Top-quark production at hadron colliders

Andrea Helen Knue∗
on behalf of the ATLAS, CDF, CMS, D0, and LHCb collaborations
Max-Planck-Institute for Physics, Munich
E-mail: aknue@cern.ch

Inclusive and differential measurements of the \( t\bar{t} \) production cross-section are presented including data from proton-antiproton (Tevatron collider) and proton-proton (LHC) collisions at various centre-of-mass energies. In addition, measurements of the \( t\bar{t} + \) vector boson production rates are presented, which were carried out at ATLAS and CMS. All results are in good agreement with theoretical predictions from Standard Model calculations.

Fourth Annual Large Hadron Collider Physics
13-18 June 2016
Lund, Sweden

∗Speaker.
1. Introduction

The top quark was discovered in 1995 [3, 4] by the CDF and D0 experiments at the Tevatron collider and is by far the heaviest quark with a mass close to the scale of electroweak symmetry breaking. It is hence supposed to have the largest coupling to the Higgs-boson. It furthermore has a lifetime which is shorter than the time neccessary to hadronise, and is therefore the only “bare” quark that can be studied. Precision measurements of the top quark production cross-sections and properties are vital to test the Standard Model (SM) predictions and search for new physics processes which could manifest themselves in increased or reduced production cross-sections. The Tevatron experiments measured the inclusive production cross-section at \( \sqrt{s} = 1.96 \) TeV to an accuracy close to the precision of the theoretical prediction. With the LHC now delivering top-quark pair events at unprecedented rates and energies, high-precision measurements are possible for inclusive and differential measurements, that allow to test the theoretical predictions in different phase spaces. Top-quark pairs are produced via the strong interaction by either \( gg \)-fusion or \( q\bar{q} \)-annihilation. At the LHC, the production mode via \( gg \) fusion strongly dominates. Top quarks decay almost exclusively into a \( W \)-boson and a \( b \)-quark. Therefore the decay of the \( W \)-boson defines the different analysis channels. If one \( W \)-boson decays into a charged lepton and a neutrino while the other \( W \) decays into two quarks, the channel is called \( \text{lepton}+\text{jets} \) in the following. If both \( W \)-bosons decay leptonically the channel is called \( \text{dilepton} \) channel and if both decay hadronically the final state is called \( \text{all-hadronic} \). The predicted \( \bar{t}t \) production cross-sections at \( \sqrt{s} = 8 \) TeV and \( 13 \) TeV are

\[
\sigma_{\bar{t}t} (8 \text{ TeV}) = 252.9^{+6.4}_{-8.6} \text{ (scale)} \pm 11.7 \text{ (PDF} + \alpha_s \text{) pb} \tag{1.1}
\]

and

\[
\sigma_{\bar{t}t} (13 \text{ TeV}) = 831.8^{+19.8}_{-29.2} \text{ (scale)} \pm 35.1 \text{ (PDF} + \alpha_s \text{) pb}. \tag{1.2}
\]

as calculated with the Top++2.0 program to next-to-next-to-leading order (NNLO) in perturbative QCD, including soft-gluon resummation to next-to-next-to-leading-log (NNLL) order (see [16] and references therein), and assuming a top-quark mass \( m_{\text{top}} = 172.5 \) GeV. The first uncertainty comes from the independent variation of the factorisation and renormalisation scales, \( \mu_F \) and \( \mu_R \), while the second one is associated to variations in the PDF and \( \alpha_s \), following the PDF4LHC prescription with the MSTW2008 68% CL at NNLO, CT10 NNLO and NNPDF2.3 5f FFN PDF sets (see [17] and references therein, and [18, 19, 20]).

