CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2017/05/24

Search for production of a Higgs boson and a single top quark in multilepton final states in proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A search for the production of a Higgs boson in association with a single top quark is presented, focusing on leptonic signatures provided by the $H \to WW$, $H \to \tau\tau$, and $H \to ZZ$ decay modes. Due to strong interference of the two main leading-order diagrams, the production cross section of this process is highly sensitive to the relative sign of the top-Higgs coupling modifier, κ_t, and the coupling modifier of vector bosons to the Higgs, κ_V. The analysis exploits signatures with two same-sign leptons or three leptons in the final state, and uses the 2016 data sample collected with the CMS detector at the LHC at a center of mass energy of 13 TeV, which corresponds to an integrated luminosity of 35.9 fb^{-1}. Multivariate techniques are used to discriminate the signal from the dominant backgrounds. The analysis yields a 95% confidence level (C.L.) upper limit on the combined $tH + t\bar{t}H$ production cross section times branching ratio of 0.64 pb, with an expected limit of 0.32 pb, for a scenario with $\kappa_t = -1.0$ and $\kappa_V = 1.0$. Values of κ_t outside the range of -1.25 to $+1.60$ are excluded at 95% C.L., assuming $\kappa_V = 1.0$.

1 Introduction

The discovery of a Higgs boson by the CMS and ATLAS experiments in 2012 [1, 2] opened a new field for exploration in the realm of particle physics. It is critical to study the couplings of this new particle to other elementary particles to test whether it is the Higgs boson as predicted by the standard model (SM). Of particular interest is the Yukawa coupling of the Higgs boson to top quarks, y_t, as the top quark is widely believed to play a special role in the mechanism of electroweak symmetry breaking due to its large mass [3]. Most measurements of the top-Higgs coupling are sensitive only to the magnitude of the coupling, rather than its sign, as processes such as Higgs production associated with top quark pairs ($t\bar{t}H$) depend only on $|y_t^2|$. Constraints on the sign of y_t can be derived from the decay rate of Higgs bosons to photon pairs [4] and from the cross section for associated production of Higgs and Z bosons via gluon fusion [5], with recent results disfavoring negative signs of the coupling [6–8]. But further measurements of the relative phase between the fermion and boson couplings of the Higgs boson are warranted, in particular in scenarios with contributions from possible new particles in the loop amplitudes [9].

The production of a single top quark in the t channel, where a Higgs boson can be radiated either from the top quark or from the exchanged W boson in the two dominant leading order diagrams (see Fig. 1) provides a unique opportunity to study the relative sign of the coupling. Any deviation from the standard model coupling structure, where the two diagrams strongly interfere negatively and thereby suppress the production cross section, can lead to a large enhancement of the event rate [10–12]. Other production modes of Higgs bosons and single top quarks are the W associated process (tHW), and the s channel. While the s channel cross section is negligible at the LHC [13], the associated tHW production is comparable and can contribute significantly [14].

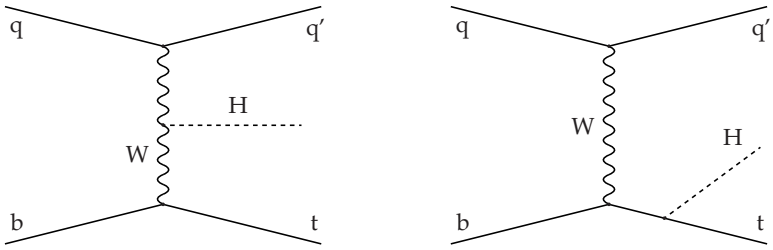

Figure 1: Dominant leading order Feynman diagrams for the production of tHq events. The Higgs boson is either radiated from the W boson (left) or the top quark (right).

Direct searches for tHq production using all relevant Higgs decay modes have previously been carried out by CMS in the 8 TeV dataset [15] and in the 2015 13 TeV dataset using the $H \to b\bar{b}$ channel [16]. In the full 2016 13 TeV dataset, a search for $t\bar{t}H$ production in multilepton final states recently produced first evidence for associated production of top quarks and Higgs bosons [17].

This note reports a search for tHq production in leptonic final states using the full 2016 LHC dataset of at 13 TeV, corresponding to an integrated luminosity of 35.9 fb^{-1}. Multilepton final states with either two same-sign leptons or three leptons target the case where the Higgs boson decays to a pair of W bosons, τ leptons, or Z bosons, and where the top quark decays leptonically. The results are interpreted as a function of the ratio of two dimensionless modifiers of Higgs couplings: that of the top-Higgs coupling, κ_t, and of the coupling of vector bosons and the Higgs, κ_V. A description of the CMS detector can be found in Ref. [18].

The analysis is designed to efficiently identify and select prompt leptons from on-shell W and Z boson decays and to reject non-prompt leptons from b quark decays and spurious lepton signatures from hadronic jets. Events are then selected in the various lepton channels, and are required to contain hadronic jets, some of which must be consistent with b quark hadronization. Finally, the signal yield is extracted by simultaneously fitting the output of two dedicated multivariate discriminants (trained to separate the tHq signal from the two dominant backgrounds) in all categories.

With respect to the 8 TeV analysis, the object selections have been adjusted for the updated LHC running conditions at 13 TeV, the lepton identification has been improved, and more powerful multivariate analysis techniques are used for the signal extraction.

2 Object Selection

The CMS particle-flow (PF) algorithm [19] provides a global event description which optimally combines the information from all sub-detectors to reconstruct and identify all individual particles in the event. The particles are classified into mutually exclusive categories: neutral and charged hadrons, photons, muons, and electrons.

Hadronic jets are reconstructed by clustering PF candidates with the anti-k_T algorithm using a distance parameter of 0.4, as implemented in the FASTJET package [20, 21]. Charged hadrons that are not consistent with the selected primary interaction vertex are discarded from the clustering. The jet energy is then corrected for the varying response of the detector as a function of transverse momentum (p_T) and pseudorapidity (η) [22]. Jets are selected for use in the analysis only if they have $p_T > 25\,\text{GeV}$ and are separated from any selected leptons by $\Delta R > 0.4$.

Jets that are likely to have originated from the hadronization of a b quark, are selected through a multivariate likelihood discriminant that uses track-based lifetime information and reconstructed secondary vertices ("combined secondary vertex" or CSV algorithm) [23]. Only jets with $|\eta| < 2.4$ (within the CMS tracker acceptance) are identified with this technique. The efficiency to correctly tag b jets and the probability to misidentify jets from light quarks or gluons are measured in data as a function of the jet p_T and η, and are used to correct for differences in the performance of the algorithm in simulated events. Two working points based on the algorithm output are used: "loose", with a b signal tagging efficiency of about 83% and a mistagging rate of about 8%; and "medium", with b efficiency of about 69% and mistagging rate of order 1% [24]. Tagging efficiencies for jets from charm quarks are about 40% (18%) for the loose (medium) working point. Separate scale factors are applied to jets originating from bottom/charm quarks and from light quarks in simulated events to match the tagging efficiencies measured in the data.

