CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2016/09/19

Measurement of the jet mass distribution in boosted $\mathrm{t\bar{t}}$ production at $\sqrt{s} = 8$ TeV

The CMS Collaboration

Abstract

A first measurement is performed of the differential $\mathrm{t\bar{t}}$ production cross section as a function of the leading jet mass in fully-merged top quark decays. Data collected with the CMS detector in pp collisions at $\sqrt{s} = 8$ TeV are used, corresponding to an integrated luminosity of 19.7 fb^{-1}. The measurement is carried out in the ℓ+jets channel, where the products of the leptonic decay are used to select $\mathrm{t\bar{t}}$ events with high Lorentz boosts. The products of the hadronic decay are reconstructed with a single Cambridge/Aachen jet with distance parameter $R = 1.2$, and transverse momentum $p_{\mathrm{T}} > 400$ GeV. The cross section, as a function of the jet mass m_{jet}, is unfolded and reported on particle level. The measurement is used to test the modelling of boosted top quark production. The peak position of the m_{jet} distribution is sensitive to the top quark mass m_{t} and the data are used to extract a value of m_{t} to assess the measurement's sensitivity.

1 Introduction

The top quark takes a special role in the standard model (SM) due to its large mass and its consequential importance in electroweak symmetry breaking. Measurements of top quark-antiquark pair ($\mathrm{t\bar{t}}$) production provide crucial information about the accuracy of the SM at and above the electroweak scale. They test the predictions of quantum chromodynamics (QCD) at high scales and can be used for the determination of fundamental parameters of the theory.

Previous differential measurements of $\mathrm{t\bar{t}}$ production cross sections [1–6] at the CERN LHC show remarkable agreement with SM predictions. However, probing very high transverse momenta of top quarks is difficult, since in this kinematic regime the top quark decay becomes highly collimated and merges into a single jet. In this boosted regime the $\mathrm{t\bar{t}}$ reconstruction efficiency deteriorates for previous, more traditional, measurements and special reconstruction techniques based on jet substructure have to be used for cross section measurements [7, 8] or searches for new physics [9–19]. A detailed understanding of jet substructure observables, most importantly the jet mass m_{jet}, is crucial for LHC analyses in boosted topologies. While measurements of the jet mass corrected to the particle level have been carried out for light quark and gluon jets [20, 21], the m_{jet} distribution for boosted top quarks has not been measured so far.

Besides testing the simulation of m_{jet} for top quark decays, the location of the peak of the m_{jet} distribution in boosted top quark production is sensitive to the top quark mass m_{t} [22]. The measurement therefore provides an alternative method for a determination of m_{t} from the boosted regime, independent of other recent measurements [23–27]. For $\mathrm{e^+e^-}$ annihilation, calculations from first principles have been performed in Soft Collinear Effective Theory (SCET) [28–31] for the dijet invariant mass distribution of boosted top quark production [32, 33], and can be extended to the LHC environment [34]. These calculations account for perturbative and non-perturbative effects and provide particle level predictions. Once predictions for the LHC become available the measurement of the m_{jet} distribution can lead to an extraction of m_{t} without the ambiguities coming from the unknown relation between m_{t} in a well defined renormalisation scheme and the top quark mass obtained from Monte-Carlo (MC) simulations m_{t} [35–38].

We present a first measurement of the differential $\mathrm{t\bar{t}}$ production cross section as a function of the leading jet mass in the boosted regime, where the leading jet includes all $\mathrm{t} \to \mathrm{bW} \to \mathrm{bq\bar{q}'}$ decay products. The measurement is carried out in the ℓ+jets channel, where the leptonic top quark decay serves as a means to select $\mathrm{t\bar{t}}$ events without biasing the m_{jet} distribution from the hadronic top quark decay. The jets used for the measurement are obtained using the Cambridge/Aachen (CA) jet-clustering algorithm [39, 40] with a distance parameter $R = 1.2$ and transverse momentum $p_{\mathrm{T}} > 400$ GeV. The m_{jet} distribution is unfolded to the particle level and compared to predictions from MC simulations. The measurement is also normalised to the total fiducial cross section and shows the expected sensitivity to the value of m_{t}. An extraction of the value of m_{t} is performed to assess the measurement's sensitivity.

2 Event reconstruction

The CMS experiment uses a particle-flow (PF) based event reconstruction [41, 42], which aggregates input from all subdetectors. This information includes charged-particle tracks from the tracking system and deposited energy from the electromagnetic and hadronic calorimeters, taking advantage of excellent granularity of the sub-systems. Particles are classified as electrons, muons, photons, charged hadrons, and neutral hadrons. Primary vertices are reconstructed using a deterministic annealing filter algorithm [43]. The vertex with the largest squared sum

of the associated track p_T values is taken to be the primary event vertex.

Electrons are reconstructed in the pseudorapidity range $|\eta| < 2.5$, by combining tracking information with energy deposits in the electromagnetic calorimeter [44, 45]. Electron candidates are required to originate from the primary event vertex. Electrons are identified using information on the shower shape, the track quality, and the spatial match between the track and electromagnetic cluster, and the fraction of total cluster energy in the hadron calorimeter. Electron candidates that are consistent with originating from photon conversions in the detector material are rejected.

Muons are detected and measured in the pseudorapidity range $|\eta| < 2.4$ using the information collected in the muon and tracker detectors [46]. Tracks from muon candidates must be consistent with a muon originating from the primary event vertex and satisfy track fit quality requirements.

Since the top-quark decay products can be collimated at high values of the top quark p_T, no isolation requirements on the leptons are imposed in either the trigger or offline selections (see Sec. 3).

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

Particle-flow candidates are clustered into jets using the FASTJET 3.0 software package [47]. Charged hadrons associated with other event vertices than the primary event vertex are removed prior to jet clustering. In the first step of the event selection jets are used obtained with the anti-k_T jet-clustering algorithm [48] with a distance parameter of 0.5 (AK5 jets). If a lepton candidate (electron or muon) is found within $\Delta R < 0.5$ of an AK5 jet, its four-momentum is subtracted from that of the jet. Large-radius jets are obtained using the Cambridge/Aachen (CA) jet-clustering algorithm [39, 40] with a distance parameter of 1.2 (CA12 jets). Similarly to AK5 jets, if a lepton candidate is found among the particle-flow candidates clustered to an CA12 jet, its four-momentum is subtracted from that of the CA12 jet. In this paper the unmodified term 'jet' will refer to these CA12 jets. All jets contain neutral particles from additional collisions in the beam crossing (pileup). The extra contribution is subtracted based on the average expectation of the pileup in the jet catchment area [49]. This is done by calculating a correction for the average offset energy density in each event as function of the number of primary vertices [50]. Jets are identified as originating from the fragmentation of a b or c quark by the combined secondary vertex algorithm [51]. The tight operating point is used, which has a misidentification probability of 0.1% for tagging light-parton jets with an average p_T of about 80 GeV. The efficiency for the tight operating point is about 50% for jet p_T in the range 50–160 GeV. Above 160 GeV the efficiency decreases gradually to about 30% for p_T values of 400 GeV [51]. All jets are required to satisfy quality selections to remove calorimeter noise and other sources of fake jets [52]. Events are required to also satisfy selection criteria to remove calorimeter noise from E_T^{miss} signals as described in Ref. [53].

The jet mass m_{jet} is calculated from the four-vectors p_i of all PF candidates clustered into a jet,

$$m_{\text{jet}}^2 = \left(\sum_{i \in \text{jet}} p_i \right)^2 , \tag{1}$$

where the pion mass is assigned to charged hadrons. The accurate reconstruction of m_{jet} for CA12 jets is studied using a sample of boosted hadronic W decays merged into a single jet, as described in Sec. 4.5.

