CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2016/03/16

Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment

The CMS Collaboration

Abstract

First results of the search for the associated production of a Higgs boson with a top quark-antiquark pair ($t\bar{t}H$) in proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV are presented. The data correspond to an integrated luminosity of 2.7 fb^{-1} recorded with the CMS experiment in 2015. Candidate $t\bar{t}H$ events are selected with criteria enhancing the lepton+jets or dileptonic decay channels of the $t\bar{t}$ system and the decay of the Higgs boson into a bottom quark-antiquark pair ($H \to b\bar{b}$). In order to increase the sensitivity of the search, selected events are split into several categories with different expected signal and background rates. In each category signal and background events are separated using a multivariate approach that combines a physics-motivated method (matrix element method, MEM) with methods from machine learning (boosted decision trees, BDTs). The result is presented in terms of the $t\bar{t}H$ signal strength modifier μ, the ratio of the observed $t\bar{t}H$ production cross section relative to the value expected for a 125 GeV standard model Higgs boson. A combined fit of multivariate discriminant templates in all categories to data results in an observed (expected) upper limit of $\mu < 2.6$ (3.6) at the 95% confidence level.

1 Introduction

The observation of a Higgs boson with a mass of approximately 125 GeV [1, 2] at the Large Hadron Collider (LHC) marked the starting point of a broad experimental program to determine the properties of the newly discovered particle. To date, the results of all measurements performed at the LHC are consistent with the expectations for a standard model (SM) Higgs boson. Decays into $\gamma\gamma$, ZZ and WW final states have been observed and there is evidence for the direct decay of the particle to fermions from the $\tau\tau$ and $b\bar{b}$ decay channels [3, 4]. The measured rates of various production and decay channels agree with the SM expectations [5, 6] and the hypothesis of a spin-0 particle is favored over other hypotheses [7, 8].

In the SM the coupling of the Higgs boson to fermions is of Yukawa type, with a coupling strength proportional to the fermion mass. Probing the coupling of the Higgs boson to the heaviest known fermion, the top quark, is hence very important for testing the SM and for constraining models of physics beyond the SM (BSM). Indirect constraints on the top–Higgs coupling are available from processes including top-quark loops, for example Higgs boson production through gluon-gluon fusion [5, 6]. On the other hand, the associated production of a Higgs boson and a top quark-antiquark pair ($t\bar{t}H$ production) is a direct probe of the top–Higgs coupling, as illustrated by the Feynman diagrams in Fig. 1. If observed it would prove the coupling of the Higgs boson to fermions with weak isospin $+1/2$ ("up-type") in addition to couplings to τ and b, which carry a weak isospin of $-1/2$ ("down-type"). The Higgs boson decay into bottom quark-antiquark pairs ($b\bar{b}$), also shown in Fig. 1, is attractive as a final state because it features the largest branching fraction of 0.58 ± 0.02 for a 125 GeV Higgs boson [9]. In addition, both $t\bar{t}H$ production and the decay $H \to b\bar{b}$ only involve third-generation quarks, which facilitates the theoretical interpretation of the results.

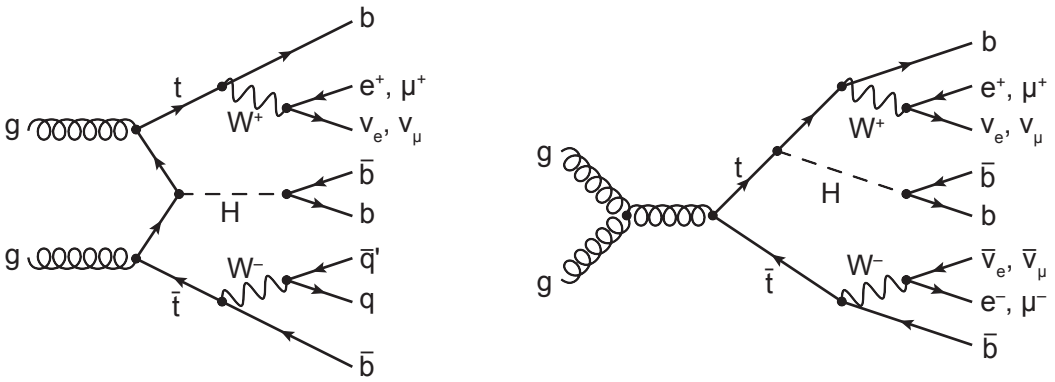

Figure 1: Exemplary leading-order Feynman diagrams for $t\bar{t}H$ production, including the subsequent decays of the top quark-antiquark pair in the lepton+jets channel (left) and the dilepton channel (right) as well as the decay of the Higgs boson into a bottom quark-antiquark pair.

