coupling constant used by PYTHIA to generate initial or final state radiation. The scale in the ISR is varied by factors of 2 and 0.5, and the scale in the FSR by factors of $\sqrt{2}$ and $1/\sqrt{2}$.

- *Matrix element – parton shower matching*: In the POWHEG matrix element to parton shower (ME-PS) matching scheme, the resummation damping factor h_{damp} is used to regulate high-p_T radiation. The nominal value is $h_{damp} = 1.58 m_t$, with m_t being the generated top mass. Uncertainties in h_{damp} are parameterized by considering alternative simulated samples with $h_{damp} = m_t$ and $h_{damp} = 2.24 m_t$.

- *Underlying event tune:* This uncertainty is estimated from alternative Monte Carlo samples with the tune CUETP8M2T4 parameters varied by $\pm 1\sigma$.

9 Cross section measurements

This section discusses the differential cross section measurements for the hadronic and lepton+jets channels.

9.1 Definition of particle and parton level

The partonic phase space to which the measurement is unfolded is constrained by the kinematic requirements of the detector-level fiducial region. Namely, for the hadronic channel, the top and anti-top partons must have $p_T > 400 \, \text{GeV}$ and $|\eta| < 2.4$, while the invariant mass of the $t\bar{t}$ system must be greater than $800 \, \text{GeV}$ in order to avoid extreme events with high top p_T and very low $m_{t\bar{t}}$.

For the lepton+jets channel, the parton level definition is different from the hadronic channel because the differential cross section is reported as a function of the hadronically decaying top quark. Therefore, the parton level phase space is confined to semileptonic events, where the p_T of the hadronically decaying top quark is greater than $400 \, \text{GeV}$, to match the fiducial requirement at detector level.

The so-called *particle level* represents the state that consists of stable particles originating from the proton-proton collision, after the hadronization process, and before the interaction of these particles with the detector. The observables computed from the particles' momenta are thought to be better defined compared to the ones computed from parton information. Also, the associated phase space is closer to the fiducial phase space of the measurement at detector level, which results in smaller theoretical uncertainties. In the context of this analysis, particle jets are reconstructed from stable particles, excluding neutrinos, with the anti-k_T algorithm of distance parameter 0.8, identical to the detector-level reconstruction. It should be noted that only particles originating from the primary interaction are considered. Subsequently, jets that are geometrically matched, within $\Delta R < 0.4$ in $\eta - \phi$, to generated leptons (i.e. from the leptonic decays of the W boson) are removed from the particle jet collection. Finally, the two particle jets with the highest p_T are considered the particle-level top-quark candidates for the hadronic channel. In order to match as closely as possible the fiducial phase space, the same kinematic cuts are applied as for detector-level events. In particular, the particle jets must have $p_T > 400 \, \text{GeV}$ and $|\eta| < 2.4$, while the mass of each jet should be in the range 120–220 GeV and the invariant mass of the two jets greater than 800 GeV. The matching efficiency between the

particle-level top-quark candidates and the original top quarks at parton level varies between 96% and 98%.

The particle level phase space for the lepton+jets channel is set up to mimic the kinematic selections at detector level. Particle-level large-R jets are selected if they fulfill $p_T > 400$ GeV, $|\eta| < 2.4$, and $105 < m_{jet} < 220$ GeV, and are then referred to as particle-level top jets. Particle-level small-R jets are selected if they have $p_T > 50$ GeV, $|\eta| < 2.4$, and are flagged as a b jet), referred to as particle-level b jets. Particle-level electrons (muons) are selected if they have $p_T > 50$ GeV and $|\eta| < 2.5(2.1)$. To fulfill the particle-level selection criteria, an event must contains at least one particle-level top jet, at least one particle-level b jet, and at least one particle-level electron or muon.

In order to quantify the overlap between the detector, particle, and partoninc phase space definitions, two fractions, $f_{1,2}$, are used. The f_1 is defined as the fraction of reconstructed events that are accompanied by a selected event at the unfolded level (parton or particle) in the same observable range, whereas f_2 is the fraction of events at the unfolded level that are accompanied by a selected reconstructed event. Figures 8 and 9 show the aforementioned fractions at parton and particle level, respectively, for the hadronic channel as a function of the leading top p_T and $|y|$. The f_1 vs p_T shows a characteristic threshold behavior due to the finite p_T resolution, while it is flat vs $|y|$. The f_2 decreases vs p_T, primarily due to the subjet b tagging inefficiency and the NN output dependence on the p_T (at high jet p_T it is more difficult to differentiate between ordinary QCD jets and boosted tops). Also, the f_2 decreases at high $|y|$ values due to the increased b tagging inefficiency at the edges of the CMS tracker subdetector.
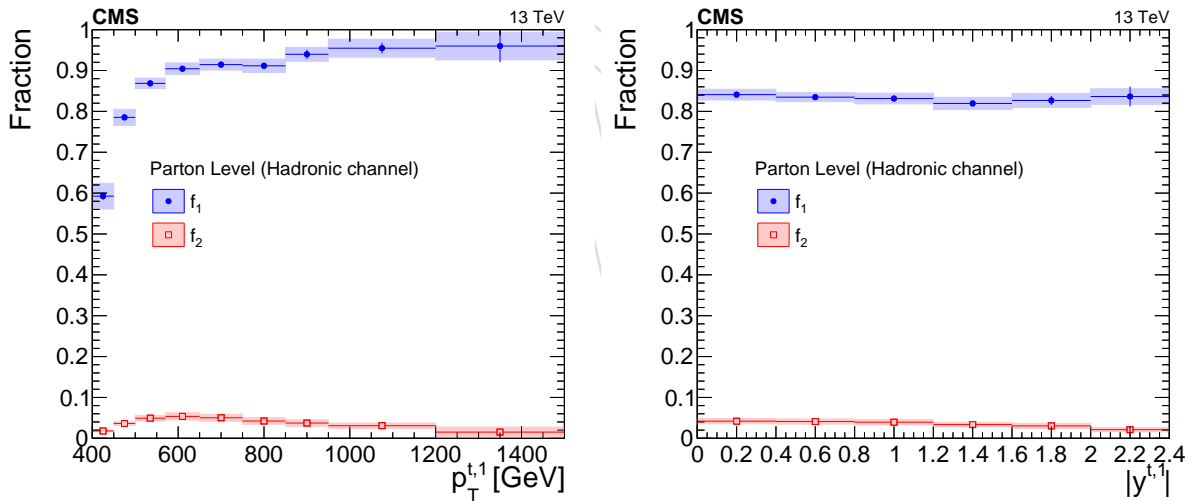
Figure 8: Simulated fractions $f_{1,2}$ for the parton-level selection in the hadronic channel as a function of the leading top p_T and $|y|$.
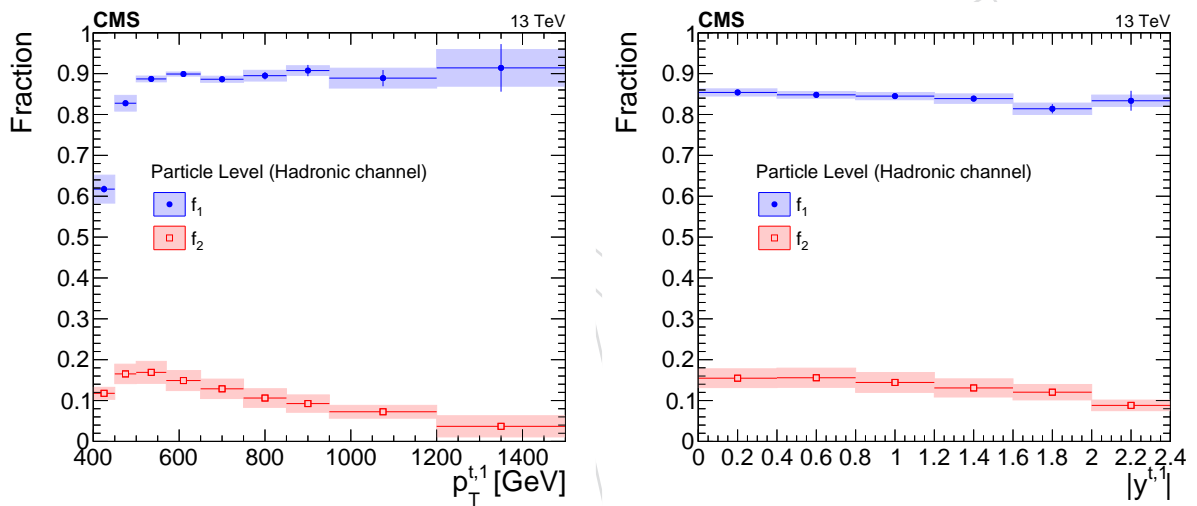
Figure 9: Simulated fractions $f_{1,2}$ for the particle-level selection in the hadronic channel as a function of the leading top p_T and $|y|$.

