

DRAFT

CMS Paper

The content of this note is intended for CMS internal use and distribution only

2019/03/26

Head Id:

Archive Id: 492779P

Archive Date: 2019/03/26

Archive Tag: trunk

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

The CMS Collaboration

Abstract

A search is presented for the production of a Higgs boson in association with a single top quark, based on data collected in 2016 by the CMS experiment at the LHC at a center-of-mass energy of 13 TeV, which corresponds to an integrated luminosity of 35.9 fb^{-1} . The production cross section for this process is highly sensitive to the absolute values of the top quark Yukawa coupling, y_t , the Higgs boson coupling to vector bosons, g_{HVV} , and, uniquely, to their relative sign. Analyses using multilepton signatures, targeting $H \rightarrow WW$, $H \rightarrow \tau\tau$, and $H \rightarrow ZZ$ decay modes, and signatures with a single lepton and a $b\bar{b}$ pair, targeting the $H \rightarrow b\bar{b}$ decay, are combined with a reinterpretation of a measurement in the $H \rightarrow \gamma\gamma$ channel to constrain y_t . For a standard model-like value of g_{HVV} , the data favor positive values of y_t and exclude values of y_t below about $-0.9 y_t^{\text{SM}}$.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	Benjamin Stieger, Pallabi Das, Nils Faltermann, Kevin Floeh, Abideh Jafari, Hamed Bakhshiansohi, Ken Bloom, Kajari Mazumdar
PDFTitle:	Search for associated production of a Higgs boson and a single top quark in proton-proton collisions at $\sqrt{s} = 13$ TeV
PDFSubject:	CMS
PDFKeywords:	CMS, higgs, top, physics

Please also verify that the abstract does not use any user defined symbols

1 Introduction

The scalar resonance discovered by the CMS and ATLAS Collaborations at the LHC [1–3] in 2012 has been found to have properties consistent with the predictions of the standard model (SM) for a Higgs boson with a mass of about 125 GeV [4]. In particular, its couplings to bosons (g_{HVV}) and fermions (y_f) corroborate an SM-like dependence on the respective masses. Furthermore, data indicate that it has zero spin and positive parity [5]. Recently, the associated production of top quark pairs with a Higgs boson ($t\bar{t}H$) and Higgs boson decays to pairs of bottom quarks have been observed [6–8], thereby directly probing the Yukawa interactions between the Higgs boson and top as well as bottom quarks for the first time. In addition to measuring the absolute strengths of Higgs boson couplings, it is pertinent to assess the possible existence of relative phases among the couplings, as well as their general Lorentz structure. Hence a broad sweep of Higgs boson production mechanisms and decay modes must be considered to reveal any potential deviations from the SM expectations.

The production rate of $t\bar{t}H$ is sensitive only to the magnitude of the top quark-Higgs boson Yukawa coupling y_t and has no sensitivity to its sign. Measurements of processes such as Higgs boson decays to photon pairs [9] or the associated production of Z and Higgs bosons via gluon-gluon fusion [10] on the other hand do have sensitivity to the sign, because of indirect effects in loop interactions. Those measurements currently disfavor a negative value of the coupling [11, 12], but rely on the assumption that only SM particles contribute to the loops [13].

In contrast, the production of Higgs bosons in association with single top quarks in proton-proton (pp) collisions proceeds via two categories of Feynman diagrams [14–17], where the Higgs boson couples either to the top quark or the W boson. The two leading order (LO) diagrams for the t channel production process (tHq) are shown in Fig. 1. Because of the interference of these diagrams, the production cross section is uniquely sensitive to the magnitude as well as the relative sign and phase of the couplings.

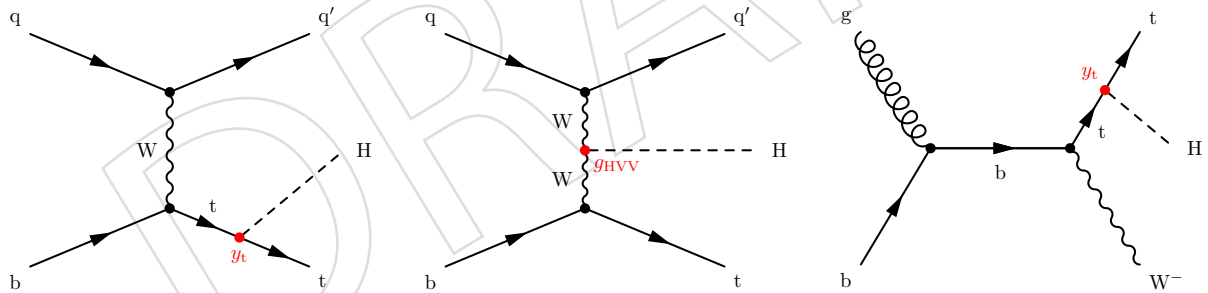


Figure 1: Leading order Feynman diagrams for the associated production of a single top quark and a Higgs boson in the t channel, where the Higgs boson couples either to the top quark (left) or the W boson (center), and one representative diagram of tHW production, where the Higgs boson couples to the top quark (right).

In the SM, the interference between these two diagrams is maximally destructive and leads to very small production cross sections of about 71, 16, and 2.9 fb for the t channel, tW process (tHW), and s channel, respectively, at a center-of-mass energy $\sqrt{s} = 13$ TeV [18, 19]. Hence measurements using the data collected at the LHC so far are not yet sensitive to the SM production. However, in the presence of new physics, there may be relative opposite signs between the t -H and W -H couplings which lead to constructive interference and enhance the cross sections by an order of magnitude or more. In such scenarios, realized, e.g., in some two-Higgs doublet models [20], tHq production would exceed that of $t\bar{t}H$ production, making it accessible with current LHC data sets. In this paper, the tHq and tHW processes are collectively referred to as tH production, while s channel production is neglected.

The event topology of tHq production is that of two heavy objects—the top quark, and the Higgs boson—in the central portion of the detector recoiling against one another, while a light-flavor quark and a soft b quark escape in the forward-backward regions of the detector. Leptonic top quark decays produce high-momentum electrons and muons that can be used to trigger the detector readout. Higgs boson decays to vector bosons or τ leptons ($H \rightarrow WW^*, ZZ^*, \text{or } \tau\tau$), which subsequently decay to light leptons ($\ell = e^\pm, \mu^\pm$), lead to a multilepton final state with comparatively small background contributions from other processes. Higgs boson decays to bottom quark-antiquark pairs ($H \rightarrow b\bar{b}$), on the other hand, provide a larger event rate albeit with challenging backgrounds from $t\bar{t}$ +jets production. In contrast, the rarer Higgs boson decays to two photons ($H \rightarrow \gamma\gamma$) result in easily triggered and relatively clean signals for both leptonic or fully hadronic top quark decays, with backgrounds mainly from other production modes of the Higgs boson. The production of tHW lacks the presence of forward activity and involves three heavy objects and therefore does not exhibit the defining features of tHq events, while closely resembling the $t\bar{t}H$ topologies, having identical final states.

The CMS Collaboration has previously searched for anomalous tHq production in pp collision data at $\sqrt{s} = 8$ TeV, assuming a negative sign of the top quark Yukawa coupling relative to its SM value, $y_t = -y_t^{\text{SM}}$, using all the relevant Higgs boson decay modes, and set limits on the cross section of this process [21]. This paper describes two new analyses targeting multilepton final states and single-lepton + $b\bar{b}$ final states, using a data set of pp collisions at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 35.9 fb^{-1} , collected in 2016. Furthermore, a previous measurement of Higgs boson properties in the $H \rightarrow \gamma\gamma$ final state at 13 TeV [22] has been reinterpreted in the context of tHq signal production and the results are included in a combination with those from the other channels.

This paper is structured as follows: the experimental setup and data samples are described in Sections 2 and 3 respectively; the two analysis channels and their event selection, background estimations, and signal extraction techniques are described in Sections 4 and 5; the reinterpretation of the $H \rightarrow \gamma\gamma$ result is described in Section 6; and the results and interpretation are given in Section 7. The paper is summarized in Section 8.

2 The CMS experiment

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T along the beam direction. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections providing pseudorapidity coverage up to $|\eta| < 3.0$. Forward calorimeters employing Cherenkov light detection extend the acceptance to $|\eta| < 5.0$. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid with a fiducial coverage of $|\eta| < 2.4$. The silicon tracker system measures charged particles within the range $|\eta| < 2.5$. The impact parameters in the transverse and longitudinal direction are measured with an uncertainty of about 10 and $30 \mu\text{m}$, respectively [23]. Tracks of isolated muons of transverse momentum $p_T \geq 100 \text{ GeV}$ and $|\eta| < 1.4$ are reconstructed with an efficiency close to 100% and a p_T resolution of about 1.3 to 2% and smaller than 6% for higher values of η . For $p_T \leq 1 \text{ TeV}$ the resolution in the central region is better than 10%. A two-level trigger system is used to reduce the rate of recorded events to a level suitable for data acquisition and storage. The first level of the CMS trigger system [24], composed of custom hardware processors, uses information from the calorimeters and muon detectors to select the most interesting events in a time interval of less than $4 \mu\text{s}$. The high-level trigger processor farm further decreases the

event rate from around 100 kHz to about 1 kHz. A more detailed description of the CMS detector, together with a definition of the coordinate system and the kinematic variables used in the analysis, can be found in Ref. [25].

A full event reconstruction is performed by the particle-flow (PF) algorithm using optimized and combined information from all the subdetectors [26]. The individual PF candidates reconstructed are muons, electrons, photons, and charged and neutral hadrons, which are then used to reconstruct higher-level objects such as jets, hadronic taus, and missing transverse momentum (p_T^{miss}). Additional quality criteria are applied to the objects to improve the selection purity.

Collision vertices are reconstructed using a deterministic annealing algorithm [27, 28]. The reconstructed vertex position is required to be compatible with the location of the LHC beam in the x - y plane. The vertex with the largest value of summed physics-object p_T^2 is considered to be the primary pp interaction (PV). Charged particles, which are subsequently reconstructed, are required to be compatible with originating from the selected PV.

The identification of muons is based on linking track segments reconstructed in the silicon tracker and in the muon system [29]. If a link can be established, the track parameters are recomputed using the combination of hits in the inner and outer detectors. Quality requirements are applied on the multiplicity of hits in the track segments, on the number of matched track segments, and on the quality of the track fit [29].

Electrons are reconstructed using an algorithm that matches tracks found in the silicon tracker with energy deposits in the electromagnetic calorimeter while limiting deposits in the hadronic calorimeter [30]. A dedicated algorithm takes into account the emission of bremsstrahlung photons and determines the energy loss [31]. A multivariate analysis (MVA) approach based on boosted decision trees (BDT) is employed to distinguish real electrons from hadrons mimicking an electron signature. Additional requirements are applied in order to remove electrons originating from photon conversions [30]. Both muons and electrons from signal events are expected to be isolated, while those from heavy-flavor decays are often situated near jets. Lepton isolation is quantified using the scalar p_T sum over PF candidates reconstructed within a cone centered on the lepton direction and shrinking with increasing lepton p_T . The effect of additional pp interactions in the same and nearby bunch crossings (pileup) on the lepton isolation is mitigated by considering only charged particles consistent with the PV in the sum, and by subtracting an estimate of the contribution from neutral pileup particles within the cone area.

