

Top pair production cross sections at LHCb

Donatella Lucchesi*

University and INFN of Padova

On behalf of the LHCb Collaboration

E-mail: donatella.lucchesi@pd.infn.it

The production and decay properties of the top-quark are sensitive to new physics in several scenarios beyond the Standard Model. LHCb is the only experiment where the top-quark can be studied in the forward region exploiting, therefore, production mechanisms different from ATLAS and CMS experiments. The first cross section measurement in the forward region has been performed at $\sqrt{s} = 7$ and 8 TeV. The most precise determination at LHCb uses data collected at $\sqrt{s} = 13$ TeV requiring final states with a b -jet and two leptons, muon and electron, coming from the W decay. A description of these measurements is provided together with comparison with theories.

Sixth Annual Conference on Large Hadron Collider Physics (LHCP2018)

4-9 June 2018

Bologna, Italy

*Speaker.

1. Introduction: top-quark reconstruction at LHCb

The top-quark production in the region covered by the LHCb detector [1] receives a higher contribution from quark-antiquark ($q\bar{q}$) annihilation than in the central region and probes higher values of Bjorken- x , where large uncertainties are present in the proton parton distribution functions (PDFs). The greater contribution from quark-initiated production results in a larger expected charge asymmetry [2, 3], which, therefore, can be measured by LHCb. A sizeable sample of reconstructed top-quark allows the precise measurements of top-quark production that can be used to constrain PDFs [4].

At LHCb, $t\bar{t}$ pairs have been reconstructed exploiting the semileptonic W boson decay mode to a muon or an electron and a neutrino in association with b -jets. Since the detector does not allow the measurement of the missing energy, only partially reconstructed final states are possible. The first observation of the top-quark production in the forward region was performed on Run 1 data collected between 2010 and 2012 at center-of-mass energies of 7 and 8 TeV. The measurement requires a final state with an high-momentum muon and a b -jet achieving a precision of 20% on the cross section [5]. This final state provides high statistics but suffers from a large background due to the production of a W boson in association with a b -jet. A more enriched top-quark data set is obtained by requiring a lepton (muon or electron) and two b -jets. Here, the yield is low, therefore the precision on the cross section is only of 40% [6].

In 2015 and 2016 LHCb has collected a data sample at a center-of-mass energy of 13 TeV. The increased energy leads to an increase of a factor of ten in the $t\bar{t}$ production cross section within the LHCb acceptance [7]. Thanks to the larger expected yield, a cleaner signal signature is possible, which includes the dilepton final state, where both top quarks decay to a W boson and a b -quark, and the W bosons decay leptonically to a lepton, one in muon and the other in electron, and a neutrino.

The above top-quark measurements have in common the lepton selection and the b -jet identification. The muons and the electrons are required to have:

- $p_T > 20$ GeV;
- pseudorapidity in the range $2.0 < \eta < 4.5$;
- $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} > 0.1$ in (η, ϕ) space, where ϕ refers to the azimuthal angle, between the two leptons when two leptons are required in the final state.

Jets are reconstructed using a particle flow algorithm and clustered using the anti- k_T algorithm with a distance parameter of 0.5. They must have p_T above 20 GeV and η between 2.2 and 4.2. The energy resolution ranges between (10 – 15)% for $p_T \in [10, 100]$ GeV. The b -jet is identified by requiring jet passing the above requirements and with at least one secondary vertex (SV) with a flight direction $\Delta R < 0.5$ with respect to the jet axis, where the flight direction is taken to be the vector joining the SV to the primary vertex. Two boosted decision trees (BDT) are trained to separate light from heavy flavor jets and b from c initiated jets using, among the others [8]:

- charge and track multiplicity;
- track kinematics properties;

- displacement of secondary vertex from the primary one;
- SV corrected mass, $M_{corr}(SV) = \sqrt{M^2 + p^2 \sin^2 \theta} + p \sin \theta$.

The efficiency for identifying a b -jet is about 65% and the probability for misidentifying a light-parton jet of 0.3%.

2. The lepton plus b -jets data sample

Events are selected by requiring the presence of either a high- p_T muon or electron and two heavy flavour b -tagged jets using the procedure described above. The lepton is required to be isolated from both jets using $\Delta R(\ell, j) > 0.5$. The sample is composed by three different signals: $t\bar{t}$, $W^\pm + b\bar{b}$ and $W^\pm + c\bar{c}$. The separation of b -jet from c -jet is made possible by the excellent LHCb b -tagging performances. The $t\bar{t}$ signal is separated from $W^\pm + b\bar{b}$ and $W^\pm + c\bar{c}$ by using a multivariate discriminant, uniform Gradient Boost (uGB), constructed with the purpose to have the output minimally correlated with the two b -jets invariant mass (m_{jj}).