The production of \( \bar{t}t + Z \) and \( \bar{t}t + W \) events has been first observed by the ATLAS [1] and CMS [2] collaborations at \( \sqrt{s} = 8 \) TeV and has been measured with an uncertainty of about 30% [30, 31]. At 13 TeV, the production cross-sections are increasing by a factor of four, allowing to decrease the dominant statistical uncertainty. The theoretical cross-sections at 13 TeV have been calculated at NLO in pertubative QCD:

\[
\sigma_{\bar{t}t+Z} = 0.84 \pm 0.09 \text{ pb (11%)} \tag{1.3}
\]

and

\[
\sigma_{\bar{t}t+W} = 0.60 \pm 0.08 \text{ pb (13%)} . \tag{1.4}
\]
2. Top-quark pair production at the Tevatron

The $t\bar{t}$ cross-section has been measured at the Tevatron at $\sqrt{s} = 1.96$ TeV. The SM prediction for this centre-of-mass energy is $7.35^{+0.29}_{-0.33}$ pb, calculated at NNLO+NNLL in perturbative QCD [7], resulting in a relative precision of 4.5%. A combination of the CDF and D0 results has been performed, taken into account all $t\bar{t}$ decay channels. The result of the combination and the individual channels can be seen in Fig. 1. Up to 8.8 fb$^{-1}$ of $p\bar{p}$ data was used in the individual analyses. All results are in good agreement with the SM expectation. The relative uncertainty of the combined result is 5.4% and therefore already challenges the uncertainty of the theoretical prediction.

$$\sigma_{t\bar{t}} = 7.26 \pm 0.13 \text{ (stat)} \pm 0.57 \text{ (syst)} \text{ pb (7.6 %)}$$

\[ (2.1) \]
3. Summary of inclusive cross-section measurements

The measurements of the inclusive $t\bar{t}$ cross-section at the Tevatron and the LHC are summarised in Fig. 2. Apart from the measurements using the large datasets at 7, 8 and 13 TeV, an additional cross-section measurement performed by CMS at $\sqrt{s} = 5.02$ TeV is included. The data used for this analysis has been taken in 2015 and has a small luminosity of 26 pb$^{-1}$. The cross-section has been extracted using a cut-and-count analysis and amounts to:

$$\sigma_{t\bar{t}} = 82 \pm 20 \text{ (stat)} \pm 5 \text{ (syst)} \pm 10 \text{ (lumi)} \text{ pb (20 %)} .$$ (3.1)

The theoretical expectation for this centre-of-mass energy is:

$$\sigma_{t\bar{t}} = 68.9^{+1.9}_{-2.3} \text{ (scale)} +^{1.4}_{-1.0} \text{ (PDF)} +^{0.4}_{-0.0} \text{ (\alpha_S)} \text{ pb (4.9 %)} \text{ (3.2)}$$

Figure 2: Summary of $t\bar{t}$ cross-section measurements performed at the Tevatron and LHC experiments as a function of centre-of-mass energies. The measurements are performed in various $t\bar{t}$ decay channels and are compared to the theoretical prediction at NNLO+NNLL in pertubative QCD [33].

A summary of the cross-section measurements performed by ATLAS and CMS at $\sqrt{s} = 8$ TeV is presented in Fig. 3. The best precision is achieved in the combination for the results of both experiments in the channel with exactly one electron and one muon with opposite sign. The relative uncertainty amounts to 3.5 %, which is already smaller than the uncertainty of the theoretical prediction (5.9%). In the following the two measurements in the $e\mu$ channel will be compared. The ATLAS measurement has been updated compared to Fig. 3 and was published in [10]. First, events with isolated electrons and muons with $E_T(p_T) > 25$ GeV are selected. The number of $b$-tagged jets is compared between data and simulation. The number of events with exactly one and exactly two $b$-tagged jets is then used to perform a simultaneous measurement of $\sigma_{t\bar{t}}$ and the efficiency for a $b$-jet to be tagged and pass the selection criteria. The systematic uncertainties are dominated by...
$t\bar{t}$ modelling and luminosity uncertainties. The updated result has a reduced luminosity uncertainty and is more precise than the combination from 2014:

$$\sigma_{t\bar{t}} = 242.9 \pm 1.7 \text{ (stat)} \pm 5.5 \text{ (syst)} \pm 5.1 \text{ (lumi)} \text{ pb (3.2 \%)} . \quad (3.3)$$

The CMS analysis in the $e\mu$ channel selects isolated leptons with $p_T > 20 \text{ GeV}$ and anti-$k_t$ jets with a radius parameter of 0.5 and $p_T > 30 \text{ GeV}$. Both a binned likelihood fit to the $p_T$ spectrum of the softest untagged jet and an event counting method have been employed. The most precise result is achieved with the binned likelihood fit and results into a relative uncertainty of 3.7 \%.