Jets are reconstructed from charged and neutral PF candidates using the anti- k_T algorithm [32, 33] with a distance parameter of 0.4, and with the constraint that the charged particles are compatible with the selected PV. Jets originating from the hadronization of b quarks are identified using the “combined secondary vertex” (CSVv2) algorithm [34], which exploits observables related to the long lifetime of b hadrons and to the higher particle multiplicity and mass of b jets compared to light-quark and gluon jets. Two working points of the CSVv2 discriminant output are used: a “medium” one, with a tagging efficiency for real b jets of 69% and a probability of wrongly tagging jets from light-flavor quarks and gluons of about 1%, and a “loose” one, with a tagging efficiency of 83% and a mistag rate for light-flavor jets of 8%. Finally, the missing transverse momentum is defined as the magnitude of the vectorial p_T sum of all PF candidates in the event.

3 Data and simulation

Collision events for this analysis are selected by the following high-level trigger algorithms. Events in the multilepton channels must pass at least one of single-lepton, dilepton, or trilepton triggers with loose identification and isolation requirements and with a minimum p_T threshold based on the lepton multiplicity in the final state. Events in the single lepton + $b\bar{b}$ channels must pass the same single-lepton triggers, or a dilepton trigger for the control region described in Section 5. The minimum p_T threshold for single lepton triggers is 24 GeV for muons and 27 GeV for electrons. For dilepton triggers, the p_T thresholds on the leading and subleading leptons are 17 GeV and 8 GeV for muons, and 23 GeV and 12 GeV for electrons, respectively. For the trilepton trigger, the third hardest lepton p_T must be greater than 5 GeV for muons and 9 GeV for electrons.

The data are compared to signal and background estimations based on Monte Carlo (MC) simulated samples and techniques based on control samples in data. All simulated samples include the response of the CMS detector based on the GEANT4 [35] toolkit and are generated with a Higgs boson mass of 125 GeV and a top quark mass of 172.5 GeV. The event generator used for the tHq and tHW signal samples is MADGRAPH5_aMC@NLO (version 2.2.2) [36] at LO precision [37] and using the NNPDF3.0 set of parton distribution functions (PDF) [38] with the PDF4LHC prescription [39, 40]. The samples are normalized to next-to-leading order (NLO) SM cross sections at 13 TeV of 71.0 and 15.6 fb for tHq and tHW, respectively [18, 19].

The Higgs boson production cross sections and branching fractions are expressed as functions of Higgs boson coupling modifiers in the kappa framework [41], where the coupling modifiers κ are defined as the ratio of the actual value of a given coupling to the one predicted by the SM. Particularly relevant for the tH case are the top quark and vector boson coupling modifiers: $\kappa_t \equiv y_t/y_t^{\text{SM}}$ and $\kappa_V \equiv g_{\text{H}VV}/g_{\text{H}VV}^{\text{SM}}$, where V stands for either W or Z bosons. The dependence of the tHq and tHW production cross sections on κ_t and κ_V are assumed to be as follows (calculated at NLO using MADGRAPH5_aMC@NLO [17–19]):

$$\begin{aligned}\sigma_{\text{tHq}} &= (2.63 \kappa_t^2 + 3.58 \kappa_V^2 - 5.21 \kappa_t \kappa_V) \sigma_{\text{tHq}}^{\text{SM}}, \\ \sigma_{\text{tHW}} &= (2.91 \kappa_t^2 + 2.31 \kappa_V^2 - 4.22 \kappa_t \kappa_V) \sigma_{\text{tHW}}^{\text{SM}}.\end{aligned}$$

Event weights are produced in the generation of both samples corresponding to 33 values of κ_t between -6.0 and $+6.0$, and for $\kappa_V = 1.0$. The tHq events are generated with the four-flavor scheme (4FS) while the tHW process uses the five-flavor scheme (5FS) to disentangle the LO interference with the $t\bar{t}$ H process [19].

The MADGRAPH5_aMC@NLO generator is also used for simulation of the $t\bar{t}$ H process and the main backgrounds: associated production of $t\bar{t}$ pairs with vector bosons, $t\bar{t}W$, $t\bar{t}Z$, at NLO [42], and with additional jets or photons, $t\bar{t} + \text{jets}$, $t\bar{t}\gamma + \text{jets}$ at LO. All the rates are normalized to next-to-next-leading order cross sections. In particular, the $t\bar{t}$ H production cross section is taken as 0.507 pb [18]. A set of minor backgrounds are also simulated with MADGRAPH5_aMC@NLO at LO, or with other generators, such as NLO POWHEG v2 [43–48]. All generated events are interfaced with PYTHIA (8.205) [49] for the parton shower and hadronization steps.

The object reconstruction in MC events uses the same algorithm as used in data. Furthermore, the trigger selection is simulated and applied for generated signal events. However, the triggering and selection efficiencies for leptons are different between data and simulation, at the level of 1%. All simulated events used in the analyses are corrected by applying small data-to-MC scale factors to improve the modeling of the data. Separate scale factors are applied to correct for the difference in trigger efficiency, lepton reconstruction and selection efficiency, as well as the b tagging efficiency and the resolution of the missing transverse momentum.

Simulated events are weighted according to the number of pileup interactions so that the distribution of additional pp interactions in the simulated samples matches that observed in data, as estimated from the measured bunch-to-bunch instantaneous luminosity and the total inelastic cross section.

4 Multilepton channels

Signal tH events where the top quark decay produces leptons and the Higgs boson decays to vector bosons or τ leptons can lead to final states containing multiple isolated, high- p_T leptons with different charge and flavor configurations. Of particular interest among these are those with three or more charged leptons or with two leptons of the same electric charge, as they appear with comparatively low backgrounds. Selecting such events in pp collisions while requiring the presence of b-tagged jets typically yields a mixture of mostly $t\bar{t}$ +jets events with nonprompt leptons and events from the associated production of $t\bar{t}$ with a vector boson ($t\bar{t}W$ and $t\bar{t}Z$) or with a Higgs boson ($t\bar{t}H$) that decay to additional prompt leptons. The analysis described in this section separates the tHq signal from these two dominant background sources by training two multivariate classifiers using features such as the forward light jet, the difference in multiplicities of jets and b-tagged jets (“b jets”), as well as the kinematic properties of the leptons. The two classifier outputs are combined into a single binned distribution, which is then fit to the data to extract the signal yield and constrain the background contributions.

4.1 Event and object selections

In the multilepton channels, events are selected with trigger algorithms involving one, two, or three leptons passing the given p_T thresholds. At the offline analysis level, a distinction is made between prompt signal leptons (from W, Z, or leptonic τ decays) and nonprompt leptons (either genuine leptons from heavy-flavor hadron decays, asymmetric γ conversions, or jets misidentified as leptons). For this purpose an MVA classifier is used [50], exploiting the properties of the jet associated with individual leptons in addition to the lepton kinematics, isolation, and reconstruction quality. The leptons are selected if they pass a certain threshold of the classifier output and are referred to as “tight” leptons, with a lower threshold defined for a relaxed selection and “loose” leptons.

The final tH event selection targets signatures with $H \rightarrow WW$ and $t \rightarrow Wb \rightarrow \ell\nu b$, which results in three W bosons, one b quark, and a light quark at high rapidity. Three mutually exclusive channels are defined based on the number of tight leptons and their flavors: exactly two same-sign leptons ($2\ell ss$), either $\mu^\pm\mu^\pm$ or $e^\pm\mu^\pm$, or exactly three leptons ($\ell\ell\ell$, $\ell = \mu$ or e). The same-sign dielectron channel suffers from larger backgrounds and does not add sensitivity and is therefore not included in the analysis. There is an additional requirement of at least one b-tagged jet (using the medium working point of the CSVv2 algorithm) and at least one light-flavor (untagged, using the loose working point) jet. The full selection is summarized in Table 1.

About one quarter of the events in the finally selected sample are from $H \rightarrow \tau\tau$ and $H \rightarrow ZZ$ decays, with the rest coming from $H \rightarrow WW$ decays, as determined from the tHq signal simulation. A significant fraction of selected events also pass the selection used in the dedicated search for $t\bar{t}H$ in multilepton channels [50]: about 50% in the dilepton channels and about 80% in the $\ell\ell\ell$ channels.

Table 1: Summary of the event selection for the multilepton channels.

Same-sign channel ($\mu^\pm\mu^\pm$ or $e^\pm\mu^\pm$)
Exactly two tight SS leptons
$p_T > 25/15$ GeV
No loose leptons with $m_{\ell\ell} < 12$ GeV
One or more b-tagged jet with $p_T > 25$ GeV and $ \eta < 2.4$
One or more untagged jets with $p_T > 25$ GeV for $ \eta < 2.4$ and $p_T > 40$ GeV for $ \eta > 2.4$
$\ell\ell\ell$ channel
Exactly three tight leptons
$p_T > 25/15/15$ GeV
No lepton pair with $ m_{\ell\ell} - m_Z < 15$ GeV
No loose leptons with $m_{\ell\ell} < 12$ GeV
One or more b-tagged jet with $p_T > 25$ GeV and $ \eta < 2.4$
One or more untagged jets with $p_T > 25$ GeV for $ \eta < 2.4$ and $p_T > 40$ GeV for $ \eta > 2.4$

4.2 Backgrounds

The background processes contributing to the signal sample can be divided into two classes, reducible and irreducible, and are estimated respectively from data and MC simulation. Irreducible physics processes, such as the associated production of an electroweak boson with a top quark pair ($t\bar{t}V$, $V = W, Z$), give rise to final states very similar to the tHq signal and are directly estimated from MC simulation. However, the dominant contribution is from the reducible background arising from nonprompt leptons, mainly from $t\bar{t}$ production. This background is suppressed to a certain extent by tightening the lepton selection criteria. The background estimation methods employed here and summarized below are identical to those used in the dedicated search for $t\bar{t}H$ in multilepton channels [50].

The yield of reducible backgrounds is estimated from the data, using a “tight-to-loose” ratio measured in a control region dominated by nonprompt leptons. The ratio represents the probability with which the nonprompt leptons that pass the looser selection can also pass the tight criteria, and is measured in categories of the lepton p_T and η . A sideband region in data which has loosely selected leptons is then extrapolated with this ratio to obtain the nonprompt background contribution.

A further background in the same-sign dilepton channels arises from events where the charge of one lepton is wrongly assigned. This can be estimated from the data, by measuring the charge misidentification probability using the Z boson mass peak in same-sign dilepton events, and weighting events with opposite-sign leptons to determine the yield in the signal region. The effect is found to be negligible for muons but sizable for electrons.

The production of WZ pairs with leptonic Z boson decays has similar leptonic features as the signal, but usually lacks the hadronic activity required in the signal selection. To determine the corresponding diboson contribution in the signal region, simulated WZ events have been used along with a normalization scale factor determined from data in an exclusive control region.

Other subdominant backgrounds are estimated from MC simulation and include additional multiboson production, such as ZZ, $W^\pm W^\pm qq$, VVV, same-sign W boson production from

double-parton scattering (DPS), associated production of top quarks with Z bosons (tZq, tZW), events with four top quarks, and $t\bar{t}$ production in association with photons and subsequent asymmetric conversions.