Muon candidates are reconstructed by combining information from the silicon tracker and the outer muon spectrometer of CMS in a global fit [25]. The quality of the spatial matching between the individual measurements in the tracker and the muon system is used to discriminate genuine prompt muons from hadrons punching through the calorimeters and from muons produced by in-flight decays of kaons and pions. In the analysis, muon candidates are considered if they have $p_T > 5\,\text{GeV}$ and $|\eta| < 2.4$. In the same-sign dilepton event categories, the relative uncertainty in the muon p_T from the fit is required to be better than 20% to ensure a high-quality charge measurement.

Electrons are reconstructed using information from the tracker and from the electromagnetic calorimeter [26]. Genuine electrons are identified by a multivariate algorithm using the shape

of the calorimetric shower and the quality of the reconstructed track. Furthermore, to reject electrons produced in photon conversions, candidates with missing hits in the innermost tracking layers or matched to a conversion secondary vertex are discarded. Electrons are selected for the analysis if they have $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ and $|\eta| < 2.5$. To suppress electrons with a misassigned electric charge in the same-sign dilepton categories, candidates are required to have consistent charge measurements from three independent observables based on the calorimeter energy deposits and the track curvature.

Electrons and muons passing the criteria described above are referred to as "loose leptons" in the following. A further discrimination between prompt signal leptons (i.e. from W and Z boson decays and from leptonic τ decays) and non-prompt and spurious leptons from b hadron decays, decays-in-flight, and photon conversions is crucial in light of the overwhelming background from $\mathrm{t\bar{t}}$ production. The small probabilities of having the second type of leptons results in a sizable number of background events since the rate of $\mathrm{t\bar{t}}$ production is much larger than the signal. To maximally exploit the available information in each event to that end, a multivariate discriminator based on a boosted decision tree (BDT) algorithm is built, taking as input not just observables related directly to the reconstructed leptons themselves, but also to the clustered energy deposits and charged particles in a cone around the lepton direction. The jet reconstruction and b-tagging algorithms are run on these, and their output is used to train the algorithm. In particular, the ratio between the lepton p_{T} and the reconstructed jet p_{T}, and the transverse momentum of the lepton with respect to the jet axis provide good separation power in addition to more traditional observables like the relative isolation of the lepton (calculated in a variable cone size depending on the lepton p_{T} [27, 28]), and the impact parameters of the lepton trajectory. The BDT algorithm is trained on prompt leptons in simulated $\mathrm{t\bar{t}H}$ signal and non-prompt leptons in $\mathrm{t\bar{t}}$ background events and validated using data in various control regions. Leptons are then selected for the final analysis if they pass a given threshold of the BDT output, and are referred to as "tight leptons" in the following.

3 Event Selection

Events are selected at the trigger level to contain either one, two, or three leptons with minimal transverse momentum thresholds for the leading lepton. The p_{T} thresholds are set at 24 GeV for muons and at 27 GeV for electrons for single-lepton triggers. For double-lepton triggers, the p_{T} thresholds on the leading and sub-leading legs are 17 and 8 GeV for muons and 23 and 12 GeV for electrons. Three-lepton triggers apply a threshold on the third hardest lepton in the event of 5 and 9 GeV for muons and electrons, respectively.

At the offline event selection level, the analysis targets the unique topology of the tHq signal with $\mathrm{H} \to \mathrm{WW}$ and $\mathrm{t} \to \mathrm{Wb} \to \ell\nu\mathrm{b}$, resulting in a state with three W bosons, one b quark, and a light spectator quark at high rapidity. Two channels are exploited, in which either all three W bosons decay leptonically, or the pair with equal electrical charge, resulting in a signature of either three charged leptons (muons or electrons), or two same-sign leptons with two light-quark jets. This selection naturally includes contributions from $\mathrm{H} \to \tau\tau$ and $\mathrm{H} \to \mathrm{ZZ}$ as well. Both the three- and two-lepton signatures are accompanied by a b quark and a light-flavor forward jet.

The main analysis strategy is to obtain a selection of events compatible with certain signal characteristics at a pre-selection level and then extract the signal contribution in a second analysis step, using multivariate discriminators against the main backgrounds of $\mathrm{t\bar{t}W^{\pm}}/\mathrm{t\bar{t}Z}$ and non-prompt leptons from $\mathrm{t\bar{t}}$. The shape of the discriminator variables is then fit to the observed data distribution to estimate the signal and background yields, simultaneously for all chan-

nels.

In the leptonic channels investigated in this analysis, the main backgrounds are expected to arise from the production of top quarks, either in the dominant $t\bar{t}$ mode, where multi-lepton and same-sign dilepton signatures can occur when a non-prompt lepton from heavy-flavor decay passes the signal selection, or in associated production with a W/Z or Higgs boson. Processes with single top quarks also contribute, mostly in the associated production with a Z boson (tZq) or when produced with both a W and a Z boson (tZW). Contributions from diboson production, while having a comparatively large cross section, can be strongly suppressed by imposing a veto on lepton pairs compatible with a Z-boson decay ("Z-veto") or by altogether vetoing additional leptons in the event. Diboson processes are further suppressed relative to processes involving top quarks when requiring b-tagged jets in the event.

An additional background in the case of same-sign dileptons arises when the charge of a lepton in events with an originally opposite-sign pair is misidentified. Furthermore, the same-sign channel receives some contribution from the associated production of two W bosons of equal charge, and two light jets, $W^{\pm}W^{\pm}qq$. Same-sign W boson pairs can also be produced in double parton scattering (DPS) processes, where each of the colliding protons gives two partons, resulting in two hard interactions.

A relatively loose selection is applied to maintain a large signal efficiency while suppressing the main backgrounds. It is summarized for both the three-lepton and same-sign dilepton channel in Tab. 1. The selections are based on the number of leptons, reconstructed invariant mass ($m_{\ell\ell}$), and b-tagged jet multiplicity, which are characteristic of the tHq process. A significant fraction of selected data events (about 50% in the dilepton channels, and about 80% in the trilepton channel) also passes the selection used in the dedicated search for $t\bar{t}H$ in multilepton channels [17].

Same-sign $\ell\ell$ channel ($\mu\mu$/eμ)	**$\ell\ell\ell$ channel**
No loose leptons with $m_{\ell\ell} < 12\,\text{GeV}$	
One or more b tagged jets	
One or more non-tagged jets	
Exactly two tight same-sign leptons	Exactly three tight leptons
$p_{\text{T}} > 25/15\,\text{GeV}$	$p_{\text{T}} > 25/15/15\,\text{GeV}$
	No lepton pair with $\lvert m_{\ell\ell} - m_{\text{Z}}\rvert < 15\,\text{GeV}$

Table 1: Summary of event selection.

