CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2017/04/04

Measurement of the inclusive $\mathrm{t\bar{t}}$ cross section at $\sqrt{s} = 5.02$ TeV

The CMS Collaboration

Abstract

The top quark pair production cross section is measured is measured in pp collisions at a center-of-mass energy $\sqrt{s} = 5.02$ TeV. The analyzed data have been collected by the CMS experiment at the CERN LHC and correspond to an integrated luminosity of 27.4 pb^{-1}. The measurement is performed by analyzing events with at least one charged lepton. The measured cross section is $\sigma_{\mathrm{t\bar{t}}} = 69.5 \pm 8.4$ pb, in agreement with the expectation from the standard model. The impact of the presented measurement on the gluon distribution function is illustrated through a QCD analysis at next-to-next-to leading order.

1 Introduction

At hadron colliders, the top quark, the heaviest elementary particle of the standard model (SM), is an important topic for investigations. The evolution of its pair-production ($\mathrm{t\bar{t}}$) cross section ($\sigma_{\mathrm{t\bar{t}}}$) as a function of center-of-mass energy is of interest for the extraction of the top-quark pole mass and can be used to constrain the gluon distribution function at large longitudinal parton momentum fraction, x. Precise measurements of the $\mathrm{t\bar{t}}$ cross section in proton-proton (pp) collisions have been published at $\sqrt{s}$ values of 7 and 8 TeV [1–4] and 13 TeV [5–8] by the CMS [9] and ATLAS [10] experiments at the LHC. In November 2015, the LHC delivered pp collisions at $\sqrt{s} = 5.02$ TeV. The fraction of $\mathrm{t\bar{t}}$ events initiated by gluon-gluon collisions grows monotonically with center-of-mass energy. It is around 73% at 5.02 TeV and, when compared to around 86% at 13 TeV, as calculated with POWHEG [11–13] at next-to-leading order (NLO) using the NNPDF3.0 NLO [14] parton distribution functions (PDFs), makes this new data set partially complementary to the higher energy samples. In addition, future measurements of the $\mathrm{t\bar{t}}$ cross section in nuclear collisions at the same nucleon-nucleon center-of-mass energy [15, 16] would profit from the availability of a reference measurement in pp collisions at 5.02 TeV, without the need to extrapolate from measurements at different $\sqrt{s}$.

In the SM, top quarks in pp collisions are mostly produced as $\mathrm{t\bar{t}}$ pairs and then each top quark decays predominantly to a W boson and a bottom (b) quark. Top-quark pair events are categorized according to the decay of the two W bosons. In $\mathrm{t\bar{t}}$ events where one W boson decays leptonically and the other hadronically (ℓ+jets channel), the final state presents a typical signature of one isolated lepton, two jets from the W boson decay, missing momentum and two b jets. On the other hand, in $\mathrm{t\bar{t}}$ events where both W bosons decay leptonically (dilepton channel), the final state contains two leptons of opposite electric charge, missing momentum, and at least two jets coming from the hadronization of the b quarks. The ℓ+jets channel presents a large branching fraction with a moderate amount of background, while the dilepton channel is characterized by the large selection purity that can be achieved in this final state, thus compensating for its smaller branching ratio.

This analysis extends the first measurement of $\sigma_{\mathrm{t\bar{t}}}$ in pp collisions at $\sqrt{s} = 5.02$ TeV [17] using $\mathrm{t\bar{t}}$ candidate events with ℓ+jets, where leptons are either electrons or muons, and dilepton ($\mathrm{e}^{\pm}\mu^{\mp}$ or $\mu^{\pm}\mu^{\mp}$) final states. In the former case, the cross section is extracted by a fit in different categories, while in the latter a simpler but more robust event-counting approach is used. The two results are then combined in the final measurement.

As previously mentioned, top-quark pair production probes the gluon distribution in the proton. In particular, $\mathrm{t\bar{t}}$ production at $\sqrt{s} = 5.02$ TeV accesses the high-x region, where the gluon distribution is poorly known. Improvements of our knowledge of the gluon distribution at high x are essential for accurate theory predictions for cross sections in the SM and interactions beyond the SM at hadron colliders. In this analysis the impact of the $\mathrm{t\bar{t}}$ cross section measurement on the uncertainty in the gluon distribution is illustrated through a QCD analysis at next-to-next-to leading order (NNLO).

This document is structured as follows. Section 2 gives a summary of the data and simulation samples used. After the discussion of the object reconstructions in Section 3 and the required triggers and event selection in Section 4, Section 5 describes the determination of backgrounds. The different sources of systematic uncertainties are discussed in Section 6. The extraction of $\sigma_{\mathrm{t\bar{t}}}$ is presented in Section 7 and the impact of the presented measurement on PDFs of the proton is discussed in Section 8. Conclusions are drawn and the summary of all results is given in Section 9.

2 Data and simulation samples

This analysis is based on data collected by the CMS detector [9] during the 5.02 TeV pp run of November 2015 [18], corresponding to an integrated luminosity of 27.4 pb^{-1}. The average number of collisions in the same bunch crossing ("pileup") is estimated online to be around 1.4 in this data set, assuming a minimum bias cross section of 65 mb.

Several Monte Carlo (MC) event generators are used to simulate signal and background events. The NLO POWHEG (v2) [12, 19] generator is used for $\mathrm{t\bar{t}}$ events, assuming a top quark mass of 172.5 GeV. For the default $\mathrm{t\bar{t}}$ MC sample, these events are passed to PYTHIA (v8.205) [20, 21] to simulate parton showering, hadronization, and the underlying event. An alternative sample is obtained by showering the events with HERWIG++ [22]. The NNPDF3.0 NLO [14] PDFs with $\alpha_S = 0.118$ are utilized in the MC calculations.

The MG5_aMC@NLO generator [23] is used to simulate W boson production with additional jets (W+jets) events and Drell–Yan quark-antiquark annihilation into lepton-antilepton pairs through Z boson or virtual photon exchange (referred to as "Z/γ^*"). The simulation includes up to two extra partons at matrix element level, and the FXFX merging procedure [24] is used to interface with PYTHIA. Low mass Z/γ^* events (20-50 GeV) are simulated with PYTHIA. Single top quark plus W boson events (tW) are simulated using POWHEG (v1) [25, 26] interfaced to PYTHIA, and are normalized to the approximate NNLO cross sections [27]. The contributions from WW and WZ (referred to as "WV") are simulated with PYTHIA, and are normalized to the NLO cross sections calculated with the MCFM program [28]. All generated events undergo a full GEANT4 [29] simulation of the detector response.

The expected yields for signal are normalized to the value of the SM prediction for the $\mathrm{t\bar{t}}$ production cross section:

$$\sigma^{\mathrm{NNLO}} = 68.9^{+1.9}_{-2.3}(\mathrm{scale}) \pm 2.3(\mathrm{PDF})^{+1.4}_{-1.0}(\alpha_S)\ \mathrm{pb} \quad (1)$$

as calculated with the TOP++ program [30] at NNLO in perturbative QCD, including soft-gluon resummation at next-to-next-to-leading-log order [31], assuming a top quark mass $m_t = 172.5$ GeV and with the NNPDF3.0 NNLO PDF set, using $\alpha_S = 0.118$ [14]. The uncertainty of 0.1% on the LHC beam energy [32] translates into an additional uncertainty of 0.22 pb on the expected cross section, with negligible impact on the acceptance of all the channels included in this analysis.

3 Object reconstruction

All the objects are reconstructed with the particle-flow (PF) algorithm [33, 34]. The PF algorithm reconstructs and identifies each individual particle using an optimized combination of information from the various elements of the CMS detector.

The electron momentum is estimated by combining the energy measurement in the ECAL with the momentum measurement in the tracker, taking into account bremsstrahlung photons spatially compatible with originating from the electron track. The momentum resolution for electrons with transverse momentum $p_T \approx 45$ GeV from $\mathrm{Z} \to \mathrm{ee}$ decays ranges from 1.7% for non-showering electrons in the barrel region to 4.5% for showering electrons in the endcaps [35]. Muon candidates are reconstructed combining the information from the muon spectrometer and the silicon tracker. This results in a relative transverse momentum resolution of 1.3–2.0% in the barrel and better than 6% in the endcaps, for muons with $20 < p_T < 100$ GeV and within

the pseudorapidity range $|\eta| < 2.4$. The p_T resolution in the barrel is better than 10% for muons with p_T up to 1 TeV [36, 37]. The energy of photons is directly obtained from the ECAL measurement, corrected for zero-suppression effects. The energy of charged hadrons is determined from a combination of their momenta measured in the tracker and the matching ECAL and HCAL energy deposits, 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 energy.

The isolation of electron and muon candidates from nearby jet activity is evaluated as follows. For each electron and muon candidate, a cone of $\Delta R = 0.3$ and $\Delta R = 0.4$, respectively, is constructed around the direction of the track at the primary event vertex (defined as the vertex with the largest value of the sum of p_T^2 over all its associated charged tracks), where ΔR is defined as $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, and $\Delta\eta$ and $\Delta\phi$ are the distances in pseudorapidity and azimuthal angle. Excluding the contribution from the lepton candidate, the scalar sum of the p_T of all particle candidates that are inside ΔR and are consistent with arising from the primary event vertex is calculated to define a relative isolation discriminant, I_{rel}, through the ratio of this sum to the p_T of the lepton candidate. The neutral-particle contribution to I_{rel} is corrected for energy deposits from pileup interactions using two different techniques for electrons and muons. For muons, half (0.5) of the momenta of the charged hadron PF candidates not originating from the primary vertex are subtracted. The 0.5 prefactor accounts for the different fraction of charged and neutral particles in the cone. For electrons, the FASTJET technique [38] is used, in which the median of the energy-density distribution of neutral particles (within the area of any jet in the event) multiplied by the geometric area of the isolation cone - scaled by a factor that accounts for the residual dependence of the average pileup deposition on the electron η - is subtracted.