Several BSM physics models predict a significantly enhanced $t\bar{t}H$ production rate while not modifying the branching fractions of Higgs boson decays by a measurable amount. For example, a number of BSM physics models predict vector-like partners of the top-quark (T) that decay into tH, bW and tZ final states [10–19]. The production and decay of $T\bar{T}$ pairs would lead to final states indistinguishable from those of $t\bar{t}H$ production. In this context, measurement of the $t\bar{t}H$ production cross section has the potential to distinguish the SM Higgs mechanism from alterative mechanisms to generate fermion masses.

Various dedicated searches for $t\bar{t}H$ production have been conducted during Run I of the LHC.

The CMS searches employ pp collision data corresponding to an integrated luminosity of $5\,\text{fb}^{-1}$ at a center-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ and $19.5\,\text{fb}^{-1}$ at $\sqrt{s} = 8\,\text{TeV}$. These searches have been performed studying Higgs boson decays to hadrons, photons, and leptons using multivariate analysis (MVA) techniques, showing a mild excess of the observed $\text{t}\bar{\text{t}}\text{H}$ signal strength relative to the SM expectation of $\mu = \sigma/\sigma_{\text{SM}} = 2.8 \pm 1.0$ [20]. A similar excess of $\mu = 2.1^{+1.4}_{-1.2}$ is observed in a search for $\text{t}\bar{\text{t}}\text{H}$ production in multilepton final states with $20.3\,\text{fb}^{-1}$ of ATLAS data at $\sqrt{s} = 8\,\text{TeV}$ [21].

The CMS search results have also entered a comprehensive test of the compatibility of the Higgs boson couplings with SM preditions [6]. The sensitivity for the $\text{t}\bar{\text{t}}\text{H}$ process in the $\text{H} \to \text{b}\bar{\text{b}}$ decay channel was further increased by employing the matrix element method (MEM) [22], resulting in an observed (expected) upper limit of $\mu = 4.3\,(3.3)$ at 95% confidence level. The current best observed (expected) upper limit on $\text{t}\bar{\text{t}}\text{H}$ production in the $\text{H} \to \text{b}\bar{\text{b}}$ decay channel amounts to $\mu = 3.4\,(2.2)$. It was obtained by the ATLAS collaboration using $20.3\,\text{fb}^{-1}$ of pp collision data at $\sqrt{s} = 8\,\text{TeV}$ [23].

The observation of $\text{t}\bar{\text{t}}\text{H}$ production is one of the major goals in Higgs boson physics for Run II. The increased center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ results in a $\text{t}\bar{\text{t}}\text{H}$ production cross section 3.9 times larger than at $\sqrt{s} = 8\,\text{TeV}$, while the cross section for the most important background, $\text{t}\bar{\text{t}}$ production, is only increased by a factor of 3.3 [24], resulting in a more favorable signal-to-background ratio. In addition, a larger fraction of events contains top quarks or Higgs bosons with transverse momenta above 200 GeV, making "boosted" jet reconstruction techniques increasingly attractive for $\text{t}\bar{\text{t}}\text{H}$ studies [25].

This document summarizes a search for $\text{t}\bar{\text{t}}\text{H}$ production performed with $2.7\,\text{fb}^{-1}$ of data recorded with the CMS detector during the first data-taking period of LHC Run II in 2015. Analysis methods established in Run I have been significantly improved, and novel methods have been added. In particular, the strengths of two multivariate techniques, one physics-motivated (matrix element method, MEM), and one from machine learning (boosted decision trees, BDT), have been combined to obtain a more powerful discriminant.