**Figure 3:** Comparison of $t\bar{t}$ cross-section measurements at $\sqrt{s} = 8 \text{ TeV}$ from the ATLAS and CMS experiments with the full NNLO+NNLL calculation. The theoretical prediction is shown for four different PDF sets [33].
When increasing the centre-of-mass energy from 8 to 13 TeV, the $t\bar{t}$ cross-section increases by a factor of 3.3 to $\sigma_{t\bar{t}} = 832^{+40}_{-46}$, with a relative uncertainty of 5.5%. The overview of the inclusive measurements from both ATLAS and CMS at 13 TeV is given in Fig. 4.

**Figure 4:** Comparison of $t\bar{t}$ cross-section measurements at $\sqrt{s} = 13$ TeV from the ATLAS and CMS experiments with the full NNLO+NNLL calculation. The theoretical prediction is shown for four different PDF sets [33].

The number of $b$-tagged jets is displayed in Fig. 5 for both ATLAS and CMS at $\sqrt{s} = 13$ TeV. Events with zero $b$-tagged jets have large background contamination and are not considered in the analyses. The analysis strategy from the ATLAS experiment is based on the analysis that was described using the 8 TeV dataset above. The dataset amounts to 3.2 fb$^{-1}$ of data collected in 2015 [12]. The fake lepton background is estimated in a dataset that has two leptons with the same charge. The precision of the measured cross-section is limited by systematic and luminosity uncertainties:

$$\sigma_{t\bar{t}} = 818 \pm 8 \text{ (stat)} \pm 27 \text{ (syst)} \pm 19 \text{ (lumi)} \text{ pb (4.2 %)}.$$ (3.4)

The largest systematic uncertainties are the matrix-element generator and the hadronisation uncertainties.
Figure 5: Number of $b$-tagged jets in events with an electron and a muon with opposite sign for the ATLAS experiment (left) [12] and the CMS experiment (right) [13]. The distributions are normalised to the NNLO+NNLL prediction which has been performed for $m_{\text{top}} = 172.5$.

The CMS measurement is performed on 2.2 fb$^{-1}$ of data taken in 2015 [13]. The measurement is performed on a dataset with at least one $b$-tagged jet using the event-counting method. The largest background in this channel originates from single-top production. The uncertainty on the measurement is strongly dominated by systematic uncertainties:

$$\sigma_{t\bar{t}} = 792 \pm 8 \text{ (stat)} \pm 37 \text{ (syst)} \pm 21 \text{ (lumi)} \text{ pb (5.6 \%)}.$$  \hspace{1cm} (3.5)

The largest systematic uncertainties are $t\bar{t}$ modelling, jet energy scale and lepton efficiency uncertainties.
4. Differential $t\bar{t}$ measurements at ATLAS and CMS

With the large datasets taken at 8 and 13 TeV it is now possible to perform cross-section measurements as function of jet multiplicity, kinematic variables and event-related quantities. In this way, theoretical predictions can be tested in different parts of the phase space. Using a Bayesian unfolding procedure, detector effects are removed and the distributions folded back to stable particle or parton level. This allows for a direct comparison between different experiments. For the measurement at parton level, the top quarks from the event record are taken after QCD radiation but before they decay further. At stable particle level, stable leptons and jets clustered from stable particles are used. This allows for generator independent studies.

4.1 Cross-section as function of top-quark $p_T$

4.2 Results resolved topology $\sqrt{s} = 8$ TeV

The $t\bar{t}$ cross-section was measured by both ATLAS and CMS as a function of the properties of the top or the $t\bar{t}$ system in the lepton+jets channel using data taken at $\sqrt{s} = 8$ TeV. This requires a full reconstruction of the $t\bar{t}$ final state.