The efficiency of the lepton selection is measured using a "tag-and-probe" method in same-flavor dilepton events enriched in Z boson candidates, following the method of Ref. [39]. The sample of $Z \to \mu^+\mu^-$ events used for muon efficiency extraction is selected by the same trigger requirement used by the main analysis (Section 4), while the $Z \to e^+e^-$ sample for electron efficiency extraction makes use of events that fired a diphoton trigger with symmetric transverse energy, E_T, thresholds of $E_T = 15$ GeV covering the full tracker acceptance. Pairs of photon candidates above such E_T threshold are accepted only if their invariant mass are above 50 GeV. The trigger selection requires a loose identification using cluster shower shapes and a selection on the hadronic over the electromagnetic energy of the photon candidates. Based on a comparison of lepton selection efficiency in data and simulation, the event yield in simulation is corrected using data-to-simulation scale factors.

Jets are reconstructed from the PF particle candidates using the anti-k_T clustering algorithm [40] with a distance parameter of 0.4. Jets closer than $\Delta R = 0.3$ to the nearest muon or electron are discarded. Jet energy corrections extracted from full detector simulation are also applied as a function of jet p_T and η [41] to data and simulation. A residual correction to the data is applied to account for the data-simulation discrepancy in the jet response.

The missing transverse momentum vector 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 p_T^{miss} and corrections to jet momenta are propagated to p_T^{miss} calculation.

4 Event selection

The event sample is selected by a loose online filter (unprescaled trigger) and further cleaned offline to remove noncollision events, such as beam-gas interactions or cosmic muon events. Collision events containing one high-p_T muon (electron) candidate are selected by requiring values of p_T (E_T) greater than 15 (40) GeV and of $|\eta|$ less than 2.5 (3.1). The measured trigger selection efficiency is higher than 90%.

In the $\mu(e)$+jets final state, events are required to have been triggered by one muon (electron) candidate, as previously described. Electrons are selected if they have $p_T > 40\,\text{GeV}$, $|\eta| < 2.5$. Further identification and isolation criteria are applied to the electron candidates. Electrons reconstructed in the ECAL barrel (endcap) are required to have $I_{\text{rel}} < 4\%$ ($< 5\%$). In addition, electron candidates in the transition region between the barrel and endcap sections of the ECAL are excluded. Muons are required to have $p_T > 25$ GeV, $|\eta| < 2.1$. Dedicated identification criteria are applied and I_{rel} is required to be $< 15\%$. Events are rejected from the analysis in case they contain extra electrons (muons) that are reconstructed using a looser set of identification criteria and have $p_T > 15\,\text{GeV}$ ($> 10\,\text{GeV}$).

Backgrounds from W+jets, QCD multijet, and Z/γ^* events can be further reduced by counting the number of jets identified as originating from b quarks ("b jets") in the selected events. The distinct signature of two b jets, expected in $t\bar{t}$ decays, is rare to occur in background processes, and thus exploited in the analysis. In addition, in the ℓ+jets final state two light jets are expected to be produced after the decay of one of the W bosons for signal events. The resonant nature of these light jets carries a distinctive hallmark with respect to the main backgrounds.

In the ℓ+jets analysis, jets are selected if they have $p_T > 30$ GeV and $|\eta| < 2.4$. The flavor of the jets is identified using a combined secondary vertex algorithm [42] with an operating point which yields a b-jet identification efficiency of 67% and a misidentification (mistag) probability of about 1% and 15% for light-flavor jets (u, d, s, and gluons) and c jets, respectively. The event selection requires at least two non-b-tagged jets to be identified as candidates from the $W \to qq'$ decay (Section 7.1). Additional jets, passing the b-identification criteria, are counted and used to classify the selected events in none (0b), exactly one (1b), or at least two ($\geq$2b) tagged jet categories. The efficiency of the b-identification algorithm is measured in-situ, simultaneously with the signal cross section.

Dilepton events are required to contain at least one muon candidate at trigger level. No requirement on the presence of electron candidates is made at trigger level due to a relatively high-p_T threshold (40 GeV) of the filter. Electrons are selected if they have $p_T > 20$ GeV, $|\eta| < 2.4$, $I_{\text{rel}} < 9\%$ (if in the barrel) or $I_{\text{rel}} < 12\%$ (if in one of the endcaps). As in the ℓ+jets channel, electrons in the transition region between the barrel and endcap sections of the ECAL are excluded. Muons are required to have $p_T > 18$ GeV, $|\eta| < 2.1$, $I_{\text{rel}} < 15\%$. Jets are considered if they have $p_T > 25$ GeV and $|\eta| < 3$. Events are subsequently selected if they have a pair of leptons passing the criteria listed above ($e^{\pm}\mu^{\mp}$ or $\mu^{\pm}\mu^{\mp}$) with opposite charge. In events with more than one pair of leptons passing the above selection, the two leptons of opposite charge that yield the highest p_T sum are selected for further study. Events with τ leptons contribute to the measurement only if they decay to electrons or muons that satisfy the selection requirements, and are included in the simulations. Candidate events with dilepton invariant masses of $M_{\ell\ell} < 20$ GeV are removed to suppress events from decays of heavy-flavor resonances and low-mass Z/γ^* processes. Dilepton events with two muons in the final state are still dominated by Z/γ^* background events. In order to suppress this contribution, events in this channel are further required to have a dilepton invariant mass outside a Z mass window, i.e. events falling within $76 < M_{\ell\ell} < 106$ GeV are vetoed. As a further requirement to reject Z/γ^* events, a cut on

the missing transverse momentum of $p_T^{miss} > 35$ GeV is imposed. Finally, the presence of at least two jets satisfying the criteria listed above is required.

Table 1 summarizes the basic selection requirements that are applied in the ℓ+jets and dilepton analysis.

Table 1: Basic selection criteria applied in the three final states considered in this analysis.

Step	Object	ℓ+jets	Dilepton	
0	Trigger	one μ (e) candidate, p_T >25 (E_T >40) GeV	one μ candidate, p_T >15 GeV	
1	Leptons	exactly 1 μ or e	$e^{\pm}\mu^{\mp}$	$\mu^{\pm}\mu^{\mp}$
2	$M_{\ell\ell}$	—	> 20 GeV	
3	Z veto	—	—	Yes
4	p_T^{miss}	—	—	> 35 GeV
5	Jets	≥ 2		

5 Background estimation

5.1 ℓ+jets final state

In the ℓ+jets analysis, all background processes are estimated from simulation, with the exception of the QCD multijet background. Due to its large cross section, there is a non-negligible probability that the latter can mimic a $t\bar{t}$ event yielding ℓ+jets in the final state, and thus pass the offline event selection. Both the contribution from hard fragmentation of c and b quarks whose hadrons decay semi-leptonically, and the contribution from fake leptons from either punch-through hadrons or collimated jets with a high electromagnetic fraction, can yield ℓ+jets-like topologies.

The estimation of the QCD multijet background is separately performed in 0b-, 1b- and $\geq$2b-jet categories using a control region where either the muon candidate fails a looser isolation requirement ($I_{rel} > 20\%$) or the electron candidate fails the identification criteria. The choice of the QCD multijet control region has been made in such a way to minimize the contamination of the signal and W+jets events, while keeping a high statistics sample for the estimation of this type of background. The initial normalization of the QCD multijet in the signal region is derived from events with $p_T^{miss} < 20\,\text{GeV}$. Events in both the signal and control region fulfilling this requirement are counted. After subtracting the expected contributions from non-QCD processes, the ratio of the number of events observed in the control region, with respect to the number of events observed in the signal region, is used as a transfer factor to normalize the QCD multijet estimate. In both the electron and muon channel, a 30% uncertainty is assigned to the estimate of the expected contribution from non-QCD processes, resulting in both a normalization and a shape uncertainty on the distributions for the QCD multijet processes. The variations are applied independently in the signal and control regions in order to find the uncertainty envelope assigned to the QCD estimate. A more accurate normalization for this contribution is obtained by the fit performed to extract the final cross section, described in Section 7.1.

5.2 Dilepton final state

Final states with two prompt leptons can also originate from background processes, primarily from Z/γ^* ($\tau^+\tau^- \to e^{\pm}\mu^{\mp}$ or $\mu^{\pm}\mu^{\mp}$), tW, and WV events. Other background sources, such as W+jets events or $t\bar{t}$ with decays into one lepton and jets, can contaminate the signal sample

if a jet is incorrectly reconstructed as a lepton, or if an event contains a lepton from the decay of b or c hadrons. These are grouped into the "Non-W/Z" category. The yields from tW and WV events are estimated from simulation, while the estimations of Z/γ^* and Non-W/Z backgrounds use data driven techniques in the $e^{\pm}\mu^{\mp}$ channel and are extracted from simulation in the $\mu^{\pm}\mu^{\mp}$ channel.