The event selection is adapted to $\text{t}\bar{\text{t}}\text{H}$ events with the decay of the Higgs boson into a $\text{b}\bar{\text{b}}$ pair and lepton+jets as well as dilepton decays of the $\text{t}\bar{\text{t}}$ pair, resulting in the final state $\ell\nu\,\text{q}\bar{\text{q}}'\,\text{b}\bar{\text{b}}$ ($\ell^+\nu\,\ell^-\bar{\nu}\,\text{b}\bar{\text{b}}$) for lepton+jets (dilepton) $\text{t}\bar{\text{t}}$ decays, where $\ell = \text{e}, \mu$. Events are split into mutually exclusive categories according to the number of reconstructed jets and the number of jets identified as coming from the hadronization of b quarks (b tagging). A further category is added with events in which "boosted" jets are identified as coming from hadronic top quark or $\text{H} \to \text{b}\bar{\text{b}}$ decays. In each category signal and background processes are separated employing BDTs which use the kinematic properties of jets and charged leptons, the b tagging probability, invariant masses and angular correlations of combinations of jets and leptons, as well as observables characterizing the event shape as inputs. The MEM discriminants are either used as input to the BDTs or the BDTs are employed to define categories in which the MEM output is used as the final discriminant. From a combined profile-likelihood fit of discriminant templates to data in all categories, an upper limit on the signal strength modifier μ is obtained.

This document is structured as follows: in Section 3, the data samples and simulated data samples are described. The basic selection of analysis objects and events is discussed in Section 4. The general analysis strategy and background estimation methods are introduced in Section 5. The influence of systematic uncertainties is studied in Section 6. Results of the studies are presented and their statistical interpretation is given in Section 7, followed by conclusions in Section 8.

2 The CMS Detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the superconducting solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [26].

3 Data and Simulation Samples

This analysis is performed using samples of proton-proton (pp) collisions at $\sqrt{s} = 13\,\text{TeV}$, which are collected with the CMS detector in 2015 and correspond to a total integrated luminosity of $2.7\,\text{fb}^{-1}$.

Monte Carlo (MC) event generators, interfaced with a detailed detector simulation, are used to model experimental effects, such as reconstruction and selection efficiencies, as well as detector resolutions. The CMS detector response is simulated using GEANT4 (v. 9.4) [27].

For the simulation of the reference $\mathrm{t\bar{t}H}$ signal sample, the next-to-leading-order (NLO) event generator POWHEG (v. 2) [28, 29] is used. The value of the Higgs boson mass is assumed to be 125 GeV, while the top quark mass value is set to 172.5 GeV. The proton structure is described by the parton distribution functions (PDF) NNPDF3.0 [30]. The generated events are subsequently processed with PYTHIA (v. 8.2) [31] for parton showering and hadronization.

Standard model backgrounds are simulated using POWHEG, MG5_aMC@NLO (v. 2.2.2) [32], or PYTHIA, depending on the process. The main background contribution originates from $\mathrm{t\bar{t}}$ production, the production of W and Z/γ^* bosons with additional jets (referred to as W+jets and Z+jets or commonly as V+jets in the following), single top quark production (tW channel), and diboson (WW, WZ, and ZZ) processes, and $\mathrm{t\bar{t}}$ production in association with a W or Z boson (referred to as $\mathrm{t\bar{t}}$ +W and $\mathrm{t\bar{t}}$ +Z or commonly as $\mathrm{t\bar{t}}$ +V in the following). Both the $\mathrm{t\bar{t}}$ and the single top quark samples are simulated with POWHEG. The V+jets and $\mathrm{t\bar{t}}$ +V samples are simulated with the NLO generator MG5_aMC@NLO, where for the V+jets samples the matching of matrix-element jets to parton showers is performed using the FxFx [33] prescription. In contrast, PYTHIA is used to simulate diboson events. Parton showering and hadronization are also simulated with PYTHIA in all the background samples. The PYTHIA CUETP8M1 tune [34, 35] is used to characterize the underlying event in both the $\mathrm{t\bar{t}H}$ signal and the background samples.

