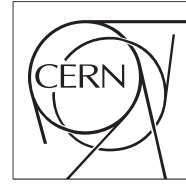**The Compact Muon Solenoid Experiment**

Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

28 November 2016 (v2, 04 December 2016)

Searches for ttH and tH with H → bb̄

Matti Schroeder for the ATLAS, CMS and Higgs collaborations.

Abstract

The associated production of a Higgs boson with a top quark-antiquark pair (ttH production) or with a single top quark (tH production) allows a direct measurement of the top-Higgs-Yukawa coupling with minimal model dependence. In this article, recent results of searches for ttH and tH production in the H → bb̄ channel performed by the ATLAS and CMS experiments are reviewed. The analyses use pp collision data collected at a centre-of-mass energy of $13\,\mathrm{TeV}$ corresponding to an integrated luminosity of up to $13.2\,\mathrm{fb}^{-1}$.

Presented at *TOP2016 9th International Workshop on Top Quark Physics*

Searches for t$\bar{\text{t}}$H and tH with H $\rightarrow$ b$\bar{\text{b}}$

MATTHIAS SCHRÖDER

ON BEHALF OF THE ATLAS AND CMS COLLABORATIONS

Karlsruher Institut für Technologie (KIT)
Institut für Experimentelle Kernphysik
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen, Germany

The associated production of a Higgs boson with a top quark-antiquark pair (t$\bar{\text{t}}$H production) or with a single top quark (tH production) allows a direct measurement of the top-Higgs-Yukawa coupling with minimal model dependence. In this article, recent results of searches for t$\bar{\text{t}}$H and tH production in the H $\rightarrow$ b$\bar{\text{b}}$ channel performed by the ATLAS and CMS experiments are reviewed. The analyses use pp collision data collected at a centre-of-mass energy of 13 TeV corresponding to an integrated luminosity of up to 13.2 fb^{-1}.

PRESENTED AT

9th International Workshop on Top Quark Physics
Olomouc, Czech Republic, September 19–23, 2016

1 Introduction

The coupling of the Higgs (H) boson to the top (t) quark is of particular interest. In the Standard Model (SM), it is of Yukawa type with a strength y_t proportional to the t-quark mass, and hence, it is exceptionally large. Thus, y_t contributes dominantly to various loop processes both in the SM and also in models of new physics. The current value of y_t is dominated by indirect constraints derived from measurements of the gluon-gluon fusion H-boson production and the $H \to \gamma\gamma$ decay rate and depends on the assumption of absence of new particle contributions to the loop amplitudes.

The associated production of a H boson with a t quark-antiquark pair ($t\bar{t}H$ production) or with a single t quark (tH production), on the other hand, allows a direct measurement of y_t with minimal model dependence, cf. Fig. 1. However, the SM cross section of both processes is relatively small with approximately 0.5 pb and 90 fb at 13 TeV centre-of-mass energy, respectively, making this a difficult measurement. The bottom (b) quark-antiquark final state of the H boson benefits from a large branching ratio. At the same time, the relatively poor jet-energy resolution and the huge combinatorial uncertainty in the event reconstruction require to use multivariate analysis methods to discriminate signal from background processes, where the signal cross section is determined with a fit of the discriminant distributions to the data.
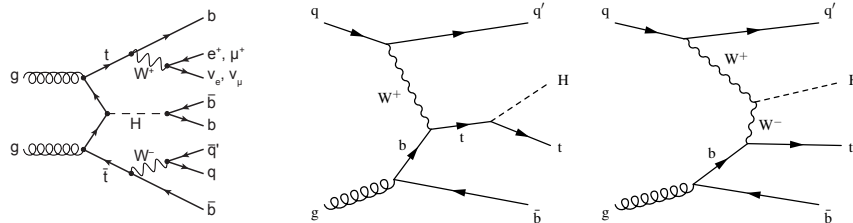

Figure 1:   Example leading-order Feynman diagrams contributing to the lepton+jets-channel $t\bar{t}H(b\bar{b})$ (left) and the t-channel $tH(b\bar{b})$ production (centre, right) [1, 2].

2 Searches for $t\bar{t}H(b\bar{b})$ production

Both CMS [3] and ATLAS [4] have performed searches for $t\bar{t}H$, $H \to b\bar{b}$, production in the dilepton and lepton+jets final states of the $t\bar{t}$ system at 13 TeV centre-of-mass energy: CMS has published the first analysis at 13 TeV using 2.7 fb^{-1} of data collected in 2015 [1], ATLAS has published an analysis using 13.2 fb^{-1} of 2015 and 2016 data [5].