A scale factor for the Z/γ^* background normalization is estimated, as in Ref. [43], from the number of events within the Z boson mass window in data and extrapolated to the number of events outside the Z boson mass window. A scale factor of $0.91 \pm 0.02(\text{stat})$ is obtained in the $e^{\pm}\mu^{\mp}$ channel, while it is found consistent with unity in the $\mu^{\pm}\mu^{\mp}$ channel. The estimation is performed using events with at least two jets and the dependence on different jet multiplicities is discussed in Section 6.

The Non-W/Z background is estimated from an extrapolation from a control region of same-sign (SS) dilepton events to the signal region of opposite-sign (OS) dileptons. The SS control region is defined using the same criteria as for the nominal signal region, except requiring dilepton pairs of the same charge. In the $e^{\pm}\mu^{\mp}$ channel, the muon isolation requirement is relaxed in order to enhance statistics. The SS dilepton events predominantly contain at least one misidentified lepton. Other SM processes with real SS pairs, such as Z/γ^*, tW, WV, and $t\bar{t}$ dilepton production give significantly smaller contributions to this category of events, and are estimated using simulation. The charge misidentification is negligible and thus ignored in events with prompt leptons. The scaling from the SS control region in data to the signal region is performed using an extrapolation factor, extracted from MC simulation, given by the ratio of the number of OS events with misidentified leptons to the number of SS events with misidentified leptons. The resulting estimate for the non-W/Z background is $1.0 \pm 0.9(\text{stat})$ events in the $e^{\pm}\mu^{\mp}$ channel, where the central value comes from the estimation using events with at least two jets. No particular dependence of this scale factor was observed for different jet multiplicities within the large statistical uncertainty.

6 Systematic uncertainties

The luminosity normalization has been independently estimated offline from the Pixel Cluster Counting method. The estimation takes into account integration and normalization uncertainties adding up to a total of $\pm 2.3\%$ [18].

The uncertainties on the electron trigger, identification and isolation efficiencies are estimated by changing the values of the data-to-simulation scale factors by one standard deviation ($\pm 1\sigma$), as obtained from the "tag-and-probe" method [39]. The uncertainty on the muon identification and isolation efficiency, including the trigger efficiency, is 3% and covers the full deviation of the scale factor from unity.

The impact of the uncertainty in the jet energy scale (JES) is estimated by changing the p_{T}- and η-dependent JES corrections by a flat factor of 2.8%. The uncertainty in jet energy resolution (JER) is estimated through η-dependent changes in the JER corrections. The uncertainty due to the use of $p_{\mathrm{T}}^{\mathrm{miss}}$ in the $\mu^{\pm}\mu^{\mp}$ channel is dominated by the unclusterd part of $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Last, a 30% uncertainty is assigned on the efficiency to tag non b-jets in the ℓ+jets analysis.

Theory uncertainties on $t\bar{t}$ production involve the systematic bias related to the missing higher-order diagrams in POWHEG, which are estimated through studies of the signal acceptance by changing the renormalization (μ_{R}) and factorization (μ_{F}) scales in POWHEG within the range $[\mu/2, 2\mu]$ ($\mu = \mu_{\mathrm{R}} = \mu_{\mathrm{F}}$). In the ℓ+jets analysis the impact of $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scales are examined

independently, while in the dilepton analysis they are varied simultaneously. In both analyses, these variations are applied independently at matrix-element (ME) and parton-shower (PS) level. The uncertainty arising from the hadronization model mainly affects the JES and the fragmentation of jets. The hadronization uncertainty is determined by comparing samples of events generated with POWHEG, where the hadronization is either modeled with PYTHIA or HERWIG++. This also includes differences in the parton shower model and the underlying event. The uncertainty from the choice of PDF is determined by reweighting the sample of simulated $t\bar{t}$ events according to the RMS of the NNPDF3.0 replica set [14]. Two extra variations of α_S are added in quadrature, to determine the total PDF uncertainty.

Furthermore, in the ℓ+jets analysis the independent variation of the μ_R/μ_F scales in the W+jets simulation is taken into account. Due to limited statistics in the simulation, an additional bin-by-bin uncertainty is assigned by generating an alternative shape to fit, where the bin prediction is varied ($\pm 1\sigma$), keeping all the other bins at their nominal expectation. The uncertainty assigned to the QCD multijet background includes the statistical uncertainty of the data, the uncertainty from the non-QCD multijet contributions subtractracted from the control region (Section 5.1) and an additional 30% normalization uncertainty. Finally, the theory uncertainty on the cross sections for the backgrounds are included and 30% is assigned given that tW, Z/γ^* and multibosons produce several jets and b-jets in the final state under study.

In the dilepton channel, an uncertainty of 30% is assumed for the cross sections of tW and WV backgrounds, to cover theory uncertainties and limited statistics in the simulation. The uncertainty on the Z/γ^* estimation is calculated from the statistical uncertainty and an additional 30% from the variation of the scale factor in the different levels of selection, resulting in uncertainties of 34% and 80% in the $e^{\pm}\mu^{\mp}$ and $\mu^{\pm}\mu^{\mp}$ channel, respectively. The systematic uncertainty in the estimated Non-W/Z background is given mainly by the statistical uncertainty of the method used. A value of 90% is estimated and 100% is assigned in the $e^{\pm}\mu^{\mp}$ and $\mu^{\pm}\mu^{\mp}$ channel, respectively.

7 Measurement of the $t\bar{t}$ cross section

7.1 ℓ+jets final state

In the ℓ+jets analysis, the $t\bar{t}$ cross section is measured in the visible phase space by means of a fit. Two variables are independently considered for the fit, which are sensitive to the resonant-behavior of the light jets produced in the W boson decay in a $t\bar{t}$ event. Given that these light jets, here denoted by j and j', are correlated at production, they are also expected to be closer in phase space, when compared to extra jets in the event. The distance ΔR can thus be used as a metric to rank all the non-b-tagged jets in the event, maximizing the probability of selecting the jets from the W boson decay in cases where more than two non b-tagged jets are found. From simulation we expect that the signal peaks at low ΔR, while the background is uniformly distributed up to $\Delta R \approx \pi$. Above that value less events are expected and background processes are predicted to dominate. The $M(j, j')$ variable also has distinctive, peaking, properties for the signal in contrast with a non-peaking background continuum. From simulation we expect that the $\min \Delta R(j, j')$ variable is robust against signal modeling uncertainties such as the choice of the μ_R/μ_F scale and jet energy scale and resolution, while the $M(j, j')$ variable tends to be more affected by such uncertainties. Consequently, the $\min \Delta R(j, j')$ variable has been chosen to extract the $t\bar{t}$ cross section.

In order to maximize the sensitivity of the analysis, the $\min \Delta R(j, j')$ distributions are simultaneously fitted in 0b-, 1b- and $\geq$2b-jet categories. The fit is separately performed in the μ+jets

and e+jets channel and the output is further combined in the ℓ+jets channel. The number of events expected and observed in each category prior to the fit is summarized in Table 2. Fair agreement is observed between data and expectations.

Table 2: Expected and observed event yields in the different categories used in the ℓ+jets analysis, prior to the fit. With the exception of the QCD multijet estimate, for which the total uncertainty is reported, the uncertainties reflect the limited statistics in the simulations.

Process	Category 0b		1b		≥2b	
	e	μ	e	μ	e	μ
tW	3.03±0.02	5.6±0.03	2.49±0.02	4.5±0.03	0.39±0.01	0.67±0.01
W+jets	776±17	1704±26	13±2	26±3	0.2±0.3	0.8±0.6
Z/γ^*	136±4	162±5	1.7±0.5	2.8±0.6	0.1±0.1	0.1±0.1
WV	0.52±0.01	1.01±0.02	<0.01	< 0.02	< 0.01	< 0.01
QCD multijet	442±132	493±148	3.6±1.1	28±8	2.5±0.8	2.0±0.8
$t\bar{t}$ signal	22.8±0.3	42.3±0.4	36.9±0.4	71.1±0.5	13.8±0.2	27.0±0.3
Total	1381±133	2408±150	57.7±2.4	131±9	16.8±0.9	31±1
Observed Data	1375	2406	61	129	19	33

The distribution for the $M(j,j')$ variable is shown in Fig. 1, while the distribution for the min $\Delta R(j,j')$ variable is shown in Fig. 2. The distributions have been combined for the μ+jets and e+jets channel to maximize the statistics and are shown for events with different b-tagged jet multiplicities. Fair agreement is observed between data and pre-fit expectations.

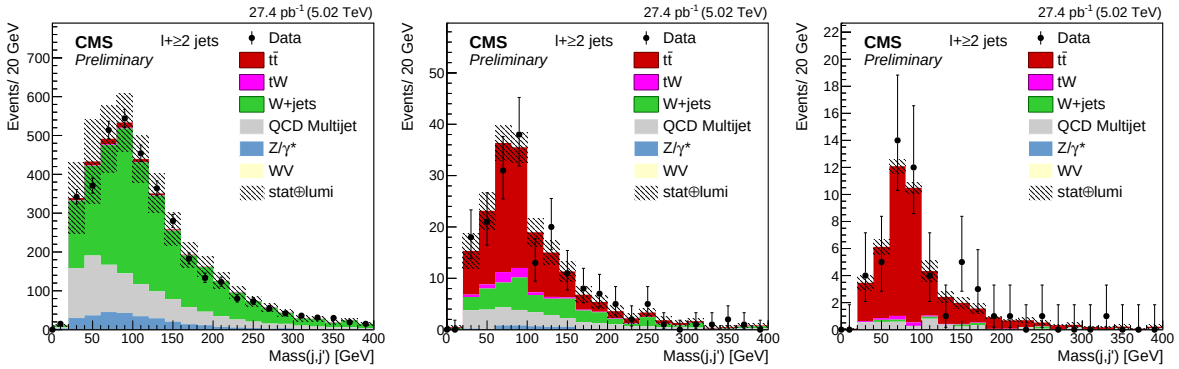


Figure 1: Distribution of the $M(j,j')$ variable for ℓ+jets events in 0b- (left), 1b- (center) and ≥2b- (right) jet categories. The distributions observed in the data are compared to the sum of the expectations for the signal and backgrounds prior to any fit. The QCD multijet background is estimated from data (cf. Section 5.1). The shaded band represents the statistical and integrated luminosity uncertainties on the expected signal and background yields.