The expected and observed event yields after the selections described in Table 1 are shown in Table 2.

Table 2: Data yields and expected backgrounds after the event selection for the three multilepton search channels in 35.9 fb^{-1} of integrated luminosity. Quoted uncertainties include statistical uncertainties reflecting the limited size of MC samples and data sidebands, and unconstrained systematic uncertainties.

Process	$\mu^\pm\mu^\pm$	$e^\pm\mu^\pm$	lll
$t\bar{t}W$	68 ± 10	97 ± 13	22.5 ± 3.1
$t\bar{t}Z/t\bar{t}\gamma$	25.9 ± 3.9	64.8 ± 9.0	32.8 ± 5.1
WZ	15.1 ± 7.7	26 ± 13	8.2 ± 2.4
ZZ	1.16 ± 0.65	2.9 ± 1.5	1.62 ± 0.87
$W^\pm W^\pm qq$	4.0 ± 2.1	7.0 ± 3.6	—
$W^\pm W^\pm$ (DPS)	2.5 ± 1.3	4.2 ± 2.2	—
VVV	3.0 ± 1.5	4.9 ± 2.5	0.42 ± 0.26
$t\bar{t}t\bar{t}$	2.3 ± 1.2	4.1 ± 2.1	1.8 ± 1.0
tZq	5.8 ± 3.6	10.7 ± 6.1	3.9 ± 2.5
tZW	2.1 ± 1.1	3.9 ± 2.0	1.70 ± 0.86
γ conversions	—	23.8 ± 7.8	7.4 ± 2.8
Nonprompt	80.9 ± 9.4	135 ± 35	26 ± 14
Charge misidentification	—	58 ± 17	—
Total background	211 ± 17	443 ± 45	106 ± 16
$t\bar{t}H$	24.2 ± 2.1	35.2 ± 2.9	18.3 ± 1.7
tHq (SM)	1.43 ± 0.12	1.92 ± 0.15	0.52 ± 0.04
tHW (SM)	0.71 ± 0.06	1.11 ± 0.09	0.62 ± 0.05
Total SM	237 ± 17	482 ± 45	126 ± 16
tHq ($\kappa_V = 1 = -\kappa_t$)	18.5 ± 1.6	27.4 ± 2.1	7.48 ± 0.58
tHW ($\kappa_V = 1 = -\kappa_t$)	7.72 ± 0.65	11.23 ± 0.91	7.38 ± 0.60
Data	280	525	127

4.3 Signal extraction

After applying the event selection of the multilepton channels, only about one percent of selected events are expected to be from tH production (assuming SM cross sections), while roughly 10% of events are from $t\bar{t}H$ production. To discriminate this small signal from the backgrounds, an MVA method is employed: a classification algorithm is trained twice with tHq events as the signal class, and either $t\bar{t}V$ (mixing $t\bar{t}W$ and $t\bar{t}Z$ according to their respective cross sections) or $t\bar{t}$ +jets as background classes. The two separate trainings allow the exploitation of the different jet and b jet multiplicity distributions, and of the different kinematic properties of the leptons in the two dominant background classes. Several machine learning algorithms were studied for potential use, and the best performance was obtained with a gradient BDT using a maximum tree depth of three and an ensemble of 800 trees [51]. Events from tHW and $t\bar{t}H$ production are not used in the training and, because of their kinematic similarity with the $t\bar{t}V$ background, tend to be classified as backgrounds.

As observed above, the features of the tHq signal can be split into three broad categories: those related to the forward jet activity; those related to jet and b-jet multiplicities; and those related to kinematic properties of leptons, as well as their total charge. A set of ten observables were used as input features to the classification training, and are listed in Table 3. The training is performed separately for the $2\ell ss$ and the $\ell\ell\ell$ channels with the same or equivalent input features.

Table 3: Input observables to the signal discrimination classifier.

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

A selection of the main discriminating input observables is shown in Figs. 2–4, comparing the data and the estimated distribution of signal and background processes.

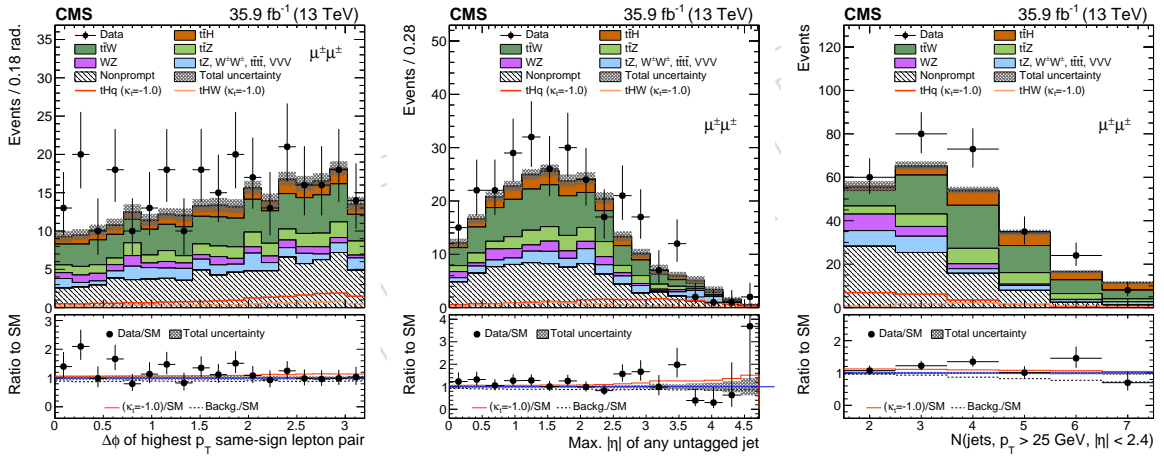


Figure 2: Distributions of discriminating observables for the same-sign $\mu^\pm\mu^\pm$ channel, normalized to 35.9 fb^{-1} , before fitting the signal discriminant to the data. The grey band represents the unconstrained (pre-fit) statistical and systematic uncertainties. In the panel below each distribution, 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 six classifier output distributions, trained against $t\bar{t}V$ and $t\bar{t}+\text{jets}$ processes for each of the three channels, are shown in Fig. 5, before a fit to the data. 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. Figure 6 shows the post-fit categorized classifier output distributions for each of the three channels, after the combined maximum likelihood fit to extract the limits, as described in Section 7.

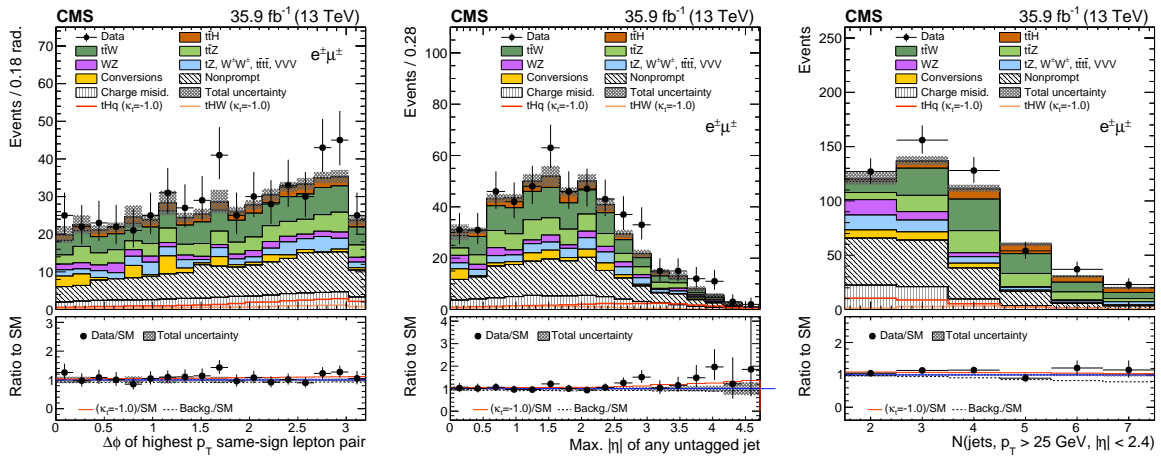


Figure 3: Distributions of discriminating observables for the same-sign $e^\pm\mu^\pm$ channel, normalized to 35.9 fb^{-1} , before fitting the signal discriminant to the data. The grey band represents the unconstrained (pre-fit) statistical and systematic uncertainties. In the panel below each distribution, 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 $\kappa_t = -1.0, \kappa_V = 1.0$.

4.4 Systematic uncertainties

The yield of signal and background events after the selection, as well as the shape of the classifier output distributions for signal and background processes, have systematic uncertainties from a variety of sources, both experimental and theoretical. Experimental uncertainties relate either to the reconstruction of physics objects or to imprecisions in estimating the background contributions. Uncertainties in the efficiency of reconstructing and selecting physics objects affect all yields and kinematic shapes taken from MC simulation, for both signal and background. Background contributions estimated from the data are not affected by these.

Uncertainties from unknown higher-order contributions to tHq and tHW production are estimated from a change in the factorization and renormalization scales of double and half the initial value, evaluated separately for each point of κ_t . The $t\bar{t}H$ component has an uncertainty of between 5.8–9.3% from scale variations and an additional 3.6% from the knowledge of PDFs and the strong coupling constant α_S [18]. Uncertainties related to the choice of the PDF set and its scale are estimated to be 3.7% for tHq and 4.0% for tHW. The effect of missing higher-order corrections to the kinematic shape of the classifier outputs is taken into account for the tH, $t\bar{t}H$, and $t\bar{t}V$ components by independent changes of the renormalization and factorization scales of double and half the nominal value.

The cross sections of $t\bar{t}Z$ and $t\bar{t}W$ production are known with uncertainties of $+9.6\%/-11.2\%$ and $+12.9\%/-11.5\%$, respectively, from missing higher-order corrections to the perturbative expansion. The corresponding values due to uncertainties in the PDFs and α_s are 3.4 and 4.0%, respectively [18].