9.2 Unfolding

In both channels the unfolded cross sections are extracted according to:

$$\frac{d\sigma_i^{\text{unf}}}{dx} = \frac{1}{\mathcal{L} \cdot \Delta x_i} \cdot \frac{1}{f_{2,i}} \cdot \sum_j \left(R_{ij}^{-1} \cdot f_{1,j} \cdot S_j \right), \tag{4}$$

where $\mathcal{L}$ is the total integrated luminosity and Δx_i is the width of the i-th bin of the observable x. The quantity R_{ij}^{-1} is the inverse of the migration matrix between the i-th and j-th bins. Due to the finite resolution of the detector, the migration matrix is non-diagonal and thus the application of an unfolding procedure is necessary. The binning of the various observables has been chosen such that the purity (fraction of reconstructed events for which the true value of the observable lies in the same bin) and the stability (fraction of true events that the reconstructed observable lies in the same bin) are well above 50%. This choice results in migration matrices with suppressed non-diagonal elements, shown for the hadronic channel in Figs. 10 and 11, and for the lepton+jets channel in Figs. 12 and 13. In order to avoid the biases introduced by the various unfolding methods with some type of regularization, the simple migration matrix inversion is used, as written in Eq. 4, at a price of a moderate increase of the statistical uncertainty.
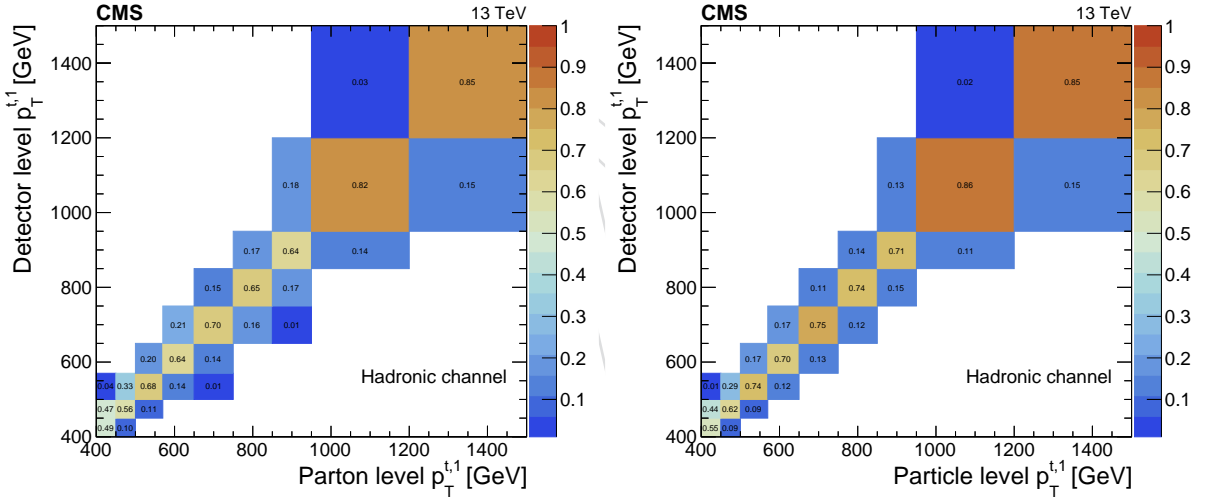
Figure 10: Simulated migration matrices for the leading top p_{T} at parton level (left) and particle level (right) in the hadronic channel. Each column is normalized to unity.

9.3 Hadronic channel

For the hadronic channel, the measurement of the differential cross sections is performed at the detector level and subsequently they are reported at the parton and particle levels.

9.3.1 Fiducial cross sections at detector level

The fiducial differential cross section is derived in bin i of the variable x from the signal yield S_i (eq. 3) in the bin as follows:

$$\frac{d\sigma_i^{\text{fid}}}{dx} = \frac{S_i}{\mathcal{L} \cdot \Delta x_i}, \tag{5}$$
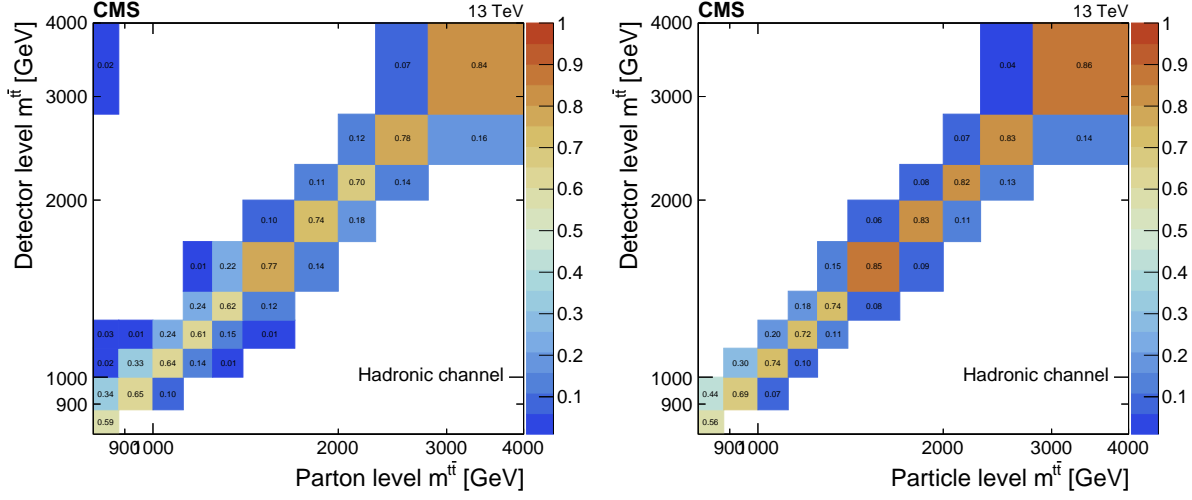
Figure 11: Simulated migration matrices for $m^{t\bar{t}}$ at parton level (left) and particle level (right) in the hadronic channel. Each column is normalized to unity.
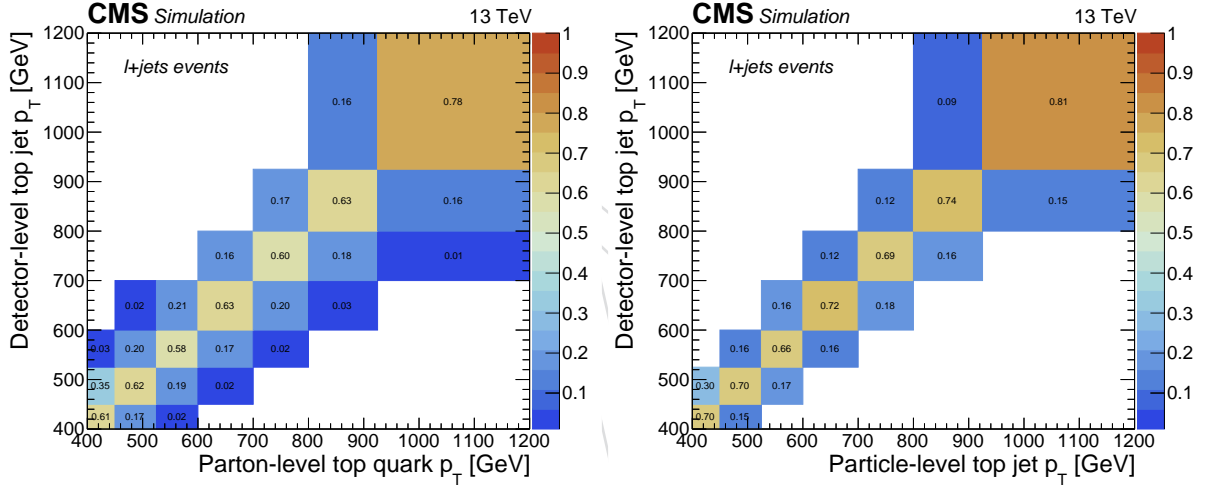
Figure 12: Simulated migration matrices for top quark p_T at parton level (left) and particle level (right) in the lepton+jets channel. Each row is normalized to unity.

where $\mathcal{L}$ is the total integrated luminosity and Δx_i is the width of the i-th bin of the observable x. Of particular interest is also the normalized cross section, computed as:

$$\frac{1}{\sigma^{\text{fid}}} \frac{d\sigma_i^{\text{fid}}}{dx} = \frac{1}{\sum_k S_k} \cdot \frac{S_i}{\mathcal{L} \cdot \Delta x_i}, \tag{6}$$

which is used to test the modeling of the differential cross section regardless of the overall normalization. In order to estimate the uncertainty in the measurement, the entire procedure of signal extraction is repeated for every source of uncertainty. Both the experimental and the theoretical uncertainties affect primarily the $t\bar{t}$ signal shape, which is used to fit the data for the QCD background normalization. As a result the total systematic uncertainty of the measurement is small and the statistical uncertainty dominates. This is true both for the absolute and the normalized cross sections.
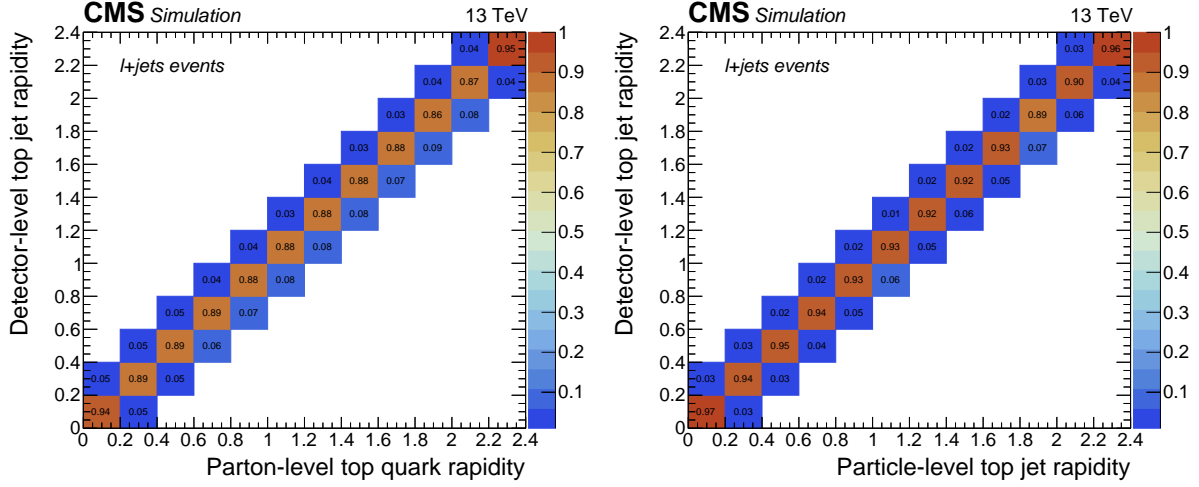
Figure 13: Simulated migration matrices for top jet rapidity at parton level (left) and particle level (right) in the lepton+jets channel. Each row is normalized to unity.