A profile likelihood ratio (PLR) method, similar to the one employed in Ref. [44], is used to perform the fit. In our analysis we consider, in addition, the b-tagging efficiency scale factor (SF_b) as a parameter of interest of the fit. The PLR is therefore written as:

$$\lambda(\mu, SF_b) = \frac{\mathcal{L}(\mu, SF_b, \hat{\hat{\Theta}})}{\mathcal{L}(\hat{\mu}, \hat{SF}_b, \hat{\Theta})} \tag{2}$$

where $\mu = \sigma/\sigma_{th}$ is the signal strength (ratio of the observed $t\bar{t}$ cross section to the expectations from theory) and Θ is a set of nuisance parameters which encode the effect on the expectations due to variations of the sources of the systematic uncertainties described in Section 6.

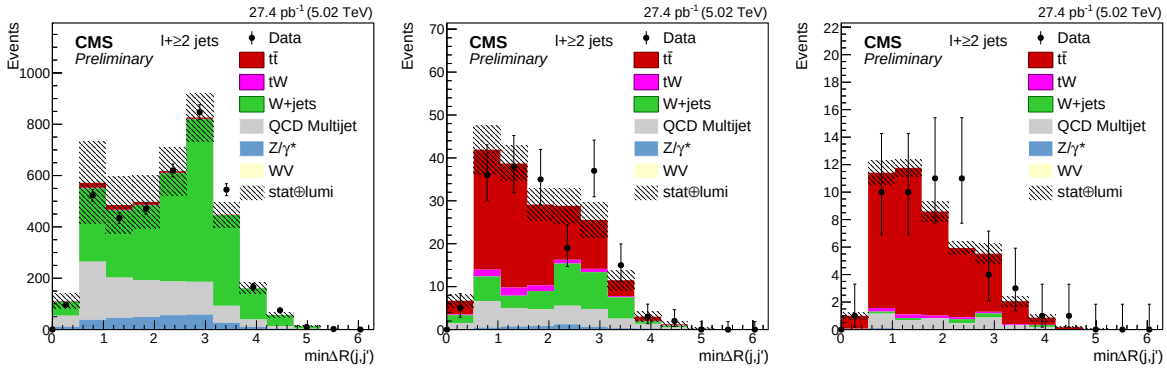


Figure 2: Distribution of the $\min \Delta R(j, j')$ variable for ℓ+jets events in 0b- (left), 1b- (center) and $\geq$2b- (right) jet categories. The distributions observed in the data are compared to the sum of the expectations for the signal and backgrounds prior to any fit. The QCD multijet background is estimated from data (cf. Section 5.1). The shaded band represents the statistical and integrated luminosity uncertainties on the expected signal and background yields.

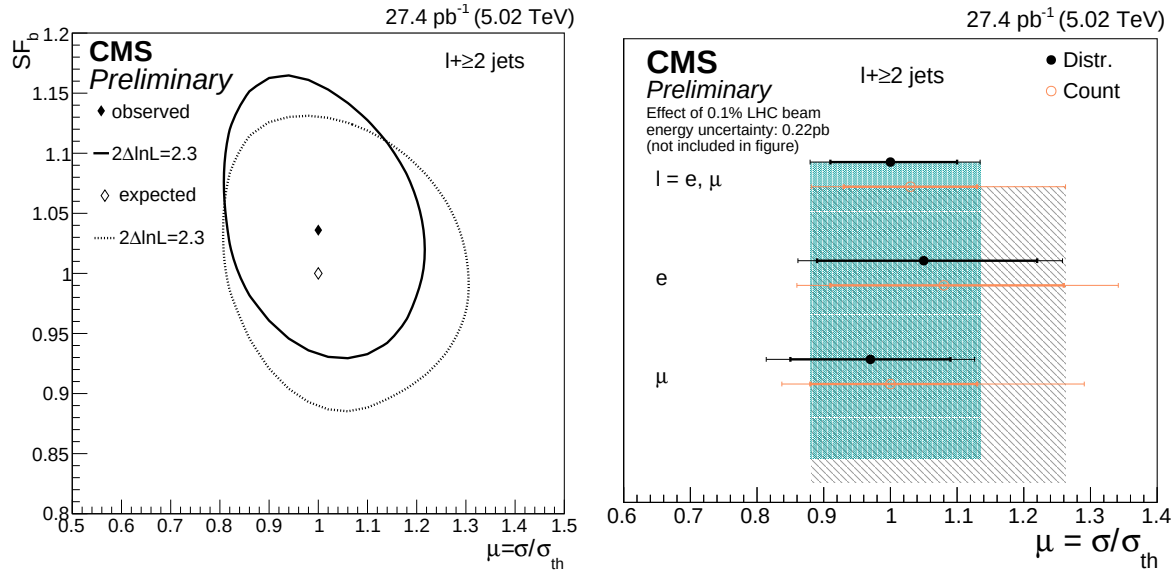
Figure 7.1 (left) shows the contours at the 68% confidence level (CL) obtained from the scan of the profile likelihood as function of μ and SF_b. The expected results are compared to the observed results and found to be in agreement within 1σ. The signal strength is obtained after profiling the b-tagging scale factor and the result is $\mu = 1.00^{+0.10}_{-0.09}\,(\text{stat})^{+0.09}_{-0.08}\,(\text{syst})$. As a cross check, the signal strength is also extracted by fitting only the total number of events observed in each of the six categories. The observed value $\mu = 1.03^{+0.10}_{-0.10}\,(\text{stat})^{+0.21}_{-0.11}\,(\text{syst})$ is in agreement with the analysis of the $\min \Delta R(j, j')$ distributions. Figure 7.1 (right) summarizes the results obtained for the signal strength fit in each channel separately from the analysis of the distributions or from event counting. In both cases the uncertainties are dominated by systematic uncertainties but with a large contribution from statistical uncertainties. In the ℓ+jets combination, the μ+jets channel is expected and observed to carry the largest weight.

In order to estimate the impact of the uncertainties on the measured signal strength, the fit is repeated after fixing one nuisance at a time to its postfit value within the postfit uncertainty ($\pm 1\sigma$). The impact on the signal strength fit is then evaluated from the difference induced in the final result from this procedure. By repeating the fits, the effect of some nuisance parameters being fixed may be reabsorbed by a variation of the ones being profiled, owing to correlations. As such, the systematic uncertainties obtained and summarized in Table 3 can only be interpreted as the observed post-fit values, and not as an absolute, orthogonalized, breakdown of uncertainties. Extra uncertainties affecting the estimation of the acceptance may be induced by the choice of the μ_R/μ_F scale at ME and PS level, PDF and hadronization model of $t\bar{t}$ signal, and thus need to be accounted for. The combined acceptance in the μ+jets and e+jets channel is estimated to be $A = 0.3014 \pm 0.0073$ with the uncertainty being dominated by the variation of the μ_R/μ_F scale (PS) and the hadronization model of $t\bar{t}$ signal. With respect to the cross-check analysis, the analysis of the distributions is less prone to the uncertainties in the QCD multijet background, jet energy resolution, and signal modeling. In both cases, the signal modeling uncertainties and the b-tagging efficiency are the leading uncertainties.

After correcting for the acceptance, the total cross section is measured to be

$$\sigma_{\text{tot}}(\text{pp} \to \text{t}\bar{\text{t}}) = 68.9 \pm 6.5\,(\text{stat}) \pm 6.1\,(\text{syst}) \pm 1.6(\text{lumi})\ \text{pb} \;=\; 68.9 \;\pm\; 9.1\ (\text{total})\ \text{pb},$$

attaining a total 13% relative uncertainty, in agreement with the SM prediction.



(a) (b)

Figure 3: Left: the 68% CL contour obtained from the scan of the likelihood, as function of the signal strength and the b-tagging scale factor in the ℓ+jets analysis. The solid (dashed) contours represent the contours observed in data (expected from simulation). The solid (hollow) point represents the observed fit result (SM expectation). Right: summary of the signal strengths separately obtained in the μ+jets and the e+jets channel, and after their combination in the ℓ+jets channel. The results of the analysis from the distributions (Distr.) are compared to those of a cross-check analysis based on event counting (Count). The inner (outer) bars correspond to the statistical (total) uncertainty of the signal strengths. Bands represent the total uncertainty of the signal strengths in the ℓ+jets channel.

Table 3: Estimated impact of each source of uncertainty on the analysis of distributions and on the cross-check from event counting. "Other backgrounds" contains the residual contribution from Z/γ^*, tW and WV events. The total uncertainty is obtained by adding in quadrature statistical, systematic and theory uncertainties. The values quoted have been symmetrized.

