

Latest measurements in single top quark with the CMS experiment

Fabio lemmi* on behalf of the CMS collaboration

Alma Mater Studiorum - Università di Bologna (Unibo) Istituto Nazionale di Fisica Nucleare (INFN) E-mail: fabio.iemmi@cern.ch

The latest measurements obtained by the CMS Collaboration in single top quark production are summarized. We present both inclusive and differential cross section measurements in different production modes, which lead to constraints on proton parton distribution functions (PDFs) and measurements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element $|V_{tb}|$ and top quark spin asymmetry A_{ℓ} ; we also show searches for rare Standard Model (SM) processes, such as the first observation of tZq associated production and the first evidence for t γ q associated production.

7th Annual Conference on Large Hadron Collider Physics - LHCP2019 20-25 May, 2019 Puebla, Mexico

*Speaker.

[©] Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

Fabio Iemmi

1. Introduction

Top quarks are mainly produced in tī pairs at the CERN Large Hadron Collider (LHC) through the strong interaction ($\sigma_{t\bar{t}}^{13} \text{ TeV} \approx 830 \text{ pb}$). Single top quark production is rarer but better understood from the theoretical point of view, since it's mediated by the weak interaction. Three single top quark production modes are possible at the LHC:

- 1. *t*-channel: exchange of a space-like W boson ($\sigma_{t-ch}^{13 \text{ TeV}} \approx 220 \text{ pb}$);
- 2. tW production: associated production with an on-shell W boson ($\sigma_{tW}^{13 \text{ TeV}} \approx 70 \text{ pb}$);
- 3. *s*-channel: exchange of a time-like W boson ($\sigma_{s-ch}^{13 \text{ TeV}} \approx 10 \text{ pb}$).

Leading order Feynman diagrams for each of these processes are reported in Figure 1.

Thank to the high integrated luminosity developed by the LHC we are now entering the precision era in *t*-channel measurements, where it is also possible to exploit the distinctive signature of a forward jet coming from the light quark recoiling against the W boson.

All the presented measurements are performed with data collected by the CMS experiment in 2016, corresponding to an integrated luminosity of 35.9 fb^{-1} , unless otherwise stated.



Figure 1: Leading order Feynman diagrams for *t*-channel (left), tW (center) and *s*-channel (right) production modes.

2. Inclusive and differential measurements

In this section we present the latest CMS results for inclusive and differential cross section measurements in different production modes.

2.1 Single t and \overline{t} cross section ratio and V_{tb}

This measurement [1] targets *t*-channel single top quark events. The flavor of the initial light quark determines the charge of the resulting top quark: up quarks in the initial state result in top quarks, while down quarks produce top antiquarks. Since charge asymmetry is present inside protons, the ratio of the cross sections for these two processes, $R_{t-ch} = \sigma_{t-ch,t}/\sigma_{t-ch,\bar{t}}$, is expected to be different from one. It also provides insight into the inner structure of protons as described by the PDFs. Moreover, a measurement of the total cross section $\sigma_{t-ch,t+\bar{t}}$ allows to perform a direct measurement of the CKM matrix element $|f_{LV}V_{tb}|$ exploiting the relation $|f_{LV}V_{tb}| = \sqrt{\frac{\sigma_{t-ch,t+\bar{t}}}{\sigma_{t-ch,t+\bar{t}}}}$, where the anomalous form factor f_{LV} takes into account the possible presence of an anomalous Wtb coupling and is assumed to be exactly one in the SM framework.

The analysis selects leptonic top quark decays, namely events containing a single isolated muon or electron in the final state. These leptons originate from the decay of the W boson coming from the top quark decay, either directly or via the $W \rightarrow \tau \nu$ and $\tau \rightarrow \ell \nu$ decays. Given the *t*-channel topology, the signal signature is one charged, isolated lepton, p_T imbalance, and two jets, one of which is b tagged. A second b jet, arising in the production process via gluon splitting (4 flavor scheme, 4FS), generally has a softer p_T spectrum and a broader pseudorapidity (η) distribution compared to the b jet coming from the top quark decay, and usually escapes detection. Events containing additional leptons are rejected.