9.3.2 Particle-level fiducial cross sections

The unfolded cross sections (see Section 9.2) at particle level are shown in Figs. 14-20.
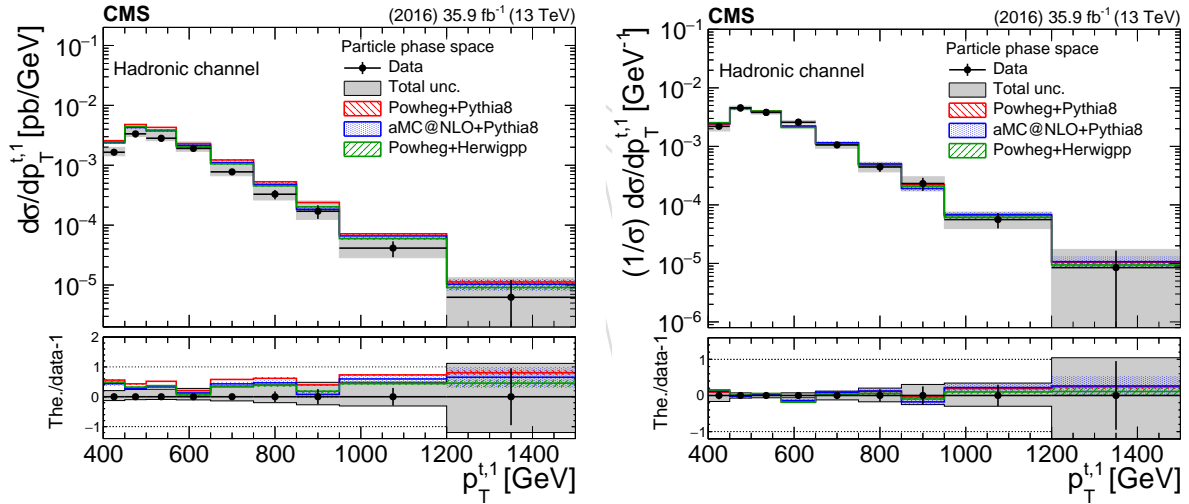
Figure 14: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of the leading top p_T in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
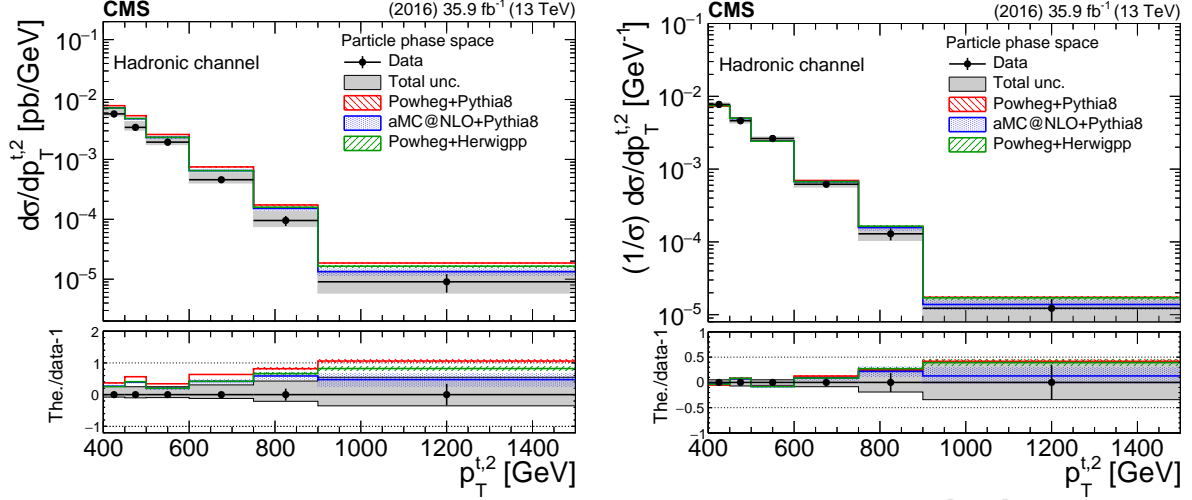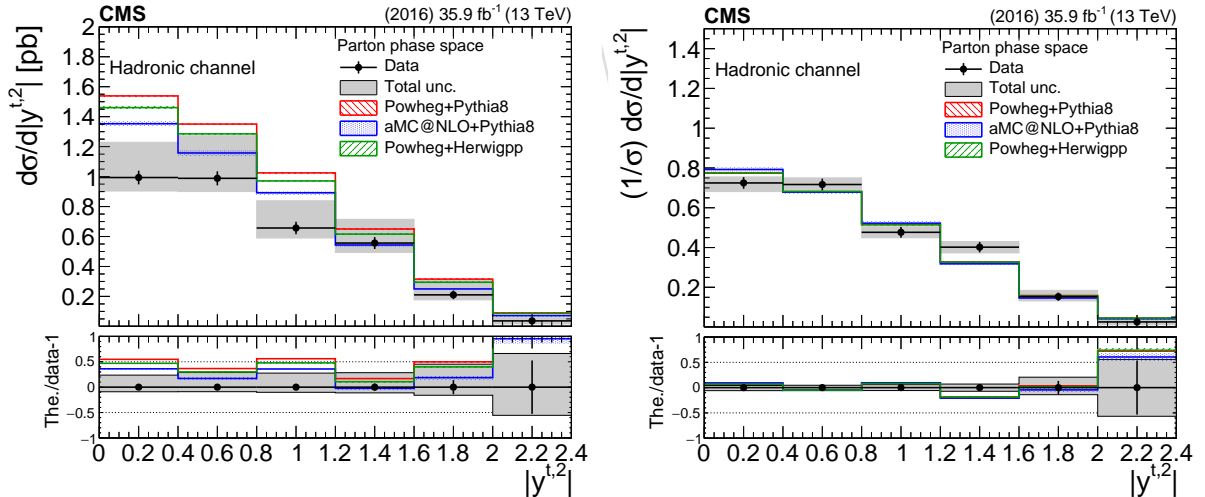
Figure 15: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of the second top p_T in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
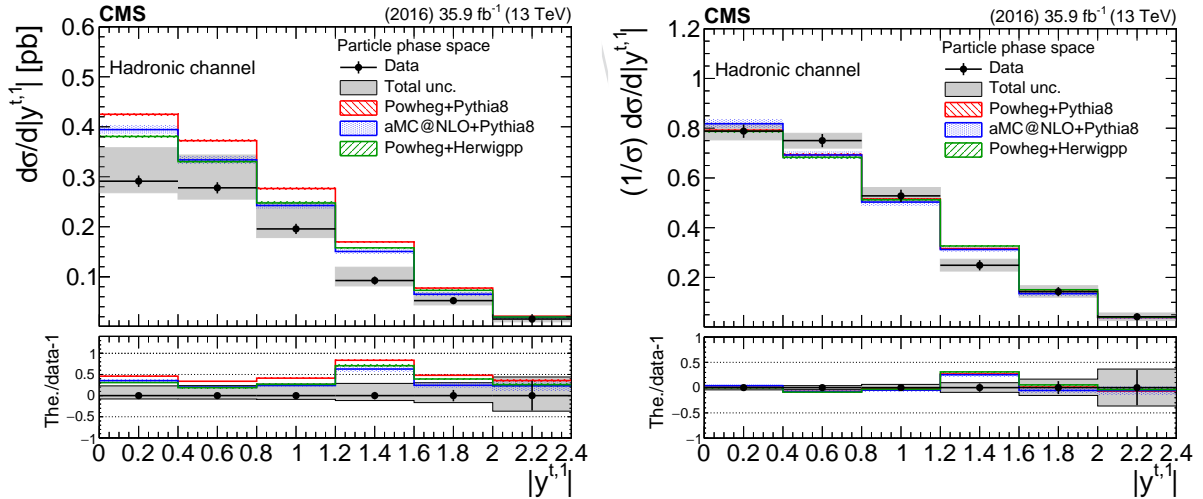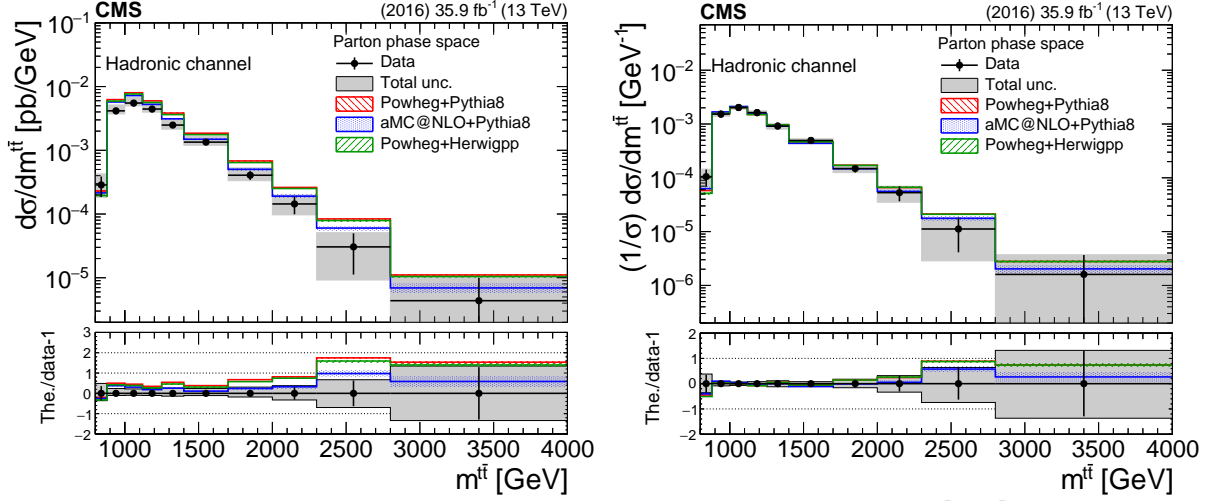