The expected and observed event yields of this selection are shown in Tab. 2. For the tH and $t\bar{t}H$ processes, the largest contribution comes from Higgs decays to WW (about 75%), followed by $\tau\tau$ (about 20%) and ZZ (about 5%). Other Higgs production modes contribute negligible event yields ($< 5\%$ of the tH +$t\bar{t}H$ yield).

4 Signal discrimination

The production cross section for the signal processes tHq, tHW, and $t\bar{t}H$ is only a few fb (even with inverted couplings, $\kappa_{\text{t}} = -1$), resulting in a small signal to background ratio even for a tight selection. A multivariate method is hence employed to train a discriminator to separate tHq signal events from background. Several methods have been studied, with the best performance obtained from a gradient boosted decision tree (BDT) classifier using a maximum tree depth of three and an ensemble of 800 trees. Two BDTs are trained separately for the same-sign dilepton and the three lepton channel on simulated events to separate the tHq process either

Process	$\ell\ell\ell$	$\mu\mu$	eμ
$t\bar{t}W^{\pm}$	22.50 ± 0.35	68.03 ± 0.61	97.00 ± 0.71
$t\bar{t}Z/t\bar{t}\gamma$	32.80 ± 1.79	25.89 ± 1.12	64.82 ± 2.42
WZ	8.22 ± 0.86	15.07 ± 1.19	26.25 ± 1.57
ZZ	1.62 ± 0.33	1.16 ± 0.29	2.86 ± 0.45
$W^{\pm}W^{\pm}qq$	–	3.96 ± 0.52	6.99 ± 0.69
$W^{\pm}W^{\pm}$(DPS)	–	2.48 ± 0.42	4.17 ± 0.54
VVV	0.42 ± 0.16	2.99 ± 0.34	4.85 ± 0.43
tttt	1.84 ± 0.44	2.32 ± 0.45	4.06 ± 0.57
tZq	3.92 ± 1.48	5.77 ± 2.24	10.73 ± 3.03
tZW	1.70 ± 0.12	2.13 ± 0.13	3.91 ± 0.18
γ conversions	7.43 ± 1.94	–	23.81 ± 6.04
Non-prompt	25.61 ± 1.26	80.94 ± 2.02	135.34 ± 2.83
Charge flips	–	–	58.20 ± 0.30
Total Background	106.05 ± 3.45	210.74 ± 3.61	443.30 ± 8.01
$t\bar{t}H$	18.29 ± 0.41	24.18 ± 0.48	35.21 ± 0.58
tHq (SM)	0.52 ± 0.02	1.43 ± 0.04	1.92 ± 0.04
tHW (SM)	0.62 ± 0.03	0.71 ± 0.03	1.11 ± 0.04
Total SM	125.48 ± 3.47	237.06 ± 3.64	481.54 ± 8.03
tHq ($\kappa_V = 1 = -\kappa_t$)	7.48 ± 0.14	18.48 ± 0.22	27.41 ± 0.27
tHW ($\kappa_V = 1 = -\kappa_t$)	7.38 ± 0.16	7.72 ± 0.17	11.23 ± 0.20
Data	**149**	**280**	**525**

Table 2: Data yields and expected backgrounds after the event pre-selection for the three channels in 35.9 fb^{-1} of integrated luminosity. Uncertainties are statistical only.

from an admixture of $t\bar{t}W^{\pm}$ and $t\bar{t}Z$ (in short $t\bar{t}V$), or from $t\bar{t}$. Events from tHW and $t\bar{t}H$ production, while counted as signal events in the interpretation, share the kinematic characteristics of $t\bar{t}V$ events and are not used in the classifier training.

Three broad categories of discriminating observables are used: related to forward jet activity; related to jet and b-jet multiplicities; and related to kinematic properties of leptons as well as their total charge. Table 3 lists the 10 input variables used in the final discriminants. Many combinations of kinematic variables were constructed to perform the discrimination between signal and background with the chosen set of variables giving the best performance in the separation. The same or equivalent input variables are found to perform well both for three lepton and same-sign dilepton channels and both for training against $t\bar{t}V$ and against $t\bar{t}$.

Number of jets with $p_T > 25\,\text{GeV}$, $\lvert\eta\rvert < 2.4$
Maximum $\lvert\eta\rvert$ of any (non-b-tagged) jet ("forward jet")
Sum of lepton charges
Number of non-b-tagged jets with $\lvert\eta\rvert > 1.0$
$\Delta\eta$ between forward light jet and leading b-tagged jet
$\Delta\eta$ between forward light jet and sub-leading b-tagged jet
$\Delta\eta$ between forward light jet and closest lepton
$\Delta\phi$ of same-sign lepton pair
Minimum ΔR between any two leptons
p_T of sub-leading (or 3^{rd}) lepton

Table 3: Input variables to the signal discrimination classifier.

The distributions for some of the BDT input variables are shown in Figures 2, 3 and 4, compar-

ing observed data and predicted yields. The analysis has been developed while blinded to the distributions of the observed events passing the signal selection.

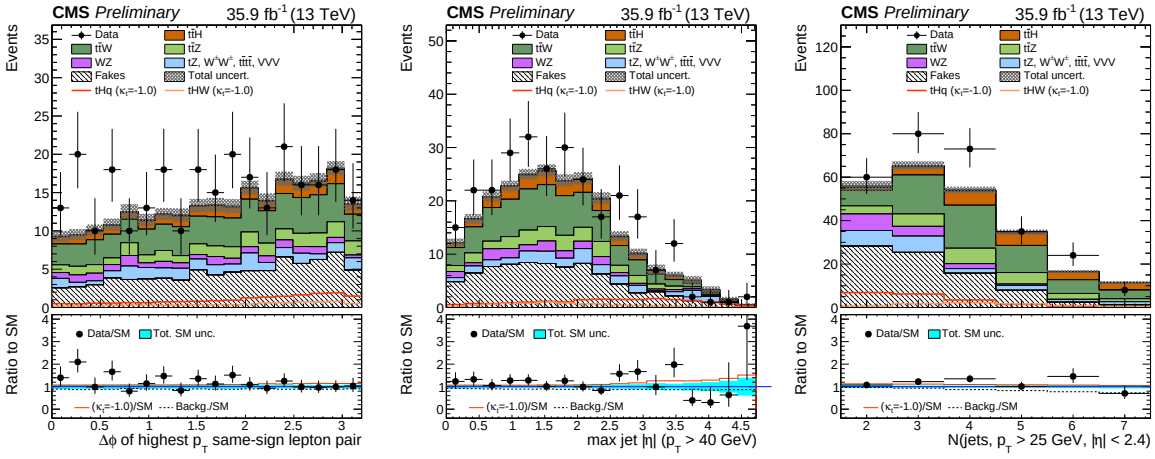

Figure 2: Distributions of discriminating variables for the event pre-selection for the same-sign $\mu\mu$ channel, normalized to 35.9 fb^{-1}, before fitting the signal discriminant to the observed data. Uncertainties are statistical and unconstrained (pre-fit) normalization systematics. The shape of the two tH signals for $\kappa_t = -1.0$ is shown, normalized to their respective cross sections for $\kappa_t = -1.0, \kappa_V = 1.0$.