Uncertainty	$\Delta\mu/\mu$	
	Distr.	Count
Statistical uncertainty	0.095	0.100
Systematic uncertainty	0.085	0.160
Experimental uncertainties		
W+jets background	0.035	0.025
QCD multijet background	0.024	0.044
"Other backgrounds"	0.013	0.013
Jet energy scale	0.030	0.031
Jet energy resolution	0.006	0.023
b-tagging efficiency	0.034	0.045
Electron efficiencies	0.011	0.028
Muon efficiencies	0.017	0.022
Theory uncertainties		
Hadronization model of $t\bar{t}$ signal	0.028	0.069
μ_R/μ_F scale of $t\bar{t}$ signal (PS)	0.044	0.115
μ_R/μ_F scale of $t\bar{t}$ signal (ME)	<0.01	<0.01
Total	0.127	0.189

7.2 Dilepton final state

Figure 4 shows the multiplicity of jets (left) and the scalar p_T sum of all jets (H_T) (right), for events passing the dilepton pair criteria in the $e^{\pm}\mu^{\mp}$ channel. Figure 5 shows the invariant mass (left) and p_T (right) of the dilepton pair, after requiring at least two jets in the event in the $e^{\pm}\mu^{\mp}$ channel. Figure 6 shows the missing transverse momentum (left) and invariant mass of the dilepton pair (right) in the $\mu^{\pm}\mu^{\mp}$ channel for events passing the dilepton criteria and Z veto, and after the missing transverse momentum requirement, respectively. Predictions take into account the efficiency corrections described in Section 4 and the background estimations discussed in Section 5.2. Agreement is observed between data and the predictions for signal and background.

The $t\bar{t}$ production cross section is measured by counting events and applying the expression

$$\sigma_{t\bar{t}} = \frac{N - N_B}{\mathcal{BR} \cdot \varepsilon \cdot \mathcal{A} \cdot \mathcal{L}}, \tag{3}$$

where N is the total number of dilepton events observed in data, N_B is the number of estimated background events, $\mathcal{A}$ is the acceptance, ε the selection efficiency, $\mathcal{BR}$ the branching fraction into the $e^{\pm}\mu^{\mp}$ or $\mu^{\pm}\mu^{\mp}$ final state, and $\mathcal{L}$ is the integrated luminosity.

Table 4 shows the total number of events observed in data, together with the total number of signal and background events expected from simulation or estimated from data, after the full set of selection criteria (Table 1). The detector, trigger and reconstruction efficiency, as estimated from data, amounts to $\varepsilon = 0.57 \pm 0.02$ (0.45 ± 0.02), while the acceptance, as estimated from MC, is found to be $\mathcal{A} = 0.52 \pm 0.01$ (0.47 ± 0.01) in the $e^{\pm}\mu^{\mp}$ ($\mu^{\pm}\mu^{\mp}$) channel. The statistical uncertainty (from MC) is included in the uncertainty of $\mathcal{A}$.

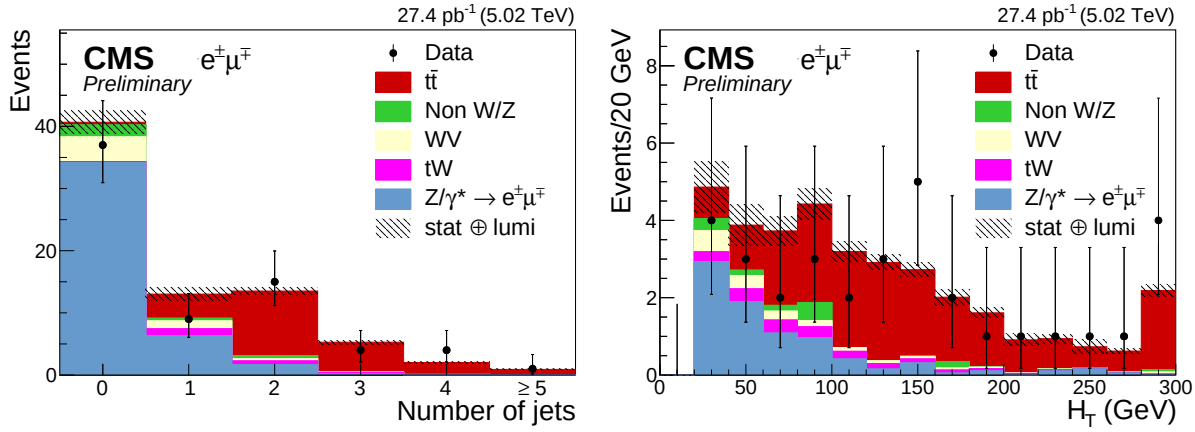


Figure 4: Distributions of the jet multiplicity (left), and scalar p_{T} sum of all jets (right) in events passing the dilepton pair criteria in the $\mathrm{e}^{\pm}\mu^{\mp}$ channel. The Z/γ^{*} and Non-W/Z backgrounds are determined from data (cf. Section 5.2). The shaded band represents the statistical and integrated luminosity uncertainties on the expected signal and background yields. The last bin of the distributions contains the overflow events.

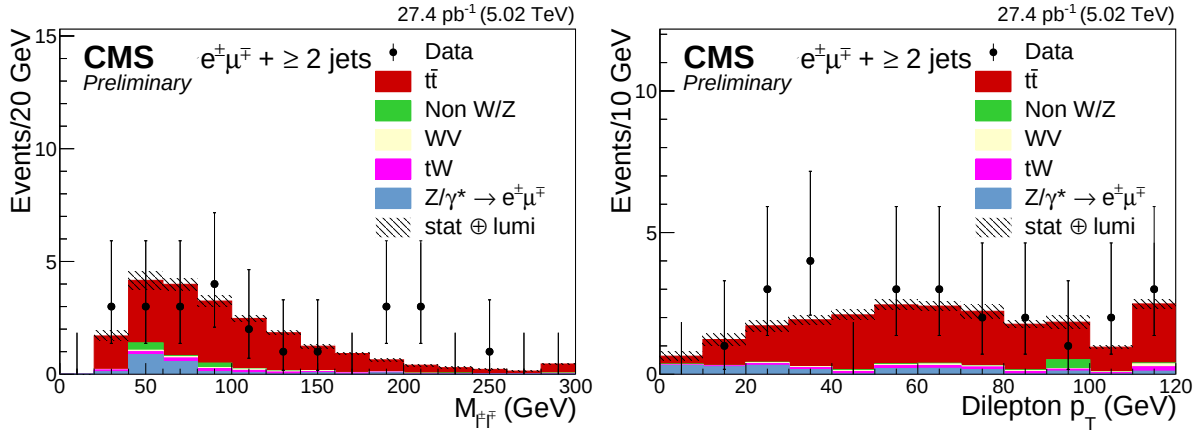


Figure 5: Distributions of the invariant mass (left) and p_{T} (right) of the dilepton pair after requiring at least two jets in the $\mathrm{e}^{\pm}\mu^{\mp}$ channel. The Z/γ^{*} and Non-W/Z backgrounds are determined from data (cf. Section 5.2). The shaded band represents the statistical and integrated luminosity uncertainties on the expected signal and background yields. The last bin of the distributions contains the overflow events.

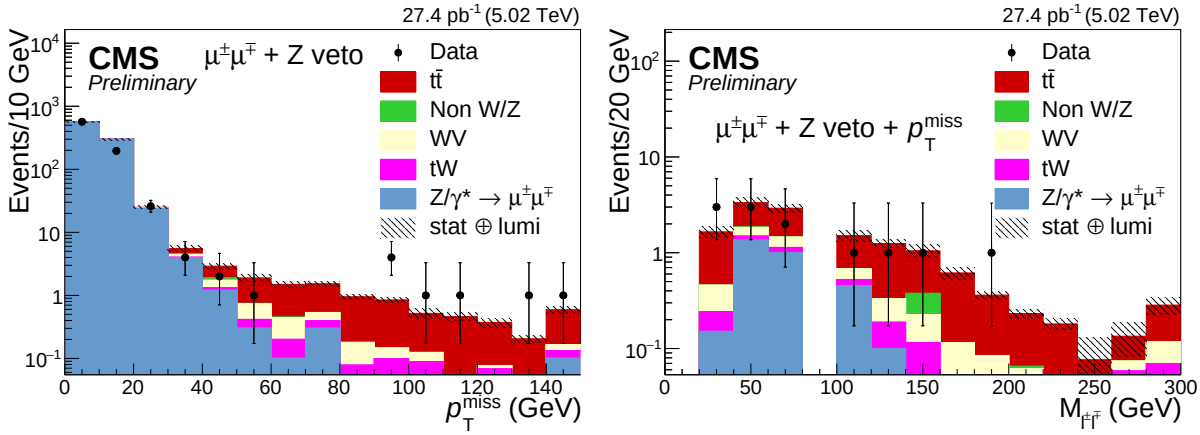


Figure 6: Distributions of the missing transverse momentum (left) in events passing the dilepton criteria and Z veto, and invariant mass (right) of the dilepton pair after the missing transverse momentum requirement in the $\mu^{\pm}\mu^{\mp}$ channel. The shaded band represents the statistical and integrated luminosity uncertainties on the expected signal and background yields. The last bin of the distributions contains the overflow events.

Table 4: Number of dilepton events obtained after applying the full selection. The results are given for the individual sources of background, $t\bar{t}$ signal, and data. The uncertainties correspond to statistical and systematic components.

