

Measurement of the single-top-quark t -channel cross section in pp collisions at $\sqrt{s} = 7$ TeV using a multivariate analysis

Dennis Klingebiel for the CMS collaboration

III. Physikalisches Institut A, RWTH Aachen University, Germany

E-mail: klingebiel@physik.rwth-aachen.de

Abstract. Electroweak top-quark production offers a unique access to search for new physics phenomena. The measurement of the t -channel cross section allows to determine the Cabibbo-Kobayashi-Maskawa matrix element $|V_{tb}|$ with high precision and without assuming its unitarity. We present the precise measurement of the t -channel single-top-quark production cross section with a multivariate analysis. Data collected with the CMS experiment in proton-proton collisions at a center-of-mass energy of 7 TeV, corresponding to 1.17 fb^{-1} for muon final states and 1.56 fb^{-1} for electron final states, are used for this measurement. The measured t -channel single-top-quark production cross section is $67.2 \pm 6.1 \text{ pb}$, compatible with the approximate next-to-next-to-leading-order Standard Model prediction. The CKM matrix element $|V_{tb}|$ is measured to be $1.020 \pm 0.046 \text{ (meas.)} \pm 0.017 \text{ (theor.)}$.

1. Introduction

Single top quarks can be produced via electroweak interactions. Three production modes are distinguished based on the virtuality of the exchanged W boson. In proton-proton collisions at a center-of-mass energy of 7 TeV, the dominant production mode is the t -channel with an approximate next-to-next-to-leading-order (NNLO) Standard Model (SM) cross section prediction of $64.6^{+2.1}_{-0.7} +1.5_{-1.7} \text{ pb}$ for a top-quark mass of $m_t = 172.5 \text{ GeV}/c^2$ [1]. Figure 1 shows the leading order Feynman diagrams for t -channel single-top-quark production. The precise

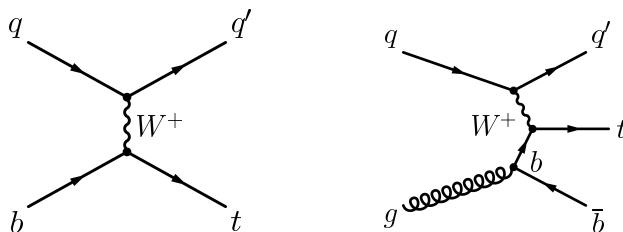


Figure 1. Feynman diagrams for t -channel single-top-quark production in leading order for the $2 \rightarrow 2$ (left) and $2 \rightarrow 3$ (right) processes.

measurement of the t -channel cross section allows a unique access to the measurement of the Cabibbo-Kobayashi-Maskawa matrix (CKM matrix) [2; 3] element $|V_{tb}|$ without assuming unitarity of the CKM matrix, as the t -channel production cross section is proportional to $|V_{tb}|^2$. In the CMS experiment [4], two complementary analysis approaches with three independent



analyses are followed to measure the t -channel single-top-quark cross section in proton-proton collisions at 7 TeV [5]. One approach exploits the pseudo-rapidity distribution of the recoil jet in a signal-enriched phase space ($|\eta_{j'}|$ analysis). The other approach exploits multivariate analysis techniques and various signal- or background-enriched categories. Either Neural Networks (NN) or Boosted Decision Trees (BDT), have been used to exploit the prior knowledge of the electroweak top-quark production mechanism and to probe the SM top-quark event topology of the candidate events. Various categories are defined on the multiplicity of jets and b-tagged jets in each event are used to enhance the signal acceptance, constrain systematic uncertainties in-situ, and to gain confidence in the background modelling. This approach allows a most precise measurement of the signal cross section. All three analyses have been combined using the best linear unbiased estimator (BLUE) [6]. In this article, we summarize the analysis strategy and the results of one of the multivariate analyses, the BDT analysis, while the $|\eta_{j'}|$ analysis is described in detail elsewhere in these conference proceedings [7]. Events with leptonic decays of the top quark into a muon or electron, a W boson, and a b quark are considered ($t \rightarrow bW \rightarrow b\ell\nu$ ($\ell = e$ or μ)). The data used in this analysis were collected with the CMS detector in proton-proton collisions at a center-of-mass energy of 7 TeV. The data corresponds to 1.17 fb^{-1} for the muon final states and 1.56 fb^{-1} for the electron final states.