The efficiency for events passing the combination of trigger requirements is measured separately for events with two or more leptons, and has an uncertainty in the range of 1–3%. Efficiencies for the reconstruction and selection of muons and electrons are measured as a function of their p_T , using a tag-and-probe method with uncertainties of 2–4% [52]. The energy scale of jets is determined using event balancing techniques and carries uncertainties of a few percent, depending on p_T and η of the jets [53]. Their impact on the kinematic distributions used in the

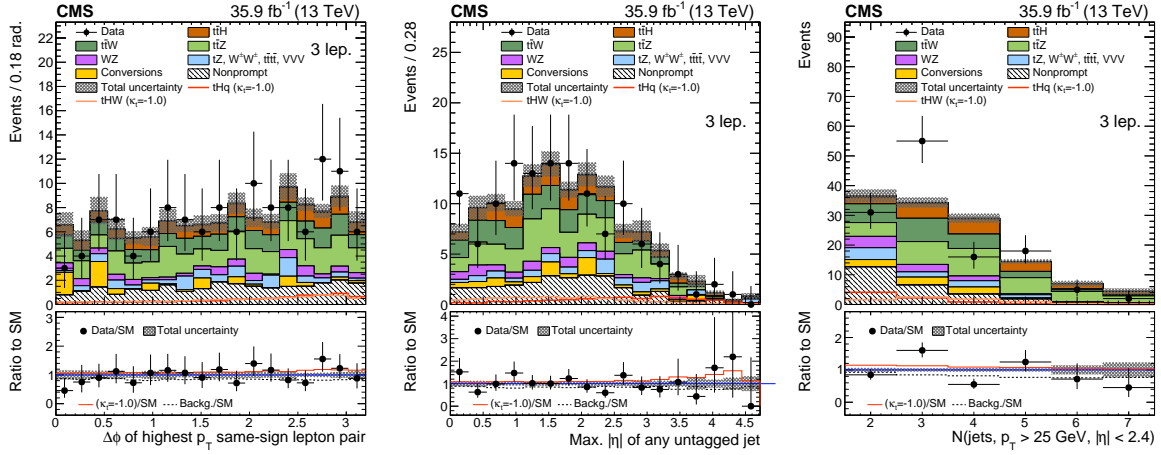


Figure 4: Distributions of discriminating observables for the three lepton channel, normalized to 35.9 fb^{-1} , before fitting the signal discriminant to the data. The grey band represents the unconstrained (pre-fit) statistical and systematic uncertainties. In the panel below each distribution, the ratio of the observed and predicted event yields is shown. The shape of the two $t\bar{t}H$ signals for $\kappa_t = -1.0$ is shown, normalized to their respective cross sections for $\kappa_t = -1.0, \kappa_V = 1.0$.

signal extraction are estimated by varying the scales within their respective uncertainty and propagating the effects to the final result, recalculating all kinematic quantities and reapplying the event selection criteria. The b tagging efficiencies are measured in heavy-flavor enriched multijet events and in $t\bar{t}$ events, with p_T - and η -dependent uncertainties of a few percent [34].

The uncertainty in the integrated luminosity is 2.5% [54] and affects the normalization of all processes modeled in simulation.

The estimate of events containing nonprompt leptons is subject to uncertainties in the determination of the tight-to-loose ratio on one hand and to the inherent bias in the selection of the control region dominated by nonprompt leptons, as tested in simulated events, on the other hand. The measurement of the lepton tight-to-loose rate has statistical as well as systematic uncertainties from the removal of residual prompt leptons in the control region, amounting to a total uncertainty of 10–40%, depending on the flavor of the leptons and their p_T and η . The validity of the method itself is tested in simulated events and contributes a small additional uncertainty both to the normalization and shape of the classifier distributions for such events.

The estimate of backgrounds from electron charge misidentification in the $e^\pm\mu^\pm$ channel carries an uncertainty of about 30% from the measurement of the misidentification probability. It is composed of a dominant statistical component from the limited event yields, and one related to the residual disagreement observed when testing the prediction in a control region.

The estimate of backgrounds from WZ production is normalized in a control region with three leptons and carries uncertainties due to its limited statistics (10%), the residual non- WZ backgrounds (20%), the b tagging rate (10%), and the theoretical uncertainties related to the flavor composition of jets produced in association with the boson pair (up to 10%). In the dilepton channels, this uncertainty is increased to 50% to account for the differences with respect to the control region.

Additional smaller backgrounds which have not yet been observed at the LHC (VVV , same-sign W boson production, tZq , tZW , $t\bar{t}t\bar{t}$) are assigned a normalization uncertainty of 50%.

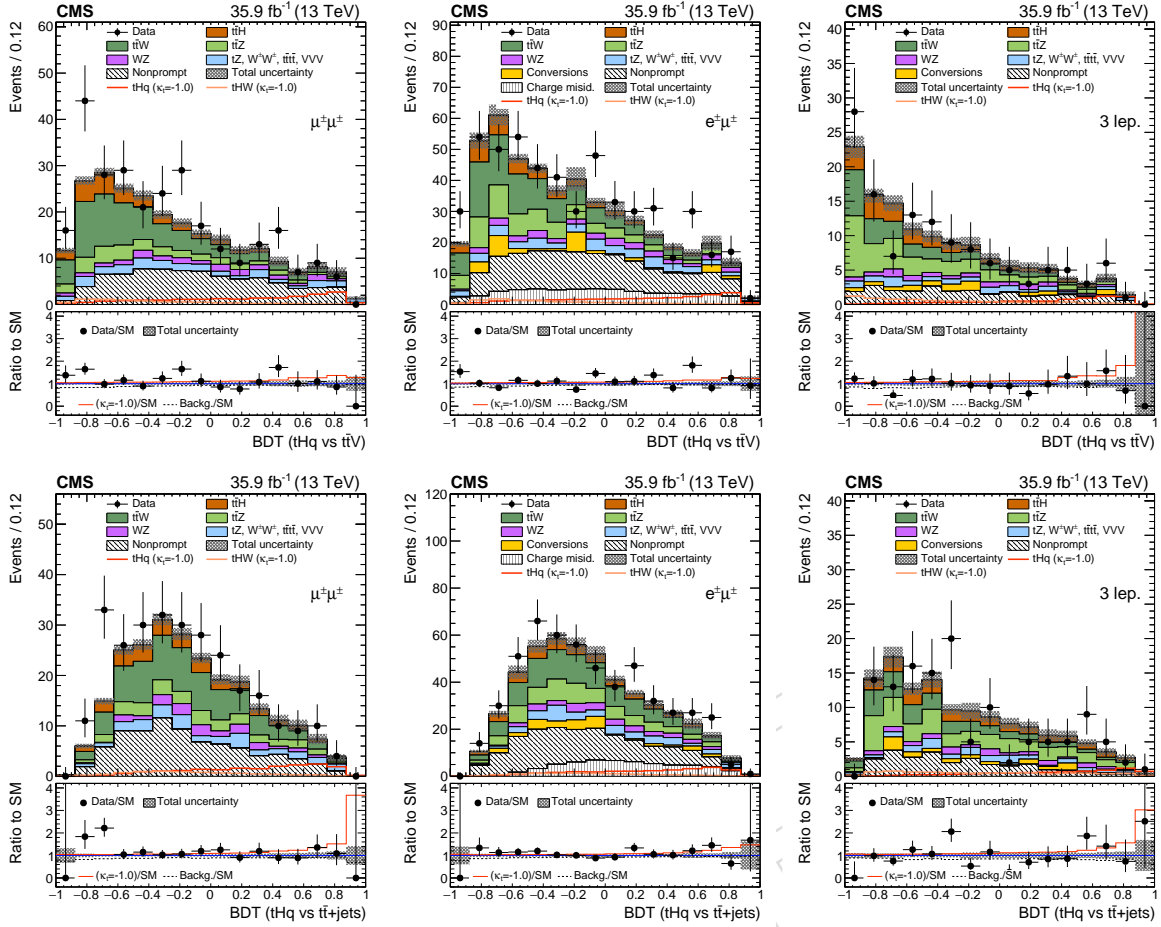


Figure 5: Pre-fit classifier outputs, for the $\mu^+\mu^-$ channel (left), $e^+\mu^-$ channel (center), and three-lepton channel (right), for training against $t\bar{t}V$ (top row) and against $t\bar{t}+\text{jets}$ (bottom row). In the box below each distribution, the ratio of the observed and predicted event yields is shown. The shape of the two $t\bar{t}H$ 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 systematic uncertainties.

Of these sources of systematic uncertainties, the ones with largest impact on the final result are found to be those related to the normalization of the nonprompt backgrounds, the scale variations for the $t\bar{t}V$ and $t\bar{t}H$ processes, and the lepton selection efficiencies.

5 Single-lepton + $b\bar{b}$ channels

Events from a $t\bar{t}H$ signal where the Higgs boson decays to a bottom quark-antiquark pair ($H \rightarrow b\bar{b}$) produce final states with at least three central b jets and a hard lepton from the top quark decay chain used for triggering. Selecting such events leads to challenging backgrounds from $t\bar{t}$ production with additional heavy-flavor quarks, which can be produced in gluon splittings from initial- or final-state radiation. The analysis described in this section uses two selections aimed at identifying signal events, with either three or four b -tagged jets, and a separate sample with opposite-sign dileptons, dominated by $t\bar{t}+\text{jets}$ events, to control $t\bar{t} + \text{heavy-flavor}$ ($t\bar{t}+\text{HF}$) events in a simultaneous fit. A multivariate classification algorithm is trained to discriminate different $t\bar{t}+\text{jets}$ background components in the control region. Ad-

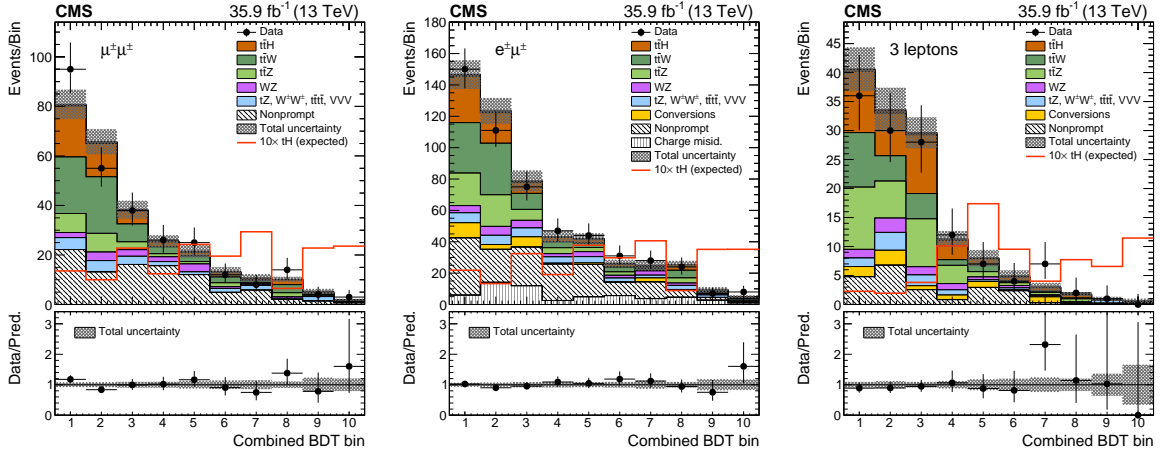


Figure 6: Post-fit categorized classifier outputs as used in the maximum likelihood fit for the $\mu^\pm\mu^\pm$ channel (left), $e^\pm\mu^\pm$ channel (center), and three-lepton channel (right), for 35.9 fb^{-1} . In the box below each distribution, the ratio of the observed and predicted event yields is shown. The shape of the tH signal is indicated with 10 times the SM.

ditional multivariate algorithms are used to optimize the jet-parton assignment used to reconstruct kinematic properties of signal and background events which, in turn, are used to distinguish these components.