Source	$e^{\pm}\mu^{\mp}$	$\mu^{\pm}\mu^{\mp}$
tW	0.92 ± 0.02	0.29 ± 0.01
Non-W/Z leptons	1.0 ± 0.9	0.04 ± 0.01
Z/γ^*	1.6 ± 0.4	1.05 ± 0.37
WV	0.44 ± 0.02	0.15 ± 0.01
$t\bar{t}$ signal	18.0 ± 0.3	6.36 ± 0.16
Total	22.0 ± 0.9	7.9 ± 0.4
Observed Data	24	7

Table 5 summarizes the magnitude of the statistical and systematic uncertainties from different sources contributing to the $t\bar{t}$ production cross section. All sources of uncertainties are added in quadrature to yield the total uncertainty.

Using the definitions above, the yields from Table 4, and the systematic uncertainties from Table 5, the inclusive cross section is measured to be:

$$\sigma_{\text{tot}}(\text{pp} \to \text{t}\bar{\text{t}}) = 76.5 \pm 18.7\,(\text{stat}) \pm 4.4\,(\text{syst}) \pm 1.8\,(\text{lumi})\ \text{pb} \;=\; 76.5 \;\pm\; 19.3\,(\text{total})\ \text{pb}$$

in the $e^{\pm}\mu^{\mp}$ and

$$\sigma_{\text{tot}}(\text{pp} \to \text{t}\bar{\text{t}}) = 59.2 \pm 28.7\,(\text{stat}) \pm 10.6\,(\text{syst}) \pm 1.4\,(\text{lumi})\ \text{pb} \;=\; 59.2 \;\pm\; 30.6\,(\text{total})\ \text{pb}$$

in the $\mu^{\pm}\mu^{\mp}$ channel, respectively. The separate total systematic uncertainty without integrated luminosity, the part attributed to the integrated luminosity, and the statistical contribution are added in quadrature to obtain the total uncertainty. The cross sections, measured with a relative uncertainty of 25% and 52%, respectively, are in agreement with the SM prediction (Eq. 1) within the large uncertainty of the measurement.

Table 5: Summary of individual contributions to the systematic uncertainty on the $\sigma_{\mathrm{t\bar{t}}}$ measurement in the dilepton channel. The absolute uncertainties $\Delta\sigma_{\mathrm{t\bar{t}}}$ in pb as well as relative uncertainties ($\Delta\sigma_{\mathrm{t\bar{t}}}/\sigma_{\mathrm{t\bar{t}}}$) are given.

	$\mathrm{e}^{\pm}\mu^{\mp}$		$\mu^{\pm}\mu^{\mp}$	
Source	$\Delta\sigma_{\mathrm{t\bar{t}}}$ (pb)	$\Delta\sigma_{\mathrm{t\bar{t}}}/\sigma_{\mathrm{t\bar{t}}}$ (%)	$\Delta\sigma_{\mathrm{t\bar{t}}}$ (pb)	$\Delta\sigma_{\mathrm{t\bar{t}}}/\sigma_{\mathrm{t\bar{t}}}$ (%)
Electron efficiencies	1.0	1.4	—	—
Muon efficiencies	2.3	3.0	3.6	6.1
Jet energy scale	1.0	1.3	0.7	1.3
Jet energy resolution	0.05	0.06	0.04	0.06
Missing transverse energy	—	—	0.4	0.7
$\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scale of $\mathrm{t\bar{t}}$ signal (PS)	0.9	1.2	1.0	1.7
$\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scale of $\mathrm{t\bar{t}}$ signal (ME)	0.1	0.2	0.6	1.1
Hadronization model of $\mathrm{t\bar{t}}$ signal	0.9	1.2	3.1	5.2
PDF	0.4	0.5	0.2	0.4
MC statistics	1.1	1.4	1.4	2.4
tW background	1.1	1.4	0.9	1.6
WV background	0.5	0.7	0.5	0.9
Z/γ^{*} background	2.1	2.7	9.1	15.4
Non W/Z background	1.9	2.5	0.4	0.7
Total systematic (w/o luminosity)	4.4	5.8	10.6	17.9
Integrated luminosity	1.8	2.3	1.4	2.3
Statistical uncertainty	18.7	24.5	28.7	48.4
Total	19.3	25.2	30.6	51.7

7.3 Combination

The individual results are combined by using the BLUE method [45]. All systematic uncertainties are considered as fully correlated across all channels, with the following exceptions: the uncertainty associated to the limited statistics in the MC is taken as uncorrelated; the electron identification is not relevant for the $\mu\mu$ channel; the b tagging and QCD multijet uncertainties are only considered for the ℓ+jets channel; also in the ℓ+jets channel, the WV and Z/γ^{*} backgrounds are not considered separately.

The inclusive cross section is measured to be:

$$\sigma_{\mathrm{tot}}(\mathrm{pp} \to \mathrm{t\bar{t}}) \;=\; 69.5 \pm 6.1\,(\mathrm{stat}) \pm 5.6\,(\mathrm{syst}) \pm 1.6\,(\mathrm{lumi})\ \mathrm{pb} \;=\; 69.5 \;\pm\; 8.4\,(\mathrm{total})\ \mathrm{pb}.$$

The weights of the individual measurements, to be understood in the sense of Ref. [46], are 81.8% for ℓ+jets, 13.5% for $e\mu$ and 4.7% for $\mu\mu$ channel.

The robustness of the combination is tested by performing an iterative variant of the BLUE method [47] and varying some assumptions on the correlations of a series of systematics. The post-fit correlations between the nuisance parameters in the ℓ+jets channel have been checked and found to introduce negligible impact.

Figure 7 presents a summary of results for $\sigma_{\mathrm{t\bar{t}}}$ at various center-of-mass energies, including the combination of the Tevatron measurements at 1.96 TeV [48] and the most precise CMS measurements at $\sqrt{s}$ = 7 and 8 TeV [1], and 13 TeV [7, 44], compared to the NNLO+NNLL predictions as a function of $\sqrt{s}$ for $\mathrm{p\bar{p}}$ and pp collisions [31] for the NNPDF3.0 [49] PDF set. Our result is also compared to the predictions for the MMHT14 [50], CT14 [51], and ABM12 [52] PDF sets.

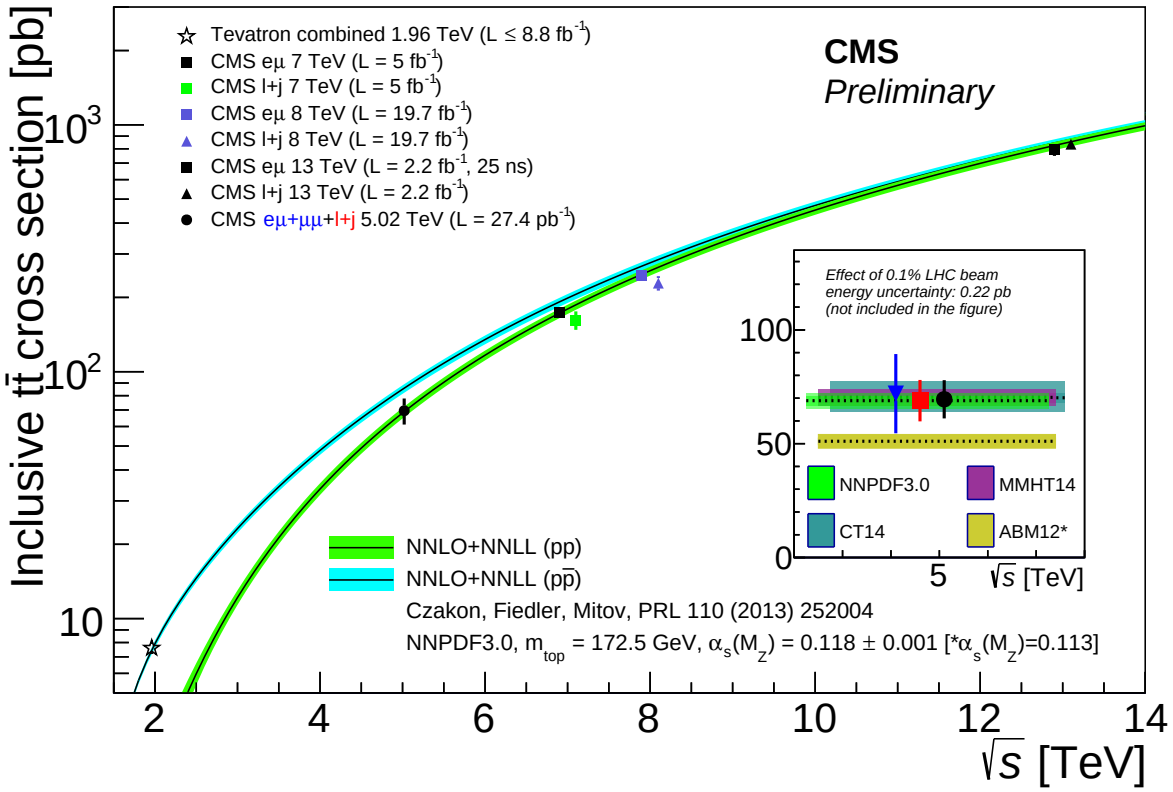


Figure 7: Top quark pair production cross section in $p\bar{p}$ and pp collisions as a function of the center-of-mass energy; the Tevatron combination at $\sqrt{s} = 1.96$ TeV is displayed, as well as CMS results at 7, 8 and 13 TeV in the dilepton and ℓ+jets channel. The measurements are compared to the NNLO+NNLL theory predictions.