3 Trigger and data sets

The data used were recorded using single lepton triggers. Events in the muon+jets channel are recorded with a trigger requiring a single muon with $p_T > 40\,\text{GeV}$ and $|\eta| < 2.1$. The efficiency for this trigger is measured in a $Z \to \mu^+\mu^-$ sample and is 95% for muons measured within $|\eta| < 0.9$, 85% if they are measured within $0.9 < |\eta| < 1.2$ and 83% for $1.2 < |\eta| < 2.1$. The trigger for electron events requires one electron with $p_T > 30\,\text{GeV}$ in conjunction with two jets that have $p_T > 100$ and 25 GeV, respectively. In both cases, no isolation requirement is applied to the leptons. A 10% increase in the signal efficiency at high transverse momenta is gained in the electron channel by including events that are triggered by a single jet with $p_T > 320\,\text{GeV}$. The events recovered by the single-jet trigger contain an electron merged in a jet, which can not be resolved at the trigger level. The resulting combined trigger efficiency is 90% for events with a leading (highest p_T) jet with $p_T < 320\,\text{GeV}$. Above this value the trigger shows a turn-on behaviour and is fully efficient above a value of 350 GeV. The total integrated luminosity associated with the datasets is $19.7\,\text{fb}^{-1}$. The efficiencies for all triggers are well modelled by the simulation.

Top-quark events, produced via the strong and electroweak interactions, are simulated using the next-to-leading-order (NLO) generator POWHEG 1.380 [54–56] with a value of $m_t = 172.5\,\text{GeV}$. The $W(\to \ell\nu)$+jets and $Z/\gamma^*(\to \ell\ell)$+jets processes are simulated using MADGRAPH 5.1 [57], where MADSPIN [58] is used for the decay of heavy resonances. Diboson production processes (WW, WZ, and ZZ) are simulated using PYTHIA 6.424 [59]. The single top quark background is simulated using POWHEG [60, 61], assuming a value of $m_t = 172.5\,\text{GeV}$. Simulated QCD multijet samples are produced with MADGRAPH, but constitute a negligible background overall. For the estimation of systematic uncertainties additional $t\bar{t}$ samples are used, generated with MC@NLO v3.41 [62] or with MADGRAPH for seven different values of m_t ranging from 166.5 to 178.5 GeV.

All the samples produced with MADGRAPH and POWHEG are interfaced to PYTHIA 6 for parton showering and fragmentation. The MLM algorithm [63] used in MADGRAPH is applied during the parton matching to avoid double counting of parton configurations. The MADGRAPH samples use the CTEQ6L [64] parton distribution functions (PDF). For the POWHEG $t\bar{t}$ sample, the CT10 [65] PDF set is utilized, whereas the single top quark processes are produced with the CTEQ6M [66] PDF set. The PYTHIA 6 Z2* tune [67, 68] is used to model the underlying event activity. Top-quark samples produced with MC@NLO use the CTEQ6M PDF set and are interfaced to HERWIG 6.520 [69] for parton showering and fragmentation and use the default HERWIG tune to model the underlying event.

The normalizations of the simulated samples are taken from the NLO+next-to-next-to-leading logarithms (NNLL) calculation for the single top quark production [70], the next-to-next-to-leading order (NNLO) calculations for $W(\to \ell\nu)$+jets and $Z/\gamma^*(\to \ell\ell)$+jets [71–73], and the NLO calculation for diboson production [74]. The normalization of the $t\bar{t}$ simulation is obtained from next-to-next-to leading order (NNLO) calculations in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [75–81].

A detailed simulation of particle propagation through the CMS apparatus and detector response is performed with GEANT4 v9.2 [82]. For all simulated samples, the hard interaction collision is overlaid with a number of simulated minimum bias collisions. The resulting events are weighted to reproduce the pileup distribution measured in data. The same event reconstruction software is used for data and simulated events. The resolutions and efficiencies for reconstructed objects are corrected to match those measured in data [44, 46, 51, 83, 84].

4 Cross section measurement

4.1 Measurement strategy

The measurement is carried out in the ℓ+jets channel which warrants the selection of a pure $t\bar{t}$ sample due to its distinct signature at high top quark boosts. The selection is based on kinematic quantities only such that the m_{jet} distribution of the hadronic decay leg is not biased. A bias would for example be introduced by selecting the leading jet based on the number of subjets or requiring a certain maximum value of the N-subjettiness [85, 86], as applied in top-tagging algorithms [87–91].

The fiducial measurement region is studied in simulation on particle level (defined by all particles with a lifetime larger than 10^{-8} s in a simulated event). The exact selection is detailed below. It relies on a boosted leptonic top quark decay, where the lepton is close in ΔR to the b jet. An additional high-p_{T} jet is selected, which is assumed to originate from the hadronic top quark decay. A veto on additional jets is employed which ensures that the hadronic decay is merged into a single jet. The jet veto is also beneficial from the theory side, as it suppresses the size of non-global logarithms present due to out-of-cone radiation. An event selection on reconstruction level objects is carried out ensuring high efficiency for events in the fiducial region. Finally, the m_{jet} distribution is unfolded for experimental effects and compared to different predictions on particle level. A measurement of the normalised m_{jet} distribution is performed as well.

4.2 Measurement phase space

The $t\bar{t}$ cross section as a function of the mass of the leading jet is unfolded to the particle level. The fiducial phase space on this level is defined by the following selection.

The measurement is performed in the ℓ+jets channel, where ℓ refers to electrons and muons from the W decay. Tau lepton decays are not considered as part of the signal. Leptons are selected if they are within $|\eta| < 2.5$ and have $p_{\text{T}} > 45\,\text{GeV}$. Jets are clustered using the CA algorithm with a distance parameter of $R = 1.2$ and have to be within a pseudo-rapidity of $|\eta| < 2.5$. The value of R has been chosen to optimize the relation between large enough statistics and a narrow enough width of the jet mass distribution. The four-momentum of the leading (highest p_{T}) lepton is subtracted from the four-momentum of a jet if the lepton is found within a distance of $\Delta R < 1.2$ to the jet. Events are selected if at least one jet with $p_{\text{T}} > 400\,\text{GeV}$ and a second jet with $p_{\text{T}} > 150\,\text{GeV}$ are found. The leading jet in p_{T} is assumed to originate from the $t \to Wb \to q\bar{q}'b$ decay, merged into a single jet. Consequently, the second jet originates from the fragmented b quark of the leptonic decay leg. In order to select events with a boosted topology, a veto on additional jets with $p_{\text{T}} > 150\,\text{GeV}$ is employed. Two additional selection criteria are introduced to ensure that the leading jet includes all particles of the hadronic decay chain. The distance ΔR between the lepton and the second jet has to be smaller than 1.2. This, together with the veto on additional jets, ensures that the top quarks were produced back-to-back in the azimuthal angle ϕ. Additionally, the invariant mass of the leading jet has to be greater than the invariant mass of the combination of the second jet and the lepton. This improves the choice of the leading jet to originate from the fully hadronic top quark decay.