2. Analysis strategy

The BDT analysis was developed in a “semi-blind” way. First, the validity of the description of physics objects kinematics and properties was checked in signal-depleted control regions. Furthermore, the agreement between data and simulation in various kinematic variables, especially the input variables for the multivariate analysis and the classifier output, was confirmed and continuously checked in signal-depleted phase space. Here, both W boson plus jets and top quark dominated control regions were used. Second, the signal region was unblinded.

The BDT analysis measures the signal cross section in 12 orthogonal categories based on the flavour of the charged lepton (e, μ), the number of selected jets $n\text{-jets} = 2, 3, \geq 4$, and the number of b-tagged jets $m\text{-btags} = 1, \geq 2$. Furthermore, the 0-tag categories are used to check the modelling of input variables and multivariate discriminator output.

The t -channel signal is simulated with the POWHEG next-to-leading-order (NLO) generator [8–10] interfaced to PYTHIA 6.4.24 [11]. Data are recorded if a trigger is fired by a single isolated muon with a transverse momentum $p_T > 17 \text{ GeV}/c$ or a single isolated electron with $p_T > 27 \text{ GeV}/c$, or in later runs by a single isolated electron with $p_T > 25 \text{ GeV}/c$ and a b-tagged jet with $p_T > 30 \text{ GeV}/c$. The simulated events are reweighted to fit the number of additional interactions (pileup) to the conditions of collected data.

The event selection requires exactly one isolated muon with $p_T > 20 \text{ GeV}/c$ and $|\eta| < 2.1$ or one isolated electron with $p_T > 30 \text{ GeV}/c$ and $|\eta| < 2.5$. Charged leptons are required to be isolated to reject non-prompt electrons, i.e. to further reduce the QCD multijet background. For electrons, an additional conversion rejection is applied. Additionally, at least two jets with $p_T > 30 \text{ GeV}/c$ and $|\eta| < 4.5$ are required. One striking feature of t -channel single-top-quark production is the very forward jet originating from the light quark recoiling against the top-quark. Most of the signal is expected to have two jets with one b-tagged jet, since the b-quark out of the initial gluon splitting (fig. 1 right) is collinear with the proton remnant, or has too low transverse momentum to be reconstructed. B-tagging of jets is applied using a track counting high purity algorithm at the tight working point [12]. Furthermore, a transverse W boson mass $M_T(W) > 40 \text{ GeV}/c^2$ [5] is required for events with a muon in the final state and $E_T > 35 \text{ GeV}$ for events with an electron in the final state.

A QCD multijet model is obtained by inverting the isolation criterion in data.

In order to discriminate between signal and background events, Boosted Decision Trees as implemented in TMVA [13] are used. A set of 11 most important variables out of the 37 variables

used in Ref. [14] is chosen based on their discrimination power. For the given analysis, this set is as minimalistic as possible and provides good statistical sensitivity. The chosen variables include kinematics of reconstructed final state and composite objects, and angular correlations between them. The full list of used input variables can be found in Ref. [5]. The description of the input variables is checked in signal-depleted control regions as described in the beginning of this section. Several BDTs are constructed with the Adaptive Boosting algorithm [13], each reweighting the event weight of misclassified events by a certain factor. With the weighted majority vote of the whole forest of trees, the BDT classifier output as one powerful discriminator is calculated.

The resulting BDT discriminator output distribution for the muon-channel in the “2-jets 1-btag” signal category is shown in fig. 2 (left). For events with two jets, the main background is W boson plus jets, while the dominant background for events with three jets is top quark pair production. The resulting BDT discriminator output distribution for the sum of the “2-jets 2-btags”, “3-jets 2-btags”, “4-jets 1-btag”, and “4-jets 2-btags” signal-depleted categories is shown in fig. 2 (right).