Figure 16: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of the leading top $|y|$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
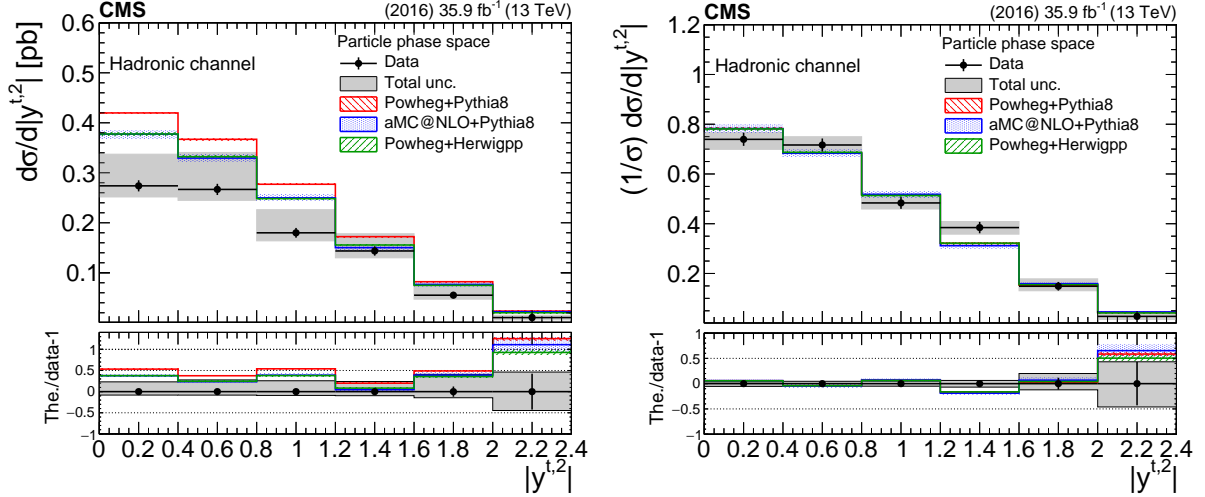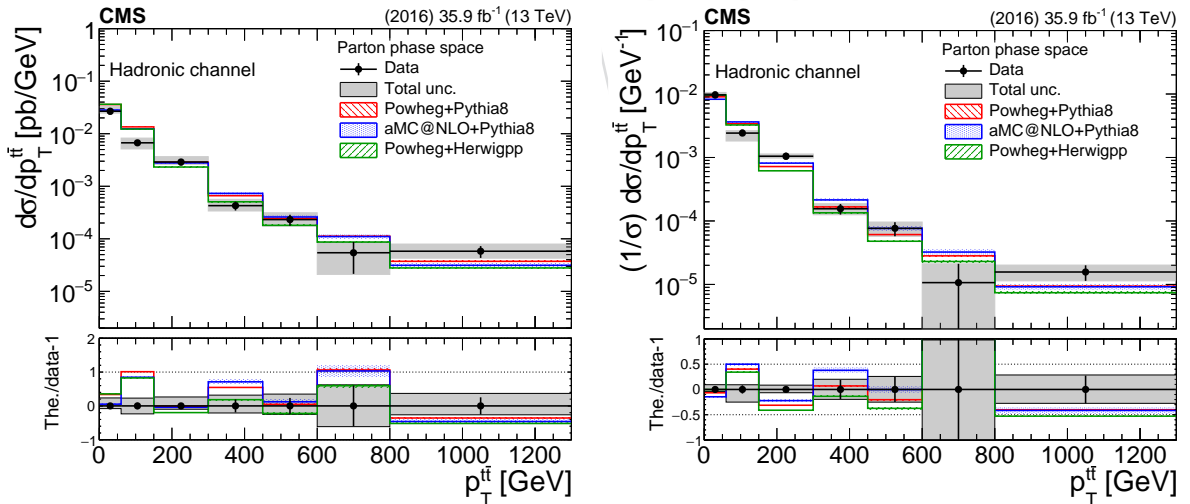
Figure 17: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of the second top $|y|$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
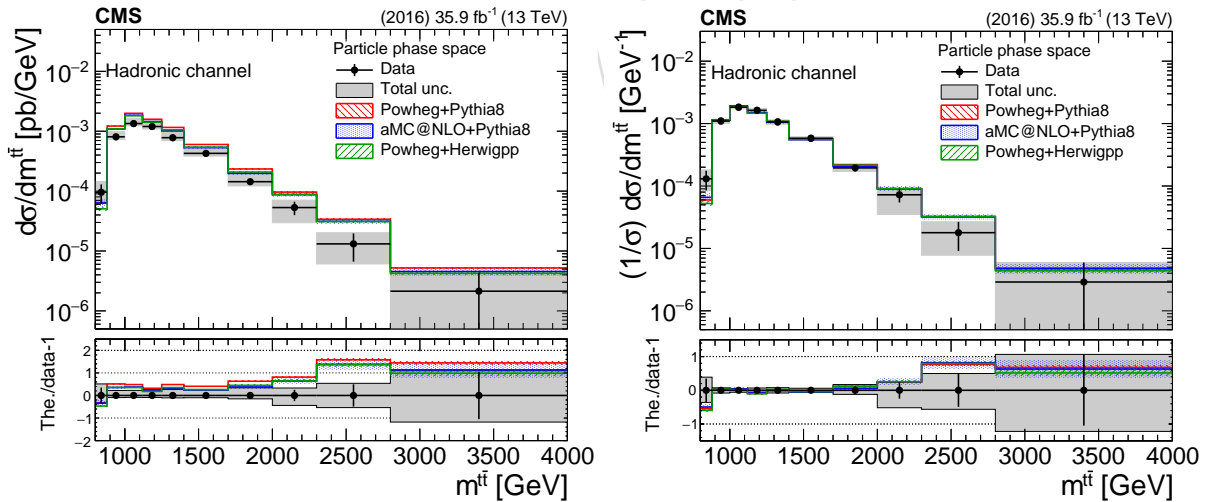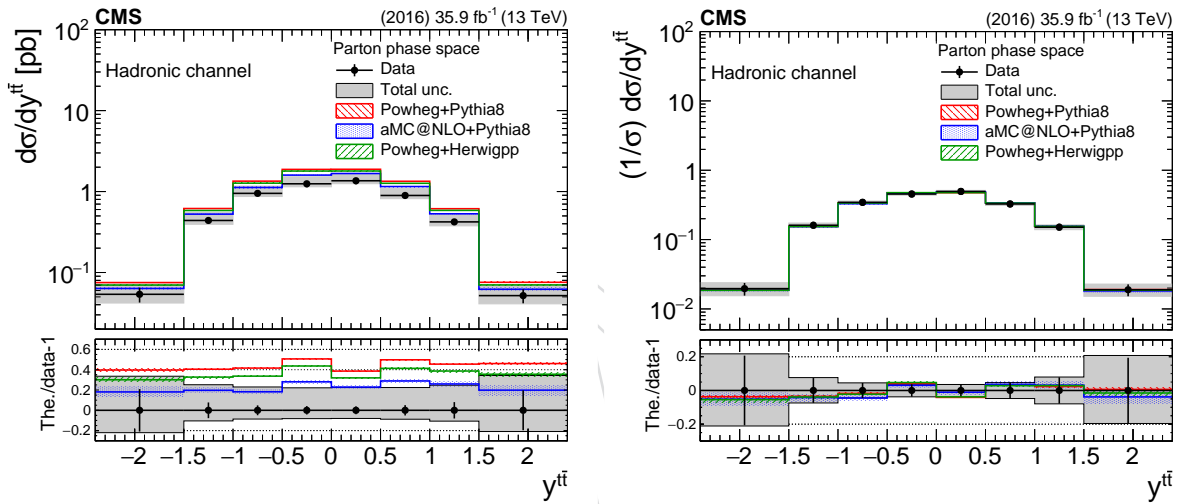