5.1 Selection

Selected events in the single-lepton + $b\bar{b}$ signal channels must pass a single-lepton trigger. Each event is required to contain exactly one muon or electron. Muon (electron) candidates are required to satisfy $p_T > 27$ (35) GeV and $|\eta| < 2.4$ (2.1), motivated by the trigger selection, and to be isolated and fulfill strict quality requirements. Events with additional leptons that have $p_T > 15$ GeV and pass less strict quality requirements are rejected. At the analysis level, the selection criteria target the $H \rightarrow b\bar{b}$ and $t \rightarrow Wb \rightarrow \ell\nu b$ decay channels. With these decays, the final state of the tHq process consists of one W boson, three b quarks, and the light-flavor quark recoiling against the top quark-Higgs boson system. In addition, a fourth b quark is expected because of the initial gluon splitting, but often falls outside the detector acceptance. The main signal region is therefore required to have either three or four b-tagged jets and at least one additional untagged jet, both defined using the medium working point. Central jets with $|\eta| < 2.4$ are required to have $p_T > 30$ GeV, while jets in the forward region ($2.4 \leq |\eta| \leq 4.7$) are required to have $p_T > 40$ GeV.

The neutrino is accounted for by requiring a minimal amount of missing transverse momentum of $p_T^{\text{miss}} > 35$ GeV in the muon channel and $p_T^{\text{miss}} > 45$ GeV in the electron channel. This renders the background from QCD multijet events negligible.

In addition to the signal regions, a control region is defined to constrain the main background contribution from top quark pair production. Events selected for this control region must pass a dilepton trigger. Each event is required to contain exactly two oppositely charged leptons, where their flavor can be any combination of muons or electrons. Two jets in each event must be b tagged. Furthermore, at least one additional jet must pass the loose b tagging requirement. Similarly to the signal regions, each event is further required to have a minimum amount of missing transverse momentum. All selection criteria are summarized in Table 4.

Table 4: Summary of event selection for the single-lepton + $b\bar{b}$ channels.

Signal region
One muon (electron) with $p_T > 27(35)$ GeV
No additional loose leptons with $p_T > 15$ GeV
Three or four medium b-tagged jets
$p_T > 30$ GeV and $ \eta < 2.4$
One or more untagged jets
$p_T > 30$ GeV for $ \eta < 2.4$ or
$p_T > 40$ GeV for $ \eta \geq 2.4$
$p_T^{\text{miss}} > 35(45)$ GeV for muons (electrons)
Control region
Two leptons: $p_T > 20/20$ GeV ($\mu^\pm\mu^\mp$)
or $p_T > 20/15$ GeV ($e^\pm e^\mp/\mu^\pm e^\mp$)
No additional loose leptons
with $p_T > 20/15$ GeV (μ^\pm/e^\pm)
Two medium b-tagged jets
$p_T > 30$ GeV and $ \eta < 2.4$
One or more additional loose b-tagged jets
$p_T > 30$ GeV and $ \eta < 2.4$
$p_T^{\text{miss}} > 40$ GeV

5.2 Backgrounds

The main background contribution in the single-lepton + $b\bar{b}$ channels arises from SM processes with multiple b quarks. The modeling and estimation of all background processes are done using samples of simulated events.

In particular, the dominant background process is top quark pair production because of the similar final state and, comparatively, a large cross section. As the modeling of the additional heavy-flavor partons in $t\bar{t}$ events is theoretically difficult, the sample of simulated $t\bar{t}$ events is further divided into different subcategories, defined by the flavor of possible additionally radiated quarks and taking into account a possible merging of b hadrons into single jets. The control region is specifically designed to separate the $t\bar{t}$ +HF and $t\bar{t}$ + light-flavor ($t\bar{t}$ +LF) components with a multivariate approach. The different categories are listed in Table 5.

Table 5: Subcategories of $t\bar{t}$ +jets backgrounds used in the analysis.

$t\bar{t}+b\bar{b}$	Two additional jets arising from b hadrons
$t\bar{t}+2b$	One additional jet arising from two merged b hadrons
$t\bar{t}+b$	One additional jet arising from one b hadron
$t\bar{t}+c\bar{c}$	The three former categories combined for c hadrons instead of b hadrons
$t\bar{t}+LF$	All events that do not meet the criteria of the other four categories

Other backgrounds contributing to the signal region are single top quark production and top quark pair production in association with electroweak bosons, namely $t\bar{t}W$ and $t\bar{t}Z$. An irreducible background for the tHq processes comes from tZq production with $Z \rightarrow b\bar{b}$. Background contributions also arise from Z +jets production, especially in the dilepton control region.

372 The expected and observed event yields for the signal and control regions are listed in Table 6.

Table 6: Data yields and expected backgrounds after the event selection for the two signal regions and in the dilepton control region. Uncertainties include both systematic and statistical components.

Process	3 tags	4 tags	Dilepton
$t\bar{t}+LF$	24100 ± 5800	320 ± 180	5300 ± 1000
$t\bar{t}+c\bar{c}$	8500 ± 4900	340 ± 260	2100 ± 1200
$t\bar{t}+b\bar{b}$	4100 ± 2300	780 ± 430	750 ± 440
$t\bar{t}+b$	4000 ± 2100	180 ± 110	770 ± 430
$t\bar{t}+2b$	2300 ± 1200	138 ± 88	400 ± 230
Single top	1980 ± 350	78 ± 26	285 ± 37
$t\bar{t}Z$	202 ± 30	32.0 ± 6.6	54.8 ± 7.3
$t\bar{t}W$	90 ± 23	4.2 ± 2.8	31.4 ± 5.9
tZq	28.3 ± 5.7	2.9 ± 2.3	—
$Z+jets$	—	—	69 ± 32
Total background	45300 ± 8300	1880 ± 550	9700 ± 1700
$t\bar{t}H$	268 ± 31	62.0 ± 9.9	48.9 ± 5.9
tHq (SM)	11.1 ± 3.3	1.3 ± 0.3	0.31 ± 0.08
tHW (SM)	7.6 ± 1.1	1.1 ± 0.3	1.4 ± 0.2
Total SM	45700 ± 8300	1940 ± 550	9700 ± 1700
tHq ($\kappa_V = 1 = -\kappa_t$)	160 ± 38	19.1 ± 5.2	3.9 ± 1.0
tHW ($\kappa_V = 1 = -\kappa_t$)	92 ± 12	13.7 ± 2.3	17.6 ± 2.2
Data	44311	2035	9065

373 5.3 Signal extraction

374 As the assignment of final state quarks to reconstructed jets is non-trivial for the multijet environment of the 3 and 4 tag signal regions, the jet-to-quark assignment is achieved with dedicated jet assignment BDTs (JA-BDTs). Each event is reconstructed under three different hypotheses: tHq signal event, tHW signal event, or $t\bar{t}+jets$ background event. Each assignment hypothesis utilizes a separate BDT, which is trained with correct and wrong jet-to-quark assignments of the respective process. When a JA-BDT is applied, all possible jet-to-quark assignments are evaluated and the one with the highest JA-BDT output value is chosen for the given hypothesis. The matching efficiency for a complete tHq event is 58 (45)% in the 3 (4) tag signal region, for a complete tHW event 38 (29)% and for a complete $t\bar{t}$ event 58 (31)%.

383 The different assignment hypotheses provide sensitive variables, which can be exploited in a further signal classification BDT (SC-BDT) to separate the tHq and tHW processes from the main background of the analysis, $t\bar{t}$ events. Global event variables that do not rely on any particular jet-to-quark assignment are used in addition to the assignment-dependent variables. The input variables used for the SC-BDT are listed in Table 7 with the result of the training illustrated in Fig. 7.

389 In addition, a dedicated flavor classification BDT (FC-BDT) is used in the dilepton region to constrain the contribution of different $t\bar{t} + X$ components. The training is performed with $t\bar{t}+LF$ as signal process and $t\bar{t}+b\bar{b}$ as background process. This FC-BDT exploits information of the number of jets per event and their response to b and c tagging algorithms. The full list of input variables is provided in Table 8 and the result of the training of the FC-BDT is shown in Fig. 8.

Table 7: Description of variables used in the SC-BDT. There are four types of variables: variables independent of any jet assignment, and variables based on objects obtained under the $t\bar{t}$, tHq, or tHW jet assignment. The natural logarithm transformation is used to smoothen and constrain broad distributions to a more narrow range.

Variable	Description
Event variables	
$\ln m_3$	Invariant mass of three hardest jets in the event
Aplanarity	Aplanarity of the event [55]
Fox–Wolfram #1	First Fox–Wolfram moment [56] of the event
$q(\ell)$	Electric charge of the lepton
$t\bar{t}$ jet assignment variables	
$\ln m(t_{\text{had}})$	Invariant mass of the reconstructed hadronically decaying top quark
CSV(W_{had} jet 1)	Output of the b tagging discriminant for the first jet assigned to the hadronically decaying W boson
CSV(W_{had} jet 2)	Output of the b tagging discriminant for the second jet assigned to the hadronically decaying W boson
$\Delta R(W_{\text{had}}$ jets)	ΔR between the two light jets assigned to the hadronically decaying W boson
tHq jet assignment variables	
$\ln p_T(H)$	Transverse momentum of the reconstructed Higgs boson candidate
$ \eta(\text{light-flavor jet}) $	Absolute pseudorapidity of light-flavor forward jet
$\ln m(H)$	Invariant mass of the reconstructed Higgs boson candidate
CSV(H jet 1)	Output of the b tagging discriminant for the first jet assigned to the Higgs boson candidate
CSV(H jet 2)	Output of the b tagging discriminant for the second jet assigned to the Higgs boson candidate
$\cos \theta(b_t, \ell)$	Cosine of the angle between the b-tagged jet from the top quark decay and the lepton
$\cos \theta^*$	Cosine of the angle between the light-flavor forward jet and the lepton in the top quark rest frame
$ \eta(t) - \eta(H) $	Absolute pseudorapidity difference of reconstructed Higgs boson and top quark
$\ln p_T(\text{light jet})$	Transverse momentum of the light-flavor forward jet
tHW jet assignment variable	
JA-BDT response	Best output of the tHW JA-BDT

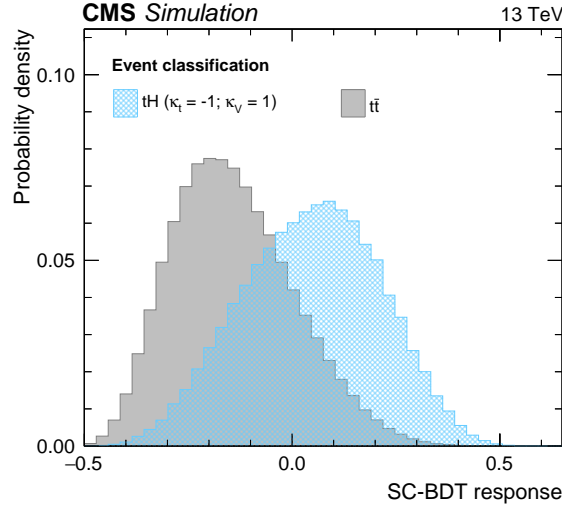
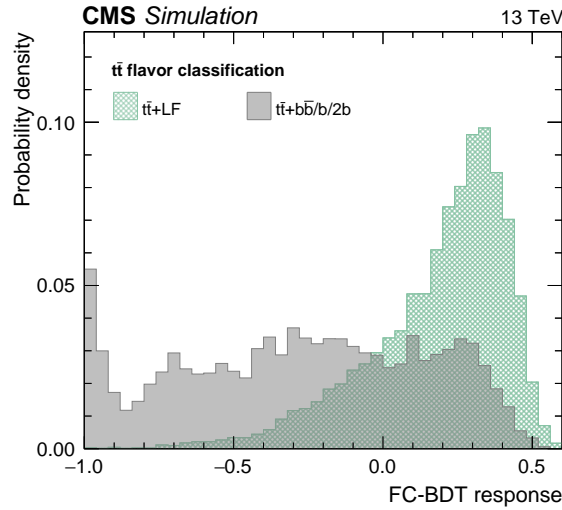


Figure 7: Output values of the SC-BDT.