The distribution of the jet mass on particle level after this selection is shown in Fig. 1. The distribution of all jets passing the particle level selection is compared to distributions of the jet mass from fully-merged and unmerged $t\bar{t}$ decays. Here, an event is called fully-merged if the maximum distance in ΔR between the leading jet on particle level and each individual parton of the hadronic top quark decay is smaller than 1.2. After the selection outlined above, jets which

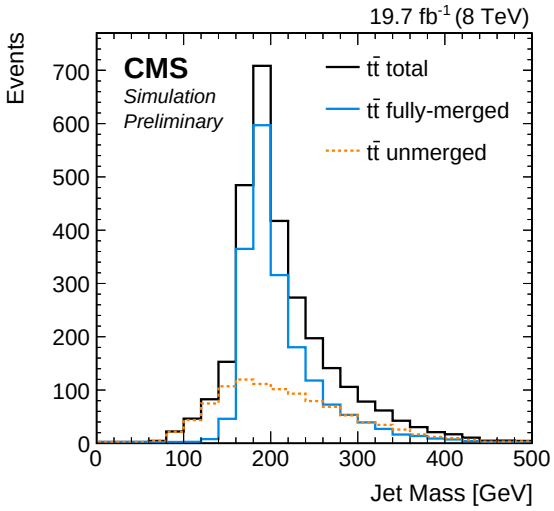

Figure 1: Simulated jet mass distribution of the leading jet in the ℓ+jets channel on particle level. The events were simulated with POWHEG+PYTHIA and are normalized to an integrated luminosity of 19.7 fb^{-1}. The total number of selected events (total, black solid line) is compared to events where the leading jet originates from the hadronic top quark decay (fully-merged, blue solid line) and events where the leading jet does not include all particles from the hadronic top quark decay (unmerged, orange dotted line).

do not originate from fully-merged hadronic top quark decays are expected to constitute about 35% of the total jets, estimated using the POWHEG+PYTHIA simulation.

4.3 Event Selection

On reconstruction level, a selection is applied to obtain a clean $\mathrm{t\bar{t}}$ sample with high p_T top quarks. The selection is based on leptons without isolation criterion. As a second step, high-p_T jets are required to obtain events kinematically similar to the selected ones on particle level. Comparable kinematics between reconstruction and particle level lead to smaller bin-to-bin migrations and therefore to smaller corrections when unfolding the data.

Selected events have to contain one muon with $p_\mathrm{T} > 45\,\mathrm{GeV}$ and $|\eta| < 2.1$ or one electron with $p_\mathrm{T} > 45\,\mathrm{GeV}$ and $|\eta| < 2.5$. Events with more than one lepton are rejected. Leptons with $2.1 < |\eta| < 2.5$ contribute only about 1% to the total cross section due to the centrality of high-p_T top quark production. The tighter selection for muons on reconstruction level than on particle level does not affect the shape of the final measured distribution.

In order to select boosted $\mathrm{t\bar{t}}$ events, at least two AK5 jets with $p_\mathrm{T} > 50\,\mathrm{GeV}$ and $|\eta| < 2.4$ are required and at least one with $p_\mathrm{T} > 150\,\mathrm{GeV}$ and $|\eta| < 2.4$. The suppression of background events from multijet production is accomplished using a two-dimensional isolation variable that is efficient at high top-quark boosts, yet reduces multijet backgrounds notably. This two-dimensional isolation requires the minimal distance between the lepton and the nearest AK5 jet $\Delta R_\mathrm{min}(\mathrm{lepton, jets}) > 0.5$ or the perpendicular component of the lepton four-vector with respect to the nearest AK5 jet $p_\mathrm{T,rel}$ to be larger than 25 GeV. In calculating these quantities only AK5 jets with $p_\mathrm{T} > 25\,\mathrm{GeV}$ are considered. The efficiency of the two-dimensional isolation requirement has been studied in data and simulation in a $\mathrm{Z}/\gamma^*(\to \ell\ell)$+jets sample [17].

A requirement of $E_\mathrm{T}^\mathrm{miss} > 20\,\mathrm{GeV}$ and $E_\mathrm{T}^\mathrm{miss} + p_\mathrm{T}^\ell > 150\,\mathrm{GeV}$ additionally reduces the contribution from multijet and $\mathrm{Z}/\gamma^*(\to \ell\ell)$+jets production, where p_T^ℓ is the lepton's transverse momentum. Given the presence of two b quarks in the events, at least one AK5 jet is required

to be identified to originate from the fragmentation of b quark, which reduces the contribution from W+jets production. The electron channel includes an additional topological section criterion to suppress the last remaining contribution from multijet production, requiring

$$|\Delta\phi(\{\text{e or jet}\}, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}) - 1.5| < E_{\mathrm{T}}^{\mathrm{miss}}/50\,\mathrm{GeV}$$

with $E_{\mathrm{T}}^{\mathrm{miss}}$ measured in GeV. This selection rejects events in which $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ points along the perpendicular component of the p_{T} of the leading jet or the lepton. After these steps the background contribution from multijet production is on a negligible level.

The selection outlined above results in a pure $\mathrm{t\bar{t}}$ sample with high selection efficiency at high top quark p_{T}. Additionally, events are selected with similar kinematics as on particle level. For each event to pass the selection, at least one CA12 jet with $p_{\mathrm{T}} > 400\,\mathrm{GeV}$ and $|\eta| < 2.5$ is required, together with one CA12 jet with $p_{\mathrm{T}} > 150\,\mathrm{GeV}$ and $|\eta| < 2.5$. Contributions from not fully-merged $\mathrm{t\bar{t}}$ events are suppressed with a veto on additional jets with $p_{\mathrm{T}} > 150\,\mathrm{GeV}$ and $|\eta| < 2.5$. The fraction of fully-merged events with a back-to-back topology is further enhanced by selecting events with a distance $\Delta R < 1.2$ between the lepton and the sub-leading CA12 jet. To ensure that the leading jet originates from the fully-merged top quark decay, its invariant mass is required to be larger than the mass of the sub-leading jet. With this selection, the reconstruction efficiency in the fiducial measurement region is 24.5%. Several of the above criteria are relaxed in the unfolding procedure to define sideband regions included in the migration matrix, hence increasing the reconstruction efficiency.

After the selection the contribution of background events from $\mathrm{t\bar{t}}$ production other than in the ℓ+jets channel is small. These events stem from the τ+jets, dilepton and all-hadronic channels and make 7.7%, 7.0% and 0.5% of the total number of selected events. These contributions are accounted for in the unfolding.

Distributions of the kinematics of the leading CA12 jet are shown in Fig. 2 for data and simulation in selected events. The mass distribution of the leading CA12 jet on reconstruction level is shown in Fig. 3 for two different p_{T} regions with $400 < p_{\mathrm{T}} < 500\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 500\,\mathrm{GeV}$.

4.4 Unfolding

The correction from reconstruction to particle level is carried out using a regularized unfolding based on a least square fit, implemented in the TUnfold [92] framework. Statistical fluctuations are suppressed by a regularization with respect to the count in each bin. The optimal regularization strength τ is determined by a minimization of the average global correlation coefficient of the output bins. Background processes like W+jets, single top quark and QCD multijet production are determined using simulation and subtracted from the data prior to the unfolding. Background processes from $\mathrm{t\bar{t}}$ production are accounted for in the unfolding by including them in the response matrix.

