PROCEEDINGS OF SCIENCE

Measurement of $t\bar{t}H$ and tH production in the $H(b\bar{b})$ channel at CMS

Valeria Botta[a,*] **on behalf of the CMS Collaboration**

[a] *RWTH Aachen University, I. Physikalisches Institut B*
Sommerfeldstr. 14, 52074 Aachen, Germany

E-mail: valeria.botta@cern.ch

This poster presents a measurement of the associated production of the Higgs boson (H) with a top quark anti-quark pair ($t\bar{t}H$) or a single top quark (tH) in the final state where the Higgs boson decays to a bottom quark anti-quark pair. Data from proton-proton collisions at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity of 138 fb^{-1} and collected by the CMS experiment at the CERN LHC between 2016 and 2018 are analysed. The observed inclusive $t\bar{t}H$ production rate relative to the standard model (SM) expectation is 0.33 ± 0.26, and a differential measurement as a function of the Higgs boson transverse momentum p_T is also performed. An upper limit on the tH production rate of 14.6 times the SM expectation is set at the 95% confidence level, with an expectation of $19.3^{+9.2}_{-6.0}$. Finally, constraints are set on the strength and CP structure of the coupling between the Higgs boson and the top quark.

The European Physical Society Conference on High Energy Physics (EPS-HEP2023)
21-25 August 2023
Hamburg, Germany

*Speaker

https://pos.sissa.it/

PoS(EPS-HEP2023)424

1. Introduction

The associated production of a Higgs boson with top quarks provides a direct probe of the top-Higgs Yukawa coupling, an essential parameter in testing the Higgs sector of the SM. Recently, $t\bar{t}H$ production has been observed by both the ATLAS and the CMS Collaborations in the combination of several Higgs boson decay modes [2, 3]. The analysis presented in this poster is the latest measurement of the $t\bar{t}H$ and tH production rates in the final state $H \to b\bar{b}$, and is based on 138 fb^{-1} of data from proton-proton collisions collected between 2016 and 2018 by the CMS experiment [1].

2. Analysis strategy

The analysis covers all final states of the $t\bar{t}$ system, including events with zero, one, or two isolated leptons (either electrons or muons). A preselection based on the number of jets and b-tagged jets is performed in each channel, requiring ≥7 jets, ≥4 b-tags in the fully-hadronic channel, ≥5 jets, ≥4 b-tags in the single-lepton channel, and ≥3 jets, ≥3 b-tags in the dileptonic channel. The largest background in the fully-hadronic channel arises from QCD multijet production, and is estimated from data. In the leptonic channels, instead, the largest background arises from $t\bar{t}$ production in association with jets, and is estimated from simulation with data-driven corrections.

Special care is needed to handle the irreducible $t\bar{t}$+b-jets component, which is notoriously difficult to model theoretically. Simulated $t\bar{t}$ events are classified at particle level according to the flavour of jets not arising from the $t\bar{t}$ decay and fulfilling the acceptance requirements $p_T > 20$ GeV and $|\eta|$<2.4. Such classification defines three orthogonal sets of $t\bar{t}$ events: the $t\bar{t}B$ component comprises events with at least one jet containing B hadrons, the $t\bar{t}C$ component events with at least one jet containing C hadrons and no jets containing B hadrons, while the $t\bar{t}$+light flavour jets ($t\bar{t}LF$) component includes all remaining events. Furthermore, among the $t\bar{t}B$ events, those with exactly two additional B hadrons close enough to be clustered inside one jet are denoted as $t\bar{t} + 2b$, targeting the regime of collinear gluon splitting, which is subject to different systematic uncertainties. The $t\bar{t}B$ background is modelled using a novel, state-of-art simulation of $t\bar{t}$+b-jets at NLO in the four flavour scheme (4 FS), produced with the POWHEG-BOX-RES event generator [4–6]. Other $t\bar{t}$ events, namely $t\bar{t}C$ and $t\bar{t}LF$, are simulated at NLO in the five flavour scheme (5 FS) using the POWHEG Monte-Carlo generator [7]. Prior to the fit, the $t\bar{t}B$ component is scaled to its expected yield from the 5 FS simulation, and the final normalisation factors for the $t\bar{t}B$ and $t\bar{t}C$ components are determined by the fit to the data. The background modelling has been extensively validated in data, and the impact of potential mismodellings of the $t\bar{t}B$ background on the extracted signal has been studied using pseudo experiments and found to be well covered by the systematic uncertainties. The uncertainty model used in the analysis is described in detail in Ref. [8].