Figure 8: Response values of the FC-BDT. The background consists of $t\bar{t}+b\bar{b}$, $t\bar{t}+1\bar{b}$, and $t\bar{t}+2\bar{b}$ events.

To determine the signal yield, the output distributions of the SC-BDT in the three and four b tag regions are fitted simultaneously with the output of the FC-BDT in the dilepton region. The SC-BDT output distributions before the fit are shown in Fig. 9 and the result of the fit is shown in Fig. 10. The pre- and post-fit distributions of the FC-BDT are shown in Fig. 11.

5.4 Systematic uncertainties

Many systematic uncertainties affect the result of the analysis, arising both from experimental and theoretical sources. All uncertainties are parametrized as nuisance parameters in the statistical inference performed in the final analysis step described in Section 7.

The uncertainty in the signal normalization due to the choice of factorization and renormalization scales is evaluated by changing their values to double and half of the nominal values. A rate uncertainty of around 5% is assigned to each process to account for the choice of PDFs, since shape variations are found to be negligible. Furthermore, for each $t\bar{t}+\text{HF}$ category, an individual 50% rate uncertainty is assigned, since the modeling of these components is theo-

Table 8: Input variables used in the training of the FC-BDT. The variables are sorted by their importance in the training within each category. In total, eight variables are used for the training of the FC-BDT.

Variable	Description
CSV(bjet 3)	Output of the b tagging discriminant for the b-tagged jet with the third-highest b tagging value in the event
$n_{\text{jets}}(\text{tight})$	Number of jets in the event passing the tight working point of the b tagging algorithm
CvsL(jet p_T 3)	Output of the charm vs. light-flavor tagging algorithm for the jet with the third-highest transverse momentum in the event
CSV(b-tagged jet 2)	Output of the b tagging discriminant for the b-tagged jet with the second-highest b tagging value in the event
CvsL(jet p_T 4)	Output of the charm vs. light-flavor tagging algorithm for the jet with the fourth-highest transverse momentum in the event
CvsB(jet p_T 3)	Output of the charm vs. bottom flavor tagging algorithm for the jet with the third-highest transverse momentum in the event
CSV(b-tagged jet 4)	Output of the b tagging discriminant for the b-tagged jet with the fourth-highest b tagging value in the event
$n_{\text{jets}}(\text{loose})$	Number of jets in the event passing the loose working point of the b tagging algorithm

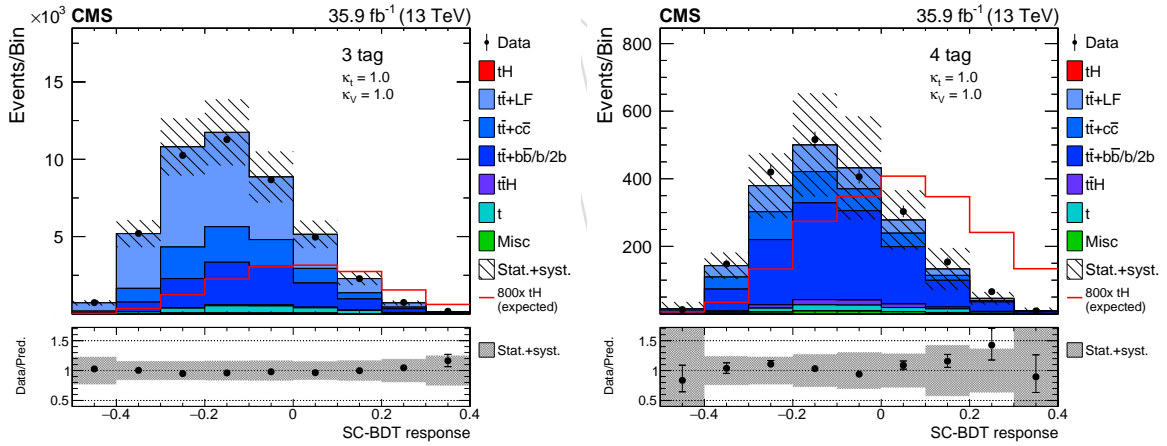


Figure 9: Pre-fit classifier outputs of the signal classification BDT for the 3 tag channel (left) and the 4 tag channel (right), for 35.9 fb^{-1} . In the box below each distribution, the ratio of the observed and predicted event yields is shown. The shape of the $t\bar{t}H$ signal is indicated with 800 times the SM.

retically difficult and cross section measurements are affected by large systematic uncertainties [57, 58].

The observed top quark p_T spectrum in $t\bar{t}$ events is found to be softer than the theoretical prediction [59]. A systematic uncertainty for this effect is derived by applying event-by-event weights that correct the disagreement.

Efficiency corrections for the selection of isolated leptons by the trigger and quality require-

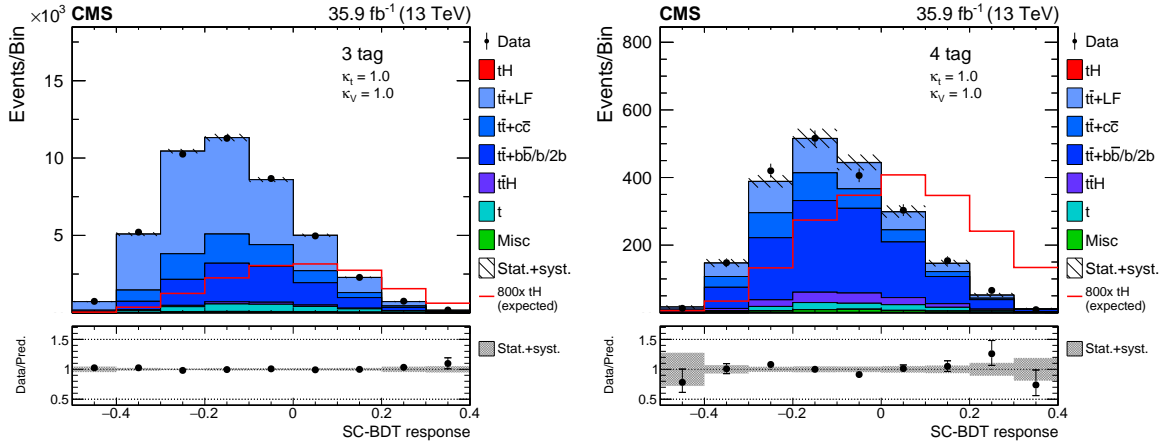


Figure 10: Post-fit classifier outputs of the signal classification BDT as used in the maximum likelihood fit for the 3 tag channel (left) and the 4 tag channel (right). In the box below each distribution, the ratio of the observed and predicted event yields is shown. The shape of the tH signal is indicated with 800 times the SM.

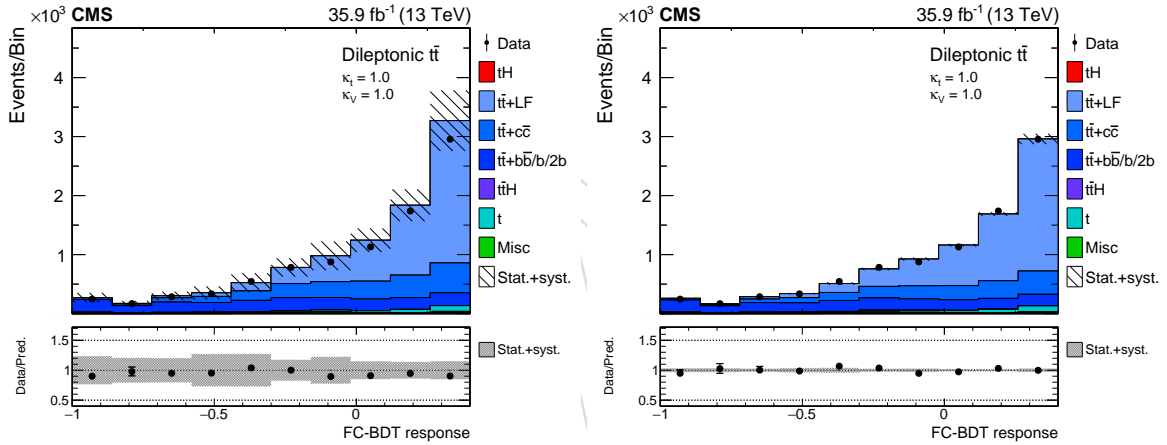


Figure 11: Pre- (left) and post-fit (right) classifier outputs of the flavor classification BDT for the dilepton selection. In the box below each distribution, the ratio of the observed and predicted event yields is shown.

ments are evaluated with a tag-and-probe method. Uncertainties in correcting the distribution of PV interactions are accounted by varying the total inelastic cross section by 4.6% [60]. The corrections applied to the jet energy scale and resolution are varied within their given uncertainties and the migration between different categories is used to determine the effect. In addition, the contribution to p_T^{miss} of unclustered particles is varied within the resolution of each particle [61]. The b tagging efficiencies for jets are measured in QCD multijet and $t\bar{t}$ enriched samples and varied within their uncertainties [34].

As for the multilepton channel, an uncertainty of 2.5% is assigned to the integrated luminosity [54] and affects the normalization of all processes.

The dominant systematic uncertainties are those related to the factorization and renormalization scales, as well as the uncertainties in the overall normalization of the $t\bar{t}$ +HF processes and the jet energy corrections.

6 Reinterpretation of the $H \rightarrow \gamma\gamma$ measurement

The standard model tHq and tHW signal processes with $H \rightarrow \gamma\gamma$ were included in previous measurements of the Higgs boson properties in the inclusive diphoton final state [22]. Events with two prompt high- p_T photons were divided into different event categories, each enriched with a particular production mechanism of the Higgs boson. The tHq and tHW processes contribute mostly to the “t \bar{t} H hadronic”, and “t \bar{t} H leptonic” categories as defined in Ref. [22], which target the t \bar{t} H process for fully hadronic top quark decays and for single-lepton or dilepton decay modes, respectively. Events in the t \bar{t} H leptonic category are selected to have at least one lepton well separated from the photons, and well reconstructed, as well as at least two jets of which at least one passes the medium b tagging requirement. The t \bar{t} H hadronic category is defined as events with no identified leptons and at least three jets, of which at least one is b tagged with the loose working point.

The signal is modeled with a sum of Gaussian functions describing the diphoton invariant mass ($m_{\gamma\gamma}$) shape derived from simulation. The background contribution is determined from the data without reliance on simulated events, using the discrete profiling method [4, 62, 63]. Different classes of models describing the falling $m_{\gamma\gamma}$ distribution in the background processes are used as input to the method. Sources of systematic uncertainties affecting the signal model and leading to migrations of signal events among the categories are considered.