Events are generally selected by requiring in the dilepton channel 2 isolated, oppositely-charged leptons (electrons or muons) and ≥ 3 jets, ≥ 2 of which are identified as coming from b quarks (b-tagged), and in the lepton+jets channel 1 isolated lepton and ≥ 4 jets, ≥ 2 of which are b-tagged. Additional, channel-dependent
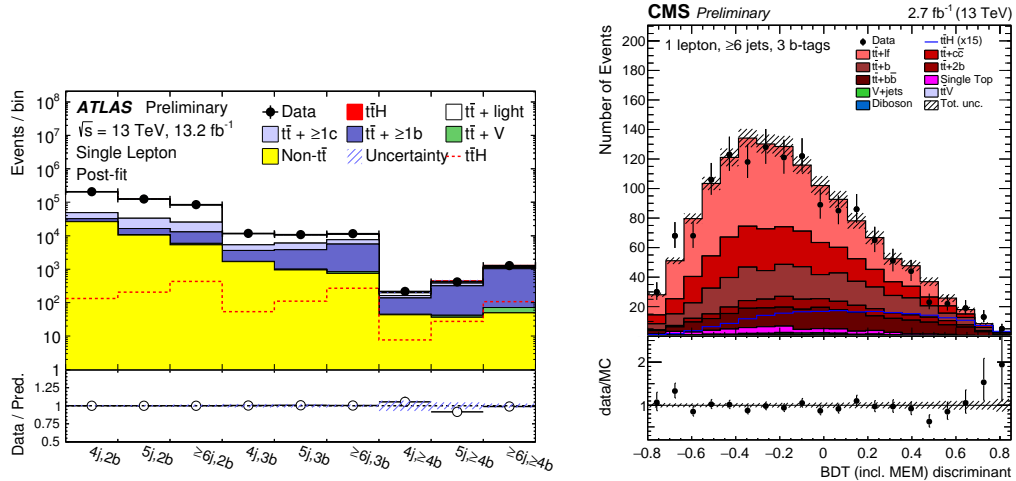
Figure 2: Predicted and observed event yields per category in the ATLAS analysis [5] (left) and BDT output in one category of the CMS analysis [1] (right) in the lepton+jets channel, after the fit to the data. The $t\bar{t}$ + jets background is divided by the flavour of the additional jets, and the $t\bar{t}H$ contribution, normalised to the best-fit value (left) and 15 times the SM expectation (right), is superimposed.

criteria are applied, such as a Z-boson-mass veto in same-flavour dilepton events.

Subsequently, the selected events are further divided into mutually exclusive categories in jet and b-tag multiplicity. Signal events are expected to contribute particularly in the categories with higher multiplicities, cf. Fig. 1 (left). The SM background consists almost entirely of $t\bar{t}$+jets events. The additional jets stem either from gluons and light-flavour quarks ($t\bar{t}$ + LF) or from heavy-flavour quarks ($t\bar{t}$ + HF), such as $t\bar{t} + b\bar{b}$ events, which represent an irreducible background. The background composition varies between the categories, cf. Fig. 2 (left) as an example. This is exploited in the final fit, which is performed simultaneously across all categories to constrain the uncertainties of the different processes. Still, especially the $t\bar{t}$ + HF processes are difficult to model, and the associated uncertainties limit the overall sensitivity.

For the CMS analysis, the $t\bar{t}$ + jets background is modelled using POWHEG, interfaced with PYTHIA8 with the CUETP8M1 tune for parton showering. The cross section is normalised to the next-to-next-leading order (NNLO) calculation with resummation to next-to-next-to-leading-logarithmic (NNLL) accuracy using the NNPDF3.0 PDF set. Events are further separated based on the flavour, defined by the hadron content, of the additional jets within acceptance that do not originate from the t-quark decays, considering also the case that two close-by B hadrons form a single jet. A rate uncertainty of 50% is assigned to each of these $t\bar{t}$ + HF processes.

For the ATLAS analysis, the $t\bar{t}$ + jets background is modelled using POWHEG and PYTHIA6 with the Perugia2012 tune, reweighted to match the t-quark and $t\bar{t}$-
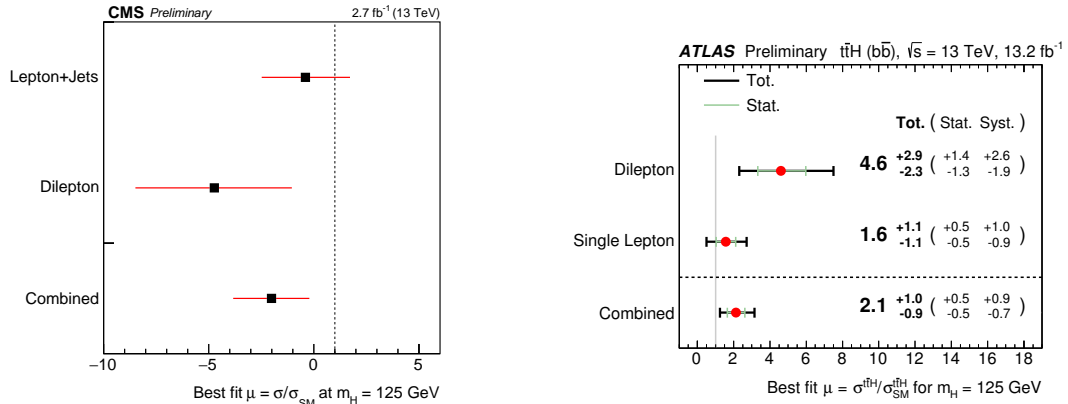
Figure 3: Best-fit values of the signal-strength modifier μ with their one standard deviation confidence intervals of the CMS [1] (left) and ATLAS [5] (right) analyses.