The events used in the analysis are divided into orthogonal categories depending on the number of selected jets and the number of b tagged jets (*n*jets-*m*tags regions). Events in the 2jets-1tag category form the signal region, while events in the 3jets-1tag and 3jets-2tags categories form control regions used to constrain the main background of the analysis, namely tī production.

In order to extract the signal, boosted decision trees (BDT) algorithms are employed to combine multiple variables into a single discriminator, enhancing the separation between signal and background processes. Important variables to achieve this task are the light quark $|\eta|$, the reconstructed top quark mass, the invariant mass of the dijet system and the distance in the $\eta - \phi$ plane between the lepton and the b tagged jet. The signal is extracted by performing a simultaneous, maximum-likelihood (ML) fit to the BDT distributions for each region, lepton flavor and lepton charge. The results of this fit are $\sigma_{t-ch,t}$ and $\sigma_{t-ch,\bar{t}}$. The ratio of the two cross sections can be calculated from these two results propagating their uncertainties using the covariance matrix coming from the fit. Nonetheless, a more elegant and straightforward way to properly account for the correlations between uncertainties in the cross sections is to repeat the fit with one of the two cross sections replaced by their ratio. This way, potential cancellations of uncertainties are taken into account directly in the fit and do not need to be computed afterwards.

Several sources of systematic uncertainty are considered, either as nuisance parameters in the ML fit (profiled uncertainties) or as nonprofiled uncertainties. The impact of the latter is determined by repeating the analysis using varied templates according to the systematic uncertainties under study instead of the nominal templates. The dominant systematic uncertainty is found to be the one concerning the signal model (PDFs, parton shower scale, etc.).

The resulting value for the ratio R_{t-ch} is found to be

$$R_{t-ch} = 1.66 \pm 0.02 \text{ (stat)} \pm 0.05 \text{ (syst)},$$

in good agreement with most of the predictions, as showed in Figure 2. The measured total cross section translates into

$$|f_{\rm LV}V_{\rm tb}| = 1.00 \pm 0.08 \,({\rm exp}) \pm 0.02 \,({\rm theo}),$$

where the first uncertainty considers all uncertainties of the measured total cross section, while the second uncertainty is derived from the uncertainty of the theoretical SM prediction. The measured value is in agreement with the assumption of $|V_{tb}| = 1$ used to compute the theoretical cross section.

2.2 ATLAS and CMS combination at $\sqrt{s} = 7$ and 8 TeV

Regarding the single top quark cross sections and $|f_{LV}V_{tb}|$ measurements, a combination of ATLAS and CMS results has been performed [2]. The *t*-channel, tW, and *s*-channel single top



Figure 2: Comparison of the measured R_{t-ch} (central, dashed line) with the next-to-leading order (NLO) predictions for different PDFs sets. For the measurement, statistical and systematic uncertainties are indicated individually by the inner and outer error bars.

quark cross section measurements by ATLAS and CMS are combined for each production mode, separately at centre-of-mass energies of 7 and 8 TeV, corresponding to integrated luminosities of 1.17 to 5.1 fb⁻¹ at $\sqrt{s} = 7$ TeV and 12.2 to 20.3 fb⁻¹ at $\sqrt{s} = 8$ TeV. A combined determination of $|f_{LV}V_{tb}|$ is also computed.

The combination is performed through iterative methods, making assumptions about correlations between similar sources of uncertainty in different measurements. All the cross section measurements are found to be in good agreement with their corresponding theoretical predictions within uncertainties, as shown in Figure 3. The combined $|f_{LV}V_{tb}|$ value from the cross section measurements at $\sqrt{s} = 7$ and 8 TeV, including uncertainties in σ_{theo} for each production mode, is found to be

$$|f_{\rm LV}V_{\rm tb}| = 1.02 \pm 0.04 \,({\rm meas}) \pm 0.02 \,({\rm theo}),$$

resulting in a relative uncertainty of 3.7%, which improves on the precision of 4.7% of the most precise individual $|f_{LV}V_{tb}|$ extraction (ATLAS, *t*-channel analysis at $\sqrt{s} = 8$ TeV [3]). A summary of the $|f_{LV}V_{tb}|$ combinations is shown in Figure 3.