After the preselection, events are categorised according to the jet and b-jet multiplicity, as depicted in Fig. 1. Further separation of signal and background events relies on feedforward artificial neural networks (ANN). In the least sensitive categories, a binary ANN classifier is used to distinguish between signal and background. In the most sensitive categories, instead, an ANN multi-classifier is trained to separate the $t\bar{t}H$ and tH signals and the different $t\bar{t}$ background components, and events are assigned to the class for which the output score is largest. The observables used to perform the final fit have been chosen carefully to optimise the sensitivity of the analysis. In particular, due the difficult separation of the $t\bar{t}H$ and $t\bar{t}B$ processes, events entering the latter categories are considered together, and a likelihood ratio observable R is built from the corresponding ANN output scores. In the control regions, instead, only the event yield is used to determine the normalisation of the $t\bar{t}B$ and $t\bar{t}C$ processes. In all other regions, the ANN output score distribution is used as input to the fit.

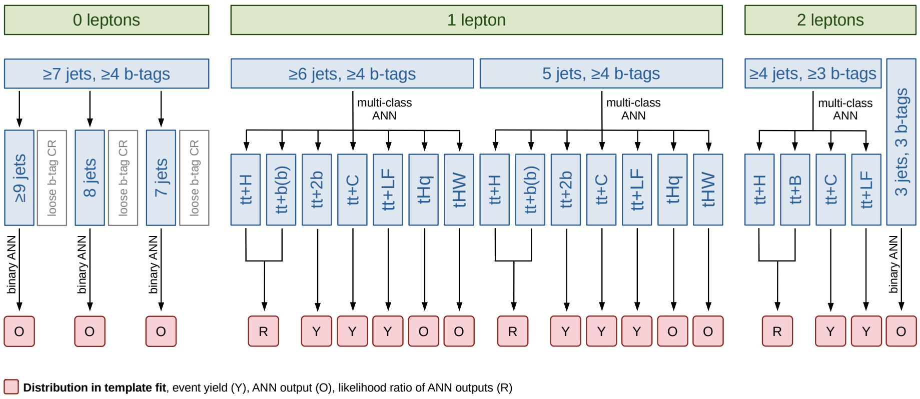

Figure 1: Overview of the analysis categories and fitted observables used for the measurement of the inclusive $t\bar{t}H$ and tH signal strengths [8].

The $t\bar{t}H$ signal strength is also measured differentially in five bins of the Higgs boson p_T, according to the scheme defined by the simplified template cross section (STXS) framework. In the fully-hadronic channel the Higgs boson p_T is reconstructed using a χ^2 method, while in the leptonic channels a dedicated multi-class ANN classifier trained on signal events is employed to assign events in the signal region to one of the defined Higgs p_T bins. The categorisation efficiencies range from 35% to 85%, depending on the category and p_T bin. The ANN output value in the node with the highest score, multiplied by the likelihood ratio value, is used as final discriminating observable in the fit.

3. Results

Signal yields are extracted from a binned maximum-likelihood fit performed simultaneously in all categories. The $t\bar{t}B$ normalisation is found to be $1.19^{+0.13}_{-0.12}$ times its prefit value, in agreement with dedicated measurements [9], while the $t\bar{t}C$ normalisation parameter is fitted to $1.07^{+0.20}_{-0.19}$. The best-fit $t\bar{t}H$ signal strength is $\mu_{t\bar{t}H} = 0.33 \pm 0.26 = 0.33 \pm 0.17\ \text{(stat.)} \pm 0.21\ \text{(syst.)}$, corresponding to a significance of 1.3σ (4.1σ expected), and is compatible with the SM prediction at the level of 2.4σ. The left plot of Fig. 2 shows the observed likelihood-ratio test statistic as a function of $\mu_{t\bar{t}H}$ and the $t\bar{t}B$ background normalisation. When fixing the $t\bar{t}H$ production rate to the SM prediction, the 95% CL upper limit on the tH signal strength is 14.6 ($19.3^{+9.2}_{-6.0}$ expected). The results of the STXS measurement are shown in Fig. 2 (right) and are fully compatible with the inclusive measurement. The results are interpreted in the κ-framework [10] to constrain the coupling strength modifiers for the Higgs boson coupling to top quarks, κ_t, and to vector bosons, κ_V. Furthermore, the CP structure of the top-Higgs coupling is probed for a potential CP-odd component, according to the model adopted in Ref. [11]. In this case, κ_V is fixed to the SM expectation. For both interpretations, the results are compatible with the SM expectation at the level of 2σ.

