

Recent CMS results on the Higgs boson produced in association with top quarks

Karim El Morabit* on behalf of the CMS Collaboration

Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

E-mail: karim.el.morabit@cern.ch

The associated production of a Higgs boson with a top quark-antiquark pair ($t\bar{t}H$) or with a single top quark (tH) allows the direct measurement of the top-Higgs Yukawa coupling. In this article, recent searches for $t\bar{t}H$ and tH production at the CMS experiment are reviewed. The analyses use pp collision data collected at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity of up to 77.4 fb^{-1} .

Sixth Annual Conference on Large Hadron Collider Physics (LHCP2018)

4-9 June 2018

Bologna, Italy

*Speaker.

1. Introduction

The probing of the coupling of the Higgs (H) boson to top (t) quarks is an important step towards the understanding of the Higgs boson. In the Standard Model (SM), the coupling is a Yukawa-type interaction with the coupling strength y_t proportional to the top quark mass. Indirect constraints on y_t can be derived from measurements of processes including virtual top quark loops such as Higgs boson production via gluon-gluon fusion. A direct measurement of y_t , on the other hand, is possible via the associated production of a Higgs boson with a top quark-antiquark pair ($t\bar{t}H$) or with a single top quark (tH), as illustrated in Fig. 1. While the $t\bar{t}H$ process is sensitive to the magnitude of y_t , the tH process is also sensitive to the sign of y_t . The dominant contributions to tH production at leading order are t-channel and associated tW production. The amplitudes for both contributions interfere depending on the couplings between the Higgs boson and the top quark or the W boson. The couplings can be expressed in terms of the coupling strengths κ_t and κ_V . In the SM, $\kappa_t = \kappa_V = 1$, resulting in destructive interference. In case of the inverted top quark coupling scenario (ITC), where $\kappa_t = -1$ and $\kappa_V = 1$, the interference is constructive. Since the SM cross sections for $t\bar{t}H$ and tH production are relatively small, approximately 500 fb and 90 fb at a center-of-mass energy of 13 TeV, respectively, this is a difficult measurement.

A top quark decays almost exclusively to a bottom quark and a W boson which subsequently decays into either a quark-antiquark pair or a lepton and a neutrino. In the following, top quark decays will be classified according to the decays of the W bosons. The Higgs boson has various decay modes including decays to bottom quark-antiquark pairs, pairs of τ leptons, pairs of photons or combinations of leptons and quarks from intermediate W or Z bosons. In order to counteract the small cross sections, searches are performed in various final states. The large number of final-state particles requires the use of sophisticated multivariate analysis methods. Unless stated otherwise, the analyses described below are performed using 35.9 fb^{-1} of data recorded in 2016 at a center-of-mass energy of 13 TeV. Furthermore, the term lepton will always stand for electrons or muons.

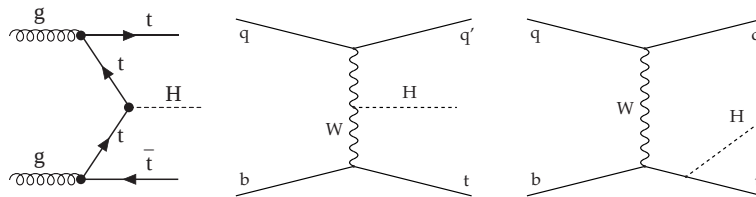


Figure 1: Examples of leading-order Feynman diagrams for $t\bar{t}H$ production (left) and tH production (center, right) [1, 2].

2. Searches for $t\bar{t}H$ production

Searches for $t\bar{t}H$ are performed in final states targeting leptonic and hadronic top quark decays and Higgs boson decays to pairs of bottom quarks, photons, τ leptons and W or Z bosons.