The background contributions that come from process like $Z + b\bar{b}$ and $Z + c\bar{c}$ production with $Z \rightarrow \mu\mu$ or $Z \rightarrow ee$, where one of the final state leptons is not reconstructed, are obtained from simulation. The QCD multi-jet background, which includes lepton misidentification and semileptonic decays of a beauty or charm hadron is determined using a data-driven method.

The data sample is split into four subsamples, according to the flavor and charge of the lepton: μ^\pm and e^\pm . A simultaneous fit to the distributions of four variables, m_{jj} , uGB and the multivariate discriminant classifier for each jet, is performed to determine the $t\bar{t}$, $W^\pm + b\bar{b}$ and $W^\pm + c\bar{c}$ yields in each sample. The production cross sections are obtained as the product of the normalisation factors determined in the fit and the expected Standard Model cross sections taking into account the efficiencies for each physics process and the integrated luminosity. Figure 1 summarises the results compared to the theory. The $t\bar{t}$ cross section has a significance of 4.9σ . The systematic uncertainties can affect signal and background yields and can alter the template shapes used in the fits. The major contribution come from the heavy-flavor tagging efficiency, which amounts to 5 – 10% depending on the jet p_T . The increased size of the next data samples will help to reduce also the systematic uncertainties in particular on the b -jet tagging by means of dedicated calibration datasets. In figure 1 the $W^\pm + b\bar{b}$ and $W^\pm + c\bar{c}$ are also shown together with the theoretical expectations. The same data set has been used also to set limits on Higgs production in particular on associated production $V + H$ with $H \rightarrow c\bar{c}$ [9].

3. The di-lepton plus b -jet data sample

The 2 fb^{-1} of data collected in Run 2 at $\sqrt{s} = 13 \text{ TeV}$ are used to reconstruct top-quark candidates by requiring that the two W bosons decay leptonically one in electron plus neutrino and the other in muon plus neutrino [10]. The selection of two leptons, performed following the criteria described above, reduces the background from single W boson production and QCD processes and the choice of different-flavor leptons suppresses the contribution from the production of Z bosons. The background coming from light-quark jets is suppressed by asking the presence of a b -jet, which is reconstructed with the standard algorithm except the BDT requirements. Electron and muon have

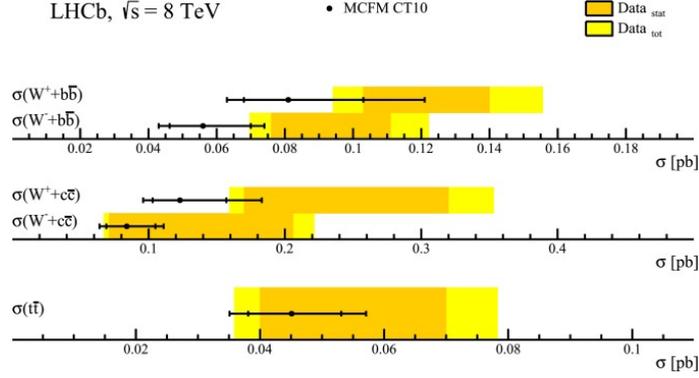


Figure 1: The black markers represent the theoretical prediction and the error bars are the uncertainties, the inner ones represent the scale and the outer ones the total errors. The outer bars (light yellow) correspond to the total uncertainties of the measured cross sections and the inner bars (dark yellow) correspond to the statistical uncertainties.

to be compatible with coming from the primary vertex, isolated and well separated, $\Delta R(e, \mu) > 0.1$. In addition the b -jet has to be separated from each lepton, $\Delta R(j, \ell) > 0.5$.