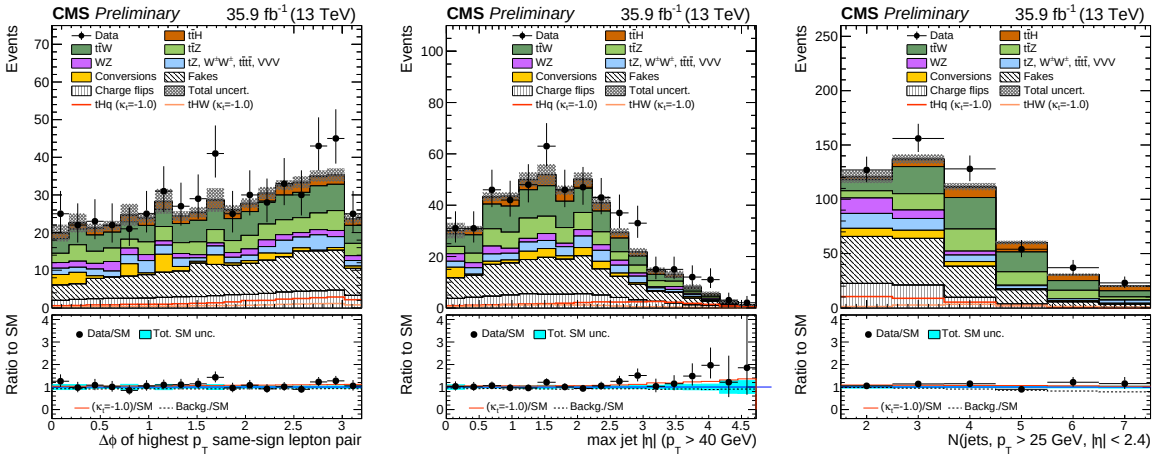

Figure 3: Distributions of discriminating variables for the event pre-selection for the same-sign eμ channel, normalized to 35.9 fb^{-1}, before fitting the signal discriminant to the observed data. Uncertainties are statistical and unconstrained (pre-fit) normalization systematics. The shape of the two tH signals for $\kappa_t = -1.0$ is shown, normalized to their respective cross sections for $\kappa_t = -1.0, \kappa_V = 1.0$.

5 Modeling of signal and background processes

The tHq and tHW signal events are generated using MG5_aMC@NLO (version 5.222) [29] at leading-order precision, using the MLM merging scheme [30] and the NNPDF3.0 PDF set [31], and are normalized to next-to-leading order cross sections. The tHq events are generated with the four-flavor scheme while the tHW process uses the five-flavor scheme to eliminate leading-order interference with the t$\bar{\text{t}}$H process [14]. Event weights are produced in the generation of both samples to allow a reshaping of observables for 51 different coupling configurations: 17

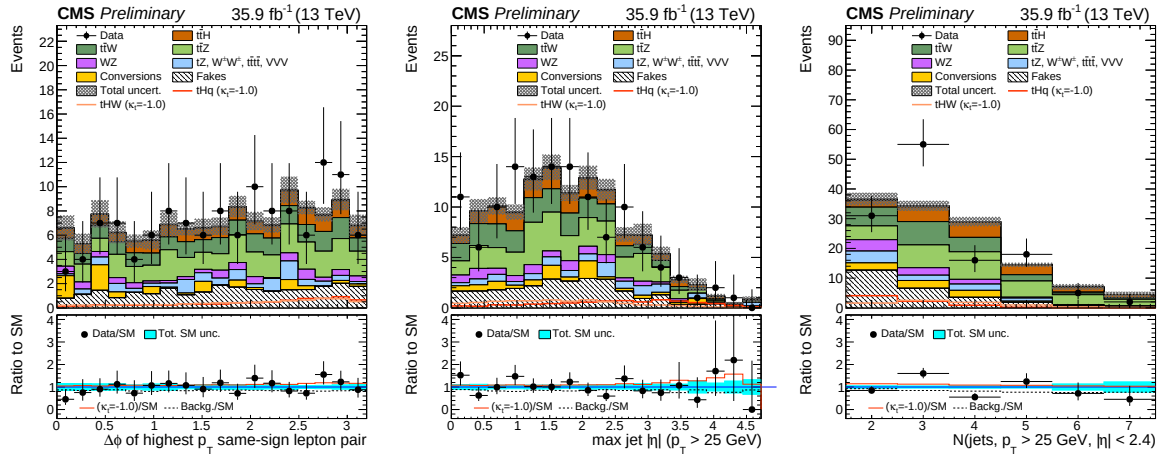

Figure 4: Distributions of discriminating variables for the event pre-selection for the three lepton channel, normalized to 35.9 fb^{-1}, before fitting the signal discriminant to the observed data. Uncertainties are statistical and unconstrained (pre-fit) normalization systematics. The shape of the two tH signals for $\kappa_t = -1.0$ is shown, normalized to their respective cross sections for $\kappa_t = -1.0, \kappa_V = 1.0$.

values of κ_t between -3.0 and $+3.0$ for three values of κ_V: $+0.5$, $+1.0$, and $+1.5$, corresponding to 33 unique values of κ_t/κ_V, ranging from -6.0 to 6.0, and therefore 33 distinct kinematic configurations.

MG5_aMC@NLO in NLO mode is used for the $\text{t}\bar{\text{t}}\text{H}$ process and the main backgrounds: $\text{t}\bar{\text{t}}\text{W}^{\pm}$, $\text{t}\bar{\text{t}}\text{Z}$, $\text{t}\bar{\text{t}}$ +jets, and $\text{t}\bar{\text{t}}\gamma$+jets, using parton-shower merging at NLO [32]. Other minor backgrounds are simulated with different generators, such as POWHEG [33–38] and MADGRAPH at leading order (LO) QCD accuracy. All generated events are interfaced to PYTHIA8 (v8.205) [39] for the parton shower and hadronization steps. Pileup interactions are simulated to reflect the observed multiplicity in data. The simulated events are weighted according to the actual pileup in data, estimated from the measured bunch-to-bunch instantaneous luminosity and the total inelastic cross section, 69.2 mb. All events are finally passed through a full simulation of the CMS detector based on GEANT4 [40], and reconstructed using the same algorithms as used for the data.