For comparison with the measured distributions, the events in the simulated samples are normalized to an integrated luminosity of $2.7\,\text{fb}^{-1}$, according to their predicted cross sections. These are taken from theoretical calculations at next-to-next-to-leading order (NNLO, for V+jets production), approximate NNLO (single top quark tW channel [36]), and NLO (diboson production [37] and $\mathrm{t\bar{t}}$ +V production [38]). The $\mathrm{t\bar{t}H}$ cross section [24, 39–42] and Higgs boson branching fractions [43–46] used in the analysis also have NLO accuracy. The $\mathrm{t\bar{t}}$ simulated sample is normalized to the full NNLO calculation with resummation to next-to-next-to-leading-logarithic accuracy (NNLL) [47–53], assuming a top quark mass value of 172.5 GeV and using the NNPDF3.0 PDF set. This sample is further separated into the following processes based on the flavor of additional jets that do not orginate from the top quark decays in the event: $\mathrm{t\bar{t}}$+$\mathrm{b\bar{b}}$,

defined at generator level as the events in which two additional b jets are generated within the acceptance requirements (cf. Section 4) and originate from one or more overlapping B hadrons; $\mathrm{t\bar{t}}$+b, for which only one additional b jet originates from a single B hadron; $\mathrm{t\bar{t}}$+2b, which corresponds to events with two additional B hadrons that are close enough in direction to produce a single b jet; $\mathrm{t\bar{t}}$+$\mathrm{c\bar{c}}$, for which events have at least one c jet within acceptance and no additional b jets; $\mathrm{t\bar{t}}$ +light flavor (lf), which correspond to events that do not belong to any of the above processes. The separation is motivated by the fact that different sub-samples originate from different physics processes and have different systematic uncertainties. A similar separation strategy has been followed by the ATLAS collaboration [23].

Effects from additional pp interactions in the same bunch crossing (pileup) are modeled by adding simulated minimum-bias events (generated with PYTHIA) to all simulated processes. The pileup multiplicity distribution in simulation is reweighted to reflect the luminosity profile of the observed pp collisions. Correction factors described in Section 4 are applied where necessary to improve the description of the data by the simulation.

4 Object and Event Selection

The event selection is consistent with the production of a Higgs boson in association with a top quark-antiquark pair. In this analysis, only the cases in which the Higgs boson decays into a bottom quark-antiquark pair are considered. In the SM, the top quark is expected to decay into a W boson and a b quark nearly 100% of the time. Hence different $\mathrm{t\bar{t}}$ decay modes can be identified according to the subsequent decays of the W bosons. Two $\mathrm{t\bar{t}}$ decay modes are considered: the lepton+jets mode ($\mathrm{t\bar{t}} \to \ell\nu\,\mathrm{q\bar{q}'}\,\mathrm{b\bar{b}}$), where one W boson decays into a charged lepton and a neutrino, and the dilepton mode ($\mathrm{t\bar{t}} \to \ell^+\nu\,\ell^-\bar{\nu}\,\mathrm{b\bar{b}}$), where both W bosons decay into a charged lepton and a neutrino. These signatures imply the presence of isolated leptons (ℓ = e, μ), missing transverse momentum owing to the neutrinos from W boson decays, and highly energetic jets originating from the final-state quarks. The heavy-quark content of the jets is identified through b tagging techniques.

At trigger level, events in the lepton+jets channel are required to contain an electron (muon) with transverse momentum (p_T) threshold of $p_\mathrm{T} > 27\,\mathrm{GeV}$ ($p_\mathrm{T} > 20\,\mathrm{GeV}$). For electrons a pseudorapidity range of $|\eta| < 2.1$ is required. Events in the dilepton channel are required to contain two leptons fulfilling p_T thresholds between 8 GeV and 17 GeV and isolation criteria. Events are reconstructed using a particle-flow (PF) technique [54, 55], which combines signals from all sub-detectors to enhance the reconstruction performance by identifying individual particle candidates in pp collisions. Charged hadrons from pileup events are subtracted event-by-event. Subsequently, the remaining neutral-hadron pileup component is subtracted at the level of jet energy corrections [56].