Figure 18: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of $m^{t\bar{t}}$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
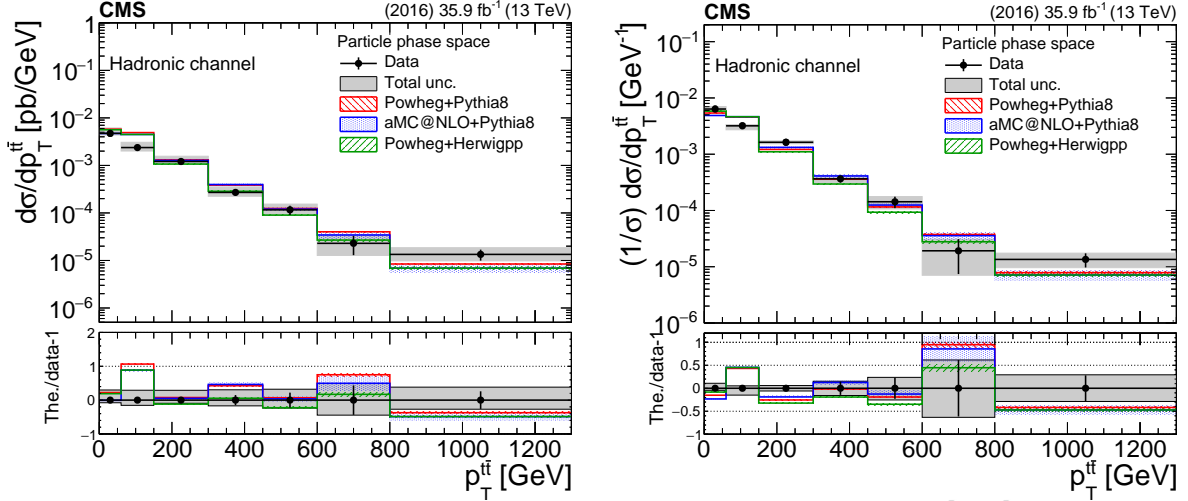
Figure 19: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of $p_T^{t\bar{t}}$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
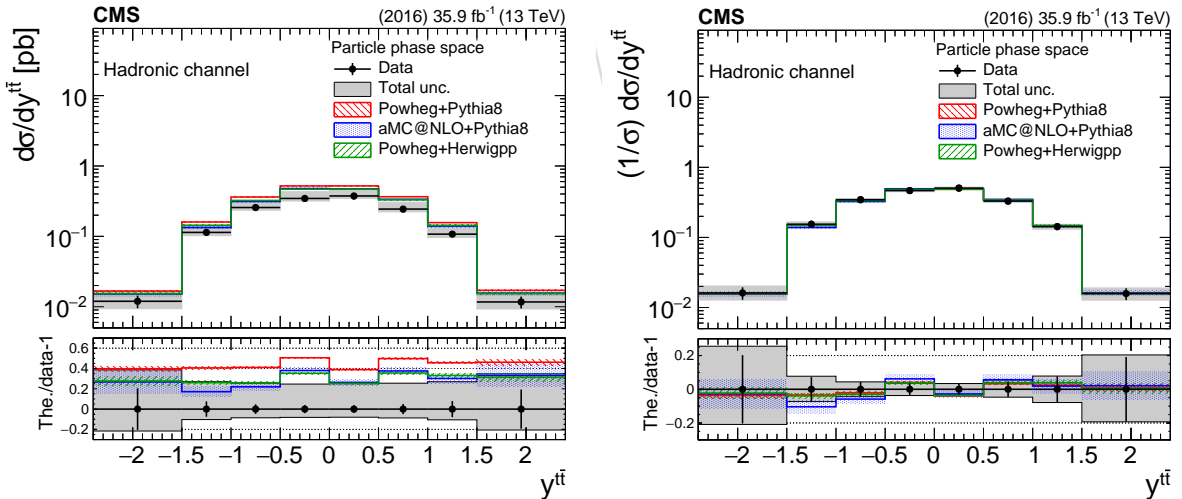

Figure 20: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of $y^{t\bar{t}}$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.

9.3.3 Parton-level cross sections

The unfolded cross sections (see Section 9.2) at parton level are shown in Figs. 21-27. Figure 28 shows a summary of the statistical and the dominant systematic uncertainties in the differential cross sections at parton and particle level as a function of the leading top p_T.
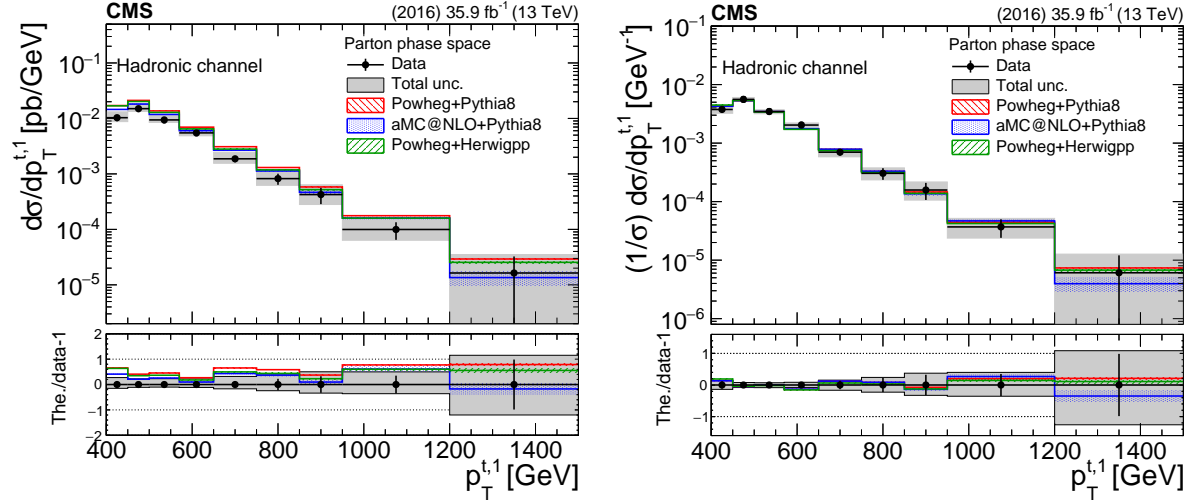
Figure 21: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of the leading top p_T in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
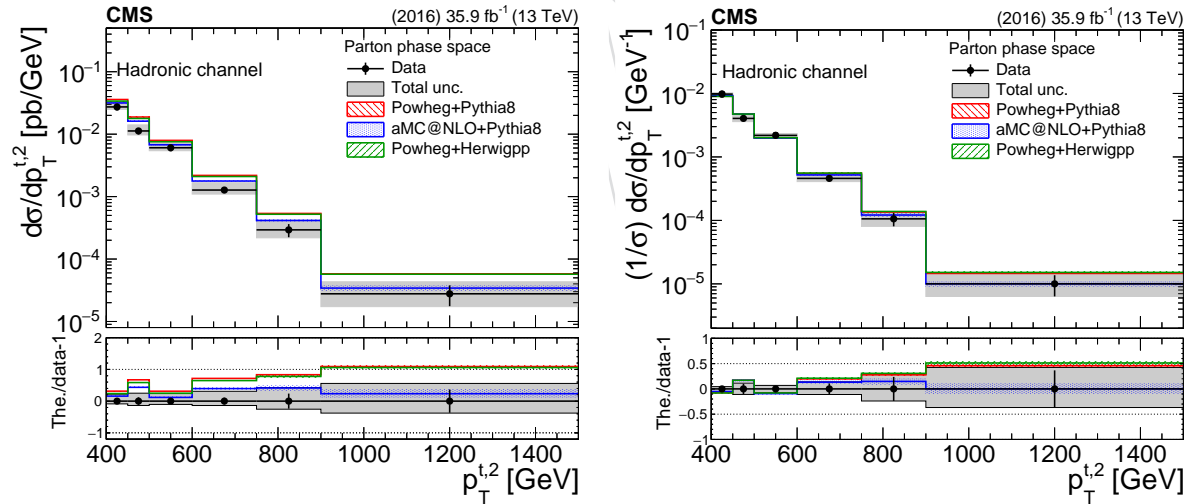
Figure 22: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of the second top p_T in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
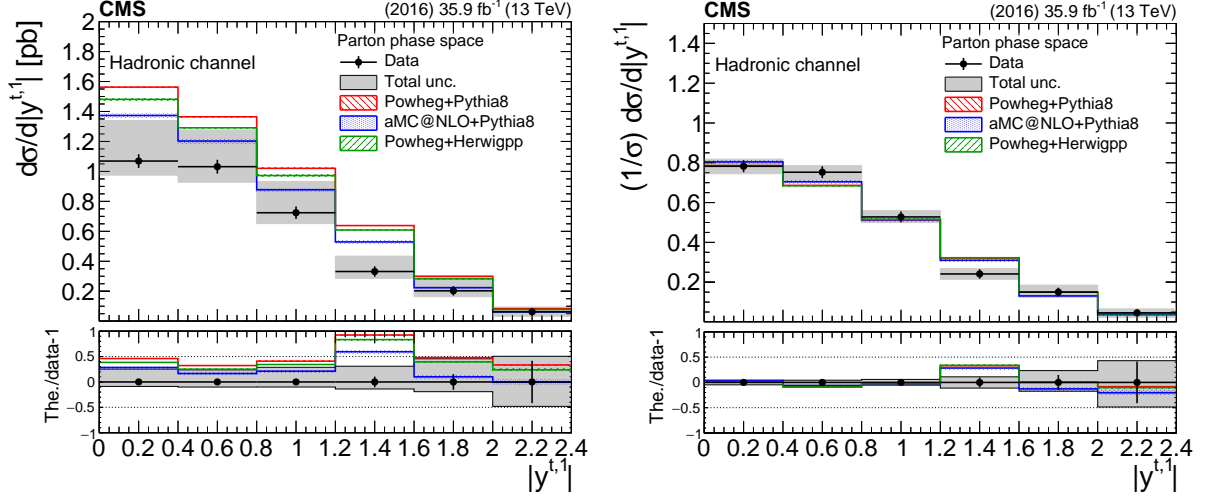
Figure 23: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of the leading top $|y|$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.

Figure 24: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of the second top $|y|$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.

Figure 25: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of $m^{t\bar{t}}$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.

Figure 26: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of $p_T^{t\bar{t}}$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.