In the analysis targeting semileptonically and dileptonically decaying top quark-antiquark pairs and Higgs boson decays to bottom quark pairs ($H \rightarrow b\bar{b}$), events are required to contain

≥ 4 jets of which ≥ 3 have to be identified as coming from b quarks (b-tagged) and exactly one or two oppositely-charged isolated leptons [3]. The background consists almost entirely of $t\bar{t}$ +jets events with an irreducible component where the additional jets originate from heavy-flavor quarks ($t\bar{t}$ +HF), such as $t\bar{t}+b\bar{b}$. The expected $t\bar{t}H$ and $t\bar{t}$ +jets contributions are predicted from simulation. The simulated $t\bar{t}$ +jets events are split into $t\bar{t}$ +light-flavor, $t\bar{t}+c\bar{c}$ and 3 different $t\bar{t}$ +HF subprocesses based on the flavor of the additional jets in acceptance. Since the $t\bar{t}$ +HF and $t\bar{t}+c\bar{c}$ subprocesses are difficult to model, they are assigned an additional 50% normalization uncertainty. To discriminate between signal and background, events are first divided into categories in lepton, jet- and b-tag-multiplicity. In the dilepton channel, a combination of boosted decision trees (BDT), trained in each category to separate $t\bar{t}H$ from $t\bar{t}$ +jets, and a matrix-element method (MEM), constructed to separate $t\bar{t}H$ from $t\bar{t}+b\bar{b}$, are used as final discriminants. In the lepton+jets channel deep neural networks (DNN), with the aforementioned MEM as input, are trained in each category to classify the events as $t\bar{t}H$ or one of the $t\bar{t}$ +jets subprocesses. The lepton+jets categories are then split into subcategories according to their most-probable process, determined by the DNN. In each of these subcategories, the DNN output corresponding to the most-probable process is used as final discriminant. The signal strength $\mu = \sigma/\sigma_{SM}$ is extracted in a simultaneous fit of the discriminant shapes in all categories resulting in a best-fit value of $\mu = 0.72 \pm 0.24(\text{stat.}) \pm 0.38(\text{syst.})$ with a total uncertainty of ± 0.45 corresponding to an observed (expected) significance of 1.6σ (2.2σ) with respect to the background-only hypothesis. Systematic uncertainties are dominated by uncertainties related to b-tagging and the $t\bar{t}$ +HF modelling.

In the analysis targeting hadronically decaying top quarks and $H \rightarrow b\bar{b}$ decays, events are required to contain no isolated lepton and ≥ 7 jets of which ≥ 2 have to be b-tagged [1]. Furthermore, events are required to have $H_T > 500$ GeV, where H_T is defined as the scalar sum of the transverse momenta of all jets. In addition, events are required to contain at least one pair of non-b-tagged jets with an invariant mass inside a window around the W boson mass. After this event selection, the background is dominated by $t\bar{t}$ +jets and QCD-multijet production. The background from QCD-multijet production is further suppressed using a quark-gluon likelihood ratio (QGL) constructed to separate jets from the hadronization of quarks, such as those coming from W boson decays, and those from the hadronization of gluons, dominant in multijet interactions and initial state radiation. All selected events with ≥ 3 b-tagged jets are categorized into 6 signal region (SR) categories according to their jet- and b-tag-multiplicities. Events with 2 b-tagged jets are used to define control regions (CR) used in the estimation of the dominant QCD-multijet background. The final discriminant is a MEM constructed to separate $t\bar{t}H$ and $t\bar{t}+b\bar{b}$ events. The expected $t\bar{t}H$ and $t\bar{t}$ +jets contributions are predicted from simulation. The shape of the discriminant for the QCD-multijet background is estimated from the final discriminant distributions of data in the control region categories after subtraction of the $t\bar{t}$ +jets and other minor backgrounds. Differences stemming from the difference in b-tag multiplicity in the SR and CR and differences in the kinematic properties of b-jets and non-b-jets are accounted for with corrections derived from data. The normalization of the QCD-multijet background is left freely floating in the final fit in the SR. The best-fit value of the signal strength is found to be $\mu = 0.9 \pm 0.7(\text{stat.}) \pm 1.3(\text{syst.})$ with a total uncertainty of ± 1.5 . An observed (expected) upper exclusion limit at the 95% confidence level (CL) is calculated as $\mu < 3.8$ (3.1). Systematic uncertainties are dominated by uncertainties related to b-tagging, the $t\bar{t}$ +HF modelling, the jet energy scale calibration and the QCD-multijet

estimation.