The electron and muon candidates are required to be sufficiently isolated from nearby jet activity as follows. For each electron (muon) candidate, a cone of $\Delta R = 0.3$ ($\Delta R = 0.4$) is constructed around the direction of the track at the event vertex, where ΔR is defined as $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, and $\Delta\eta$ and $\Delta\phi$ are the distances in pseudorapidity and azimuthal angle. Excluding the contribution from the lepton candidate, the scalar sum of the p_T of all particle candidates inside ΔR consistent with arising from the chosen primary event vertex is calculated to define a relative isolation discriminant, I_rel, through the ratio of this sum to the p_T of the lepton candidate. Electron candidates are selected if they have values of $I_\mathrm{rel} < 0.15$, while muons are selected if they fulfill $I_\mathrm{rel} < 0.15$ in the lepton+jets channel and $I_\mathrm{rel} < 0.25$ in the dilepton channel. In addition, electrons from identified photon conversions are rejected. To further increase the pu-

rity of muons originating from the primary interaction and to suppress misidentified muons or muons from decay-in-flight processes, additional quality criteria, such as a minimal number of hits associated with the muon track, are required in both the silicon tracker and the muon system. The neutral component from pileup events is subtracted event-by-event based on the average transverse energy deposited by neutral particles in the event, which is removed from the transverse energy in the isolation cone.

For the lepton+jets channel, events are selected containing exactly one energetic, isolated lepton (e or μ), which is required to have $p_T > 25\,\text{GeV}$ or $p_T > 30\,\text{GeV}$ in the case of the μ or e, respectively, and $|\eta| < 2.1$ (but excluding electrons within a small region of $|\eta|$ between the barrel and endcap sections of the ECAL). For the dilepton channel, events are required to have a pair of oppositely charged energetic leptons (e^+e^-, $\mu^+\mu^-$, $\mu^\pm e^\mp$). The leading lepton is required to have $p_T > 20\,\text{GeV}$ and the subleading lepton $p_T > 15\,\text{GeV}$, and both leptons are required to fulfill $|\eta| < 2.4$. The invariant mass of the selected lepton pair is required to be larger than 20 GeV to suppress events from heavy-flavor resonance decays and low-mass Drell–Yan processes. In the same-flavor channels, events are rejected if the dilepton invariant mass is within the region $76\,\text{GeV} < m^{\ell\ell} < 106\,\text{GeV}$, thereby suppressing further contribution from Z+jets processes.

Jets are reconstructed from the PF particle candidates using the anti-k_T clustering algorithm [57] with a distance parameter of 0.4, optimized for the running conditions at the higher center-of-mass energy in LHC Run II. The jet energy is corrected for pileup in a manner similar to that used to find the energy within the lepton isolation cone. Jet energy corrections are also applied as a function of jet p_T and η [58] to data and simulation. Events in the lepton+jets channel are required to have at least four reconstructed jets with $p_T > 30\,\text{GeV}$ and $|\eta| < 2.4$. In the dilepton channels, at least three jets with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.4$ are required, from which the two leading jets must satisfy $p_T > 30\,\text{GeV}$.

Jets originating from the hadronization of b quarks are identified using a combined secondary vertex algorithm (CSV) [59], which provides a b tagging discriminant by combining identified secondary vertices and track-based lifetime information. A discriminant value is chosen such that the efficiency for tagging jets from b (c) quarks is ≈70% (20%), while the probability of tagging jets originating from light-flavor quarks (u, d, or s) or gluons is around 1%. The shape of the CSV discriminant distribution in simulation is corrected by scale factors to better describe the efficiency observed in the data [60]. This correction is derived separately for light-flavor and b jets from a "tag-and-probe" approach using control samples enriched in events with a Z boson and exactly two jets, and $t\bar{t}$ events in the eμ channel with no additional jets. At least two b-tagged jets are required.

In the lepton+jets channel, a dedicated reconstruction of the hadronically decaying top quark and the H boson is performed, targeting at the case of high p_T where their decay products are strongly collimated and potentially clustered within one jet (referred to as 'boosted' in the following). The reconstruction is based on the HEP Top Tagger [61–63] and subjet-filterjet algorithm [64] techniques, respectively. All PF particle candidates in the event except the one identified as lepton are clustered into large-radius jets with the Cambridge/Aachen algorithm [65] with a cone-size parameter of 1.5. The presence of the $t \to bqq$ and $H \to b\bar{b}$ decay products within the jet is inferred by successively reverting the clustering sequence and searching for a significant decrease of the invariant mass when splitting into two subjets, and the subjets eventually obtained are assigned to the t and $H \to b\bar{b}$ final state particles depending on their CSV discriminant value and their p_T. For the final selection, the boosted t and H candidates are required to have $p_T > 200\,\text{GeV}$ and to fulfil a multivariate selection criterion constructed

from kinematic and secondary-vertex information of the subjets, which assesses the agreement of the candidate with the boosted object hypothesis.