8 QCD analysis

To illustrate the impact of the the $\mathrm{t\bar{t}}$ production cross section measurements at $\sqrt{s} = 5.02$ TeV on the PDFs, these data are used in a QCD analysis at NNLO together with the combined measurements of neutral- and charged-current cross sections of deep inelastic electron[1]-proton scattering (DIS) at HERA [53] and the CMS measurement of the muon charge asymmetry in W boson production [54]. The latter data set is used in order to improve the constraint on the light quark distributions.

The version 2.0.0 of the open-source QCD fit framework for PDF determination xFitter [55, 56] is used with the partons evolved by using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi equations [57–62] at NNLO, as implemented in the QCDNUM 17-01/13 program [63]. The treatment and the choices for the central values and variations of c and b quark masses, strong coupling constant and the fraction of the strange content of the proton follow that of earlier CMS analyses e.g. Ref. [54]. The scales μ_r and μ_f are set to Q, which denotes the four-momentum transfer in case of the DIS data, the mass of the W boson in case of the muon charge asymmetry, and the mass of the top quark in case of top-quark production, respectively.

The systematic uncertainties for all the three channels of top-quark pair production and their correlations are treated the same way as in the combination, as described in Section 7.3. The theoretical predictions for the top-quark pair production are obtained at NNLO by using the Hathor calculation [64], assuming the top-quark mass of $m_t = 172.5$ GeV. The bin-to-bin correlations of the experimental uncertainties for the muon charge asymmetry and for the inclusive DIS cross sections are taken into account. The theoretical predictions for the muon charge asymmetry are obtained as described in Ref. [54].

The procedure for the determination of the PDFs follows the approach used in the QCD analysis Ref. [54] and results in a 14-parameter fit. The parameterized PDFs are the gluon distribution, $x\mathrm{g}$, the valence-quark distributions, $x\mathrm{u}_v$, $x\mathrm{d}_v$, and the u-type and d-type anti-quark distributions, $x\overline{U}$, $x\overline{D}$. The relations $x\overline{U} = x\overline{\mathrm{u}}$ and $x\overline{D} = x\overline{\mathrm{d}} + x\overline{\mathrm{s}}$ are assumed at the initial scale of the QCD evolution $Q_0^2 = 1.9$ GeV^2:

$$
\begin{aligned}
x\mathrm{g}(x) &= A_{\mathrm{g}} x^{B_{\mathrm{g}}} \cdot (1-x)^{C_{\mathrm{g}}} \cdot (1 + D_{\mathrm{g}} x), && (4)\\
x\mathrm{u}_v(x) &= A_{\mathrm{u}_v} x^{B_{\mathrm{u}_v}} \cdot (1-x)^{C_{\mathrm{u}_v}} \cdot (1 + D_{\mathrm{u}_v} x + E_{\mathrm{u}_v} x^2), && (5)\\
x\mathrm{d}_v(x) &= A_{\mathrm{d}_v} x^{B_{\mathrm{d}_v}} \cdot (1-x)^{C_{\mathrm{d}_v}}, && (6)\\
x\overline{U}(x) &= A_{\overline{U}} x^{B_{\overline{U}}} \cdot (1-x)^{C_{\overline{U}}} \cdot (1 + E_{\overline{U}} x^2), && (7)\\
x\overline{D}(x) &= A_{\overline{D}} x^{B_{\overline{D}}} \cdot (1-x)^{C_{\overline{D}}}. && (8)
\end{aligned}
$$

The normalization parameters $A_{\mathrm{u_v}}$, $A_{\mathrm{d_v}}$, and A_{g} are determined by the QCD sum rules, the B parameters are responsible for small-x behavior of the PDFs, and the parameters C describe the shape of the distribution as $x \to 1$. Additional constraints $B_{\overline{\mathrm{U}}} = B_{\overline{\mathrm{D}}}$ and $A_{\overline{\mathrm{U}}} = A_{\overline{\mathrm{D}}}(1 - f_{\mathrm{s}})$ are imposed, with f_{s} being the strangeness fraction, $f_{\mathrm{s}} = \bar{\mathrm{s}}/(\bar{\mathrm{d}} + \bar{\mathrm{s}})$, which is fixed to $f_{\mathrm{s}} = 0.31 \pm 0.08$ as in Ref. [65], consistent with the determination of the strangeness fraction by using the CMS measurements of W + c production as in Ref. [66].

Using the top-quark production cross sections allows the release of an additional parameter in Eq. 5, as compared to the analysis in Ref. [54]. The global and partial χ^2 values for each data set are listed in Table 6, where the χ^2 values illustrate a general agreement among all the data

[1] If not specified otherwise, "electron" indicates both, electron and positron in the context of HERA data.

sets. The somewhat high $\chi^2/n_{\rm dof}$ values for the combined DIS data are very similar to those observed in Ref. [53], where they are investigated in detail.

Table 6: Partial χ^2 per number of data points, $n_{\rm dp}$, and the global χ^2 per degrees of freedom, $n_{\rm dof}$, as obtained in the QCD analysis of DIS data, the CMS muon charge asymmetry and the inclusive cross sections of top-quark pair production at $\sqrt{s}$=5.02 TeV. For HERA measurements, the energy of the proton beam is listed for each data set, with electron energy being $E_{\rm e} = 27.5$ GeV.

Data sets	Partial $\chi^2/n_{\rm dp}$
HERA1+2 neutral current, e^+p, $E_{\rm p} = 920$ GeV	449/377
HERA1+2 neutral current, e^+p, $E_{\rm p} = 820$ GeV	71/70
HERA1+2 neutral current, e^+p, $E_{\rm p} = 575$ GeV	224/254
HERA1+2 neutral current, e^+p, $E_{\rm p} = 460$ GeV	218/204
HERA1+2 neutral current, e^-p, $E_{\rm p} = 920$ GeV	218/159
HERA1+2 charged current, e^+p, $E_{\rm p} = 920$ GeV	43/39
HERA1+2 charged current, e^-p, $E_{\rm p} = 920$ GeV	53/42
CMS $W^\pm$ muon charge asymmetry $\mathcal{A}(\eta_\mu)$, $\sqrt{s} = 8$ TeV	2.4/11
CMS inclusive $t\bar{t}$ 5.02 TeV, $e^\pm\mu^\mp$	1.03/1
CMS inclusive $t\bar{t}$ 5.02 TeV, $\mu^\pm\mu^\mp$	0.01/1
CMS inclusive $t\bar{t}$ 5.02 TeV, ℓ+jets	0.70/1
Correlated χ^2	100
Global $\chi^2/n_{\rm dof}$	1387/1145

The experimental uncertainties of the measurements are propagated to the extracted QCD fit parameters using the Monte-Carlo (MC) method [67, 68]. In this method, 400 replicas of pseudo-data are generated, with cross sections allowed to vary within the statistical and systematic uncertainties. For each of them, the PDF fit is performed and the uncertainty is estimated as the RMS around the central value. In Fig. 8, the gluon distribution with its relative uncertainties is shown, as obtained in the QCD analyses with and without top-quark production cross sections. Moderate reduction of the uncertainty in the gluon distribution at high x is observed, once the top-quark pair production cross sections are included into the fit. The uncertainties in the valence quark distributions remain unaffected. All changes of the central values of the PDFs are well within the fit uncertainties.

In addition, possible effects of variations of the model input parameters and the initial PDF parametrization are investigated in the same way as in the similar analysis of Ref. [54]. These model and parametrization uncertainties are evaluated for the cases when top quark-pair production cross sections are included in or excluded from the fit, respectively, and no changes are observed.

In summary, the top-quark cross section measurements provide additional constraints on the gluon distribution at high x, however their impact is small due to large experimental uncertainties.

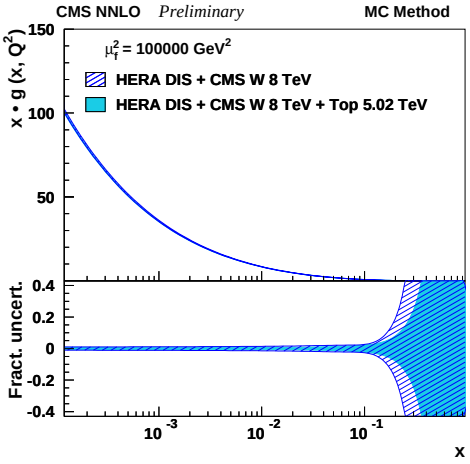


Figure 8: The gluon distribution as functions of x at the scale of 100000 GeV^2. The results of the fit including top-quark measurements (shaded band), and without those (hatched band) are compared. The fit uncertainties as obtained by using the MC method,are shown. In the bottom panel, the relative fractional uncertainties are presented.

9 Summary

In summary, the first measurement of the $\mathrm{t\bar{t}}$ production cross section in pp collisions at $\sqrt{s}$ = 5.02 TeV is presented for events with one or two leptons and at least two jets using a data sample corresponding to an integrated luminosity of 27.4 pb^{-1}. The final measurement is obtained as the combination of the measurements in the individual channels. The result is 69.5 pb, with a total relative uncertainty of 12%, which is consistent with the SM prediction. The impact of the measured $\mathrm{t\bar{t}}$ cross section on the proton PDFs is studied in a QCD analysis at NNLO and a moderate decrease of the uncertainty in the gluon distribution at high fractions x of the proton momentum carried by the gluon is observed.