In the analysis targeting Higgs boson decays to pairs of W bosons, Z bosons and τ leptons ($H \rightarrow \tau\tau/WW/ZZ$), events are selected to be consistent with one lepton coming from a decaying top quark while additional leptons and τ come from the decays of the Higgs boson or the other top quark [4]. For this requirements on the jet, b-tagged jet, lepton and identified hadronically decaying τ lepton (τ_h) multiplicities are imposed. Events with four leptons with an invariant mass around the Higgs boson mass are vetoed to be orthogonal to a dedicated analysis described below. Selected events are subsequently sorted into six categories of lepton charges as well as lepton and τ_h multiplicities. Reducible backgrounds mostly arise from misidentification of leptons or τ_h or mismeasurement of lepton charges. They are estimated from data by using fake-factor methods applied to dedicated control regions. Irreducible backgrounds consist mostly of $t\bar{t}+V$ and VV +jets production and are estimated from simulation. To separate signal and background, dedicated discriminating variables are constructed in five of the six analysis categories, while in the category with four leptons event counting is performed. Depending on the category the discriminants are a combination of a MEM, constructed to separate the signal from $t\bar{t}+Z$ and $t\bar{t}$ events, and BDTs trained to separate $t\bar{t}H$ from $t\bar{t}$ +jets or $t\bar{t}+V$. The best-fit value of the signal strength is found to be $\mu = 1.23_{-0.43}^{+0.45} [_{-0.25}^{+0.26}(\text{stat.})_{-0.35}^{+0.37}(\text{syst.})]$ corresponding to an observed (expected) significance of 3.2σ (2.8σ) with respect to the background-only hypothesis. The systematic uncertainties are dominated by uncertainties related to jet energy calibration, the τ_h identification, the non-prompt lepton estimation as well as the modelling of the $t\bar{t}H$ and $t\bar{t}+W/Z$ processes.

The analysis of $t\bar{t}H$ with Higgs bosons decaying to a pair of photons is part of an inclusive $H \rightarrow \gamma\gamma$ analysis [5]. In the $t\bar{t}H$ channel of this analysis, events are required to contain 2 non-prompt photons consistent with the signal hypothesis and ≥ 2 jets of which ≥ 1 have to be b-tagged. Selected events are also required to either pass a cut on the output of a BDT trained to identify $t\bar{t}H$ events or to contain at least one lepton. The signal strength is extracted by fitting the invariant diphoton mass distributions expected from $t\bar{t}H$ and the background prediction to data. The $t\bar{t}H$ distribution is predicted using simulated events. The background model is taken from a parametric fit to the data distribution. The fit of all analysis categories (including the non- $t\bar{t}H$ categories) results in a best-fit value for the $t\bar{t}H$ signal strength modifier of $\mu_{t\bar{t}H} = 2.2_{-0.8}^{+0.9}$.

The search for $t\bar{t}H$ with Higgs bosons decaying to four leptons via a pair of intermediate Z bosons is part of an inclusive $H \rightarrow ZZ^* \rightarrow 4l$ analysis using 41.5 fb^{-1} of data recorded in 2017 at a center-of-mass energy of 13 TeV [7]. In the $t\bar{t}H$ channel, selected events are required to contain 2 pairs of isolated, well-identified oppositely charged leptons (electrons or muons) compatible with coming from a Higgs boson decay and ≥ 4 jets of which ≥ 1 has to be b-tagged. The predicted signal and the irreducible background from ZZ production are taken from simulation while the reducible background from misidentified leptons is derived from control regions in data. The signal strength is extracted by fitting the invariant four-lepton mass distribution expected from $t\bar{t}H$ and the background prediction to data. The combination with the analogous analysis performed using 35.9 fb^{-1} of data recorded in 2016 results in $\mu = 0.0_{-0.0}^{+0.5}$ [6].