The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is defined as the projection on the plane perpendicular to the beams of the negative vector sum of the momenta of all reconstructed particles in an event. Its magnitude is referred to as $E_{\mathrm{T}}^{\mathrm{miss}}$. In the dilepton same-flavor channels, events are required to fulfill $E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$.

Signal $\mathrm{t\bar{t}H}$ events are generally characterized by having more jets and more b-tags than the background processes. Events are then divided into categories based on the number of jets, the number of b-tagged jets, and the presence of boosted objects to improve the sensitivity of the analyses. For the lepton+jets channel, events are separated into the following seven categories: if both a boosted t and H are selected, events are assigned to the 'boosted category', else events are assigned to either of the ≥ 6 jets, 2 b-tags; 4 jets, 3 b-tags; 5 jets, 3 b-tags; ≥ 6 jets, 3 b-tags; 4 jets, ≥ 4 b-tags; 5 jets, ≥ 4 b-tags and ≥ 6 jets, ≥ 4 b-tags categories. For the dilepton channel, events are divided into five categories: 3 jets, 2 b-tags; 3 jets, 3 b-tags; ≥ 4 jets, 2 b-tags; ≥ 4 jets, 3 b-tags and ≥ 4 jets, ≥ 4 b-tags.

Tables 1 and 2 show the predicted and observed event yields after the event selection in the lepton+jets and dilepton channels, respectively. The tables are sub-divided into the different jet and b-tag categories used in each channel. The expected and observed yields agree well in all final states across the different categories of jets and b-tags.

5 Analysis Strategy and Background Estimation

Boosted decision trees and a matrix element method are used to further improve the signal-to-background separation in all channels of the analysis. Both techniques are combined into one single discriminant, which exploits the strenghts of both methods. This is a new feature of this analysis compared to previous CMS results [20, 22].

The BDTs utilize information related to object kinematics, event shape, the CSV b-tag discriminant, variables specific to the boosted object reconstruction, and in some cases also the output of the MEM described below. A separate BDT is trained for each category, resulting in eight BDTs in the lepton+jets and five in the dilepton channel. The training is performed using simulated $\mathrm{t\bar{t}H}$ and $\mathrm{t\bar{t}}$+jets events as signal and background, respectively, which are weighted to achieve equal yields of signal and background events in each category. In order to avoid overoptimization, the signal and background events are split in half: one half is used to perform the training, and the other half is used in the final analysis to monitor the performance and derive the expected limits. The specific BDT method used is the stochastic Gradient Boost, available as part of the TMVA package in ROOT [66]. The choice of BDT input variables as well as the tree architecture is optimized separately in each category with a procedure based on the particle swarm algorithm [67]. A description of the input variables is provided in Appendices A and B for each category of the lepton+jets and the dilepton channel, respectively.

Within the MEM, each event is assigned a probability density value computed from the four-momenta of the reconstructed particles, which is based on the differential cross section of the signal or background process. The MEM discriminant is constructed as ratio of the probability density values of the signal and background hypothesis. The deployed algorithm is an improved version of the method described in [22]. The probability density functions are constructed at LO, assuming gluon-gluon fusion production both for signal and background processes. The $\mathrm{t\bar{t}} + \mathrm{b\bar{b}}$ matrix elements have been found to provide similar discrimination power

Table 1: $t\bar{t}H$ and background event yields for the lepton+jets categories. The processes and the separation of the $t\bar{t}$ +jets sample are described in Section 3.