The inputs to Ref. [22] from the t \bar{t} H categories are used here in a combination with the multi-lepton and single-lepton + $b\bar{b}$ channels to put constraints on the coupling modifier κ_t and on the production cross section of tH events. The coupling modifiers κ_t and κ_V affect both the tH and t \bar{t} H production cross sections, as well as the Higgs boson decay branching fraction into two photons through the interference of bosonic and fermionic loops. Changes in the kinematic properties of the tH signal arising from the modified couplings are taken into account by considering their effect on the signal acceptance and selection efficiency. Figure 12 shows the modified tHq and tHW selection efficiencies including acceptances for the two relevant categories of the $H \rightarrow \gamma\gamma$ measurement as a function of the ratio of coupling modifiers κ_t/κ_V . The signal diphoton mass shape is found to be independent of κ_t/κ_V .

The dependence of the signal acceptance and efficiency on κ_t/κ_V is implemented in the same statistical framework as that of Ref. [22], modifying the signal only in the t \bar{t} H categories.

7 Results and interpretation

The different discriminator output distributions in the multilepton and single-lepton + $b\bar{b}$ channels and the $\gamma\gamma$ invariant mass distributions in the diphoton channel are compared to the data in a combined maximum likelihood fit for various assumptions on the signal kinematics and normalizations, and are used to derive constraints on the signal yields.

The event selections in the different channels are mutually exclusive, therefore allowing a straightforward combination. Common systematic uncertainties such as the integrated luminosity normalization, the b tagging uncertainties, and the theoretical uncertainties related to the signal modeling are taken to be fully correlated among different channels.

A profile likelihood scan is performed as a function of the coupling modifier κ_t , which affects the production cross sections of the three signal components tHq, tHW, and t \bar{t} H, as well as the Higgs boson branching fractions. Effects on Higgs boson decays via fermion and boson loops to $\gamma\gamma$, $Z\gamma$, and gluon-gluon final states also affect the branching fractions in other channels. Furthermore, the kinematic properties of the two tH processes and thereby the shape

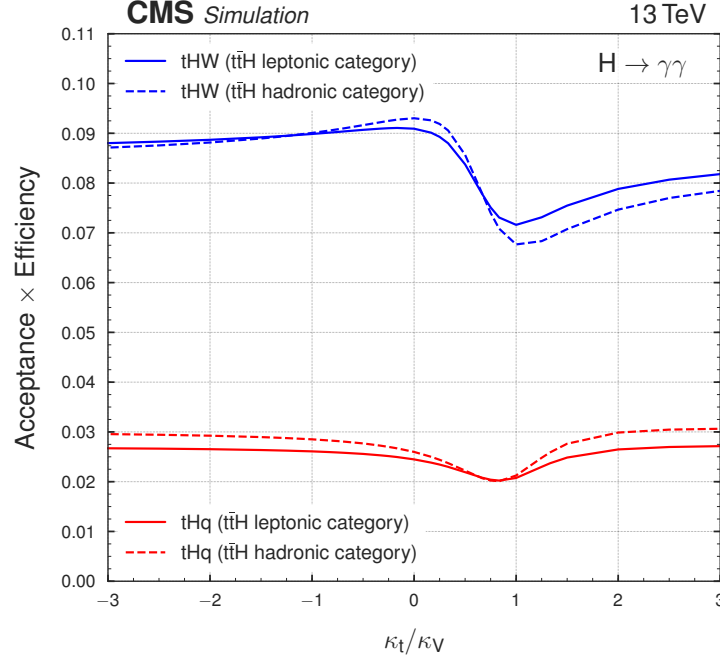


Figure 12: Acceptance and selection efficiency for the tHq (red) and tHW (blue) signal processes as a function of κ_t/κ_V , for the $t\bar{t}H$ leptonic (solid lines) and $t\bar{t}H$ hadronic categories (dashed lines) of the $H \rightarrow \gamma\gamma$ measurement.

of the classifier outputs entering the fit depend on the value of κ_t . Systematic uncertainties are included in the form of nuisance parameters in the fit and are treated via the frequentist paradigm, as described in Refs. [64, 65]. Uncertainties affecting the normalization are constrained either by Γ -function distributions, if they are statistical in origin, or by log-normal probability density functions. Systematic uncertainties that affect both the normalization and shape in the discriminating observables are included in the fit using the technique detailed in Ref. [66], and represented by Gaussian probability density functions.

Table 9 shows the impact of the most important groups of nuisances parameters on the $tH + t\bar{t}H$ signal yield. Pre-fit systematic uncertainties of the same groups are shown for comparison.

Table 9: Summary of the main sources of systematic uncertainty. $\Delta\mu/\mu$ corresponds to the relative change in $tH + t\bar{t}H$ signal yield induced by varying the systematic source within its associated uncertainty.

Source	Uncertainty [%]	$\Delta\mu/\mu$ [%]
e, μ selection efficiency	2–4	17
b tagging efficiency	2–15	6
Jet energy calibration	2–15	3
Forward jet modeling	10–35	3
Integrated luminosity	2.5	10
Reducible background estimate	10–40	14
Theoretical sources	≈ 10	14
$t\bar{t} + HF$ normalization	≈ 50	7
PDFs	2–6	8

To derive constraints on κ_t for a fixed value of $\kappa_V = 1.0$, a scan of the likelihood ratio $\mathcal{L}(\kappa_t)/\mathcal{L}(\hat{\kappa}_t)$

is performed, where $\hat{\kappa}_t$ is the best fit value of κ_t . Figure 13 shows the negative of twice the logarithm of this likelihood ratio ($-2\Delta \ln(\mathcal{L})$), for scans on the data, and for an Asimov data set [67] with SM expectations for $t\bar{t}H$ and tH . On this scale, a 95% confidence interval covers values below 3.84, while standard deviations are at values of 1, 4, 9, 16, etc. The expected performance for an SM-like signal is to favor a value of $\kappa_t = 1.0$ over one of $\kappa_t = -1.0$ by more than four standard deviations, and to exclude values outside of about -0.5 and 1.6 at 95% confidence level (CL). In the combined scan, the data slightly favors a positive value of κ_t over a negative one, by about 1.5 standard deviations, while excluding values outside the ranges of about $[-0.9, -0.5]$ and $[1.0, 2.1]$ at 95% CL. The sensitivity is driven by the $\gamma\gamma$ channel at negative values of the coupling modifiers and by the multilepton channels at positive values.

An excess of observed over expected events is seen both in the multilepton and $\gamma\gamma$ channels, with a combined significance of about two standard deviations. Consequently, floating a signal strength modifier (defined as the ratio of the fitted signal cross section to the SM expectation) of a combined $tH + t\bar{t}H$ signal yields a best fit value of 2.00 ± 0.53 under the SM hypothesis. These results are in agreement with those from the dedicated $t\bar{t}H$ searches [6], as expected, since they share a large fraction of events with the data set used here.

To establish limits on tH production, a different signal strength parameter is introduced for the combination of tHq and tHW , not including $t\bar{t}H$. A maximum likelihood fit for this signal strength is then performed based on the profile likelihood test statistic [64, 65] at fixed points of κ_t . Upper limits on the signal strength are then derived using the CL_s method [68, 69] and using asymptotic formulae for the distribution of the test statistic [67]. They are multiplied by the κ_t -dependent tH production cross section times the combined Higgs boson branching fractions to $WW^* + \tau\tau + ZZ^* + b\bar{b} + \gamma\gamma$ and are shown in Fig. 14. Limits for the SM and for a scenario with $\kappa_t = -1.0$ for the individual channels are shown in Table 10. The $t\bar{t}H$ contribution is kept fixed to its κ_t -dependent expectation. The fiducial cross section for SM-like tH production is limited to about 1.9 pb, with an expected limit of 0.9 pb, corresponding, respectively, to about 25 and 12 times the expected cross section times branching fraction in the combination of the channels explored. The visible discrepancy between observed and expected limits around $\kappa_t = 0.0$ is caused by the fact that the predicted $t\bar{t}H$ cross section vanishes while the data favors even larger than expected yields for $t\bar{t}H$ in both the $\gamma\gamma$ and multilepton channels.

8 Summary

Events from proton-proton collisions at $\sqrt{s} = 13$ TeV compatible with the production of a Higgs boson (H) in association with a single top quark (t) have been studied to derive constraints on the magnitude and relative sign of Higgs boson couplings to top quarks and vector bosons. Dedicated analyses of multilepton final states and final states with single leptons and a pair of bottom quarks are combined with a reinterpretation of a measurement of Higgs bosons decaying to two photons for the final result. For standard model-like Higgs boson couplings to vector bosons, the data favor a positive value of the top quark Yukawa coupling, y_t , by 1.5 standard deviations and exclude values outside the ranges of about $[-0.9, -0.5]$ and $[1.0, 2.1]$ times y_t^{SM} at the 95% confidence level. An excess of events compared with expected backgrounds is observed, but it is still compatible with the standard model expectation for $tH + t\bar{t}H$ production.

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS

Table 10: Expected and observed 95% CL upper limits on the tH production cross section times $H \rightarrow WW^* + \tau\tau + ZZ^* + b\bar{b} + \gamma\gamma$ branching fraction for a scenario of inverted couplings ($\kappa_t = -1.0$, top rows) and for an SM-like signal ($\kappa_t = 1.0$, bottom rows), in pb. The expected limit is calculated on a background-only data set, i.e., without tH contribution, but including a κ_t -dependent contribution from the $t\bar{t}H$ production. The $t\bar{t}H$ normalization is kept fixed in the fit, while the tH signal strength is allowed to float. Limits can be compared to the expected product of tH cross sections and branching fractions of 0.83 and 0.077 pb for the inverted top quark Yukawa coupling and for the SM, respectively.

Scenario	Channel	Observed	Expected
$\kappa_t = -1$	$b\bar{b}$	4.98 pb	$2.52^{+1.29}_{-0.81}$ pb
	$\gamma\gamma$	0.84 pb	$0.88^{+0.46}_{-0.28}$ pb
	$\mu^\pm\mu^\pm + e^\pm\mu^\pm + \ell\ell\ell$	0.85 pb	$0.77^{+0.36}_{-0.24}$ pb
	Combined	0.74 pb	$0.53^{+0.24}_{-0.16}$ pb
$\kappa_t = 1$ (SM-like)	$b\bar{b}$	6.88 pb	$3.19^{+1.64}_{-1.02}$ pb
	$\gamma\gamma$	3.68 pb	$2.03^{+1.05}_{-0.67}$ pb
	$\mu^\pm\mu^\pm + e^\pm\mu^\pm + \ell\ell\ell$	1.36 pb	$1.18^{+0.53}_{-0.35}$ pb
	Combined	1.94 pb	$0.92^{+0.40}_{-0.27}$ pb

institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: BMBWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); NKFI (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); MES (Latvia); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MOS (Montenegro); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR, and NRC KI (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI, and FEDER (Spain); MOSTR (Sri Lanka); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the F.R.S.-FNRS and FWO (Belgium) under the "Excellence of Science - EOS" - be.h project n. 30820817; the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Lendület ("Momentum") Program and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the New National Excellence Program ÚNKP, the NKFI research grants 123842, 123959, 124845, 124850 and 125105 (Hungary); the Council of Science and Industrial Research, India; the HOMING PLUS program of the Foundation for Polish Science,