Furthermore, the trigger selection is simulated and applied for generated signal events. Residual differences in the trigger efficiency between data and MC are studied and corrected for, using the measured trigger efficiencies of the data.

5.1 Signal modeling

Systematic uncertainties on the signal selection efficiency arise from correction factors applied to the simulated events to better match the measured detector performance and also from theoretical uncertainties in the modeling of the signal process.

Scale factors applied to correct for data/MC differences in the trigger efficiency, lepton reconstruction and identification performance, and lepton selection efficiency carry a combined uncertainty of about 5% per lepton. The impact of the uncertainty in the signal selection efficiency from jet energy corrections is evaluated by varying the correction factors within their uncertainty and propagating the effect to the final result by recalculating all kinematic quantities. Effects on the overall normalization of event yields and on the shape of kinematic properties are both taken into account. Jet energy resolution effects have negligible impact on this

analysis. Correction factors for data/MC differences in the b-tagging performance are applied depending on the p_T and η, and on the flavor of the jet, and their effect on the signal efficiency is evaluated by varying the factors within their measured uncertainty and recalculating the overall event scale factors.

The uncertainties from unknown higher orders of tHq and tHW production are estimated from a change in the Q^2 scale of double and half the initial value, evaluated for each point of κ_t and κ_V. The $t\bar{t}H$ signal component has an uncertainty of about $+5.8/-9.2\%$ from Q^2 scale variations and a further 3.6% from the knowledge of PDFs and α_S [41].

Uncertainties related to the choice of PDF set and its scale are estimated to be about 3.7% for tHq and about 4.0% for tHW.

5.2 $t\bar{t}V$, WZ, and ZZ backgrounds

Backgrounds from associated production of $t\bar{t}$ pairs and electroweak bosons ($t\bar{t}W^{\pm}$ and $t\bar{t}Z$) are estimated directly from simulated events, which are corrected for data/MC differences and inefficiencies in the same way as signal events. Their production cross sections are calculated at next-to-leading order of QCD and EWK, with theoretical uncertainties from unknown higher orders of 12% for $t\bar{t}W^{\pm}$ and 10% for $t\bar{t}Z$. Further uncertainties arise from the knowledge of PDFs and α_S of about 4% each for $t\bar{t}W^{\pm}$ and $t\bar{t}Z$.

Diboson production with leptonic Z decays and additional jet radiation in the final state can lead to signatures very similar to that of the signal. Due to the larger cross section, the main contribution arises from WZ production. Inclusive production cross sections for both WZ and ZZ have been measured at the LHC and agree well with the NLO calculations.However, the good agreement of the cross section measurements in the inclusive phase space does not necessarily hold in the signal region of this analysis, which requires the presence of hadronic jets, including b jets. Therefore, a dedicated control region dominated by WZ production is used to constrain the overall normalization of this process. It is defined by the presence of at least three leptons, of which one opposite-sign pair must be compatible with a Z boson decay. Furthermore, at least two jets are required, with a veto on jets that pass the loose b tag selection to ensure exclusivity with the signal selection. A scale factor is then extracted from the predicted distribution of WZ events in the control region, and the observed data, keeping other processes fixed. Finally, this factor is used to scale the diboson prediction in the signal selection.

The majority of diboson events passing the signal selection contain jets from light quarks and gluons that are incorrectly tagged as b jets, making this estimate mainly sensitive to the experimental uncertainty in the mis-tag rate rather than the theoretical uncertainty in the jet flavor composition. The overall uncertainty assigned to the diboson prediction is estimated from the statistical uncertainty due to the limited sample size in the control region (30%), the residual background in the control region (20%), the uncertainties on the b-tagging rate (10–40%), and from the knowledge of PDFs and the theoretical uncertainties of the extrapolation (up to 10%).

5.3 Non-prompt and charge mis-identified leptons

The main contribution to the overall event yield in the signal selection, and one that can be reduced up to a certain point by tighter lepton selections, comes from processes with comparatively large cross sections in which one of the leptons is produced inside a jet (i.e. it is non-prompt). These are mostly real leptons from b hadron decays but also contain hadronic jets misidentified as leptons. The yield of such events is estimated from a loose-to-tight extrapolation, in which a looser lepton selection is defined and the rate at which such leptons

enter the tighter selection is measured in a control region and then used to extrapolate from a sideband with loose leptons to the signal selection with tight leptons.

The probability of a non-prompt lepton candidate passing a given loose selection to also pass the tight signal requirement is measured in a sample dominated by non-prompt leptons, as a function of p_{T} and $|\eta|$ and separately for muons and electrons. The definitions of loose and tight leptons are given in Sec. 2. Two event samples are defined for the measurement of tight-to-loose ratios: one dominated by QCD multijet events, collected using single lepton triggers at relatively high p_{T} thresholds; and one dominated by $\mathrm{Z}+$jets events, where the two high p_{T} leptons from the Z decay can be used to trigger the events without biasing the p_{T} spectrum of a third lepton at low transverse momentum. The QCD-dominated sample is then used to extract ratios for lepton candidates with p_{T} above 30 GeV, whereas the ratios for low p_{T} leptons are determined in the $\mathrm{Z}+$jets sample. For both regions, contributions from prompt leptons, mainly from W and $\mathrm{Z}+$jets or from WZ and ZZ events, respectively, are first suppressed by vetoing additional leptons in the selection, and the residual contamination is then subtracted using the transverse mass as a discriminating variable.

A sideband control region is then defined by relaxing the lepton selection criteria to "loose" (see Sec. 2), while keeping all other selections equivalent to the full signal selection. By weighting events in this expanded selection with a factor dependent on the measured tight-to-loose ratios, a fully data-driven estimation for the contribution of non-prompt leptons to the signal selection can be obtained. In events where just one of the two leptons fails the tight criteria, the applied event weight is $f/(1-f)$ (where f is the tight-to-loose ratio measured as described above), while events where both leptons fail the tight criteria are weighted by $-f_1 f_2/[(1-f_1)(1-f_2)]$. The resulting prediction of the event yield in the signal selection carries an uncertainty of 30–50%, arising from the statistical uncertainty in the measurement of the tight-to-loose ratios, and from a systematic uncertainty derived by comparing alternative methods of subtracting prompt lepton backgrounds and from testing the closure of the method in simulated background events.

Similarly, background from events where the charge of one of the leptons is wrongly assigned—relevant only in the same-sign dilepton channels—are determined by measuring the charge mis-assignment probability in a sample of same-sign dilepton event compatible with a Z boson decay and weighting events with opposite-sign leptons in the signal selection. The charge mis-assignment probability is found to be negligible for this analysis for muons, whereas for electrons it ranges from about 0.02% in the barrel section ($|\eta| < 1.48$) up to about 0.4% in the detector endcaps ($1.48 < |\eta| < 2.5$). It is measured separately in these two regions, and additionally as a function of the electron p_{T}. A systematic uncertainty of 30% is assigned to the prediction from the statistical uncertainty of the probability measurement and from testing the performance of the method on simulated events.