After all the requirements, 44 events are selected. The major background contribution to this sample comes from events where at least one of the leptons is the result of the misidentification of a hadron. This background can arise from QCD multijet production, vector boson plus jet events, or from $t\bar{t}$ production itself. A control sample obtained by requiring two same sign leptons is used to evaluate this background contribution with amount to 3.9 ± 1.9 events. Minor contributions come from Z +jets and W boson +top-quark events for a total of 5.6 ± 2.0 events. Data is well described by the expected signal and the background as shown in figure 2 where the invariant mass of the muon, electron and b -jet is shown. The cross section, $\sigma_{t\bar{t}}$, is measured in the fiducial region

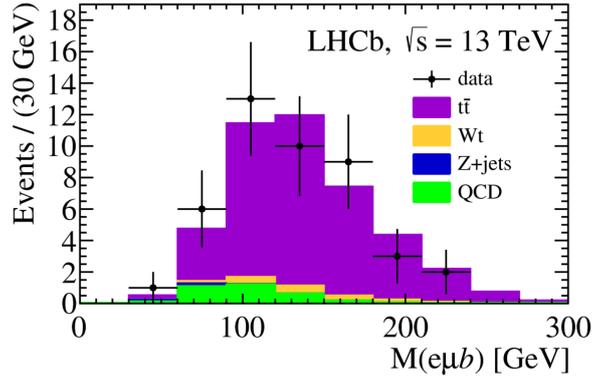


Figure 2: The combined invariant mass of the muon, electron and b -jet in data compared to the expected contributions.

defined by the p_T , η , and ΔR requirements placed on the muon, electron and b -jet candidates, and is calculated using the formula

$$\sigma_{t\bar{t}} = \frac{N - N_{bkg}}{\mathcal{L} \cdot \epsilon} \cdot \mathcal{F}_{res}$$

where N is the total number of candidates selected in data, N_{bkg} is the sum of the expected background contributions, \mathcal{F}_{res} is a resolution factor that accounts for migrations in to and out of the fiducial region, ε is the efficiency to reconstruct and select the signal events, and \mathcal{L} is the integrated luminosity of the data sample. The integrated luminosity is $\mathcal{L} = 1.93 \pm 0.07 \text{ fb}^{-1}$. The overall efficiency is given by the product of the reconstruction and selection contributions,

$$\varepsilon = \varepsilon_{\mu}^{rec} \cdot \varepsilon_{\mu}^{id} \cdot \varepsilon_{\mu}^{trg} \cdot \varepsilon_e^{rec} \cdot \varepsilon_e^{id} \cdot \varepsilon^{jet} \cdot \varepsilon^{tag} \varepsilon^{sel}$$

where the terms are ordered from left to right such that the efficiency is evaluated for candidates passing the stage to the left. $\varepsilon_{\mu,e}^{rec}$ and $\varepsilon_{\mu,e}^{id}$ are the efficiency to reconstruct and identify respectively for muon and electron, while ε_{μ}^{trg} refers to the efficiency to trigger the event on the muon candidate. The efficiency to reconstruct and tag the jet are given by ε^{jet} and ε^{tag} respectively, and the efficiency of the additional selection requirements is given by ε^{sel} . Each term is determined using a combination of Monte Carlo simulated data and collider data, with the last one used in particular to validate the Monte Carlo based results. The efficiency to tag a jet is 0.556 ± 0.056 and for the additional selections 0.564 ± 0.023 while all the other efficiencies are around 90%. The largest contribution to \mathcal{F}_{res} arises from the momentum resolution of the electron due to bremsstrahlung and it is determined from simulation, $\mathcal{F}_{res} = 1.207 \pm 0.006$.

Using the formula and the inputs described, the $t\bar{t}$ cross section is

$$\sigma_{t\bar{t}} = 126 \pm 19(stat) \pm 16(syst) \pm 5(lumi) \text{ fb}$$

where the first uncertainty is statistical, the second is systematic, and the third is due to the luminosity determination. A summary of the systematic uncertainties is reported in table 1 from where it is evident that the knowledge of the jet tagging efficiency dominates. The overall precision is at the level of 20% and it is dominated by the statistical error.

Source	%
trigger	2.0
muon reconstruction	1.1
electron reconstruction	2.8
muon identification	0.8
electron identification	1.3
jet reconstruction	1.6
event selection	4.0
jet tagging	10.0
background	5.1
acceptance	0.5
total	12.7

Table 1: Summary of the systematic uncertainties expressed as a percentage of the measured cross section.

The result is compared to the theoretical predictions obtained from the Monte Carlo simulations: aMC@NLO [11], POWHEG [12, 13, 14] and MCFM [15, 16]. For each calculation different sources of uncertainty are considered: description of the PDFs, choice of renormalisation and factorisation scales, the value of the strong coupling constant used in the calculation. The uncertainty

due to the choice of the top-quark mass is expected to be small and is not considered further. The total theoretical uncertainty is determined by combining the individual uncertainties. A comparison of the measured cross section with the predictions in the case where the fiducial requirements are placed on the final state muon, electron and b -jet, is shown on the top of figure 3. The measured cross section in the top-quark fiducial region, obtained by extrapolating from the measured fiducial region using predictions from aMC@NLO and requiring that both top quarks have a rapidity between 2.0 and 5.0 and a transverse momentum greater than 10 GeV, is compared with the predictions on the bottom of figure 3. The cross section in the top-quark fiducial region is compared also to predictions from MCFM. The measured cross section in both fiducial regions is in agreement with the theoretical predictions.