The response matrix is evaluated using $\mathrm{t\bar{t}}$ events simulated with POWHEG interfaced to PYTHIA. The response matrix is obtained for two regions in p_{T} of the leading jet, for $400\,\mathrm{GeV} < p_{\mathrm{T}} < 500\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 500\,\mathrm{GeV}$. This division is necessary to account for the shape of the p_{T}-spectrum in the unfolding process. The response matrix includes three additional sideband regions to account for migrations in and out of the measurement phase space. These are obtained for a lower leading jet p_{T} of $300\,\mathrm{GeV} < p_{\mathrm{T}} < 400\,\mathrm{GeV}$, a lower second-leading jet p_{T} of $100\,\mathrm{GeV} < p_{\mathrm{T}} < 150\,\mathrm{GeV}$ and a higher third jet veto of $150\,\mathrm{GeV} < p_{\mathrm{T}} < 200\,\mathrm{GeV}$. Events that are reconstructed but do not pass the particle level selections are also included in the response matrix. The electron and muon channels are combined, and the combined distribution

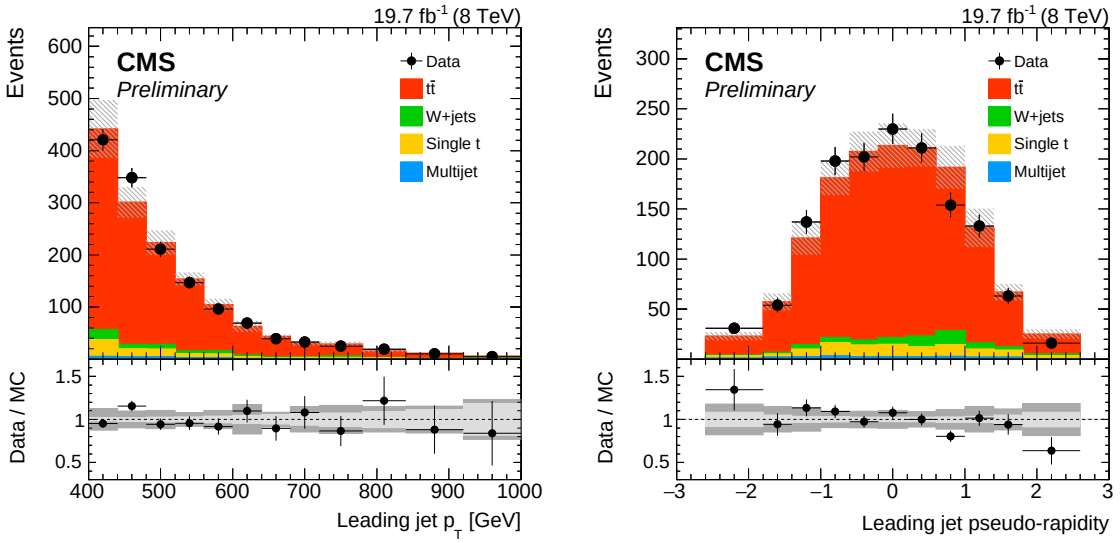

Figure 2: Distributions of p_{T} (left) and η (right) of the leading CA12 jet in data and simulation. The combination of the electron and muon channels is shown. The $\mathrm{t\bar{t}}$ sample is scaled such that the number of events in simulation matches the number of events observed in data. The uncertainty band includes statistical and experimental systematic uncertainties, where the statistical (light grey) and total (dark grey) uncertainties are shown separately in the ratio.

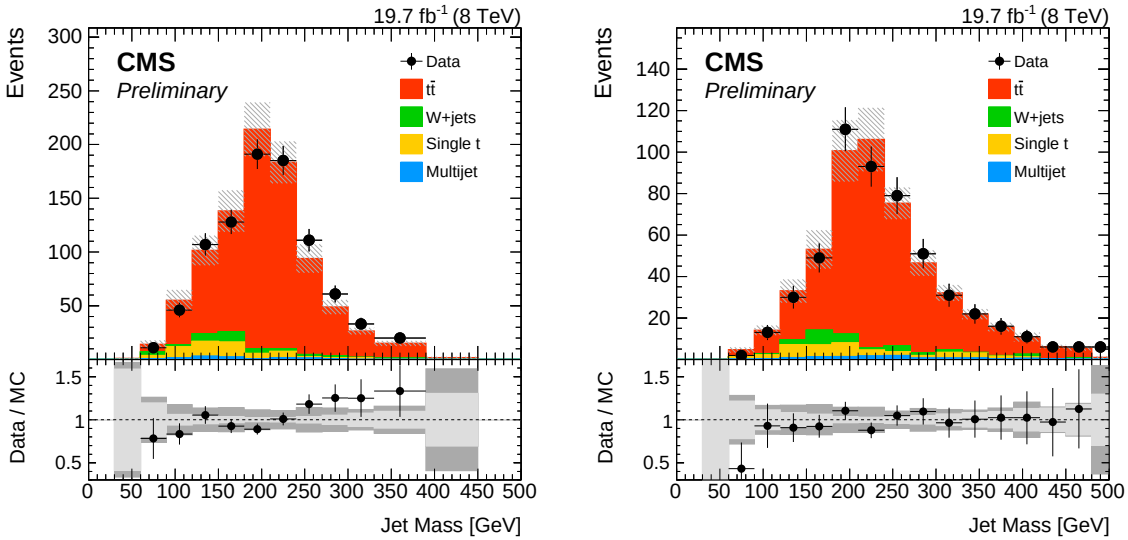

Figure 3: Distribution of the invariant mass of the leading CA12 jet in data and simulation. The combination of the electron and muon channels is shown. The distributions are shown for two different p_{T} bins for $400 < p_{\mathrm{T}} < 500\,\mathrm{GeV}$ (left) and $p_{\mathrm{T}} > 500\,\mathrm{GeV}$ (right). The $\mathrm{t\bar{t}}$ sample is scaled such that the number of events in simulation matches the number of events observed in data. The uncertainty band includes statistical and experimental systematic uncertainties, where the statistical (light grey) and total (dark grey) uncertainties are shown separately in the ratio.

is unfolded to ensure sufficient statistics in the unfolding procedure. The electron and muon channels are also unfolded separately and the results are tested for consistency.

4.5 Uncertainties

Statistical Uncertainties

Statistical uncertainties on the unfolding output arise from three different sources. The first and dominant source are statistical uncertainties of the input data. Secondly, uncertainties are taken into account from limited statistics in the simulated response matrix and the third source are statistical uncertainties on the simulation of the background processes. After the unfolding a total statistical uncertainty is obtained for each measurement bin, including the effects of all three different sources. It is correlated between different individual measurement bins.

Experimental uncertainties

Systematic uncertainties due to experimental effects are evaluated by variations of calibration factors or corrections of efficiencies within their corresponding uncertainties. The resulting covariance matrix of the unfolded measurement is computed using error propagation. The uncertainties are evaluated by unfolding pseudo-data simulated with MADGRAPH+PYTHIA. Pseudo-data are preferred over real data because of smaller statistical fluctuations in the estimation of systematic uncertainties. The up or down variation yielding the larger overall uncertainty is quoted as uncertainty on the measurement. The following sources of experimental systematic uncertainties are considered.

Jet energy corrections (JECs) are applied dependent on p_{T} and η of the individual jets. The JECs were derived using anti-k_{T} jets with $R = 0.7$ (AK7) [84] and their application on CA12 jets has been tested. The uncertainties on the JECs are increased from the AK7 values by factors of two to four depending on p_{T}, to account for residual differences due to the larger jet size used in this analysis. When varying the JECs within their uncertainties to estimate the effect on the measured cross section only the momenta of jets are changed. The jet mass is kept fixed to avoid double-counting of uncertainties when including the uncertainty on the jet mass scale. Jet energy resolution (JER) smearing is applied as an η dependent correction to all jets in simulation. The corrections are varied within their uncertainty to estimate the systematic uncertainty related to the JER smearing. The uncertainties are found to be small compared to the ones from the JECs. The jet mass scale and corresponding uncertainty for CA12 jets has been studied in events containing an hadronic W decay reconstructed in a single jet in $\mathrm{t\bar{t}}$ production. The ratio of the reconstructed jet mass peak positions in data and simulation is $\mu_{\mathrm{Data}}/\mu_{\mathrm{MC}} = 1.015 \pm 0.012$. No correction to the jet mass scale is applied, but an uncertainty of 1.5% is assigned corresponding to the difference of the peak positions. The widths of the jet mass distributions are about 15 GeV, consistent between data and simulation. Corrections due to the b tagging efficiency are applied as p_{T} dependent scale factors for each jet flavour. The corresponding systematic uncertainty is obtained by a variation of the scale factors within their uncertainty. Pile-up correction factors are applied to match the number of primary interactions to the instantaneous luminosity profile in data. The uncertainty is obtained by a variation of the minimum bias cross section by $\pm 5\%$. Trigger and lepton identification scale factors are used to correct for differences in the lepton selection efficiency between data and simulation. The corresponding uncertainties are computed by a variation of the scale factors within their uncertainties.