For the full event reconstruction in ATLAS, the so-called pseudo-top algorithm is used. First the $p_T$ of the neutrino is calculated. This is done by using the four-momentum of the charged lepton and the missing transverse energy and imposing a constraint of the reconstructed $W$-mass. The $b$-tagged jet closest to the lepton is added to reconstruct the leptonically decaying top. The two untagged jets that have an invariant mass closest to the $W$ mass are taken and together with the second $b$-tagged jet they form the hadronic top-quark candidate. The reconstructed distributions are unfolded to parton and stable particle level using POWHEG+PYTHIA6 as baseline MC generator.

The CMS experiment uses a kinematic fitting algorithm for the reconstruction of the $t\bar{t}$ final state. Up to five jets are used as input to the algorithm. Jet-parton assignments where untagged jets are in the position of a $b$-quark or $b$-tagged jets are in the position of a light quark are rejected. This allows to significantly reduce the number of possible jet-parton assignments. A $\chi^2$ value is calculated for each permutation. Events where the best $\chi^2$ value is higher than a defined threshold are vetoed. This allows to remove badly reconstructed events. The baseline MC generator used by CMS is MADGRAPH5+PYTHIA6.

Figure 6 shows the unfolded top $p_T$ spectrum from ATLAS (upper left) and CMS (upper right) and the comparison of both experiments with a NNLO calculation (lower plot). Both experiments observe a higher top momentum in MC than in data, with an increasing deviation at high momentum. Furthermore, large differences between the chosen MC generator setups are observed. When comparing the data with the NNLO prediction however, the agreement is improved for both experiments, with CMS still seeing a small disagreement.
Figure 6: Differential cross-section as function of the top-quark $p_T$ in the resolved topology, using data taken at 8 TeV. The measurements are performed in the fiducial phase space [21] (left) and the full phase space [22] (right). In the lower distribution, the distributions are unfolded to parton level both for the ATLAS and CMS data and compared to the full NNLO calculation and different MC generator predictions [33].
4.3 Results using boosted topologies at $\sqrt{s} = 8$ TeV

For large top-quark momenta, the top-quark decay products are strongly collimated. For the hadronically decaying top-quark, the three jets may even overlap in the detector. This strong collimation allows to identify the top quark as one jet with large radius (large-$R$ jet). Additional cuts are applied to increase the efficiency of identifying a large-$R$ jet as originating from a top quark (top-tagging).

Both ATLAS and CMS have measured the $t\bar{t}$ cross-section as function of the transverse momentum of a top-tagged large-$R$ jet at particle and parton level. The results are displayed in Fig. 7. For the ATLAS result, large-$R$ jets with a radius parameter of 1.0 are used. After the removal of pileup effects (trimming), the hardest top-tagged jet with $p_T > 300$ GeV and $m > 100$ GeV is taken as hadronic top candidate. The uncertainties on particle level vary between 13 and 29% and are dominated by the jet-energy-scale uncertainty of the large-$R$ jets.

Figure 7: Differential cross-section as function of the top-quark $p_T$ in the boosted topology, using data taken at 8 TeV. The distributions are unfolded to stable particle level (left) and parton level (right). The results from ATLAS are shown in the upper plots [25] and the results from CMS in the lower plots [26].
CMS uses large-$R$ jets with a radius of 0.8. The hardest top-tagged jet with $p_T > 400$ GeV and $140 < m < 250$ GeV is used as hadronic top-quark candidate. The predictions using POWHEG+PYTHIA6 and MADGRAPH5+PYTHIA6 show a large overestimation of the cross-section. The results from both experiments see the same trend in the transverse top momentum as observed in the resolved channel.

### 4.4 Results covering resolved and boosted topologies at $\sqrt{s} = 13$ TeV

The CMS experiment also published a measurement of the $t\bar{t}$ cross-section as function of the leading top-quark $p_T$ in the all-hadronic channel, using 2.5 fb$^{-1}$ of 13 TeV data [28]. The measurement has been performed in the resolved and boosted topology. For the resolved topology, events with at least 6 anti-$k_t$ jet with radius of 0.4 are selected. Furthermore, at least two of the jets have to be $b$-tagged. The event reconstruction is done with a kinematic fitting method. Additional selection criteria on the reconstructed top-quark mass and the spatial separation of the two $b$-jets are applied. In the boosted channel, large-$R$ jets with a radius of 0.8 and $p_T > 200$ GeV are selected. The event has to have at least 2 large-$R$ jets that both should contain a $b$-tagged jet with smaller radius. The leading top-tagged large-$R$ jet has to fulfill $p_T > 450$ GeV and $150 < m < 200$ GeV. For both topologies the same trend as in the lepton+jets and the dilepton channel is observed, as can be seen in Fig. 8.