The discussed analyses using 35.9 fb^{-1} of data recorded in 2016 at a center-of-mass energy of 13 TeV are combined with analogous analyses from LHC Run-I using up to 5.1 fb^{-1} at 7 TeV and up to 19.7 fb^{-1} at 8 TeV [8]. In this combination, systematic uncertainties are correlated among the individual analyses where appropriate. Furthermore, non- $t\bar{t}H$ Higgs boson process are treated

as background processes normalized to their SM prediction. The combined fit results in a best-fit value for the signal strength of $\mu = 1.26_{-0.26}^{+0.31} = 1.26_{-0.16}^{+0.16}(\text{stat})_{-0.15}^{+0.17}(\text{exp})_{-0.13}^{+0.14}(\text{th.bkg})_{-0.07}^{+0.15}(\text{th.sig})$. This constitutes the first observation of $t\bar{t}H$ production with an observed (expected) significance of 5.2σ (4.2σ).

3. Searches for tH production

Two searches for tH production are performed targeting final states with the top quark decaying leptonically and either $H \rightarrow \tau\tau/\text{WW}/\text{ZZ}$ or $H \rightarrow b\bar{b}$ decays.

In the $H \rightarrow \tau\tau/\text{WW}/\text{ZZ}$ analysis selected events are required to contain ≥ 1 b-tagged jets, ≥ 1 non-b-tagged jets and exactly two same-sign leptons or exactly three leptons. The background mostly consists of irreducible contributions from $t\bar{t}+W/Z$, ZZ and WZ production, estimated from simulation, and reducible contributions from events with non-prompt leptons or leptons with misidentified charges, which are estimated from control regions in data. To discriminate between signal and background, two BDTs are trained separately for the categories with two or three lepton events. The BDTs are trained to either separate tHq from $t\bar{t}$ +jets or $t\bar{t}+W/Z$ events and are subsequently combined in a one-dimensional discriminant.

In the $H \rightarrow b\bar{b}$ analysis, events in the SR are required to contain one lepton and ≥ 4 jets of which ≥ 3 have to be b-tagged [9]. The background is dominated by $t\bar{t}$ +jets and is estimated from simulation. A CR enriched in dileptonically decaying $t\bar{t}$ +jets events is constructed to constrain uncertainties related to the modelling of the $t\bar{t}$ +jets prediction in the final fit to data. In the SR, the discrimination between signal and background is achieved by first reconstructing events under the tHq , tHW and $t\bar{t}$ hypothesis with the help of dedicated BDTs. Next, BDTs using features derived from these event reconstructions are used as final discriminants separating between signal and background.

In both analyses, exclusion limits at the 95% CL on the combined production cross sections for tHq , tHW and $t\bar{t}H$ production times the branching ratios for the targeted Higgs boson decay modes are calculated as a function of the ratio of the coupling strength modifiers κ_t/κ_V . The results for both analyses are shown in Fig. 2. In the $H \rightarrow \tau\tau/\text{WW}/\text{ZZ}$ analysis values of κ_t/κ_V outside of $(-1.25, +1.6)$ are excluded at the 95% CL. Observed upper limits on the combined signal strength of $\mu_{tHq+tHW+t\bar{t}H} < 3.1$ and $\mu_{tHq+tHW+t\bar{t}H} < 1.4$ are found for the SM and ITC case, respectively. The best-fit value of the combined signal strength in the SM case is found to be $\mu_{tHq+tHW+t\bar{t}H} = 1.8 \pm 0.3(\text{stat.}) \pm 0.6(\text{syst.})$ corresponding to an observed (expected) significance of 2.7σ (1.5σ) with respect to the background-only hypothesis. For the ITC case $\mu_{tHq+tHW+t\bar{t}H} = 0.7 \pm 0.4$ corresponding to an observed (expected) significance of 1.7σ (2.5σ) is found. In the $H \rightarrow b\bar{b}$ analysis, observed (expected) upper limits on the tH signal strength of $\mu_{tH} < 89.5$ (41.4) and $\mu_{tH} < 5.83$ (2.94) are found for the SM and ITC cases, respectively.