2.3 Differential cross sections and charge ratios, spin asymmetry A_ℓ

(

The CMS collaboration also performed a measurement [4] of differential cross sections for *t*-channel single top quark production in the 4FS (see Figure 4), targeting leptonic top quark decays. In addition, charge ratios, defined as the ratio of the single top quark cross section to the sum of the single top quark and antiquark cross sections, are measured differentially. These quantities are sensitive to proton PDFs and can be used to set constraints on them. Furthermore, differential angular distributions can be used to assess the electroweak coupling structure of the Wtb vertex, where a vector-axial vector (V-A) coupling is predicted in the SM. A powerful observable to investigate the nature of the coupling structure in the *t*-channel is the top quark polarization angle $\cos \theta_{pol.}^*$, defined as

$$\cos \theta_{\text{pol.}}^* = \frac{\vec{p}_{q'}^{(\text{top})} \cdot \vec{p}_{\ell}^{(\text{top})}}{|\vec{p}_{q'}^{(\text{top})}| \cdot |\vec{p}_{\ell}^{(\text{top})}|},$$



Figure 3: Comparison of the measurements performed by ATLAS and CMS and their combination with the theoretical predictions (left panel); $|f_{LV}V_{tb}|$ combined measurements along with theoretical and experimental uncertainties (right panel).

where the momenta of the charged lepton, $\ell \in \{e, \mu\}$, and of the spectator quark, q', are calculated in the top quark rest frame. The normalized differential cross section as a function of $\cos \theta_{\text{pol.}}^*$ at parton level is related to the top quark spin asymmetry A_{ℓ} and to the top quark polarization *P*, as

$$\frac{1}{\sigma} \frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*_{\mathrm{pol.}}} = \frac{1}{2} \left(1 + 2A_\ell \cos\theta^*_{\mathrm{pol.}} \right), \qquad A_\ell = \frac{1}{2} P \cdot \alpha_\ell,$$

where α_{ℓ} is the so-called spin-analyzing power of the charged lepton.



Figure 4: Feynman diagram for single top quark production in the *t*-channel assuming the 4FS.

Given the *t*-channel topology, the signal signature requires one charged, isolated lepton, p_T imbalance, and two jets, one of which is b tagged. Events containing additional leptons are rejected. To classify signal and control regions, different orthogonal event categories are defined, denoted *n*jets-*m*tags, based on the number of jets and b tagged jets. The total number of jets *n* can range from 2 to 3, while the number of b tagged jets can range from 0 to 2. The 2jets-1tag category forms the signal region, while the remaining regions are used to assess the background modeling.

The number of *t*-channel single top quark events in data is determined from a ML fit to the W boson transverse mass distribution $m_T(W)$ and two BDT discriminants distributions. The first BDT, labelled BDT_{*t*-ch}, is trained to separate *t*-channel events from background events, namely t \bar{t} , W+jets and multijet events. Since the BDT_{*t*-ch} discriminant shapes for the t \bar{t} and

W+jets are found to be similar, to obtain sensitivity in the fit to both backgrounds individually, a second BDT, labelled $BDT_{t\bar{t}/W}$, is trained to classify events for these two processes only. The ML fit is performed using the following distributions, which are also reported in Figure 5:

- BDT_{*t*-ch} in the 2jets-1tag region for events with m_T(W) > 50 GeV and BDT_{*t*-ch} > 0, a region of the phase space which is enriched in signal events;
- BDT_{tī/W} in the 2jets-1tag region for events with $m_T(W) > 50$ GeV and BDT_{t-ch} < 0, a region of the phase space enriched in tī and W+jets events but depleted of signal and multijet events;
- m_T(W) in the 2jets-1tag region for events with m_T(W) < 50 GeV, a region of the phase space which is particularly sensitive to the amount of multijet events;
- m_T(W) in the 3jets-2tags region, a region of the phase space which provides additional sensitivity to the main tt background.