system p_T spectra of the NNLO calculation, including resummation at NNLL accuracy. Events with ≥ 1 additional b jets are reweighted to a SHERPA+OPENLOOPS prediction at next-to-leading order, using the CT10 4-flavour-scheme PDF set. Uncertainties on the $t\bar{t}+\mathrm{HF}$ processes are assigned based on the differences observed to a MADGRAPH5 aMC@NLO prediction, and the overall normalisation of the contributions from $t\bar{t}+\geq 1$ b or c jets events is left freely floating in the final fit.

Both analyses employ multivariate methods to combine in each category the information of several variables, such as kinematic properties or invariant masses of combinations of jets and leptons, into a final discriminating variable, cf. Fig. 2 (right). The signal cross section σ is determined with a binned fit of the background and signal discriminant distributions to the data, where the uncertainties, which affect the rate and the shape of the distributions, are taken into account via nuisance parameters.

At CMS, boosted decision trees (BDTs) and, in several categories, also a matrix-element-method (MEM) classifier are used. The latter is a likelihood of the event kinematics under the signal or background hypothesis, taking into account response and acceptance effects, which is constructed to separate against the important $t\bar{t}+b\bar{b}$ background. Depending on the category, the MEM variable is an input to the BDT or events are further separated into two sub-categories with low and high BDT-output and the MEM is used as final discriminant in each sub-category. In a novel approach, also dedicated techniques are applied to reconstruct events in which the H boson and the hadronically decaying t quark are fairly boosted, resulting in reduced combinatorics in the jet assignment and thus a better event reconstruction efficiency.

At ATLAS, a two-staged multivariate approach is used in the signal-enriched categories. A BDT is trained to assign the jets to the partons from the H-boson and t-quark decays under the signal hypothesis. Based on this event reconstruction,

3

additional separating variables, such as the invariant mass of the b jets from the H-boson decay, are computed. They are used together with reconstruction-independent variables as input to a classification BDT or an artificial neural network that separate signal from background events. In the signal-depleted categories in the lepton+jets channel (dilepton channel) the scalar sum of the jet (and lepton) p_{T} is used, which aims at constraining the systematic uncertainties of the background model.

CMS obtains a signal strength $\mu = \sigma/\sigma_{\mathrm{SM}}$ relative to the SM expectation of $\mu = -2.0^{+1.8}_{-1.8}$ with $2.7\,\mathrm{fb}^{-1}$ of data and ATLAS of $\mu = 2.1^{+1.0}_{-0.9}$ with $13.2\,\mathrm{fb}^{-1}$ of data, cf. Fig. 3, which are compatible with the SM expectation within 1.7 standard deviations. These correspond to observed (expected) upper limits on μ at the 95% confidence level (C.L.) of 2.6 ($3.6^{+1.6}_{-1.1}$) for CMS and 4.0 ($1.9^{+0.9}_{-0.5}$) for ATLAS.

3 Search for tH(b$\overline{\mathrm{b}}$) production

At leading order, tH production occurs predominantly via t-channel and associated tW production. In both cases, the H boson can be emitted either from the t quark or the intermediate W boson, cf. Fig. 1. The amplitudes of both contributions interfere depending on the coupling of the H boson to the t quark and to the W boson, expressed hereafter as coupling strengths κ_{t} and κ_{V} relative to the SM expectation, respectively. Hence, tH production is sensitive to both the magnitude and the sign of y_{t}. The interference is destructive in the SM, but can in general also be constructive resulting in an enhanced tH production cross section, e.g. by a factor ten for $\kappa_{\mathrm{t}} = -1$ and $\kappa_{\mathrm{V}} = +1$ (inverted top coupling scenario, ITC).