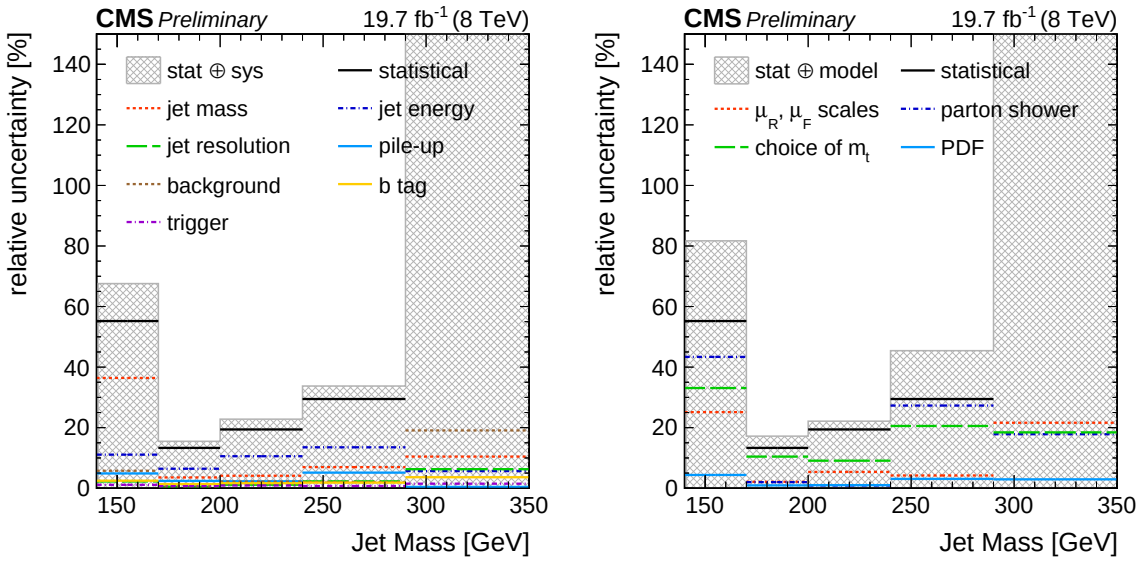

Figure 4: Statistical uncertainties compared to individual experimental systematic uncertainties (left) and statistical uncertainties compared to uncertainties originating from the modelling of $t\bar{t}$ production (right) for the differential cross section measurement. The respective total uncertainty is shown as cross-hatched regions. The statistical and total uncertainties in the last bin are around 340% and exceed the vertical scale. The size of the horizontal bars represents the bin widths.

Normalization Uncertainties

The effects of normalization uncertainties on the background processes are calculated by varying the amount of background subtracted prior to the unfolding, and propagating the effect to the unfolding output. The uncertainty on the W+jets cross section is taken to be 19%, as obtained from a measurement of W+heavy flavour production [93]. An uncertainty of 23% is applied to the single top quark production cross section [94]. For QCD multijet production an uncertainty of 100% is assumed.

Uncertainties affecting the overall normalization are added in quadrature to the total uncertainty after the unfolding. An uncertainty of 2.6% is applied due to the measurement of the integrated luminosity [95].

Modelling uncertainties

The unfolding is tested for dependencies on the simulation of $t\bar{t}$ production using different simulation programs. The effect on the measurement is estimated using one simulation as pseudo-data, input to the unfolding, and another for the estimation of the migration matrix. The unfolding output is then compared to the particle level distribution from the simulation used as pseudo-data. Differences between the unfolded result and the particle level distribution are considered as modelling uncertainties.

The uncertainty due to the choice of the MC generator is estimated by unfolding pseudo-data simulated with MADGRAPH+PYTHIA using a response matrix evaluated using POWHEG+PYTHIA. The effect due to the choice of the parton shower is estimated using MC@NLO+HERWIG.

The dependency on the choice of the top quark mass $m_{\text{top}}^{\text{MC}}$ used in the simulation for correcting the data is thoroughly tested. While the unfolded measurement is largely independent on the choice of $m_{\text{top}}^{\text{MC}}$, residual effects due to the kinematics of the leptons and jets can lead to

Table 1: Summary of the selection defining the fiducial measurement region.

Leptons	$p_{\mathrm{T}}^{\ell} > 45\,\mathrm{GeV}$	$\lvert\eta^{\ell}\rvert < 2.5$
CA12 jets	$p_{\mathrm{T},1} > 400\,\mathrm{GeV}$ $p_{\mathrm{T},2} > 150\,\mathrm{GeV}$ $p_{\mathrm{T,veto}} > 150\,\mathrm{GeV}$	$\lvert\eta\rvert < 2.5$
Event	$\Delta R(\ell, \mathrm{jet2}) < 1.2$ $m(\mathrm{jet1}) > m(\mathrm{jet2} + \ell)$	

additional uncertainties. This uncertainty is evaluated using events as pseudo-data simulated with MADGRAPH+PYTHIA for a large range of values of m_{t}, from 166.5 to 178.5 GeV. This large range is considered as no measurement of the top quark mass in this kinematic regime exists and a stable result, independent of the specific choice of m_{t} is therefore crucial. For this test, the migration matrix is obtained using MADGRAPH+PYTHIA with a value of $m_{\mathrm{t}} = 172.5\,\mathrm{GeV}$. The envelope of the uncertainty obtained due to the variations of m_{t} is used to define an additional modelling uncertainty.

The uncertainty due to missing higher orders in the simulation is estimated by varying the choice of the renormalisation and factorisation scales. Events simulated with POWHEG+PYTHIA and scales varied by factors of 2 and 0.5 are unfolded using a migration matrix obtained from the same simulation with the nominal choice of scales. The uncertainty on the measurement is obtained by using the maximum variation.

Uncertainties due to the PDFs are evaluated using the eigenvectors of the CT10 PDF set with the POWHEG+PYTHIA simulation. The resulting differences on the migration matrix are propagated to the measurement. The individual uncertainties due to each eigenvector are scaled to 68% confidence level and are added in quadrature [65].

Summary of uncertainties

A summary of the relative uncertainties relevant for this measurement is shown in Fig. 4. The largest contribution to the total uncertainty comes from statistical uncertainties. Experimental systematic uncertainties are smaller by several factors than uncertainties due to the modelling of $\mathrm{t\bar{t}}$ production. These uncertainties are expected to improve with larger datasets at higher centre-of-mass energies.