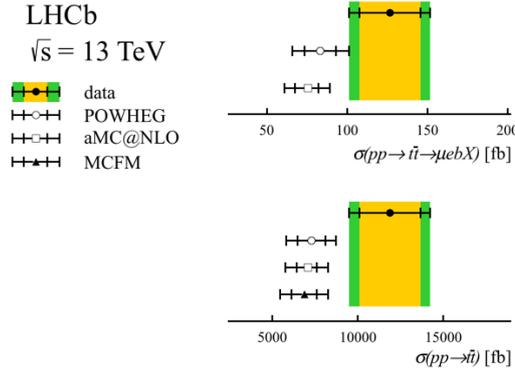


Figure 3: Comparison of the measured cross sections with the predictions from the aMC@NLO, POWHEG and MCFM generators. For the data, the inner error band represents the statistical uncertainty, and the outer the total, while for the theoretical predictions, the inner band represents the scale uncertainty and the outer represents the total. The prediction is shown (above) for the muon, electron and jet fiducial, and (below) for the top fiducial region.

4. Conclusions

The cross section for top-quark production in the forward region has been measured in different final states at LHCb where the acceptance is complementary to the ATLAS and CMS. The results are statistically limited and they will be updated in Run 2 exploiting the 6 fb^{-1} that the experiment is planning to collect. In Run 3 more than 50 fb^{-1} are expected that will allow measurements at LHCb to reach percent level statistical uncertainty, bringing LHCb in the game of precision top-quark physics measurements. The $t\bar{t}$ asymmetry will be the next to come.

References

- [1] LHCb Coll., A. A. Alves Jr. et al., The LHCb detector at the LHC, JINST 3 (2008) S08005.
- [2] A. L. Kagan, J. F. Kamenik, G. Perez, and S. Stone, Top LHCb physics, Phys. Rev. Lett. 107 (2011) 082003, arXiv:1103.3747.
- [3] R. Gauld, Leptonic top-quark asymmetry predictions at LHCb, Phys. Rev. D91 (2015) 054029, arXiv:1409.8631.

- [4] R. Gauld, Feasibility of top-quark measurements at LHCb and constraints on the large- x gluon PDF, JHEP 02 (2014) 126, arXiv:1311.1810.
- [5] LHCb Coll., R. Aaij et al., First observation of top-quark production in the forward region, Phys. Rev. Lett. 115 (2015) 112001, arXiv:1506.00903.
- [6] LHCb Coll., R. Aaij et al., Measurement of the $t\bar{t}$, $W + b\bar{b}$ and $W + c\bar{c}$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV, Phys. Lett. B767 (2017) 110, arXiv:1610.08142.
- [7] R. Gauld, Measuring top-quark production asymmetries at LHCb, LHCb-PUB-2013- 009.
- [8] LHCb Coll., R. Aaij et al., Identification of beauty and charm quark jets at LHCb, JINST 10 (2015) P06013, arXiv:1504.07670.
- [9] LHCb Coll., R. Aaij et al., Search for $H^0 \rightarrow b\bar{b}$ or $c\bar{c}$ in association with a W or Z boson in the forward region of pp collisions, LHCb-CONF-2016-006
- [10] LHCb Coll., R. Aaij et al., Measurement of forward top pair production in the dilepton channel in pp collisions at $\sqrt{s} = 13$ TeV, JHEP 08 (2018) 174, arXiv:1803.05188.
- [11] 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, arXiv:1405.0301.
- [12] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, JHEP 11 (2004) 040, arXiv:hep-ph/0409146.
- [13] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with parton shower simulations: the POWHEG method, JHEP 11 (2007) 070, arXiv:0709.2092.
- [14] 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, arXiv:1002.2581.
- [15] J. M. Campbell and R. K. Ellis, An update on vector boson pair production at hadron colliders, Phys. Rev. D60 (1999) 113006, arXiv:hep-ph/9905386.
- [16] J. M. Campbell, R. K. Ellis, and C. Williams, Vector boson pair production at the LHC, JHEP 07 (2011) 018, arXiv:1105.0020.