CMS has performed a search for tH production in the H $\to$ b$\overline{\mathrm{b}}$ final state with a leptonically decaying t quark, using $2.3\,\mathrm{fb}^{-1}$ of data at $13\,\mathrm{TeV}$. Events are selected requiring 1 isolated lepton and ≥ 3 b-tagged jets, targeting the b quarks from the H boson and t quark decay. Events are further divided into two exclusive signal regions, one with 3 and one with 4 b-tagged jets as well as 1 additional non b-tagged jet in each case, targeting the fourth b jet, which often falls out of acceptance, and the additional light-flavour jet, cf. Fig. 1. The dominant SM background in both signal regions stems from t$\overline{\mathrm{t}}$+jets production, which is modelled as for the CMS t$\overline{\mathrm{t}}$H search.

Events are reconstructed under both the signal and the t$\overline{\mathrm{t}}$ + jets hypothesis, assigning the jets to the final-state quarks depending on the output of two dedicated BDTs that take into account b-tagging and jet kinematic information. Based on the specific event reconstruction, separating variables, such as the p_{T} of the H boson, are computed. These, together with reconstruction-independent variables, are used as input to a classification BDT that separates signal and background, cf. Fig. 4 (left).

Limits on the signal cross section are obtained for 51 different points in the $(\kappa_{\mathrm{t}}, \kappa_{\mathrm{V}})$ parameter space by fitting the corresponding BDT output distributions to the data, simultaneously in the two signal regions. Since the kinematic properties of the signal
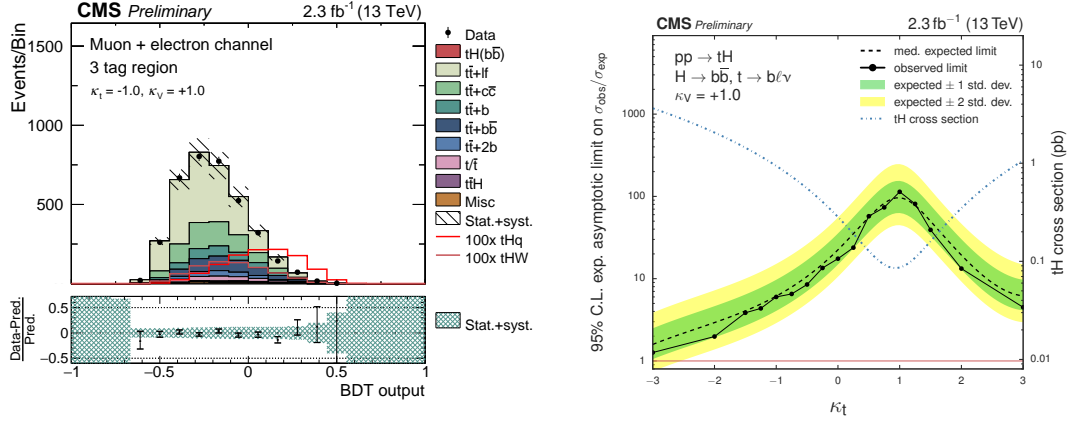
Figure 4: Classification BDT output for the ITC case (left), after the fit to the data. The $t\bar{t}$ + jets background is divided by the flavour of the additional jets, and the tH contributions in both considered production channels, normalised to 100 times the expectation, are superimposed. Also shown are the observed and expected upper limits on the signal production cross section as a function of κ_t for $\kappa_V = +1$ (right) [2].

events depend on κ_t and κ_V, dedicated BDT trainings are performed for each point. The results for $\kappa_V = +1$ are shown in Fig. 4 (right), further results for $\kappa_V = +0.5$ and $+1.5$ have been derived [2]. For the SM and the ITC scenario, observed (expected) upper limits at 95% C.L. of 113.7 ($98.6^{+60.6}_{-34.6}$) times the SM expectation and 6.0 ($6.4^{+3.7}_{-2.2}$) times the ITC expectation, respectively, are obtained.

4   Summary

First searches by CMS and ATLAS for $t\bar{t}H(b\bar{b})$ production at 13 TeV result in a signal strength of $\mu = -2.0^{+1.8}_{-1.8}$ with 2.7 fb^{-1} and $\mu = 2.1^{+1.0}_{-0.9}$ with 13.2 fb^{-1} of data. A 95% C.L. upper limit on tH(b$\bar{b}$) production with inverted couplings $\kappa_t = -1$ of 6.0 times the expectation is observed by CMS using 2.3 fb^{-1} of data. The achieved sensitivities are close to or surpass the ones at 8 TeV with only a fraction of the data.

References

[1] CMS Collaboration, CMS-PAS-HIG-16-004 (2016).

[2] CMS Collaboration, CMS-PAS-HIG-16-019 (2016).

[3] CMS Collaboration, JINST 3 (2008) S08004.

[4] ATLAS Collaboration, JINST 3 (2008) S08003.

[5] ATLAS Collaboration, ATLAS-CONF-2016-080 (2016).