Figure 27: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of $y^{t\bar{t}}$ in the hadronic channel. The bottom panel shows the ratio (theory - data)/data. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
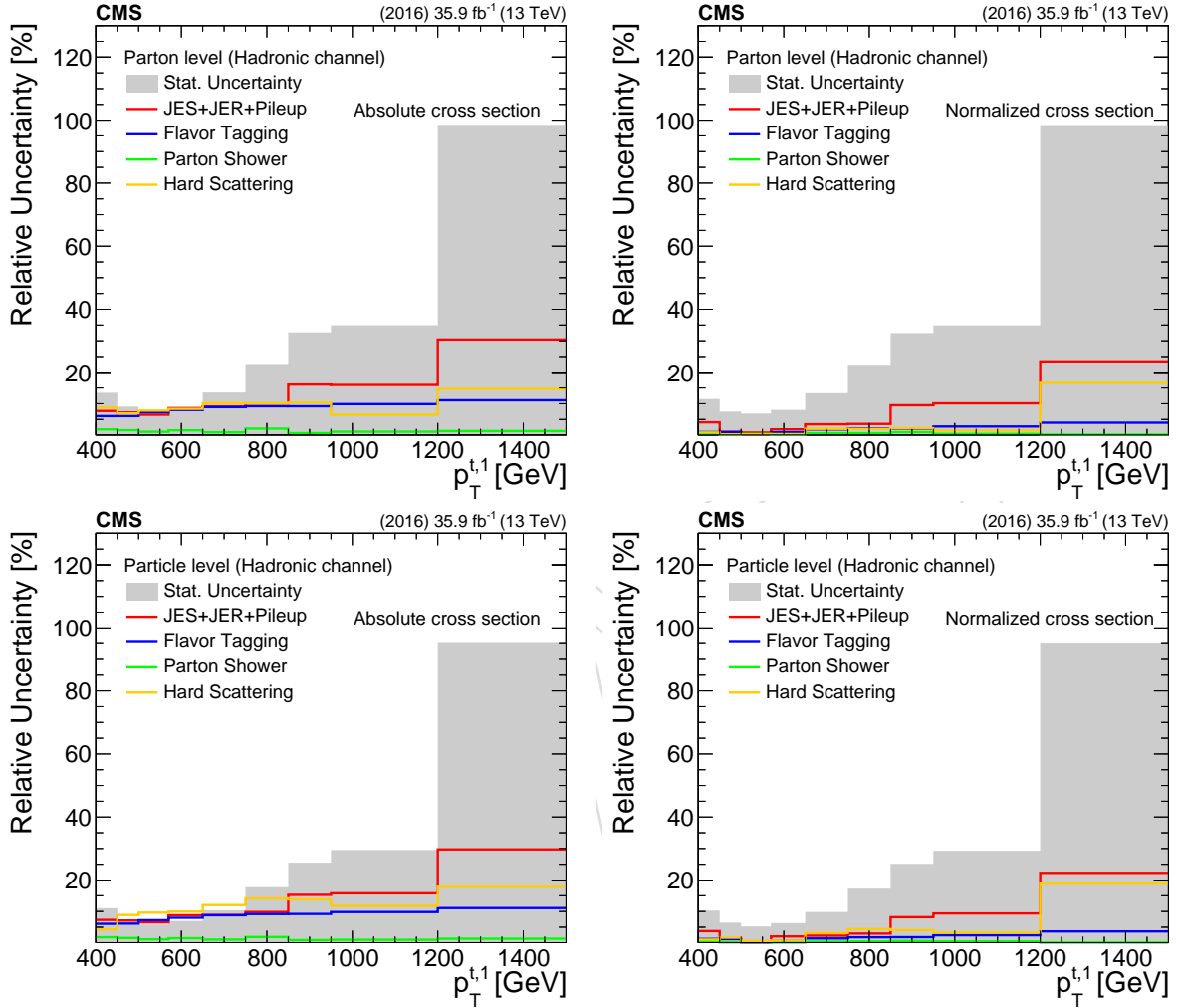
Figure 28: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of the leading top p_T in the hadronic channel. The grey band shows the statistical uncertainty, while the solid lines show the systematic uncertainties grouped in four categories: a) uncertainty due to pileup and the jet energy scale and resolution of the large-R jets, b) uncertainty due to flavor tagging of the subjets, c) uncertainty due to the modelling of the parton shower, and d) uncertainty due to the modelling of the hard scattering.

9.4   Lepton+jets channel

The differential $t\bar{t}$ cross section is measured as a function of the p_T and $|y|$ of the hadronically-decaying top quark. The measurement is performed at particle level within a region of phase space following the event selection criteria, as well as at parton level within the full phase space. Semi-leptonic $t\bar{t}$ events are selected at parton level, and the properties of the hadronically decaying top quark are defined to represent the true top quark p_T.

The differential cross section is extracted from the signal-dominated 1t1b category. The measured signal distribution is determined by subtracting the estimated background contributions from the distribution in data, using the posterior normalizations from the fit given in Table 3. To account for signal reconstruction efficiencies and bin migrations, an unfolding procedure based on a least squares minimization with Tikhonov regularization, as implemented in the TUNFOLD framework, is applied [59]. A non-regularized unfolding was found to be optimal for the measurement. The unfolding relies on response matrices that map the p_T and $|y|$ distributions for the t tagged jet to corresponding properties for either the particle-level t jet candidate or the parton-level top quark.

The systematic uncertainty in the unfolded measurement receives contributions from experimental and theoretical sources, discussed in Section 8. The posteriori values from the likelihood fit are used for the t tagging efficiency, background normalizations, and lepton efficiencies, while the a priori values are used for the remaining uncertainties. For each systematic variation that affects the shape of the p_T or $|y|$ distribution, a separate response matrix is created that is used to unfold the data. The resulting uncertainties are added in quadrature to obtain the total uncertainty for the unfolded distribution. The theoretical and experimental uncertainties are accounted for separately.

The muon and electron channels are combined before the unfolding, by merging the measured muon and electron distributions as well as the response matrices. The background contributions are also merged before subtracting these from the measured distributions, with the exception of the muon and electron multijet backgrounds that are treated as separate background sources.