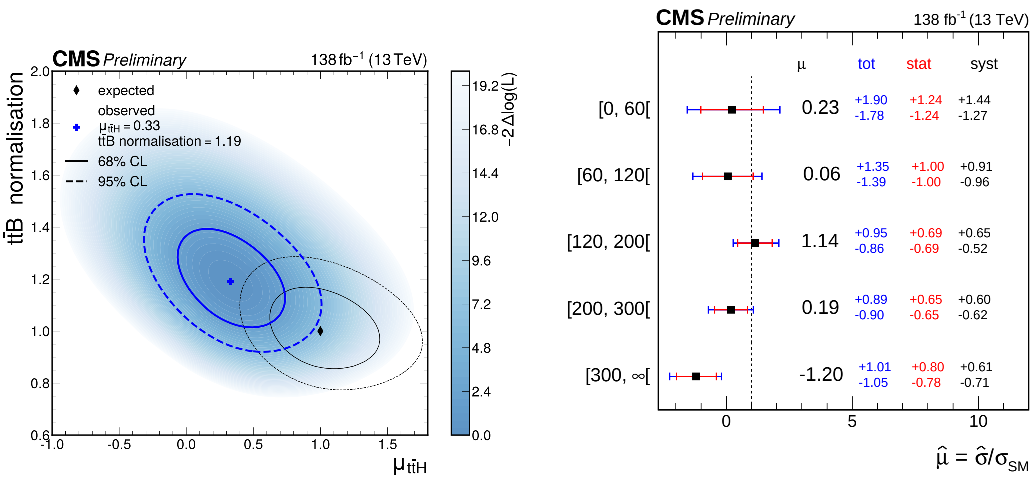

Figure 2: Observed likelihood ratio test statistic as a function of the $t\bar{t}H$ and of the $t\bar{t}B$ background normalisation (left), and observed $t\bar{t}H$ signal strength in bins of the Higgs boson p_T defined in the STXS framework (right) [8].

References

[1] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) S08004.

[2] ATLAS Collaboration, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector", Phys. Lett. B **784**(2018) 159.

[3] CMS Collaboration, "Observation of $t\bar{t}H$ production", Phys. Rev. Lett. **120** (2018) 231801.

[4] T. Jezo et al., "New NLOPS predictions for $t\bar{t}$+b-jet production at the LHC", Eur. Phys. J. C **78** (2018) 502.

[5] F. Cascioli et al., "NLO matching for $t\bar{t}b\bar{b}$ production with massive b-quarks", Phys. Lett. B **734** (2014) 210.

[6] F. Buccioni et al., "NLO QCD predictions for $t\bar{t}b\bar{b}$ production in association with a light jet at the LHC", JHEP **12** (2019) 015.

[7] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", JHEP **11** (2007) 070.

[8] CMS Collaboration, "Measurement of the $t\bar{t}H$ and tH production rates in the $H \to b\bar{b}$ decay channel with 138 fb^{-1} of proton-proton collision data at $\sqrt{s}$=13 TeV", CMS-PAS-HIG-19-011.

[9] CMS Collaboration, "Inclusive and differential cross section measurements of $t\bar{t}b\bar{b}$ production in the lepton+jets channel at $\sqrt{s}$ = 13 TeV with the CMS detector", CMS-PAS-TOP-22-009.

[10] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs cross sections: 3. Higgs properties", CERN-2013-004.

[11] A. V. Gritsan et al., "Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques", Phys. Rev. D **94** (2016) 055023.