Figure 5: Distributions used in the ML fit to extract the signal yields. BDT_{*t*-ch} distribution in the 2jets-1tag region (upper left panel); BDT_{$t\bar{t}/W$} distribution in the 2jets-1tag region (upper right panel); m_T(W) distribution in the 2jets-1tag region (lower left panel); m_T(W) distribution in the 3jets-2tags region (lower right panel).

In order to compute differential cross sections, individual fits are performed for each observable that is later used in the unfolding. In each fit, the signal and background yields can vary independently across the unfolding intervals. The likelihood to be maximized in such fits can be expressed as a product of Poisson-like terms:

$$\ln L = \sum_{k}^{\text{region interval bins}} \sum_{j} \sum_{i}^{k} \left(d_{kji} \cdot \ln p_{kji} - p_{kji} \right),$$

where *d* denotes the number of observed events in data and *p* the estimated prediction. Summation over *k* denotes the 16 regions (4 distributions \times lepton flavor \times lepton charge), *j* denotes the interval in the unfolding observable and *i* denotes the distribution bin. The obtained yields are used to perform unfolding at parton and particle level using the TUNFOLD algorithm [5].

Several systematic uncertainties are taken into account, the uncertainty on the top quark mass being the dominant one.

As a result, the differential charge ratios are found to be in agreement with the theoretical predictions within uncertainties. Also, the 4FS prediction for the differential cross sections is found to be better in agreement with the predictions with respect to the five flavor scheme (5FS, where the b quark is expected to come from the proton), especially for the top quark and W boson transverse momenta distributions. Finally, performing a χ^2 -based fit to the normalized, differential cross section as a function of the top quark polarization angle, the top quark spin asymmetry A_{ℓ} is extracted, which is found to be

$$A_{\ell} = 0.439 \pm 0.032 \text{ (stat + exp)} \pm 0.053 \text{ (syst)},$$

which is in good agreement with the theoretical expectations. These results are summarized in Figure 6.



Figure 6: Differential charge ratio as a function of the parton-level top quark p_T (upper left panel); differential cross section as a function of the parton-level W boson p_T (upper right panel); normalized, differential cross section as a function of the parton-level top quark polarization angle used to extract A_ℓ (lower panel).

2.4 tW production cross section

The CMS Collaboration also produced the fist measurement of the tW production cross section at 13 TeV [6]. The analysis is performed using the $e^{\pm}\mu^{\mp}$ dilepton channel, in which both W bosons, either produced in association with the top quark or from the decay of the top quark, decay leptonically. Based on the tW topology (see Figure 7), this leads to a final state composed of two, oppositely charged leptons, p_T imbalance and a b tagged jet.



Figure 7: One possible Feynman diagram for the tW associated production.

Events are classified into three orthogonal categories: a signalenriched region with 1jet-1tag, and two background-dominated regions with two jets: a 2jets-1tag region and a 2jets-2tags region. A multivariate method is used to discriminate the tW signal from the main background of the measurement, namely tt. Several observables are combined into a single discriminator using a BDT technique. Two BDTs are trained, the first one, BDT_{1j-1b} in the 1jet-1tag region and the second one, BDT_{2j-1b}, in the 2jets-1tag region. The signal is extracted by performing a ML fit to the BDT distributions in the corresponding regions, and to the subleading jet p_T distribution in the 2jets-2tags region. The latter distribution, fitted in a tt-enriched region, is used to constrain this background. The distributions used in the fit are shown in Figure 8.

Several systematic uncertainties are taken into account; the dominating ones regard the trigger and lepton identification efficiencies.

The measured cross section for tW production is found to be

$$\sigma_{\rm tW} = 63.1 \pm 1.8 \; ({\rm stat}) \pm 6.4 \; ({\rm syst}) \pm 2.1 \; ({\rm lumi}) \; {\rm pb}_{\rm st}$$

which is in agreement with the SM predictions.

3. Rare single top quark production

In this section we present two new measurements by the CMS Collaboration concerning rare SM processes involving single top quark production: the fist observation of tZq associated production and the first evidence for $t\gamma q$ associated production.