The unfolded cross section results are shown in Figs. 29 and 30. The measurements are compared to the prediction from POWHEG +PYTHIA 8. The breakdown of the sources of systematic uncertainty are shown in Fig. 31.
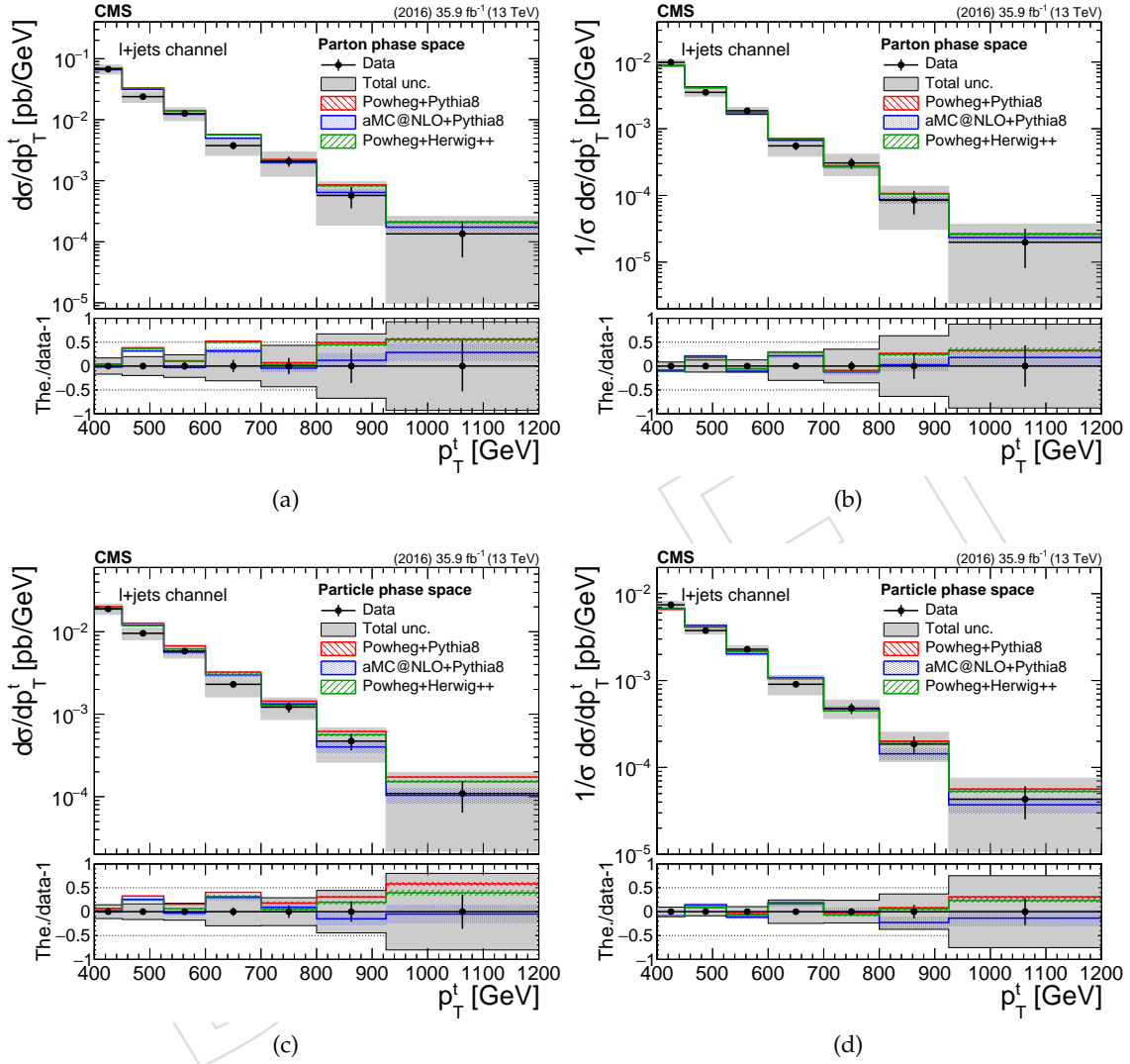
Figure 29: Differential cross section measurements at parton (top row) and particle level (bottom row), absolute (left column) and normalized (right column), as a function of the t jet p_T for the lepton+jets analysis. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
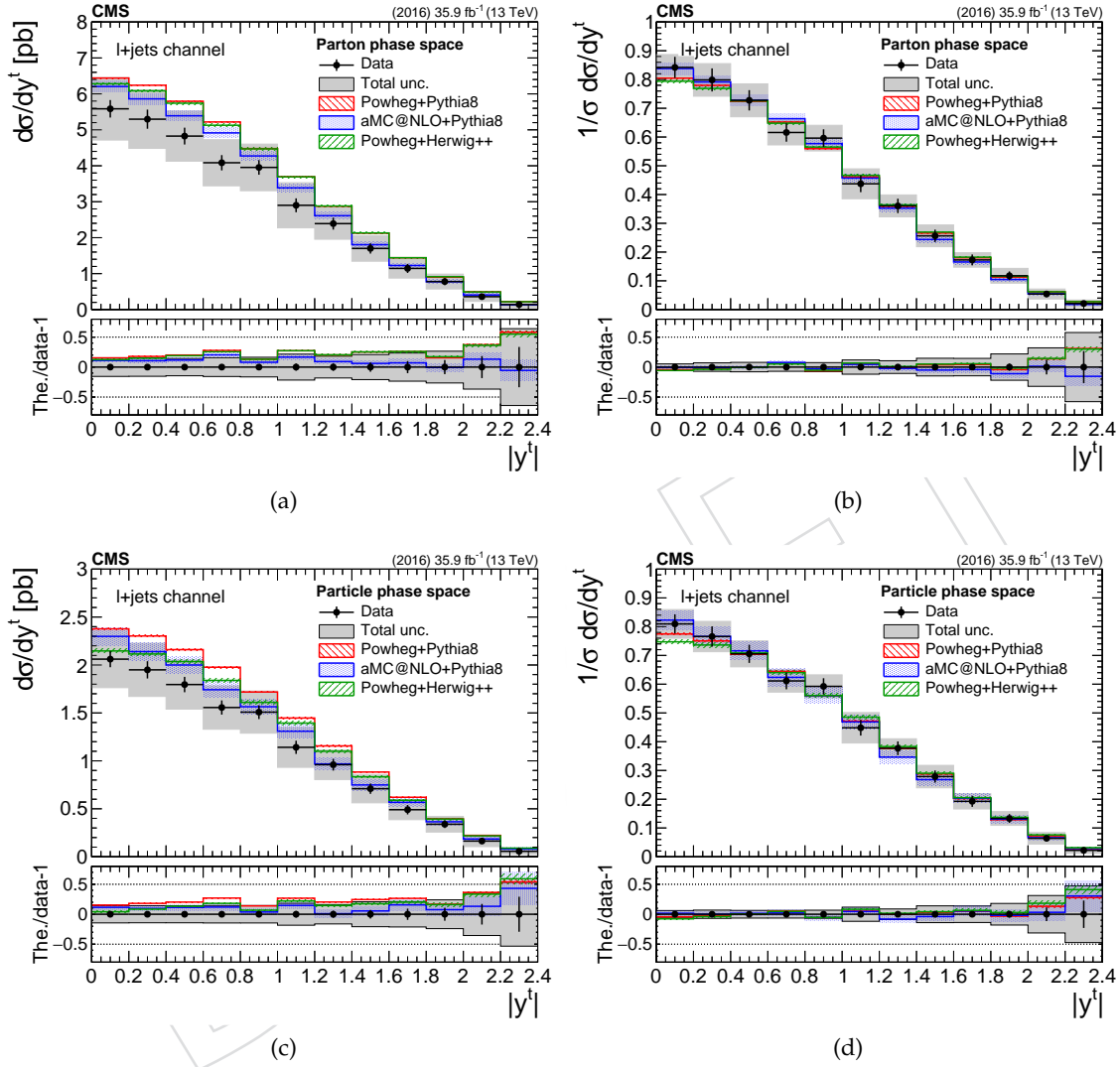
Figure 30: Differential cross section measurements at parton (top row) and particle level (bottom row), absolute (left column) and normalized (right column), as a function of the t jet $|y|$ for the lepton+jets analysis. The uncertainties on the data markers are the total statistical uncertainty, while the grey band shows the total statistical and systematic uncertainty.
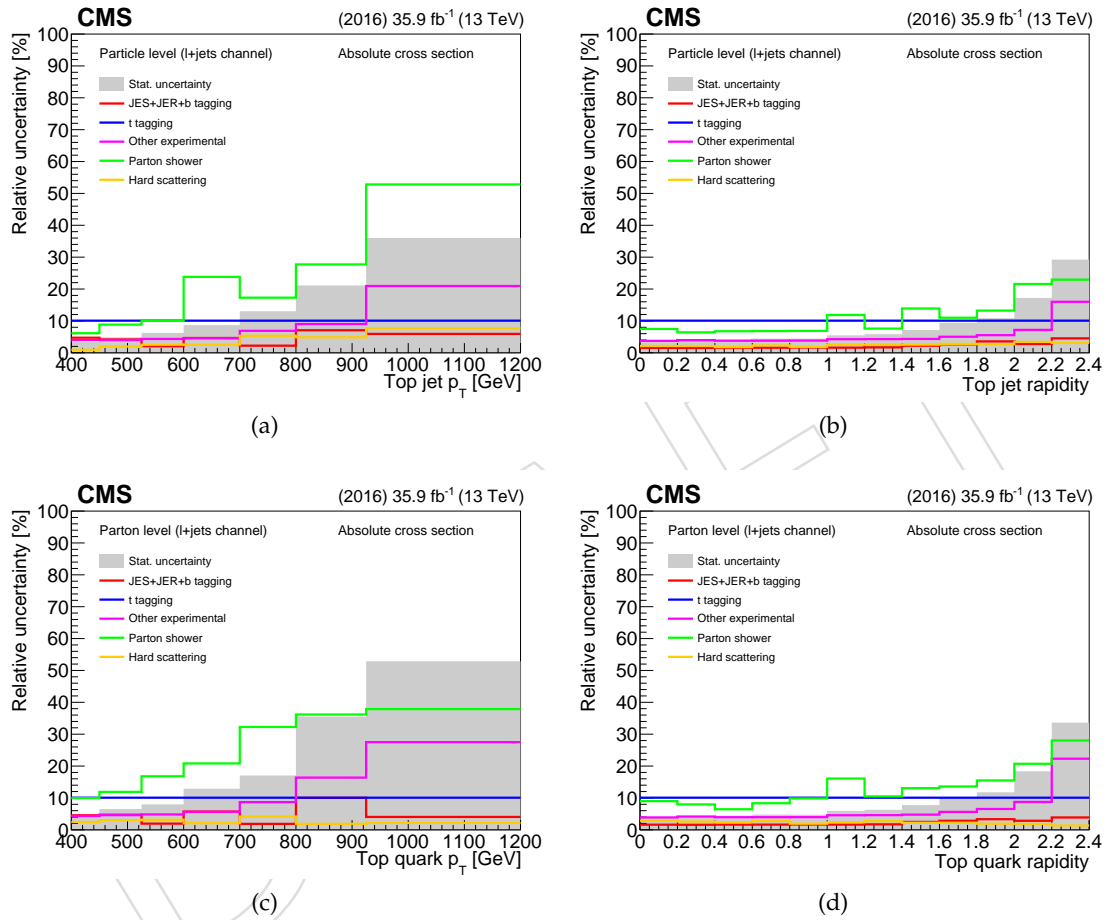
Figure 31: Breakdown of sources of systematic uncertainty affecting the differential cross section measurements in the lepton+jets analysis at particle level (top row) and at parton level (bottom row).

9.5 Discussion

The unfolded cross sections at parton and particle levels presented in Section 9 reveal some interesting features. First, there is a significant overestimate of the inclusive cross section by 35 (25)% in the hadroninc (l+jets) channel, which is also in agreement with previous results [16]. In terms of differential cross section shapes (normalized), there is very good agreement between data and theory predictions for the leading top (hadronic channel) and the hadronically decaying top (l+jets channel) p_T, while the cross section as a function of the second top p_T (hadronic channel) appears to be softer in data for the POWHEG predictions with MC@NLO providing the best description. The rapidity distributions are described well by the theory predictions in both channels. Finally, the measured $t\bar{t}$ kinematic distributions are in agreement with the theory prediction, with a possible deviation in the $m^{t\bar{t}}$ variable, where POWHEG tends to produce a harder spectrum and MC@NLO is fully consistent with the data.