6 Results

After applying the event pre-selection on the dataset, 280 events are observed in the same-sign $\mu\mu$ channel, 525 in the same-sign eμ channel, and 149 events in the trilepton channel ($\ell\ell\ell$). The events are then sorted into ten categories depending on the output of the two BDT classifiers according to an optimized binning strategy, resulting in a one-dimensional histogram with ten bins. The expected signal and background shapes for this distribution are then fit to the observed data in a maximum likelihood fit, simultaneously for all three channels and separately for the signal shapes for each of the 33 $\kappa_{\mathrm{t}}/\kappa_{\mathrm{V}}$ coupling configuration points.

In each point, the tH and $t\bar{t}H$ production cross sections and the Higgs decay branching ratios are modified with the Higgs-top (κ_t) and Higgs-vector boson (κ_V) coupling strength. The Higgs-tau coupling strength modifier (κ_τ) is assumed to be equal to κ_t. All other parameters are assumed to be at the values predicted by the standard model. This implies that the combined signal shape is uniquely defined by the ratio of κ_t/κ_V. In the fit, the tH and $t\bar{t}H$ signal are then floated with a common signal strength modifier (defined as the ratio to the expected cross section) to produce a 95% confidence level (C.L.) upper limit on the observed $tH + t\bar{t}H$ cross section times the combined branching ratio of $H \to WW^* + \tau\tau + ZZ^*$.

The pre-fit BDT output distributions are shown in Fig. 5, whereas Fig. 6 shows the post-fit categorized BDT output distributions obtained in the maximum likelihood fit to extract the limits.

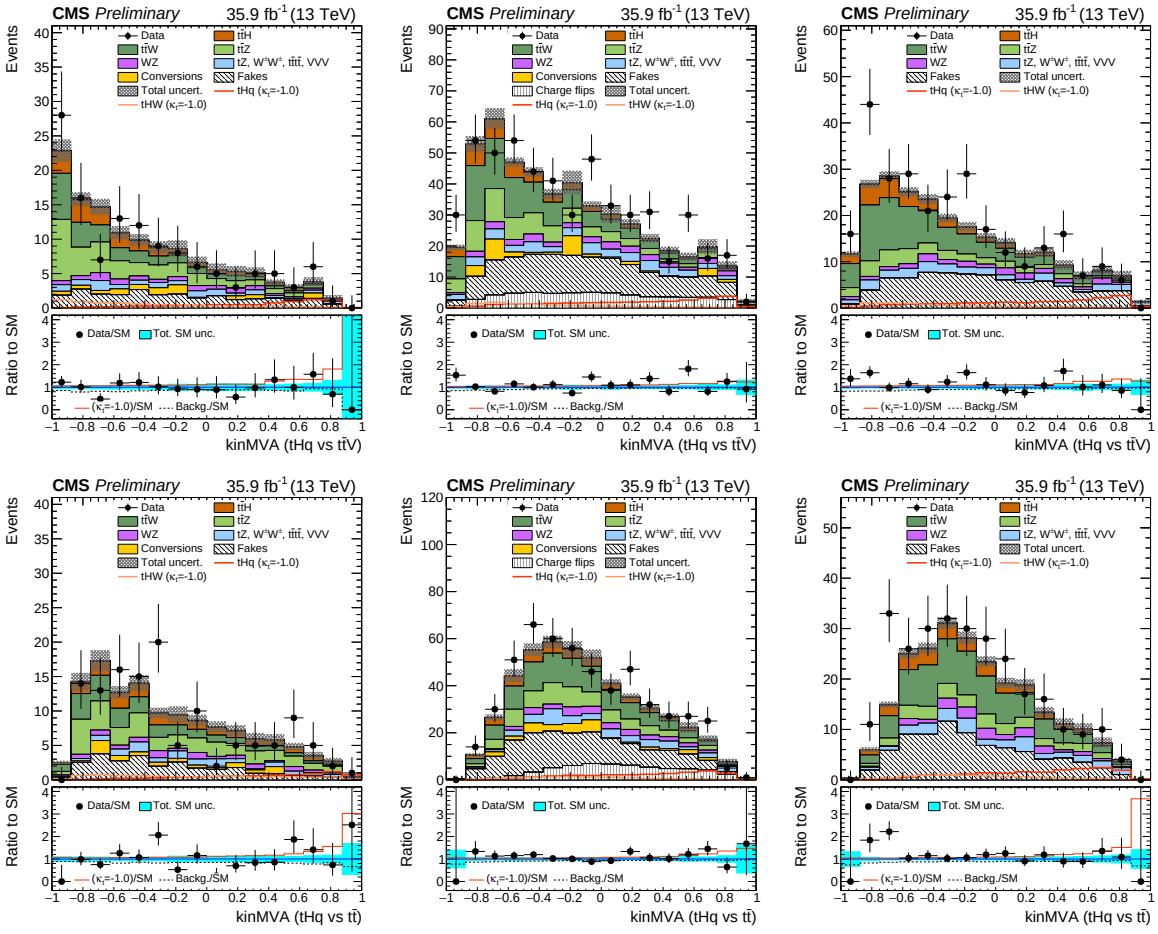

Figure 5: Pre-fit BDT classifier outputs, for the three-lepton channel (left), eμ (center), and $\mu\mu$ (right), for 35.9 fb^{-1}, for training against $t\bar{t}V$ (top row) and against $t\bar{t}$ (bottom row). In the box below each distribution, the ratio of the observed and predicted event yields is shown. The shape of the two tH signals for $\kappa_t = -1.0$ is shown, normalized to their respective cross sections for $\kappa_t = -1.0, \kappa_V = 1.0$. The grey band represents the unconstrained (pre-fit) statistical and systematical uncertainties.