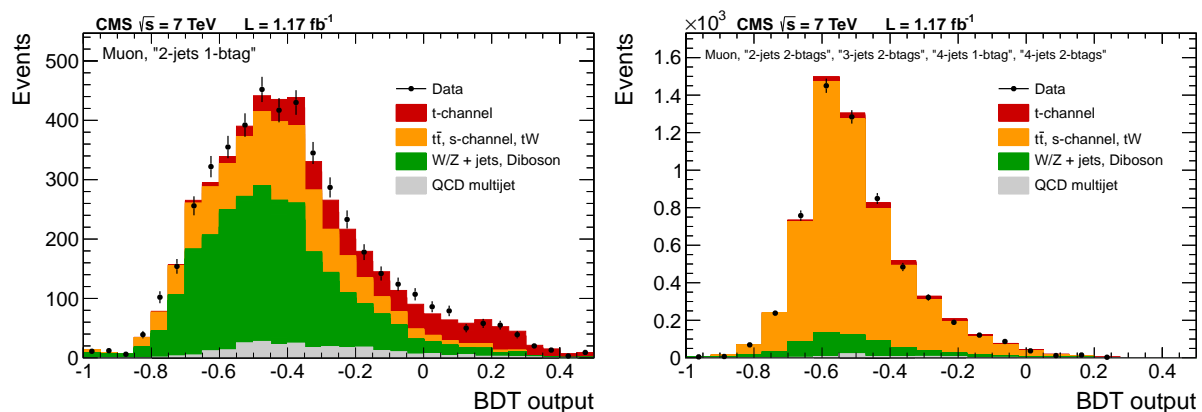


Figure 2. Distributions of the BDT discriminator output for events in a signal category “2-jets 1-btag” with a muon final state (left), and for a signal-depleted control regions, the sum of muon-final-state events in the categories “2-jets 2-btags”, “3-jets 2-btags”, “4-jets 1-btag”, and “4-jets 2-btags” (right). Simulated signal and background contributions are scaled to the best fit results.

For the statistical evaluation, a bayesian core method is used. The implementation of the statistical model and its evaluation is done with the THETA package [15]. Experimental uncertainties and uncertainties of theoretical predictions on the background production cross sections are represented as nuisance parameters. Variations of the jet energy scale, jet energy resolution, b-tagging efficiency, muon and electron trigger and reconstruction efficiencies, pileup conditions, E_T modelling and luminosity are considered as experimental uncertainties. The nuisance parameters itself are marginalized using a Markov Chain Monte Carlo technique. Correlations of nuisance parameters are taken into account. An over-constraint of the statistical model might arise from the limited knowledge of the W boson plus jets rate in the dedicated phase space of this analysis. Thus, a total of 18 nuisance parameters for the W boson plus jets processes are introduced. These depend on the flavour of the additional partons (b, c, or udsg partons) and the number of reconstructed jets in an event. Gaussian or log-normal priors are used for nuisance parameters and a flat prior for the signal strength. The median of the resulting signal strength posterior is used as the central measurement result. The [16%, 84%] quantiles are used as the uncertainty interval.

Theoretical uncertainties due to the factorization and renormalization scale, matching

thresholds for samples simulated with MADGRAPH, parton distribution function, and description of the signal modelling are externalized, i.e. not modelled as additional nuisance parameters within the statistical model. Instead, their impact on the measured cross section is evaluated by generating pseudo-experiments with varied templates and re-doing the default signal extraction procedure, in order to avoid that the statistical model artificially over-constrains those systematic uncertainties. An over-constraint might arise when uncertainties cannot be fully parametrized within the statistical model, as it is often the case for theoretical uncertainties. The dominant systematic uncertainties are the signal modelling, the W boson plus jets background rate prediction, and the simulation statistics.

The CKM matrix element $|V_{tb}|$ is calculated using the approach of Feldman and Cousins [16].