References

[1] CMS Collaboration, "Measurement of the $\mathrm{t\bar{t}}$ production cross section in the eμ 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.

[2] CMS Collaboration, "Measurements of the $\mathrm{t\bar{t}}$ production cross section in lepton+jets final states in pp collisions at 8 TeV and ratio of 8 to 7 TeV cross sections", arXiv:1602.09024. *submitted to Eur. Phys. J. C*.

[3] ATLAS Collaboration, "Measurement of the top quark pair production cross-section with ATLAS in the single lepton channel", *Phys. Lett. B* **711** (2012) 244, doi:10.1016/j.physletb.2012.03.083, arXiv:1201.1889.

[4] 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.

[5] CMS Collaboration, "Measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. Lett.* **116** (2016), no. 5, 052002, doi:10.1103/PhysRevLett.116.052002, arXiv:1510.05302.

[6] CMS Collaboration, "Measurement of the inclusive and differential $\mathrm{t\bar{t}}$ production cross sections in lepton + jets final states at 13 TeV", CMS Physics Analysis Summary CMS-PAS-TOP-15-005, 2015.

[7] CMS Collaboration, "Measurement of the $\mathrm{t\bar{t}}$ production cross section using events in the eμ final state in pp collisions at $\sqrt{s} = 13$ TeV", arXiv:1611.04040.

[8] ATLAS Collaboration, "Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector", *Phys. Lett. B* **761** (2016) 136, doi:10.1016/j.physletb.2016.08.019, arXiv:1606.02699.

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

[10] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[11] 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.

[12] 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.

[13] 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.

[14] F. Demartin et al., "Impact of parton distribution function and α_S uncertainties on Higgs boson production in gluon fusion at hadron colliders", *Phys. Rev. D* **82** (2010) 014002, doi:10.1103/PhysRevD.82.014002, arXiv:1004.0962.

[15] D. d'Enterria, K. Krajczar, and H. Paukkunen, "Top-quark production in proton-nucleus and nucleus-nucleus collisions at LHC energies and beyond", *Phys. Lett.* **B746** (2015) 64–72, doi:10.1016/j.physletb.2015.04.044, arXiv:1501.05879.

[16] CMS Collaboration, "Projections for Heavy Ions with HL-LHC", CMS Physics Analysis Summary CMS-PAS-FTR-13-025, 2013.

[17] CMS Collaboration, "First measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-TOP-16-015, 2016.

[18] CMS Collaboration, "CMS Luminosity Calibration for the pp Reference Run at $\sqrt{s} = 5.02$ TeV", Technical Report CMS-PAS-LUM-16-001, CERN, Geneva, 2016.

[19] 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.

[20] 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.

[21] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[22] M. Bähr et al., "HERWIG++ physics and manual", *Eur. Phys. J. C* **58** (2008) 639, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.

[23] 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.

[24] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[25] 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].

[26] 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.

[27] N. Kidonakis, "Top Quark Production", in *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)*, p. 139. Verlag Deutsches Elektronen-Synchrotron, Hamburg, 2014. arXiv:1311.0283. doi:10.3204/DESY-PROC-2013-03/Kidonakis.

[28] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.* **205** (2010) 10, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492.

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

[30] 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.

[31] M. Czakon, P. Fiedler and A. Mitov, "Total top-quark pair-production cross section at hadron colliders through $O(\alpha_S^4)$", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[32] J. Wenninger and E. Todesco, "Large Hadron Collider momentum calibration and accuracy", Technical Report CERN-ACC-2017-0007, CERN, Geneva, Feb, 2017.

[33] 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.

[34] 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.

[35] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\,\mathrm{TeV}$", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[36] CMS Collaboration 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.

[37] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", *JINST* **8** (2013) P11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.

[38] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[39] CMS Collaboration, "Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.

[40] 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.

[41] 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.

[42] 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.

[43] CMS Collaboration, "First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett.* **B695** (2011) 424–443, doi:10.1016/j.physletb.2010.11.058, arXiv:1010.5994.

[44] CMS Collaboration, "A measurement of the inclusive $\mathrm{t\bar{t}}$ production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV, using events with one isolated charged lepton and at least one jet", Technical Report CMS-PAS-TOP-16-006, 2016.

[45] L. Lyons, D. Gibaut, and P. Clifford, "How to combine correlated estimates of a single physical quantity", *Nucl. Instrum. Meth.* **A270** (1988) 110, doi:10.1016/0168-9002(88)90018-6.

[46] A. Valassi and R. Chierici, "Information and treatment of unknown correlations in the combination of measurements using the BLUE method", *Eur. Phys. J.* **C74** (2014) 2717, doi:10.1140/epjc/s10052-014-2717-6, arXiv:1307.4003.

[47] L. Lista, "The bias of the unbiased estimator: a study of the iterative application of the BLUE method", *Nucl. Instrum. Meth.* **A764** (2014) 82–93, doi:10.1016/j.nima.2014.07.021, 10.1016/j.nima.2014.11.054, arXiv:1405.3425. [Erratum: Nucl. Instrum. Meth.A773,87(2015)].

[48] CDF and D0 Collaborations, "Combination of measurements of the top-quark pair production cross section from the Tevatron Collider", *Phys. Rev. D* **89** (2014) 072001, doi:10.1103/PhysRevD.89.072001, arXiv:1309.7570.

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

[50] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs", *Eur. Phys. J.* **C75** (2015), no. 5, 204, doi:10.1140/epjc/s10052-015-3397-6, arXiv:1412.3989.

[51] S. Dulat et al., "New parton distribution functions from a global analysis of quantum chromodynamics", *Phys. Rev.* **D93** (2016), no. 3, 033006, doi:10.1103/PhysRevD.93.033006, arXiv:1506.07443.

[52] S. Alekhin, J. Blumlein, and S. Moch, "The ABM parton distributions tuned to LHC data", *Phys. Rev.* **D89** (2014), no. 5, 054028, doi:10.1103/PhysRevD.89.054028, arXiv:1310.3059.

[53] ZEUS and H1 Collaboration, "Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data", *Eur. Phys. J. C* **75** (2015) 580, doi:10.1140/epjc/s10052-015-3710-4, arXiv:1506.06042.

[54] CMS Collaboration, "Measurement of the differential cross section and charge asymmetry for inclusive pp $\to W^{\pm} + X$ production at $\sqrt{s} = 8$ TeV", *Eur. Phys. J.* **C76** (2016), no. 8, 469, doi:10.1140/epjc/s10052-016-4293-4, arXiv:1603.01803.

[55] S. Alekhin et al., "HERAFitter, Open source QCD fit project", *Eur. Phys. J. C* **75** (2015) 304, doi:10.1140/epjc/s10052-015-3480-z, arXiv:1410.4412.

[56] XFitter Collaboration, http://www.xfitter.org.

[57] V. N. Gribov and L. N. Lipatov, "Deep inelastic ep scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438.

[58] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language", *Nucl. Phys. B* **126** (1977) 298, doi:10.1016/0550-3213(77)90384-4.

[59] G. Curci, W. Furmanski, and R. Petronzio, "Evolution of parton densities beyond leading order: The non-singlet case", *Nucl. Phys. B* **175** (1980) 27, doi:10.1016/0550-3213(80)90003-6.

[60] W. Furmanski and R. Petronzio, "Singlet parton densities beyond leading order", *Phys. Lett. B* **97** (1980) 437, doi:10.1016/0370-2693(80)90636-X.

[61] S. Moch, J. A. M. Vermaseren, and A. Vogt, "The three-loop splitting functions in QCD: the non-singlet case", *Nucl. Phys. B* **688** (2004) 101, doi:10.1016/j.nuclphysb.2004.03.030, arXiv:hep-ph/0403192.

[62] A. Vogt, S. Moch, and J. A. M. Vermaseren, "The three-loop splitting functions in QCD: the singlet case", *Nucl. Phys. B* **691** (2004) 129, doi:10.1016/j.nuclphysb.2004.04.024, arXiv:hep-ph/0404111.

[63] M. Botje, "QCDNUM: Fast QCD evolution and convolution", *Comput. Phys. Commun.* **182** (2011) 490, doi:10.1016/j.cpc.2010.10.020, arXiv:1005.1481.

[64] M. Aliev et al., "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR", *Comput. Phys. Commun.* **182** (2011) 1034–1046, doi:10.1016/j.cpc.2010.12.040, arXiv:1007.1327.

[65] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC", *Eur. Phys. J. C* **63** (2009) 189, doi:10.1140/epjc/s10052-009-1072-5, arXiv:0901.0002.

[66] CMS Collaboration, "Measurement of the muon charge asymmetry in inclusive $pp \to W + X$ production at $\sqrt{s} = 7$ TeV and an improved determination of light parton distribution functions", *Phys. Rev. D* **90** (2014) 032004, doi:10.1103/PhysRevD.90.032004, arXiv:1312.6283.

[67] W. T. Giele and S. Keller, "Implications of hadron collider observables on parton distribution function uncertainties", *Phys. Rev. D* **58** (1998) 094023, doi:10.1103/PhysRevD.58.094023, arXiv:hep-ph/9803393.

[68] W. T. Giele, S. A. Keller, and D. A. Kosower, "Parton distribution function uncertainties", (2001). arXiv:hep-ph/0104052.