3.1 First observation of tZq associated production

In order to achieve the first observation of the production of a single top quark in association with a Z boson (tZq associated production), the CMS Collaboration exploited the 2016 and 2017 data sets, corresponding to an integrated luminosity of 77.4 fb⁻¹. The analysis [7] target is the leptonic top quark and Z boson decays. Given the tZq topology (see Figure 9), the signal signature requires three charged, isolated leptons, p_T imbalance, and two jets one of which is b tagged. The Z boson candidate is formed taking a pair of opposite sign, same flavor leptons with an invariant mass lying in a mass window of 30 GeV width centered on the world-average Z boson mass. The top candidate is formed from the b tagged jet, the p_T imbalance and the lepton not forming the Z candidate.



Figure 8: BDT_{1j-1b} output (upper left panel); BDT_{2j-1b} output (upper right panel); subleading jet p_T in the 2jets-2tags region (lower panel).



Figure 9: One possible Feynman diagram for the tZq associated production.

This search crucially depends on efficiently distinguishing leptons originating from the decay of electroweak bosons from the so-called nonprompt leptons (genuine leptons produced in hadron decays, jets misidentified as leptons, etc.). The reach of previous analyses by CMS [8] was severely limited by the relative contribution from nonprompt lepton background and by the uncertainty in its prediction. Taking this into account, BDTs are set up to maximally discriminate between prompt and nonprompt leptons. The BDTs exploit the properties of the jet closest to the lepton in the $\eta - \phi$ space, the lepton relative isolation, the impact parameters with respect to the primary vertex, etc. A tight requirement is placed on the BDT output resulting in a selection efficiency improved by 12 (8) % for electrons (muons) with respect to the

previous CMS analysis. The BDTs are also used to estimate the nonprompt lepton background. Leptons that pass the aforementioned BDT selection criteria are referred to as "tight leptons", while "loose leptons" are a combined set of tight leptons and leptons that pass looser criteria. Data events containing exactly three loose leptons, where at least one of them fails to be a tight lepton, are reweighted with the probability for a loose nonprompt lepton to be a tight nonprompt lepton ("tight-to-loose" ratio method) to obtain a sample used to estimate the nonprompt lepton background.

Events are divided into three signal regions (SRs) based on the number of jets and b tagged jets. Events with a total of two or three jets, one of which is b tagged, form the SR-2/3j-1b, which

contains most of the tZq events. Events with four or more jets, one of which is b tagged, form the SR-4j-1b, while SR-2b contains events with two or more b tagged jets. In each of these categories, a dedicated BDT is trained to extract the tZq signal. A simultaneous ML fit to the BDT distributions (shown in Figure 10) is performed to measure the tZq associated production cross section.

Several systematic uncertainties are taken into account, the uncertainty on the nonprompt lepton background normalization being the dominant one.

The measured product of the cross section and branching fraction for the tZq associated production is found to be

$$\sigma_{tZq}\mathfrak{B}(Z \to \ell^+ \ell^-) = 111 \pm 13(\text{stat})^{+11}_{-9}(\text{syst}) \text{ fb},$$

in agreement with the SM prediction. The significance of this measurement is well above five standard deviations, resulting in the first observation of the tZq associated production.



Figure 10: BDT output in the SR-2/3j-1b (upper left panel); SR-4j-1b (upper right panel); SR-2b (lower panel).

3.2 First evidence for $t\gamma q$ associated production



Finally, the CMS Collaboration produced the first evidence for the production of a single top quark in association with a photon (t γ q associated production). The analysis [9] targets the muon decay channel for the top quark since it enables a good signal selection efficiency with low background contamination. The final state includes possible contributions from $W \rightarrow \tau v_{\tau}$ where the τ decays to $\mu v_{\mu} v_{\tau}$. Given the t γ q topology (see Figure 11), the signal signature is one isolated muon, p_{T} imbalance, one isolated photon, and two jets one of which is b tagged. Events containing additional muons are rejected.

Figure 11: One possible Feynman diagram for the $t\gamma q$ associated production.