4. Summary

Recent searches by CMS for $t\bar{t}H$ and tH production at 13 TeV using up to 77.4 fb^{-1} of data are presented. A combination of different searches at 13 TeV with searches at 7 TeV and 8 TeV results in the observation of $t\bar{t}H$ production with an observed (expected) significance of 5.2σ (4.2σ). A

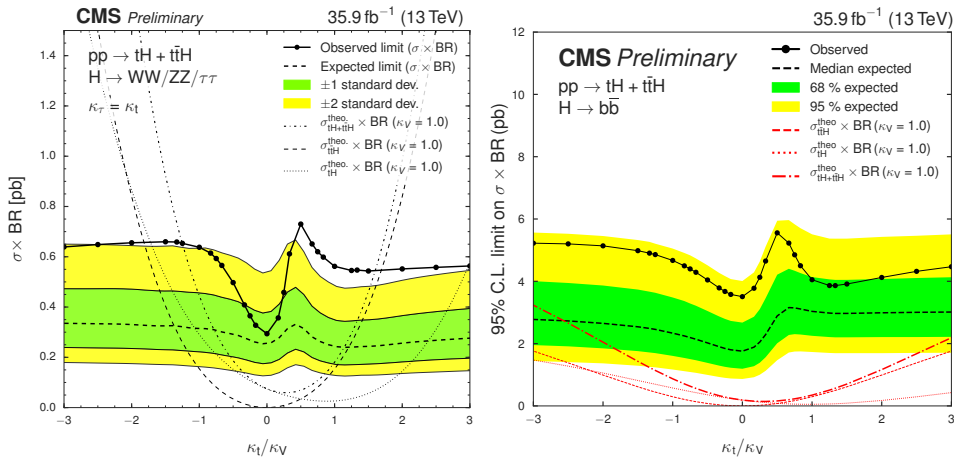


Figure 2: Observed and expected 95% CL upper limit on the $tH+t\bar{t}H$ cross section times the Higgs boson decay branching fractions for the analyses targeting Higgs bosons decays to WW , ZZ or $\tau\tau$ (left) and pairs of bottom quark (right) for different values of the coupling ratio κ_t/κ_V [2, 9].

95% CL upper limit of $\mu_{tH} < 5.83$ for tH production with inverted top quark couplings is observed in a search targeting $H \rightarrow b\bar{b}$ decays. In a search for tH production with $H \rightarrow \tau\tau/WW/ZZ$ decays a 95% CL upper limit of $\mu_{tHq+tHW+t\bar{t}H} < 3.1$ is observed.

References

- [1] A.M. Sirunyan et al. , *Search for $t\bar{t}H$ production in the all-jet final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* 06 (2018) 101, [[arXiv:1803.06986](https://arxiv.org/abs/1803.06986) [hep-ex]]
- [2] CMS Collaboration, *Search for production of a Higgs boson and a single top quark in multilepton final states in proton collisions at $\sqrt{s} = 13$ TeV*, CMS-PAS-HIG-17-005, <https://cds.cern.ch/record/2264553>
- [3] A.M. Sirunyan et al. , *Search for $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV*, [[arXiv:1804.03682](https://arxiv.org/abs/1804.03682) [hep-ex]], *subm. to JHEP*
- [4] A.M. Sirunyan et al. , *Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV*, *JHEP* 08 (2018) 066, [[arXiv:1803.06986](https://arxiv.org/abs/1803.06986) [hep-ex]]
- [5] A.M. Sirunyan et al. , *Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, [[arXiv:1804.02716](https://arxiv.org/abs/1804.02716) [hep-ex]], *subm. to JHEP*
- [6] A.M. Sirunyan et al. , *Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* 11 (2017) 047, [[arXiv:1706.09936](https://arxiv.org/abs/1706.09936) [hep-ex]]
- [7] CMS Collaboration, *Measurements of properties of the Higgs boson in the four-lepton final state at $\sqrt{s} = 13$ TeV*, CMS-PAS-HIG-18-001, <https://cds.cern.ch/record/2621419>
- [8] A.M. Sirunyan et al. , *Observation of $t\bar{t}H$ production*, *Phys. Rev. Lett.* 120 (2018) 231801, [[arXiv:1804.02610](https://arxiv.org/abs/1804.02610) [hep-ex]]
- [9] CMS Collaboration, *Search for the $tH(H \rightarrow b\bar{b})$ process in pp collisions at $\sqrt{s} = 13$ TeV and study of Higgs boson couplings*, CMS-PAS-HIG-17-016, <https://cds.cern.ch/record/2621422>