The observed 95% C.L. upper limits on the $tH + t\bar{t}H$ signal cross sections for the case of inverted couplings ($\kappa_t/\kappa_V = -1.0$) for the individual channels are, respectively, 1.00, 0.84, and 0.70 pb for the $\mu\mu$, eμ, and trilepton final states, corresponding to 2.3, 1.9, and 1.6 times the respective

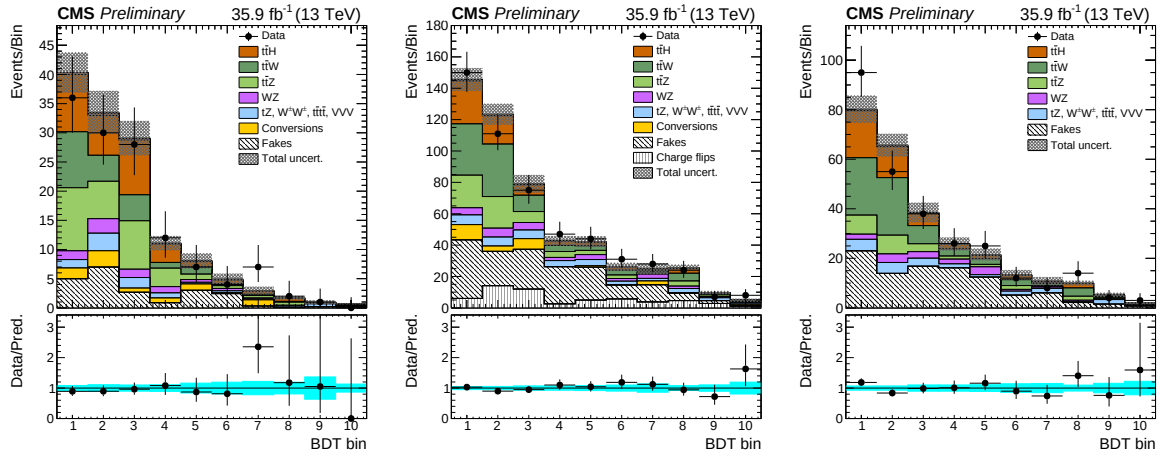

Figure 6: Post-fit categorized BDT classifier outputs as used in the maximum likelihood fit for the three-lepton channel (left), eμ (center), and $\mu\mu$ (right), for 35.9 fb^{-1}. In the box below each distribution, the ratio of the observed and predicted event yields is shown.

expected cross sections for $\kappa_V = 1.0$. The combination of all three channels yields a limit of 0.64 pb on a signal shape expected for $\kappa_t/\kappa_V = -1.0$, corresponding to 1.4 times the expected $tH + t\bar{t}H$ cross section with $\kappa_t = -1.0, \kappa_V = 1.0$. In the standard model scenario ($\kappa_t/\kappa_V = 1.0$), the observed upper limit on the cross section times branching ratio is 0.56 pb, corresponding to 3.1 times the expected SM cross section of $tH + t\bar{t}H$. Table 4 summarizes the results for the inverted coupling scenario and for the standard model.

The best-fit combined signal strength for the standard model hypothesis is $1.8 \pm 0.3(\text{stat.}) \pm 0.6(\text{syst.})$, corresponding to an observed significance of 2.7σ (1.5σ expected) over a background-only hypothesis. For a scenario of inverted couplings ($\kappa_t = -1 = -\kappa_V$), the best fit signal strength is 0.7 ± 0.4, corresponding to a significance of 1.7σ (2.5σ expected), whereas the fit prefers a signal strength compatible with 0 for a scenario with $\kappa_t = 0$ (where the $t\bar{t}H$ component vanishes).

The expected and observed limits are shown in Fig. 7 as a function of κ_t/κ_V. Comparing the observed upper limit with the theoretical prediction of the $tH + t\bar{t}H$ cross section times BR for $\kappa_V = 1.0$ constrains the allowed range of coupling configurations κ_t/κ_V to between about -1.25 and $+1.60$.

The sensitivity of the analysis is limited by systematic uncertainties, predominantly by those concerning the normalizations of the main background components (the non-prompt lepton estimation, the scale uncertainties for $t\bar{t}W^{\pm}$ and $t\bar{t}Z$), as well as by the uncertainties on the measured lepton efficiency.

7 Conclusions

A search for the production of a Higgs boson in association with a single top quark has been presented, using the CMS detector and the full 2016 data sample of pp collisions at 13 TeV, corresponding to an integrated luminosity of 35.9 fb^{-1}. Three channels have been analyzed, targeting the Higgs boson decaying to a pair of W or Z bosons, or two τ leptons and the leptonic decay of the top: two same-sign leptons ($\mu\mu$, eμ) and three leptons. This process can benefit from a greatly enhanced production cross section in the case of anomalous top-Higgs

Scenario	Channel	Obs. Limit (pb)	Exp. Limit (pb) Median	$\pm1\sigma$	$\pm2\sigma$
$\kappa_t/\kappa_V = -1$	$\mu\mu$	1.00	0.58	[0.42, 0.83]	[0.31, 1.15]
	eμ	0.84	0.54	[0.39, 0.76]	[0.29, 1.03]
	$\ell\ell\ell$	0.70	0.38	[0.26, 0.56]	[0.19, 0.79]
	Combined	**0.64**	**0.32**	[0.22, 0.46]	[0.16, 0.64]
$\kappa_t/\kappa_V = 1$ (SM-like)	$\mu\mu$	0.87	0.41	[0.29, 0.58]	[0.22, 0.82]
	eμ	0.59	0.37	[0.26, 0.53]	[0.20, 0.73]
	$\ell\ell\ell$	0.54	0.31	[0.22, 0.43]	[0.16, 0.62]
	Combined	**0.56**	**0.24**	[0.17, 0.35]	[0.13, 0.49]

Table 4: Expected and observed 95% C.L. upper limits on the $tH + t\bar{t}H$ production cross section times $H \to WW^* + \tau\tau + ZZ^*$ branching ratio for a scenario of inverted couplings ($\kappa_t/\kappa_V = -1.0$, top rows) and for a standard-model-like signal ($\kappa_t/\kappa_V = 1.0$, bottom rows), in pb. The expected limit is calculated on a background-only MC dataset.