3. Results

The t -channel single-top-quark production cross section in the BDT analysis is measured to be

$$\begin{aligned}\sigma_{t\text{-ch.}} &= 66.6^{+7.0}_{-6.6} \text{ (stat. + syst. + lum.)}^{+6.4}_{-3.5} \text{ (theor.) pb (muons),} \\ \sigma_{t\text{-ch.}} &= 66.4^{+8.4}_{-7.9} \text{ (stat. + syst. + lum.)}^{+5.4}_{-5.4} \text{ (theor.) pb (electrons),}\end{aligned}$$

and the combination of both channels is

$$\sigma_{t\text{-ch.}} = 66.6 \pm 4.0 \text{ (stat.)} \pm 3.3 \text{ (syst.)}^{+3.9}_{-3.3} \text{ (theor.)} \pm 1.5 \text{ (lum.) pb.}$$

where “stat.” refers to statistical uncertainties, “syst.” to systematic uncertainties, “lum.” to uncertainties of the luminosity measurement, and “theor.” to externalized theoretical uncertainties as defined in the previous section.

The results of this analysis are combined with the two other analyses $|\eta_{j'}|$ and NN [5] using the BLUE method, which is applied iteratively. The obtained χ^2 corresponds to a p -value of 0.90, i.e. the individual analyses are compatible with each other. The correlations of statistical and systematic uncertainties between the three analyses are taken into account [5]. Correlations are varied, resulting in negligible changes of the combined cross section.

The combined measurement of the single-top-quark t -channel cross section is

$$\sigma_{t\text{-ch.}} = 67.2 \pm 3.7 \text{ (stat.)} \pm 3.0 \text{ (syst.)} \pm 3.5 \text{ (theor.)} \pm 1.5 \text{ (lum.) pb ,}$$

which is in agreement with the approximate next-to-next-to-leading-order (NNLO) SM prediction of $\sigma_{t\text{-ch.}}^{\text{th}} = 64.6^{+2.1}_{-0.7}{}^{+1.5}_{-1.7}$ pb [1].

The CKM matrix element is measured to be $|f_{LV}V_{tb}| = \sqrt{\sigma_{t\text{-ch.}}/\sigma_{t\text{-ch.}}^{\text{th}}} = 1.020 \pm 0.046 \text{ (meas.)} \pm 0.017 \text{ (theor.)}$, assuming only that the top-quark decays into a W boson and a b quark ($t \rightarrow bW$) with a branching ratio of 1. f_{LV} is a general anomalous coupling form factor. Assuming SM electroweak couplings, i.e. $f_{LV} = 1$ and $|V_{tb}| \leq 1$, we measure $0.92 < |V_{tb}| \leq 1$ at 95% confidence level.

4. Conclusions

Two analysis approaches have been followed to measure the t -channel single-top-quark production cross section with the CMS experiment. The results of all analyses are compatible with each other. The t -channel single-top-quark production cross section is measured in pp collisions to be 67.2 ± 6.1 pb at a center-of-mass energy of $\sqrt{s} = 7$ TeV, compatible with the approximate NNLO Standard Model prediction of $64.6^{+2.1}_{-0.7}{}^{+1.5}_{-1.7}$ pb [1]. The CKM matrix element $|V_{tb}|$ is measured to be $1.020 \pm 0.046 \text{ (meas.)} \pm 0.017 \text{ (theor.)}$.

Figure 3 shows the comparison of this t -channel cross section measurement [5] with dedicated measurements at the Tevatron [17; 18], ATLAS [19], and with expectations from Quantum chromodynamics. The SM expectation is computed at NLO with MCFM in the 5-flavour scheme [20] and at approximate NNLO [1].