10 Summary

A measurement of the $t\bar{t}$ production cross section for high-p_T top quarks at 13 TeV proton-proton collisions has been presented. The measurement is performed with events where either one or both top quarks decay hadronically, and where the hadronic decay products cannot be resolved but are instead clustered in a single large-R jet with $p_T > 400$ GeV. The hadronic final state thus contains two large-R jets, while the lepton+jets final state is identified through the presence of an electron/muon, a b tagged jet from the leptonically decaying top quark, missing transverse momentum from the escaping neutrino, and a single t tagged large-R jet. The purpose of the measurement is to utilize a much larger dataset collected with the CMS detector in order to explore a wider phase space of the $t\bar{t}$ production and elucidate the discrepancies with the theory predictions that have been reported in earlier publications. The cross section is reported differentially, for the hadronic channel, as a function of the leading and second top p_T and $|y|$, and as a function of the invariant mass, p_T, and rapidity of the $t\bar{t}$ system, unfolded to the parton and particle levels, absolute and normalized. For the lepton+jets channel, the differential cross section is measured as a function of the p_T and $|y|$ of the hadronically decaying top quark, at particle level within a fiducial phase space and at parton level. The results are compared to theoretical predictions from the POWHEG matrix-element generator, interfaced with PYTHIA 8 or HERWIG ++ for the underlying event and parton shower, and from the MC@NLO matrix-element generator, interfaced with PYTHIA 8. All the models overpredict significantly the absolute cross section in the phase space of the measurement (up to 40%), while they are able to describe consistently the differential shapes in most of the variables. The most notable discrepancy can be seen for the hadronic channel in the invariant mass of the $t\bar{t}$ system and the p_T of the second leading top jet, where the theoretical models predict a higher cross section at high mass and p_T values, respectively. However, in order to estimate the significance of the discrepancy, more data are needed.

Acknowledgements

References

693 [1] ATLAS Collaboration, "Measurements of top-quark pair differential cross-sections in the
694   lepton+jets channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector", *JHEP* **11**
695   (2017) 191, `doi:10.1007/JHEP11(2017)191`, `arXiv:1708.00727`.

696 [2] ATLAS Collaboration, "Measurements of top-quark pair differential cross-sections in the
697    $e\mu$ channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector", *Eur. Phys. J.* **C77**
698   (2017), no. 5, 292, `doi:10.1140/epjc/s10052-017-4821-x`, `arXiv:1612.05220`.

699 [3] ATLAS Collaboration, "Measurement of lepton differential distributions and the top
700   quark mass in $t\bar{t}$ production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector",
701   *Eur. Phys. J.* **C77** (2017), no. 11, 804, `doi:10.1140/epjc/s10052-017-5349-9`,
702   `arXiv:1709.09407`.

703 [4] ATLAS Collaboration, "Measurement of top quark pair differential cross-sections in the
704   dilepton channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV with ATLAS", *Phys. Rev.* **D94**
705   (2016), no. 9, 092003, `doi:10.1103/PhysRevD.94.092003`, `arXiv:1607.07281`.

706 [5] ATLAS Collaboration, "Differential top-antitop cross-section measurements as a function
707   of observables constructed from final-state particles using pp collisions at $\sqrt{s} = 7$ TeV in
708   the ATLAS detector", *JHEP* **06** (2015) 100, `doi:10.1007/JHEP06(2015)100`,
709   `arXiv:1502.05923`.

710 [6] ATLAS Collaboration, "Measurements of top quark pair relative differential
711   cross-sections with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J.* **C73** (2013), no. 1,
712   2261, `doi:10.1140/epjc/s10052-012-2261-1`, `arXiv:1207.5644`.

713 [7] CMS Collaboration, "Measurement of normalized differential $t\bar{t}$ cross sections in the
714   dilepton channel from pp collisions at $\sqrt{s} = 13$ TeV", *JHEP* **04** (2018) 060,
715   `doi:10.1007/JHEP04(2018)060`, `arXiv:1708.07638`.

716 [8] CMS Collaboration Collaboration, "Measurements of normalised multi-differential
717   cross sections for top quark pair production in pp collisions at $\sqrt{s} = 13$ TeV and
718   simultaneous determination of the strong coupling strength, top quark pole mass and

parton distribution functions", Technical Report CMS-PAS-TOP-18-004, CERN, Geneva, 2018.

[9] CMS Collaboration, "Measurements of $t\bar{t}$ differential cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons", *Submitted to: JHEP* (2018) `arXiv:1811.06625`.

[10] 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.* **D95** (2017), no. 9, 092001, `doi:10.1103/PhysRevD.95.092001`, `arXiv:1610.04191`.

[11] CMS Collaboration, "Measurements of differential cross sections of top quark pair production as a function of kinematic event variables in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **06** (2018) 002, `doi:10.1007/JHEP06(2018)002`, `arXiv:1803.03991`.

[12] CMS Collaboration, "Measurement of double-differential cross sections for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV and impact on parton distribution functions", *Eur. Phys. J.* **C77** (2017), no. 7, 459, `doi:10.1140/epjc/s10052-017-4984-5`, `arXiv:1703.01630`.

[13] CMS Collaboration, "Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J.* **C75** (2015), no. 11, 542, `doi:10.1140/epjc/s10052-015-3709-x`, `arXiv:1505.04480`.

[14] CMS Collaboration, "Measurement of the differential cross sections for top quark pair production as a function of kinematic event variables in pp collisions at $\sqrt{s}$=7 and 8 TeV", *Phys. Rev.* **D94** (2016), no. 5, 052006, `doi:10.1103/PhysRevD.94.052006`, `arXiv:1607.00837`.

[15] CMS Collaboration, "Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J.* **C73** (2013), no. 3, 2339, `doi:10.1140/epjc/s10052-013-2339-4`, `arXiv:1211.2220`.

[16] ATLAS Collaboration, "Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector", *Phys. Rev.* **D98** (2018), no. 1, 012003, `doi:10.1103/PhysRevD.98.012003`, `arXiv:1801.02052`.

[17] ATLAS Collaboration, "Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions using the ATLAS detector", *Phys. Rev.* **D93** (2016), no. 3, 032009, `doi:10.1103/PhysRevD.93.032009`, `arXiv:1510.03818`.

[18] CMS Collaboration Collaboration, "Measurement of the $t\bar{t}$ production cross section at 13 TeV in the all-jets final state", Technical Report CMS-PAS-TOP-16-013, CERN, Geneva, 2016.

[19] CMS Collaboration, "Measurement of the integrated and differential $t\bar{t}$ production cross sections for high-p_t top quarks in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev.* **D94** (2016), no. 7, 072002, `doi:10.1103/PhysRevD.94.072002`, `arXiv:1605.00116`.

[20] CMS Collaboration Collaboration, "Measurement of the differential $t\bar{t}$ cross section with high-p_T top-quark jets in the all-hadronic channel at $\sqrt{s} = 8$ TeV", Technical Report CMS-PAS-TOP-16-018, CERN, Geneva, 2017.

[21] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[22] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[23] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[24] S. Frixione, P. Nason, and G. Ridolfi, "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[25] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[26] 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.

[27] S. Alioli, S. O. Moch, and P. Uwer, "Hadronic top-quark pair-production with one jet and parton showering", *JHEP* **01** (2012) 137, doi:10.1007/JHEP01(2012)137, arXiv:1110.5251.

[28] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].

[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] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J.* **C53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[31] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[32] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A brief introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[33] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[34] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2015) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[35] CMS Collaboration 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", Technical Report CMS-PAS-TOP-16-021, CERN, Geneva, 2016.

[36] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[37] 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, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[38] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in the framework of the FEWZ simulation code", *Phys. Rev. D* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.

[39] N. Kidonakis, "Top Quark Production", (2014). arXiv:1311.0283.

[40] M. Bahr et al., "Herwig++ Physics and Manual", *Eur. Phys. J.* **C58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.

[41] S. Gieseke, C. Rohr, and A. Siodmok, "Colour reconnections in Herwig++", *Eur. Phys. J.* **C72** (2012) 2225, doi:10.1140/epjc/s10052-012-2225-5, arXiv:1206.0041.

[42] 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.

[43] 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.

[44] 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.

[45] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[46] J. Thaler and K. Van Tilburg, "Identifying boosted objects with N-subjettiness", *JHEP* **03** (2011) 015, doi:10.1007/JHEP03(2011)015, arXiv:1011.2268.

[47] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions", *Nucl. Phys.* **B406** (1993) 187–224, doi:10.1016/0550-3213(93)90166-M.

[48] J. Thaler and K. Van Tilburg, "Maximizing Boosted Top Identification by Minimizing N-subjettiness", *JHEP* **02** (2012) 093, doi:10.1007/JHEP02(2012)093, arXiv:1108.2701.

[49] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", *JHEP* **09** (2013) 029, doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.

[50] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC", *Phys. Rev. Lett.* **100** (2008) 242001, `doi:10.1103/PhysRevLett.100.242001`, `arXiv:0802.2470`.

[51] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop", *JHEP* **05** (2014) 146, `doi:10.1007/JHEP05(2014)146`, `arXiv:1402.2657`.

[52] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms", *JHEP* **08** (1997) 001, `doi:10.1088/1126-6708/1997/08/001`, `arXiv:hep-ph/9707323`.

[53] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", in *Proceedings of the Workshop on Monte Carlo Generators for HERA Physics, Hamburg, Germany*, p. 270. 1998. `arXiv:hep-ph/9907280`.

[54] 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`.

[55] J. Therhaag, "TMVA Toolkit for multivariate data analysis in ROOT", *PoS* **ICHEP2010** (2010) 510, `doi:10.22323/1.120.0510`.

[56] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling", *eConf* **C0303241** (2003) MOLT007, `arXiv:physics/0306116`. [,186(2003)].

[57] ATLAS Collaboration, "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC", *Phys. Rev. Lett.* **117** (2016), no. 18, 182002, `doi:10.1103/PhysRevLett.117.182002`, `arXiv:1606.02625`.

[58] CMS Collaboration Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[59] 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`.