4.6 Results

The fiducial particle-level $\mathrm{t\bar{t}}$ cross section is measured differentially as a function of the leading jet mass in the ℓ+jets channel. The following selection defines the fiducial measurement region (cf. Sec. 4.2): one lepton with $p_{\mathrm{T}}^{\ell} > 45\,\mathrm{GeV}$ and $\lvert\eta^{\ell}\rvert < 2.5$, two CA12 jets with $\lvert\eta\rvert < 2.5$ and $p_{\mathrm{T},1} > 400\,\mathrm{GeV}$ and $p_{\mathrm{T},2} > 150\,\mathrm{GeV}$, no other CA12 jets with $p_{\mathrm{T,veto}} > 150\,\mathrm{GeV}$ in $\lvert\eta\rvert < 2.5$, a small distance between the subleading jet and the lepton $\Delta R(\ell, \mathrm{jet2}) < 1.2$; and the leading jet mass m_{jet} has to be greater than the mass of the sum of the four-vectors of the subleading jet and the lepton. The fiducial measurement region is summarized in Tab. 1.

The measured cross section in this fiducial region is shown in Fig. 5. The data are compared to simulated distributions obtained with POWHEG+PYTHIA, MADGRAPH+PYTHIA and MC@NLO+HERWIG. The measurement gives a total cross section of $103.5 \pm 11.2_{\mathrm{stat}} \pm 10.9_{\mathrm{sys}} \pm 9.4_{\mathrm{model}}\,\mathrm{fb}$. The predicted cross sections from the MADGRAPH+PYTHIA and POWHEG+PYTHIA $\mathrm{t\bar{t}}$ simulations, assuming a total $\mathrm{t\bar{t}}$ cross section of 252.9 pb, are 160 fb and 134 fb respectively. The predictions exceed the data by about 20–30%, consistent with previously measured $\mathrm{t\bar{t}}$ cross

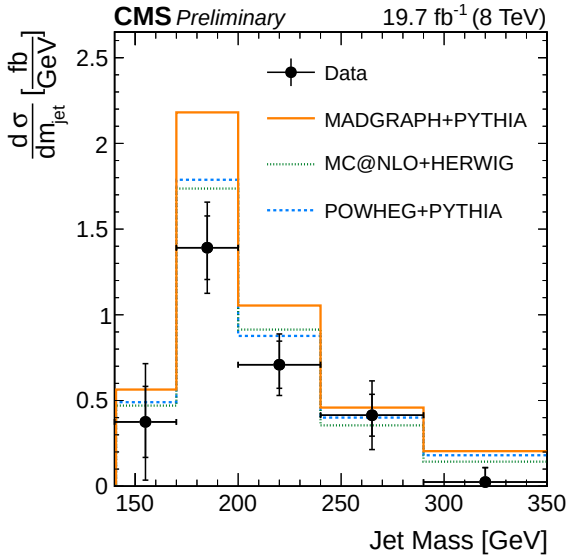

Figure 5: Fiducial particle-level differential $t\bar{t}$ production cross section as a function of the leading jet mass. The combined cross sections obtained from the electron and muon channels are compared to the predictions from MADGRAPH+PYTHIA, POWHEG+PYTHIA and MC@NLO+HERWIG. The data are shown with statistical (inner bars) and total (outer bars) uncertainties.

sections at high top quark p_T [7, 8]. A similar trend is observed when comparing the data to the prediction from MC@NLO+HERWIG.

The normalised differential cross section $\frac{1}{\sigma}\frac{d\sigma}{dm_{jet}}$ is obtained by dividing by the total cross section in the m_{jet} range from 140 to 350 GeV. The result is shown in Fig. 6 together with the prediction of MADGRAPH+PYTHIA using three different values of m_t. The data are well described by the simulation, showing the overall good modelling of the top quark jet mass once the disagreement due to the overall cross section at high top quark p_T is eliminated by the normalisation. The sensitivity of the measurement to m_t is clearly visible, albeit compromised by the uncertainties.

5 Sensitivity to the top quark mass

Calculations of m_{jet} for $t\bar{t}$ production from first principles are not available yet for the LHC. Still, a determination of m_t using the normalised particle-level cross section measurement provides a proof-of principle for the feasibility of the method, a cross check with other determinations of m_t and an estimate of the current measurement's sensitivity. The value of m_t is determined using the normalised measurement since only the shape of the m_{jet} distribution can be reliably calculated. Correlations are taken into account using the full covariance matrix of the measurement. Theory predictions are obtained from MADGRAPH+PYTHIA with different values of m_t. A template fit is performed based on values of χ^2, which are evaluated as

$$\chi^2 = d^T V^{-1} d \tag{2}$$

where d is the vector of differences between measured normalised cross section values and predictions and V is the covariance matrix which includes statistical, experimental systematic, modelling and theory uncertainties. Theory uncertainties are calculated by varying the scales

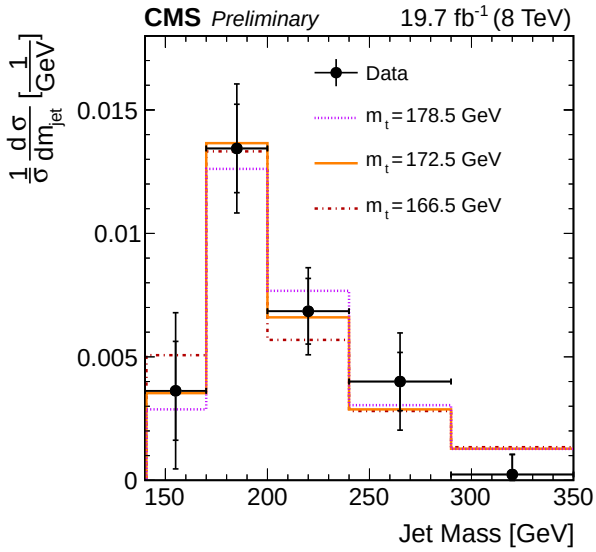

Figure 6: Normalised particle-level differential $t\bar{t}$ production cross section in the fiducial region as a function of the leading jet mass. The measurement is compared to predictions from MADGRAPH+PYTHIA using three different values of m_t. The data are shown with statistical (inner bars) and total (outer bars) uncertainties.

μ_R and μ_F by factors of 0.5 and 2 in the MADGRAPH+PYTHIA simulation. The resulting uncertainty is added to the covariance matrix. The resulting χ^2 values for different values of m_t are fitted by a second-order polynomial to determine the minimum and the uncertainty, defined by a change in χ^2 by one. The result is

$$m_t = 171.8 \pm 9.5 \text{ GeV} \quad (3)$$

$$= 171.8 \pm 5.4\,(\text{stat}) \pm 3.0\,(\text{syst}) \pm 5.5\,(\text{model}) \pm 4.6\,(\text{theory}) \text{ GeV} \quad (4)$$

with a minimum value of $\chi^2_{\text{min}} = 2.2$ for four degrees of freedom. The individual statistical, systematic, modelling and theory uncertainties are obtained by subsequently adding the corresponding uncertainties. The value obtained represents the first determination of m_t from boosted $t\bar{t}$ production calibrated to the MADGRAPH+PYTHIA simulation. It is consistent with recent direct determinations of m_t using MC event generators [23, 25–27], extractions of m_t from cross section measurements [24, 96, 97] and with indirect constraints from the electroweak fit [98]. The result shows the feasibility of the method and will provide an important ingredient for studies of the relation between the value of m_t obtained using MC event generators and the top-quark mass obtained from first-principle calculations.

6 Summary

A first measurement of the differential $t\bar{t}$ cross section as a function of the leading jet mass m_{jet} in the boosted regime has been performed. The measurement is carried out in a fiducial region with fully-merged hadronic top quark decays and is corrected to the particle level. The shape of the m_{jet} distribution agrees with predictions from simulations, showing the overall good modelling of the jet mass for top quarks. The total cross section for m_{jet} between 140 – 350 GeV is 103.5 ± 18.2 fb, which is 20–30% lower than predicted due to the softer top quark p_T spectrum observed in data than in simulation [7, 8].