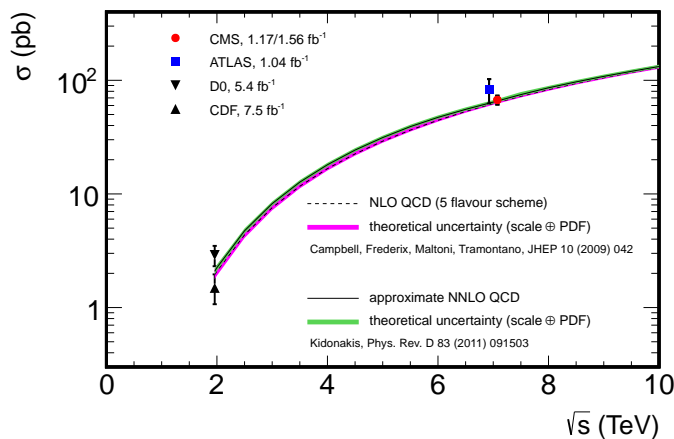


Figure 3. t -channel single-top-quark cross section vs. centre-of-mass energy. The error band (width of the curve) of the SM expectation calculation includes variations of the cross section due to the top-quark mass uncertainty [21], the parton distribution function [22], and the coherent variations of factorisation and renormalisation scales by a factor of two or 0.5.

References

- [1] Kidonakis N 2011 *Phys. Rev. D* **83** 091503 (*Preprint* 1103.2792)
- [2] Cabibbo N 1963 *Phys. Rev. Lett.* **10** 531
- [3] Kobayashi M and Maskawa T 1973 *Prog. Theor. Phys.* **49** 652
- [4] Chatrchyan S *et al.* (CMS) 2008 *JINST* **03** S08004
- [5] Chatrchyan S *et al.* (CMS Collaboration) 2012 *JHEP* **1212** 035 (*Preprint* 1209.4533)
- [6] Lyons L, Gibaut D and Clifford P 1988 *Nucl. Instr. and Meth. A* **270** 110
- [7] Merola M 2013 *these conference proceedings*
- [8] Alioli S, Nason P, Oleari C and Re E 2010 *JHEP* **06** 043 (*Preprint* 1002.2581)
- [9] Alioli S, Nason P, Oleari C and Re E 2009 *JHEP* **09** 111 (*Preprint* 0907.4076)
- [10] Frixione S, Nason P and Oleari C 2007 *JHEP* **11** 070 (*Preprint* 0709.2092)
- [11] Sjöstrand T, Mrenna S and Skands P Z 2006 *JHEP* **05** 026 (*Preprint* hep-ph/0603175)
- [12] CMS Collaboration (CMS) 2011 b-jet identification in the CMS experiment CMS Physics Analysis Summary CMS-PAS-BTV-11-004 URL <http://cdsweb.cern.ch/record/1427247>
- [13] Hoecker A, Speckmayer P, Stelzer J, Therhaag J, von Toerne E and Voss H 2007 TMVA: Toolkit for Multivariate Data Analysis (*Preprint* physics/0703039)
- [14] Chatrchyan S *et al.* (CMS) 2011 *Phys. Rev. Lett.* **107** 091802 (*Preprint* 1106.3052)
- [15] Müller T, Ott J and Wagner-Kuhr J 2012 THETA – a framework for template-based statistical modeling and inference *Preprint* IEKP-KA-CMS/2012-1 URL <http://www-ekp.physik.uni-karlsruhe.de/pub/web/cms/iekp-ka-cms2012-1.pdf>
- [16] Feldman G J and Cousins R D 1998 *Phys. Rev. D* **57** 3873 (*Preprint* physics/9711021)
- [17] Abazov V M *et al.* (D0) 2011 *Phys. Lett. B* **705** 313 (*Preprint* 1105.2788)
- [18] Aaltonen T *et al.* (CDF) 2009 *Phys. Rev. Lett.* **103** 092002 (*Preprint* 0903.0885)
- [19] Aad G *et al.* (ATLAS) 2012 *Phys. Lett. B* **717** 330 (*Preprint* 1205.3130)
- [20] Campbell J M, Frederix R, Maltoni F and Tramontano F 2009 *JHEP* **10** 042 (*Preprint* 0907.3933)
- [21] CDF and D0 Collaborations 2009 Combination of CDF and D0 measurements of the single top production cross section (*Preprint* 0908.2171)
- [22] Campbell J M, Huston J W and Stirling W J 2007 *Rept. Prog. Phys.* **70** 89 (*Preprint* hep-ph/0611148)