The inclusive cross-section in the all-hadronic channel is:

$$\sigma_{t\bar{t}} = 834 \pm 25 \text{(stat)} +^{118}_{-104} \text{(syst)} \pm 23 \text{(lumi)} \text{ pb (14.7 \%)} .$$  \hspace{1cm} (4.1)

![Figure 8: Left: Differential cross-section as a function of the leading top-quark $p_T$ unfolded to parton level [28] using data taken at $\sqrt{s} = 13$ TeV with the CMS experiment. Right: Ratio of unfolded data distribution with the Powheg+Pythia8 prediction.](image)

### 5. Cross-section as function of top and $t\bar{t}$ quantities

A $t\bar{t}$ cross-section measurement using 2.2-2.3 fb$^{-1}$ of data taken at $\sqrt{s} = 13$ TeV was performed by the CMS collaboration. Both the lepton+jets and dilepton channel were used. The unfolded distributions are shown in Fig. 9 for the hadronic top-$p_T$ (top), and the $p_T$ (middle) and invariant mass (bottom) of the reconstructed $t\bar{t}$ system. POWHEG+PYTHIA8 was used as baseline...
MC generator. For the top quark $p_T$, the same trend as observed in the measurements discussed above is observed here. Larger deviations can also be seen for the shape and the normalisation of the invariant mass and transverse momentum of the $t\bar{t}$ system. For the dilepton channel, a comparison with the NNLO prediction is displayed. The data is in good agreement with this prediction.

![Figure 9: Cross-section as function of the hadronic top-$p_T$ (top), and the $p_T$ (middle) and invariant mass (bottom) of the reconstructed $t\bar{t}$ system in the lepton+jets [23] and dilepton [24] channels using 13 TeV collision data taken with the CMS experiment.](image-url)
5.1 Cross-section as function of jet multiplicity

In Fig. 10, data and MC are compared for the jet multiplicity distributions for both ATLAS (left) and CMS (right) using data taken at $\sqrt{s} = 13$ TeV. For the ATLAS measurement, a disagreement between data and the POWHEG+PYTHIA6 generator is observed, while the other generator predictions agree well with data within the uncertainties. The CMS data distribution shows less jets than predicted by all MC generator setups.

Figure 10: Jet multiplicity distribution unfolded back to stable particle level using data taken at 13 TeV by the ATLAS [27] (left) and CMS [24] (right) experiments. The distributions show the additional number of jets that do not belong to the $t\bar{t}$ system. The data distributions are compared with different NLO+PS setups.

6. Measurement of $t\bar{t}+V$ production

The production of $t\bar{t}+Z$ and $t\bar{t}+W$ events has been first observed by the ATLAS and CMS collaborations at $\sqrt{s} = 8$ TeV and have been measured with an uncertainty of about 30% [30, 31]. At 13 TeV, ATLAS is using a binned profile likelihood fit to extract the cross-section. Signal depleted control regions are included in the fit in order to constrain systematic uncertainties in the fit to data. The events are split into signal enriched and signal depleted regions depending on the number of leptons, jets and $b$-tagged jets. For the analysis in the trilepton and the same-sign dimuon channel, the matrix method is used to estimate the fake lepton background from data. The diboson background is also estimated from data in a dedicated control region. For the tetralepton analysis, the fake lepton background is taken from simulation. Scale factors are estimated in the $t\bar{t}$ and $Z$+jets control regions and applied to the simulation. The resulting cross-sections are:

$$\sigma_{t\bar{t}+Z} = 0.9 \pm 0.3 \text{(stat)} \pm 0.1 \text{(syst)} \text{pb (35\%)}$$  \hspace{1cm} (6.1)

and

$$\sigma_{t\bar{t}+W} = 1.5 \pm 0.7 \text{(stat)} \pm 0.3 \text{(syst)} \text{pb (51\%)}$$  \hspace{1cm} (6.2)
The $\sigma_{t\bar{t}+Z}$ cross-section is in good agreement with the NLO prediction. The $\sigma_{t\bar{t}+W}$ cross-section agrees with the theoretical prediction within 1.1 standard deviations. The uncertainties on the measured cross-sections are large and are dominated by statistical uncertainties.