The peak position of the m_{jet} distribution exhibits sensitivity to the top quark mass m_t. This can

be used for an independent determination of m_t in the boosted regime, with the prospect of a more reliable resemblance between the pole-mass (or m_t in any well-defined renormalisation scheme) and the top quark mass parameter m_t in full-scale event generators.

The normalised particle-level measurement of m_{jet} is used to extract a value of m_t to estimate the current sensitivity of the data. The value obtained, $m_t = 171.8 \pm 9.5$ GeV is consistent with the current LHC+Tevatron average, 173.34 ± 0.27 (stat) ± 0.71 (syst) GeV [99], albeit with a much larger uncertainty. New data at higher centre-of-mass energies with higher integrated luminosities will lead to an improvement in the statistical uncertainty. Larger statistics can also lead to improvements on the experimental systematic uncertainties, most notably on the jet mass scale, which is expected to improve with smaller jet distance parameters. Additionally, improvements on the modelling uncertainty are expected due to better constraints of the simulation in the boosted regime. A reduction of the theory uncertainty is foreseeable with the availability of higher order calculations.

References

[1] CMS Collaboration, "Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **73** (2013) 2339, doi:10.1140/epjc/s10052-013-2339-4, arXiv:1211.2220.

[2] ATLAS Collaboration, "Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$TeV using the ATLAS detector", *Phys. Rev. D* **90** (2014) 072004, doi:10.1103/PhysRevD.90.072004, arXiv:1407.0371.

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

[4] 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.

[5] CMS Collaboration, "Measurement of the $\mathrm{t\bar{t}}$ production cross section in the all-jets final state in pp collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J. C* **76** (2016) 128, doi:10.1140/epjc/s10052-016-3956-5, arXiv:1509.06076.

[6] ATLAS Collaboration, "Measurement of top quark pair differential cross-sections in the dilepton channel in pp collisions at $\sqrt{s}$ = 7 and 8 TeV with ATLAS", (2016). arXiv:1607.07281.

[7] 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. D* **93** (2016) 032009, doi:10.1103/PhysRevD.93.032009, arXiv:1510.03818.

[8] CMS Collaboration, "Measurement of the integrated and differential t-tbar production cross sections for high-pt top quarks in pp collisions at $\sqrt{s}$ = 8 TeV", *Submitted to: Phys. Rev. D.* (2016) arXiv:1605.00116.

[9] CMS Collaboration, "Search for anomalous $\mathrm{t\bar{t}}$ production in the highly-boosted all-hadronic final state", *JHEP* **1209** (2012) 029, doi:10.1007/JHEP09(2012)029, arXiv:1204.2488.

[10] ATLAS Collaboration, "A search for $\mathrm{t\bar{t}}$ resonances in lepton+jets events with highly boosted top quarks collected in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *JHEP* **1209** (2012) 041, doi:10.1007/JHEP09(2012)041, arXiv:1207.2409.

[11] ATLAS Collaboration, "Search for resonances decaying into top-quark pairs using fully hadronic decays in pp collisions with ATLAS at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2013) 116, doi:10.1007/JHEP01(2013)116, arXiv:1211.2202.

[12] ATLAS Collaboration, "Search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 4.7 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. D* **88** (2013) 012004, doi:10.1103/PhysRevD.88.012004, arXiv:1305.2756.

[13] CMS Collaboration, "Searches for new physics using the $t\bar{t}$ invariant mass distribution in pp collisions at $\sqrt{s}$=8 TeV", *Phys. Rev. Lett.* **111** (2013) 211804, doi:10.1103/PhysRevLett.111.211804, arXiv:1309.2030. [Erratum ibid. 112 (2014) 119903 doi:10.1103/PhysRevLett.112.119903].

[14] ATLAS Collaboration, "Search for $W' \to tb \to qqbb$ decays in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector", *Eur. Phys. J. C* **75** (2015) 165, doi:10.1140/epjc/s10052-015-3372-2, arXiv:1408.0886.

[15] CMS Collaboration, "Search for vector-like T quarks decaying to top quarks and Higgs bosons in the all-hadronic channel using jet substructure", *JHEP* **06** (2015) 080, doi:10.1007/JHEP06(2015)080, arXiv:1503.01952.

[16] ATLAS Collaboration, "A search for $t\bar{t}$ resonances using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *JHEP* **08** (2015) 148, doi:10.1007/JHEP08(2015)148, arXiv:1505.07018.

[17] CMS Collaboration, "Search for resonant $t\bar{t}$ production in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **93** (2016) 012001, doi:10.1103/PhysRevD.93.012001, arXiv:1506.03062.

[18] CMS Collaboration, "Search for the production of an excited bottom quark decaying to tW in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JHEP* **01** (2016) 166, doi:10.1007/JHEP01(2016)166, arXiv:1509.08141.

[19] ATLAS Collaboration, "Search for the production of single vector-like and excited quarks in the Wt final state in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector", *JHEP* **02** (2016) 110, doi:10.1007/JHEP02(2016)110, arXiv:1510.02664.

[20] ATLAS Collaboration, "Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment", *JHEP* **05** (2012) 128, doi:10.1007/JHEP05(2012)128, arXiv:1203.4606.

[21] CMS Collaboration, "Studies of jet mass in dijet and W/Z + jet events", *JHEP* **05** (2013) 090, doi:10.1007/JHEP05(2013)090, arXiv:1303.4811.

[22] A. H. Hoang and I. W. Stewart, "Top Mass Measurements from Jets and the Tevatron Top-Quark Mass", *Nucl. Phys. Proc. Suppl.* **185** (2008) 220, doi:10.1016/j.nuclphysbps.2008.10.028, arXiv:0808.0222.

[23] ATLAS Collaboration, "Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets and $t\bar{t} \to$ dilepton channels using $\sqrt{s} = 7$ TeV ATLAS data", *Eur. Phys. J. C* **75** (2015) 330, doi:10.1140/epjc/s10052-015-3544-0, arXiv:1503.05427.

[24] ATLAS Collaboration, "Determination of the top-quark pole mass using $t\bar{t}$ + 1-jet events collected with the ATLAS experiment in 7 TeV pp collisions", *JHEP* **10** (2015) 121, doi:10.1007/JHEP10(2015)121, arXiv:1507.01769.

[25] 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.

[26] CMS Collaboration, "Measurement of the top quark mass using charged particles in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **93** (2016) 092006, doi:10.1103/PhysRevD.93.092006, arXiv:1603.06536.

[27] ATLAS Collaboration, "Measurement of the top quark mass in the $t\bar{t} \to$ dilepton channel from $\sqrt{s} = 8$ TeV ATLAS data", *Phys. Lett. B* **761** (2016) 350, doi:10.1016/j.physletb.2016.08.042, arXiv:1606.02179.

[28] C. W. Bauer, S. Fleming, and M. E. Luke, "Summing Sudakov logarithms in $B \to X_s + \gamma$ in effective field theory", *Phys. Rev. D* **63** (2000) 014006, doi:10.1103/PhysRevD.63.014006, arXiv:hep-ph/0005275.

[29] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, "An Effective field theory for collinear and soft gluons: Heavy to light decays", *Phys. Rev. D* **63** (2001) 114020, doi:10.1103/PhysRevD.63.114020, arXiv:hep-ph/0011336.