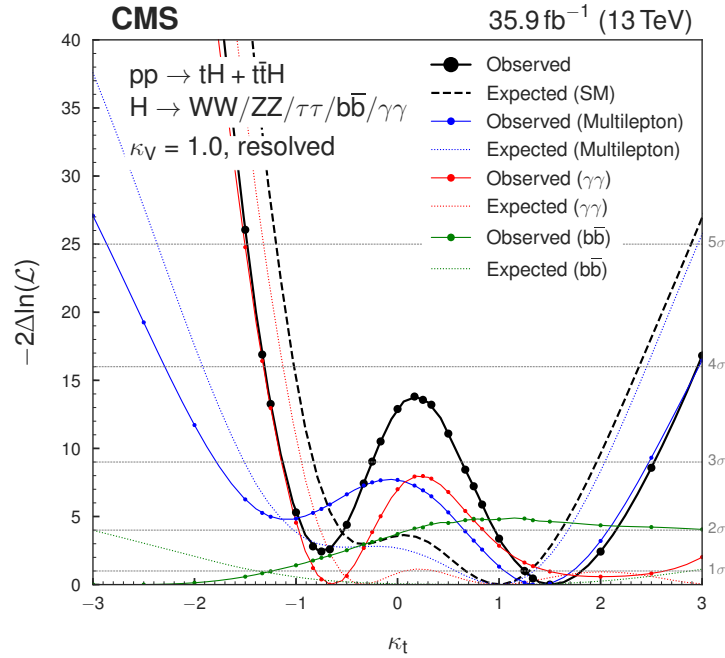


Figure 13: Scan of $-2\Delta\ln(\mathcal{L})$ versus κ_t for the data (black line) and the individual channels (blue, red, and green), compared to Asimov data sets corresponding to the SM expectations (dashed lines).

cofinanced from European Union, Regional Development Fund, the Mobility Plus program of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia María de Maeztu, grant MDM-2015-0509 and the Programa Severo Ochoa del Principado de Asturias; the Thalís and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

References

- [1] 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. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [2] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [3] CMS Collaboration, “Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV”, *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

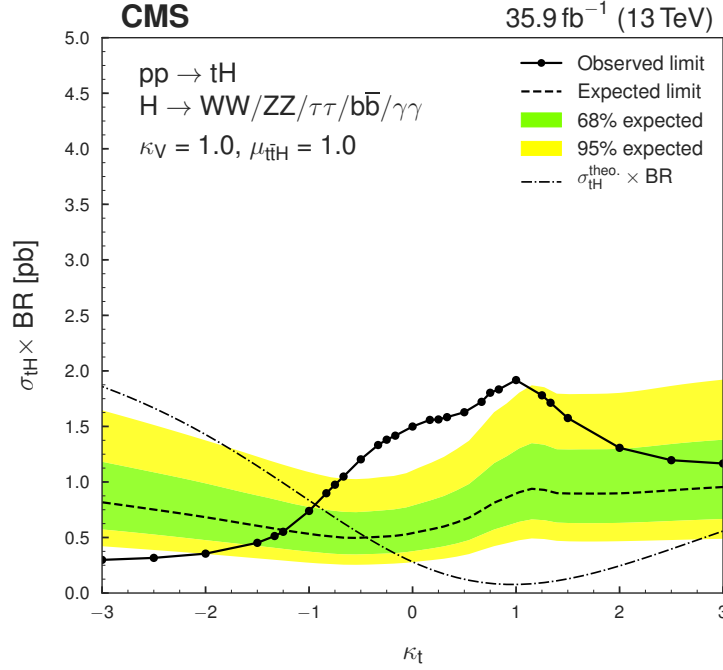


Figure 14: Observed and expected 95% CL upper limit on the $t\bar{H}$ cross section times combined $H \rightarrow WW^* + \tau\tau + ZZ^* + b\bar{b} + \gamma\gamma$ branching fraction for different values of the coupling ratio κ_t . The expected limit is calculated on a background-only data set, i.e., without $t\bar{H}$ contribution, but including a κ_t -dependent contribution from $t\bar{t}H$. The $t\bar{t}H$ normalization is kept fixed in the fit, while the $t\bar{H}$ signal strength is allowed to float.

- [4] ATLAS and CMS Collaborations, “Combined measurement of the Higgs boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments”, *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.
- [5] CMS Collaboration, “Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV”, *Phys. Rev. D* **92** (2015) 012004, doi:10.1103/PhysRevD.92.012004, arXiv:1411.3441.
- [6] CMS Collaboration, “Observation of $t\bar{t}H$ production”, *Phys. Rev. Lett.* **120** (2018) 231801, doi:10.1103/PhysRevLett.120.231801, arXiv:1804.02610.
- [7] 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) 173, doi:10.1016/j.physletb.2018.07.035, arXiv:1806.00425.
- [8] CMS Collaboration, “Observation of Higgs boson decay to bottom quarks”, *Phys. Rev. Lett.* **121** (2018) 121801, doi:10.1103/PhysRevLett.121.121801, arXiv:1808.08242.
- [9] S. Biswas, E. Gabrielli, and B. Mele, “Single top and Higgs associated production as a probe of the $Ht\bar{t}$ coupling sign at the LHC”, *JHEP* **01** (2013) 088, doi:10.1007/JHEP01(2013)088, arXiv:1211.0499.
- [10] 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.

- [11] 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. C* **75** (2015) 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.
- [12] ATLAS and CMS Collaborations, “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”, *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.
- [13] J. Ellis and T. You, “Updated global analysis of Higgs couplings”, *JHEP* **06** (2013) 103, doi:10.1007/JHEP06(2013)103, arXiv:1303.3879.
- [14] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, “Associated production of Higgs and single top at hadron colliders”, *Phys. Rev. D* **64** (2001) 094023, doi:10.1103/PhysRevD.64.094023, arXiv:hep-ph/0106293.
- [15] M. Farina et al., “Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson”, *JHEP* **05** (2013) 022, doi:10.1007/JHEP05(2013)022, arXiv:1211.3736.
- [16] P. Agrawal, S. Mitra, and A. Shivaji, “Effect of anomalous couplings on the associated production of a single top quark and a Higgs boson at the LHC”, *JHEP* **12** (2013) 077, doi:10.1007/JHEP12(2013)077, arXiv:1211.4362.
- [17] 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) 267, doi:10.1140/epjc/s10052-015-3475-9, arXiv:1504.00611.
- [18] LHC Higgs cross section working group, “Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector”, (2016). arXiv:1610.07922.
- [19] F. Demartin et al., “tWH associated production at the LHC”, *Eur. Phys. J. C* **77** (2017) 34, doi:10.1140/epjc/s10052-017-4601-7, arXiv:1607.05862.
- [20] A. Celis, V. Ilisie, and A. Pich, “LHC constraints on two-Higgs doublet models”, *JHEP* **07** (2013) 053, doi:10.1007/JHEP07(2013)053, arXiv:1302.4022.
- [21] 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.
- [22] CMS Collaboration, “Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV”, (2018). arXiv:1804.02716. Submitted to JHEP.
- [23] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [24] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.
- [25] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

- [26] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [27] E. Chabanat and N. Estre, “Deterministic annealing for vertex finding at CMS”, in *Computing in high energy physics and nuclear physics. Proceedings, Conference, CHEP’04, Interlaken, Switzerland, September 27-October 1, 2004*, p. 287. 2005.
- [28] Frühwirth, R. and Waltenberger, W. and Vanlaer, P., “Adaptive vertex fitting”, *J. Phys. G* **34** (2007) N343, doi:10.1088/0954-3899/34/12/N01.
- [29] 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.
- [30] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV”, *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [31] CMS Collaboration, “Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV”, *JINST* **10** (2015) P08010, doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.
- [32] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_T jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [33] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
- [34] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- [35] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [36] 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.
- [37] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [38] NNPDF Collaboration, “Parton distributions for the LHC run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [39] M. Botje et al., “The PDF4LHC working group interim recommendations”, (2011). arXiv:1101.0538.
- [40] S. Alekhin et al., “The PDF4LHC working group interim report”, (2011). arXiv:1101.0536.
- [41] LHC Higgs cross section working group, “Handbook of LHC Higgs cross sections: 3. Higgs properties”, (2013). arXiv:1307.1347.

- [42] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- [43] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [44] 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.
- [45] 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.
- [46] E. Re, “Single-top W -channel production matched with parton showers using the POWHEG method”, *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- [47] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG: s - and t -channel contributions”, *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].
- [48] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, “ W^+W^- , WZ and ZZ production in the POWHEG BOX”, *JHEP* **11** (2011) 078, doi:10.1007/JHEP11(2011)078, arXiv:1107.5051.
- [49] T. Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [50] CMS Collaboration, “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, doi:10.1007/JHEP08(2018)066, arXiv:1803.05485.
- [51] H. Voss, A. Höcker, J. Stelzer, and F. Tegenfeldt, “TMVA, the toolkit for multivariate data analysis with ROOT”, in *XIth International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT)*, p. 40. 2007. arXiv:physics/0703039. [PoS(ACAT)040]. doi:10.22323/1.050.0040.
- [52] CMS Collaboration, “Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.
- [53] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- [54] CMS Collaboration, “CMS luminosity measurement for the 2016 data-taking period”, CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- [55] V. D. Barger, J. Ohnemus, and R. J. N. Phillips, “Event shape criteria for single lepton top signals”, *Phys. Rev. D* **48** (1993) 3953, doi:10.1103/PhysRevD.48.R3953, arXiv:hep-ph/9308216.

- [56] G. C. Fox and S. Wolfram, "Observables for the analysis of event shapes in e^+e^- annihilation and other processes", *Phys. Rev. Lett.* **41** (1978) 1581, doi:10.1103/PhysRevLett.41.1581.
- [57] CMS Collaboration, "Measurement of $t\bar{t}$ production with additional jet activity, including b quark jets, in the dilepton decay channel using pp collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J. C* **76** (2016) 379, doi:10.1140/epjc/s10052-016-4105-x, arXiv:1510.03072.
- [58] CMS Collaboration, "Measurements of $t\bar{t}$ cross sections in association with b jets and inclusive jets and their ratio using dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **776** (2018) 355, doi:10.1016/j.physletb.2017.11.043, arXiv:1705.10141.
- [59] CMS Collaboration, "Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J. C* **75** (2015) 542, doi:10.1140/epjc/s10052-015-3709-x, arXiv:1505.04480.
- [60] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV", *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161, arXiv:1802.02613.
- [61] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- [62] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies, "Handling uncertainties in background shapes: the discrete profiling method", *JINST* **10** (2015) P04015, doi:10.1088/1748-0221/10/04/P04015.
- [63] CMS Collaboration, "Observation of the diphoton decay of the Higgs boson and measurement of its properties", *Eur. Phys. J. C* **74** (2014) 3076, doi:10.1140/epjc/s10052-014-3076-z, arXiv:1407.0558.
- [64] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.
- [65] CMS Collaboration, "Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **710** (2012) 26, doi:10.1016/j.physletb.2012.02.064, arXiv:1202.1488.
- [66] J. S. Conway, "Incorporating nuisance parameters in likelihoods for multisource spectra", (2011). arXiv:1103.0354.
- [67] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].
- [68] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.
- [69] A. L. Read, "Presentation of search results: The CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.