The analysis from the CMS collaboration is based on 12.3 fb$^{-1}$ of data. Due to the larger dataset and reduced systematic uncertainties compared to the ATLAS measurement presented above, the cross-sections are measured at higher precision:

$$\sigma_{t\bar{t}+Z} = 0.70^{+0.16}_{-0.15} \text{ (stat)}^{+0.14}_{-0.12} \text{ (syst)} \text{ pb (30%) (6.3)}$$

and

$$\sigma_{t\bar{t}+W} = 0.98^{+0.23}_{-0.22} \text{ (stat)}^{+0.22}_{-0.18} \text{ (syst)} \text{ pb (32%) (6.4)}$$

The results are in good agreement with the SM predictions.

7. Top-quark production at LHCb

In addition to the measurements performed by the ATLAS and CMS collaborations, top quark production can also be measured with the LHCb experiment in the forward region. The production of top quarks in the forward region has a larger contribution from the $q\bar{q}$ and the $gq$ processes and allows to constrain gluon PDFs. About 75% of the events are coming from top-quark pair production and 25% from single top production. Both channels were measured together in the $\mu + b$ final state using data taken at $\sqrt{s} = 7$ and 8 TeV [29]. Muons with $p_T > 25$ GeV and $2 \geq \eta \geq 4.5$ and jets with $p_T$ between 50–100 GeV and $2 \geq \eta \geq 4.2$ are selected. The transverse momentum of the $\mu + b$ system has to be larger than 20 GeV. Jets are identified as originating from a $b$-hadron using a boosted decision tree discriminant. The cross-section and the charge asymmetry is measured by performing a profile likelihood fit to the $p_T$ spectrum of the $\mu + b$ system.
The cross-sections at 7 and 8 TeV were measured to be:

$$\sigma(7 \text{ TeV}) = 239 \pm 53 \text{ (stat)} \pm 33 \text{ (syst)} \pm 24 \text{ (theory)} \text{ pb (26 \%)} \ .$$  \hspace{1cm} (7.1) \\

and

$$\sigma(8 \text{ TeV}) = 289 \pm 43 \text{ (stat)} \pm 40 \text{ (syst)} \pm 29 \text{ (theory)} \text{ pb (23 \%)} \ .$$  \hspace{1cm} (7.2)

and are in good agreement with the theoretical calculation from MCFM \[32\]. In summary, top quark production has been observed by LHCb in the forward region with a significance of 5.4 standard deviations.

8. Summary

High precision inclusive $t\bar{t}$ cross-sections from both Tevatron and LHC experiments were presented. The total uncertainties in the resolved topology are strongly dominated by modelling uncertainties. In the boosted topologies, the jet-energy-scale uncertainty of the large-$R$ jets is the dominant uncertainty. The uncertainties on the inclusive cross-sections are already smaller than the ones on the theoretical predictions. Furthermore, top quark production has been observed at the LHCb experiment in the forward detector region. In addition, differential cross-sections were performed by the ATLAS and CMS collaborations as function of quantities of the reconstructed top-quarks and $t\bar{t}$ systems, the jet multiplicity and event level quantities. The NLO+parton shower generators predict on average a higher transverse top momentum than observed in data. This was observed both by the ATLAS and CMS collaborations at 8 and at 13 TeV. The comparison of a full NNLO calculation with data has however shown a better agreement. With the new data taken in 2016, the experiments will be able to look at double-differential distributions and further scrutinize the differential distributions in the boosted topology.
References

[23] The CMS Collaboration, CMS PAS TOP-16-008
[31] The CMS Collaboration, EPJC 74 (2014) 3060
[33] The ATLAS Collaboration,
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/TOP/