[30] C. W. Bauer and I. W. Stewart, "Invariant operators in collinear effective theory", *Phys. Lett. B* **516** (2001) 134, doi:10.1016/S0370-2693(01)00902-9, arXiv:hep-ph/0107001.

[31] C. W. Bauer, D. Pirjol, and I. W. Stewart, "Soft collinear factorization in effective field theory", *Phys. Rev. D* **65** (2002) 054022, doi:10.1103/PhysRevD.65.054022, arXiv:hep-ph/0109045.

[32] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, "Jets from massive unstable particles: Top-mass determination", *Phys. Rev. D* **77** (2008) 074010, doi:10.1103/PhysRevD.77.074010, arXiv:hep-ph/0703207.

[33] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, "Top Jets in the Peak Region: Factorization Analysis with NLL Resummation", *Phys. Rev. D* **77** (2008) 114003, doi:10.1103/PhysRevD.77.114003, arXiv:0711.2079.

[34] A. H. Hoang, A. Pathak, P. Pietrulewicz, and I. W. Stewart, "Hard Matching for Boosted Tops at Two Loops", *JHEP* **12** (2015) 059, doi:10.1007/JHEP12(2015)059, arXiv:1508.04137.

[35] S. Moch et al., "High precision fundamental constants at the TeV scale", arXiv:1405.4781.

[36] A. H. Hoang, "The Top Mass: Interpretation and Theoretical Uncertainties", in *7th International Workshop on Top Quark Physics (TOP2014) Cannes, France, September 28-October 3, 2014*. 2014. arXiv:1412.3649.

[37] G. Corcella, "Interpretation of the top-quark mass measurements: a theory overview", *PoS* **TOP2015** (2016) 037, arXiv:1511.08429.

[38] M. Butenschoen et al., "Top Quark Mass Calibration for Monte Carlo Event Generators", arXiv:1608.01318.

[39] 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.

[40] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, p. 270. 1998. arXiv:hep-ph/9907280.

[41] CMS Collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[42] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[43] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[44] 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.

[45] CMS Collaboration, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV", *JINST* **8** (2013) P09009, doi:10.1088/1748-0221/8/09/P09009, arXiv:1306.2016.

[46] 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.

[47] 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.

[48] 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.

[49] M. Cacciari, G. P. Salam, and G. Soyez, "The catchment area of jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.

[50] 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.

[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, "Jet Performance in pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-003, 2010.

[53] CMS Collaboration, "Missing transverse energy performance of the CMS detector", *JINST* **6** (2011) P09001, doi:10.1088/1748-0221/6/09/P09001, arXiv:1106.5048.

[54] 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.

[55] 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.

[56] 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.

[57] J. Alwall et al., "MadGraph 5: going beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[58] 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.

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

[60] 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:10.1007/JHEP02(2010)011, 10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: JHEP02,011(2010)].

[61] 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.

[62] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", *JHEP* **06** (2002) 029, doi:10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244.

[63] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-quark production in hadronic collisions", *JHEP* **01** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.

[64] P. M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables", *Phys. Rev. D* **78** (2008) 013004, doi:10.1103/PhysRevD.78.013004, arXiv:0802.0007.

[65] 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.

[66] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195.

[67] CMS Collaboration, "Study of the underlying event at forward rapidity in pp collisions at $\sqrt{s}$ = 0.9, 2.76, and 7 TeV", *JHEP* **04** (2013) 072, doi:10.1007/JHEP04(2013)072, arXiv:1302.2394.

[68] 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.

[69] G. Corcella et al., "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)", *JHEP* **01** (2001) 010, doi:10.1088/1126-6708/2001/01/010, arXiv:hep-ph/0011363.

[70] N. Kidonakis, "NNLL threshold resummation for top-pair and single-top production", *Phys. Part. Nucl.* **45** (2014) 714, doi:10.1134/S1063779614040091, arXiv:1210.7813.

[71] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", *Comput. Phys. Commun.* **182** (2011) 2388, doi:10.1016/j.cpc.2011.06.008, arXiv:1011.3540.

[72] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "W Physics at the LHC with FEWZ 2.1", *Comput. Phys. Commun.* **184** (2013) 208, doi:10.1016/j.cpc.2012.09.005, arXiv:1201.5896.

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

[74] J. M. Campbell, R. K. Ellis, and K. Williams, "Vector boson pair production at the LHC", *JHEP* **07** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[75] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, "Hadronic top-quark pair production with NNLL threshold resummation", *Nucl. Phys. B* **855** (2012) 695, doi:10.1016/j.nuclphysb.2011.10.021, arXiv:1109.1536.

[76] M. Cacciari et al., "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation", *Phys. Lett. B* **710** (2012) 612–622, doi:10.1016/j.physletb.2012.03.013, arXiv:1111.5869.

[77] P. Bärnreuther, M. Czakon, and A. Mitov, "Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$", *Phys. Rev. Lett.* **109** (2012) 132001, doi:10.1103/PhysRevLett.109.132001, arXiv:1204.5201.

[78] M. Czakon and A. Mitov, "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels", *JHEP* **12** (2012) 054, doi:10.1007/JHEP12(2012)054, arXiv:1207.0236.

[79] M. Czakon and A. Mitov, "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction", *JHEP* **01** (2013) 080, doi:10.1007/JHEP01(2013)080, arXiv:1210.6832.

[80] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{S})$", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[81] 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.

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

[83] 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.

[84] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *Submitted to: JINST* (2016) arXiv:1607.03663.

[85] 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.

[86] 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.

[87] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, "Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks", *Phys. Rev. Lett.* **101** (2008) 142001, doi:10.1103/PhysRevLett.101.142001, arXiv:0806.0848.

[88] CMS Collaboration, "A Cambridge-Aachen (C-A) Based Jet Algorithm For Boosted Top-Jet Tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, 2009.

[89] T. Plehn, G. P. Salam, and M. Spannowsky, "Fat Jets for a Light Higgs", *Phys. Rev. Lett.* **104** (2010) 111801, doi:10.1103/PhysRevLett.104.111801, arXiv:0910.5472.

[90] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, "Stop Reconstruction with Tagged Tops", *JHEP* **1010** (2010) 078, doi:10.1007/JHEP10(2010)078, arXiv:1006.2833.

[91] T. Lapsien, R. Kogler, and J. Haller, "A new tagger for hadronically decaying heavy particles at the LHC", arXiv:1606.04961.

[92] 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.

[93] CMS Collaboration, "Measurement of the production cross section of the W boson in association with two b jets in pp collisions at sqrt(s) = 8 TeV", arXiv:1608.07561.

[94] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}$ =8 TeV", *Phys. Rev. Lett.* **112** (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.

[95] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", CMS Physics Analysis Summary CMS-PAS-LUM-13-001, 2013.

[96] ATLAS Collaboration, "Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b -tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector", *Eur. Phys. J. C* **74** (2014) 3109, doi:10.1140/epjc/s10052-014-3109-7, arXiv:1406.5375.

[97] CMS Collaboration, "Measurement of the t-tbar production cross section in the e-mu channel in proton-proton collisions at sqrt(s) = 7 and 8 TeV", *JHEP* **08** (2016) 029, doi:10.1007/JHEP08(2016)029, arXiv:1603.02303.

[98] Gfitter Group Collaboration, "The global electroweak fit at NNLO and prospects for the LHC and ILC", *Eur. Phys. J. C* **74** (2014) 3046, doi:10.1140/epjc/s10052-014-3046-5, arXiv:1407.3792.

[99] ATLAS, CDF, CMS, D0 Collaboration, "First combination of Tevatron and LHC measurements of the top-quark mass", arXiv:1403.4427.