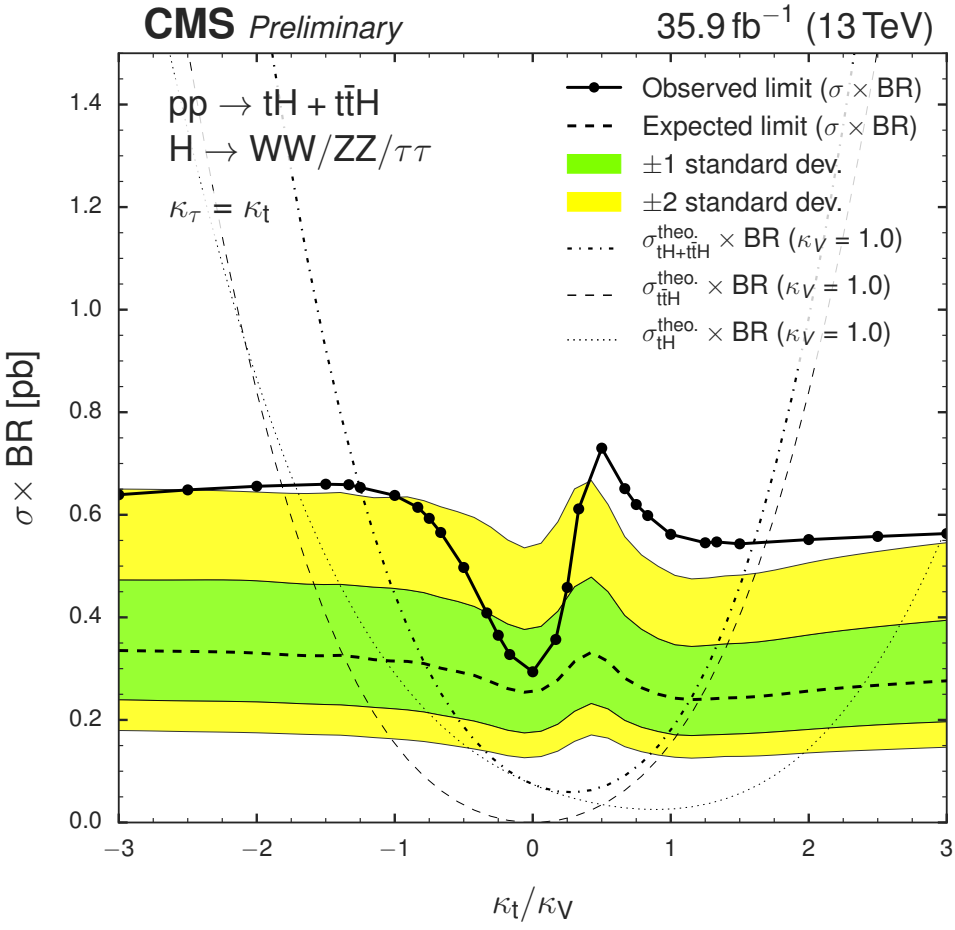

Figure 7: Observed and expected 95% C.L. upper limit on the $tH + t\bar{t}H$ cross section times $H \to WW^* + \tau\tau + ZZ^*$ branching fraction for different values of the coupling ratio κ_t/κ_V. The expected limit is derived from a background-only MC dataset.

couplings, and the results are used to constrain these couplings.

Combining the results from all three channels yields a 95% confidence level (C.L.) upper limit

on the production cross section times branching ratio of events containing a SM Higgs boson of 0.56 pb, with an expected limit of 0.24 pb.

Values of the ratio of Higgs-top coupling modifier κ_t and Higgs-vector boson coupling modifier κ_V are outside the range -1.25 to $+1.60$ are excluded at 95% C.L.

References

[1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[2] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett.* **B716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[3] CMS Collaboration, "Measurement of the top quark mass using proton-proton data at $\sqrt{s}$ = 7 and 8 TeV", *Phys. Rev.* **D93** (2016), no. 7, 072004, doi:10.1103/PhysRevD.93.072004, arXiv:1509.04044.

[4] S. Biswas, E. Gabrielli, F. Margaroli, and B. Mele, "Direct constraints on the top-Higgs coupling from the 8 TeV LHC data", *JHEP* **07** (2013) 73, doi:10.1007/JHEP07(2013)073.

[5] B. Hespel, F. Maltoni, and E. Vryonidou, "Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM", *JHEP* **06** (2015) 065, doi:10.1007/JHEP06(2015)065, arXiv:1503.01656.

[6] ATLAS Collaboration, "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC", *Phys. Lett.* **B726** (2013) 88–119, doi:10.1016/j.physletb.2014.05.011, 10.1016/j.physletb.2013.08.010, arXiv:1307.1427. [Erratum: Phys. Lett.B734,406(2014)].

[7] CMS Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J.* **C75** (2015) 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.

[8] ATLAS, CMS Collaboration, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", arXiv:1606.02266. Submitted to *JHEP*.

[9] J. Ellis and T. You, "Updated Global Analysis of Higgs Couplings", *JHEP* **06** (2013) 103, doi:10.1007/JHEP06(2013)103, arXiv:1303.3879.

[10] G. Bordes and B. van Eijk, "On the associate production of a neutral intermediate mass Higgs boson with a single top quark at the LHC and SSC", *Phys. Lett. B* **299** (1993) 315, doi:10.1016/0370-2693(93)90266-K.

[11] T. M. P. Tait and C. P. Yuan, "Single top quark production as a window to physics beyond the standard model", *Phys. Rev. D* **63** (2000) 014018, doi:10.1103/PhysRevD.63.014018, arXiv:hep-ph/0007298.

[12] M. Farina et al., "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson", *JHEP* **1305** (2013) 022, doi:10.1007/JHEP05(2013)022, arXiv:1211.3736.

[13] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, "Higgs production in association with a single top quark at the LHC", *Eur. Phys. J. C* **75** (2015), no. 6, 267, doi:10.1140/epjc/s10052-015-3475-9.

[14] F. Demartin et al., "tWH associated production at the LHC", *Eur. Phys. J.* **C77** (2017) 34, doi:10.1140/epjc/s10052-017-4601-7, arXiv:1607.05862.

[15] CMS Collaboration, "Search for the associated production of a Higgs boson with a single top quark in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JHEP* **06** (2016) 177, doi:10.1007/JHEP06(2016)177, arXiv:1509.08159.

[16] CMS Collaboration, "Search for $H \to b\bar{b}$ in association with a single top quark as a test of Higgs boson couplings at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-16-019, 2016.

[17] CMS Collaboration, "Search for Higgs boson production in association with top quarks in multilepton final states at $\sqrt{s}$ = 13 TeV", CMS Physics Analysis Summary CMS-PAS-HIG-17-004, 2017.

[18] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[19] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss}", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[20] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[21] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_T jet-finder", *Phys. Lett.* **B641** (2006) 57–61, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[22] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[23] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[24] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, 2016.

[25] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s}$ = 7 TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[26] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s}$ = 8 TeV", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[27] K. Rehermann and B. Tweedie, "Efficient Identification of Boosted Semileptonic Top Quarks at the LHC", *JHEP* **03** (2011) 059, doi:10.1007/JHEP03(2011)059, arXiv:1007.2221.

[28] CMS Collaboration, "Search for SUSY in same-sign dilepton events at $\sqrt{s}$=13 TeV", *Eur. Phys. J.* **C76** (2016) 439, doi:10.1140/epjc/s10052-016-4261-z, arXiv:1605.03171.

[29] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[30] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J.* **C53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[31] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[32] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[33] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **0411** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[34] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[35] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[36] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J.* **C71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[37] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **0909** (2009) 111, doi:10.1007/JHEP02(2010)011, 10.1088/1126-6708/2009/09/111, arXiv:0907.4076.

[38] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, "W^+W^-, WZ and ZZ production in the POWHEG BOX", *JHEP* **1111** (2011) 078, doi:10.1007/JHEP11(2011)078, arXiv:1107.5051.

[39] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[40] J. Allison et al., "GEANT4 developments and applications", *IEEE Trans. Nucl. Sci.* **53** (2006) 270, doi:10.1109/TNS.2006.869826.

[41] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", arXiv:1610.07922.