In order to separate signal from background, a multivariate classification using a BDT is performed. Important variables used to achieve this are the η of the light flavor jet, the η

of che muon candidate, the distance in the space $\eta - \phi$ between the light flavor jet and the photon, etc. The BDT output distribution is parameterized as

$$F(x) = C_{\text{sig}}S_{\text{sig}}(x) + C_{\text{t\bar{t}}}S_{\text{t\bar{t}}}(x) + C_{W\gamma j}S_{W\gamma j}(x) + C_{Z\gamma j}S_{Z\gamma j}(x) + C_{\text{misid}}S_{\text{misid}}(x) + C_{\text{B}}S_{\text{B}}(x),$$

where x is the BDT output, $S_{sig}(x)$, $S_{t\bar{t}}(x)$, $S_{W\gamma j}(x)$, $S_{Z\gamma j}(x)$, $S_{misid}(x)$, and $S_B(x)$ are the templates for signal, $t\bar{t}+\gamma$, $W\gamma+jets$, $Z\gamma+jets$, misidentified photon backgrounds and other minor backgrounds, while the quantities C_i are the corresponding normalizations. The different templates $S_i(x)$ are taken from simulation except for $t\bar{t}+\gamma$ and misidentified photons. The former is estimated from data using a control region defined by requiring exactly two b tagged jets and keeping all the other selection criteria the same; the latter is estimate from data measuring the p_T -dependent probability for a jet to be reconstructed as a photon. In order to extract the signal cross section, a simultaneous ML fit is performed on the BDT distributions in the signal region and in the $t\bar{t}+\gamma$ control region (see Figure 12).

Several systematic uncertainties are taken into account, the uncertainties on jet energy scale (JES), jet energy resolution (JER) and signal modelling being the dominant ones.

The measured product of the cross section and branching fraction for the $t\gamma q$ associated production is found to be

$$\sigma_{t\gamma q}\mathfrak{B}(t \to \mu \nu_{\mu} b) = 115 \pm 17(\text{stat}) \pm 30(\text{syst}) \text{ fb},$$

in agreement with the SM prediction. An excess of events above the expected background is observed with a *p*-value of 4.27×10^{-6} , corresponding to a significance of 4.4 standard deviations, making this measurement the first evidence for the t γ q associated production.

4. Conclusions

The CMS Collaboration has performed several single top quark measurements, both inclusively and differentially, as well as searches for rare SM processes. All the measurements are in agreement with the SM predictions within uncertainties. Most of the analyses are still systematicsdominated, but an improvement in the experimental uncertainties is expected exploiting the full Run2 data set and more.



Figure 12: BDT output distribution for data and SM predictions in the signal region.

References

- [1] CMS Collaboration, *Measurement of the single top quark and antiquark production cross section in the t channel and their ratio in proton-proton collisions at* $\sqrt{s} = 13$ TeV, submitted to PLB, arXiv:1812.10514.
- [2] ATLAS and CMS Collaborations, Combination of single-top-quark production cross section measurements and $|f_{LV}V_{tb}|$ determinations at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments, JHEP 05 (2019) 088, arXiv:1902.07158.
- [3] ATLAS Collaboration, Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector, Eur. Phys. J. C 77 (2017) 531, arXiv:1702.02859.
- [4] CMS Collaboration, Measurement of differential cross sections and charge ratios for t-channel single top quark production at 13 TeV, CMS-PAS-TOP-17-023, 2017.
- [5] S. Schmitt, TUnfold: An algorithm for correcting migration effects in high energy physics, JINST 7 (2012) T10003, arXiv:1205.6201.
- [6] CMS Collaboration, Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV, JHEP **10** (2018) 117, arXiv:1805.07399.
- [7] CMS Collaboration, Observation of single top quark production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **122** (2019) 132003, arXiv:1812.05900.
- [8] CMS Collaboration, Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV, Phys. Lett. B **779** (2018) 358, arXiv:1712.02825.
- [9] CMS Collaboration, Evidence for the associated production of a single top quark and a photon in proton-proton collisions at $\sqrt{s} = 13$ TeV, PRL **121** (2018) 121802, arXiv:1808.02913.