Process	≥ 6 jets, 2 b-tags	4 jets, 3 b-tags	5 jets, 3 b-tags	≥ 6 jets, 3 b-tags
$t\bar{t}$+lf	5359.3 ± 1226.3	2026.1 ± 651.4	1000.2 ± 352.9	589.5 ± 199.7
$t\bar{t} + c\bar{c}$	1722.2 ± 849.5	363.2 ± 190.9	368.1 ± 191.3	396.6 ± 209.5
$t\bar{t}$+b	393.7 ± 188.2	203.1 ± 92.5	199.6 ± 90.8	170.8 ± 81.4
$t\bar{t}$+2b	165.2 ± 81.2	78.9 ± 38.0	87.2 ± 40.7	97.3 ± 46.8
$t\bar{t} + b\bar{b}$	226.4 ± 113.2	75.8 ± 35.3	114.1 ± 52.3	183.7 ± 86.7
Single Top	283.0 ± 49.0	115.3 ± 30.8	76.2 ± 19.5	47.5 ± 12.7
V+jets	130.5 ± 35.2	38.6 ± 17.8	22.8 ± 10.4	13.6 ± 6.4
$t\bar{t}$+V	43.5 ± 8.2	4.3 ± 1.2	6.4 ± 1.8	10.0 ± 2.7
Diboson	2.8 ± 1.3	2.1 ± 1.3	0.9 ± 0.5	0.2 ± 0.3
Total bkg	8326.7 ± 1788.6	2907.4 ± 836.5	1875.5 ± 534.7	1509.1 ± 423.7
$t\bar{t}H$	29.6 ± 2.1	7.4 ± 1.0	10.9 ± 1.2	16.7 ± 2.1
Data	7185	2793	1914	1386
S/B	0.0036	0.0026	0.0059	0.011
Data/B	0.9 ± 0.2	1.0 ± 0.3	1.0 ± 0.3	0.9 ± 0.3

Process	4 jets, ≥ 4 b-tags	5 jets, ≥ 4 b-tags	≥ 6 jets, ≥ 4 b-tags	boosted
$t\bar{t}$+lf	17.8 ± 10.8	17.7 ± 10.9	17.6 ± 11.3	45.1 ± 9.4
$t\bar{t} + c\bar{c}$	11.6 ± 8.2	22.1 ± 15.4	35.9 ± 24.9	21.8 ± 12.0
$t\bar{t}$+b	8.4 ± 4.4	14.8 ± 7.7	20.0 ± 10.9	10.3 ± 5.5
$t\bar{t}$+2b	3.5 ± 1.9	6.9 ± 3.7	12.3 ± 6.9	12.3 ± 6.6
$t\bar{t} + b\bar{b}$	10.1 ± 4.9	28.8 ± 13.9	73.4 ± 36.6	17.0 ± 8.4
Single Top	2.5 ± 1.1	4.3 ± 1.4	5.5 ± 2.0	7.0 ± 1.7
V+jets	1.0 ± 0.8	0.9 ± 0.8	1.4 ± 0.7	2.5 ± 0.8
$t\bar{t}$+V	0.3 ± 0.1	0.7 ± 0.3	1.6 ± 0.6	0.9 ± 0.3
Diboson	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0	0.1 ± 0.1
Total bkg	55.2 ± 23.0	96.5 ± 37.6	167.6 ± 65.7	117.0 ± 24.9
$t\bar{t}H$	0.9 ± 0.2	2.7 ± 0.6	5.9 ± 1.4	2.2 ± 0.3
Data	75	104	150	104
S/B	0.017	0.028	0.035	0.019
Data/B	1.4 ± 0.5	1.1 ± 0.4	0.9 ± 0.4	0.9 ± 0.2

Table 2: $t\bar{t}H$ and background event yields for dilepton categories. The processes and the separation of the $t\bar{t}$ +jets sample are described in Section 3.

	3 jets, 2 b-tags	3 jets, 3 b-tags	≥ 4 jets, 2 b-tags	≥ 4 jets, 3 b-tags	≥ 4 jets, ≥ 4 b-tags
$t\bar{t}$+lf	2558.6 ± 542.7	26.6 ± 10.5	2271.6 ± 505.0	60.3 ± 25.6	0.9 ± 0.8
$t\bar{t} + c\bar{c}$	220.9 ± 103.4	22.7 ± 13.6	478.4 ± 234.4	78.4 ± 45.4	3.4 ± 2.9
$t\bar{t}$+b	65.4 ± 28.5	21.4 ± 10.2	126.2 ± 57.7	52.2 ± 25.1	2.7 ± 1.6
$t\bar{t}$+2b	16.9 ± 7.6	6.6 ± 3.1	42.9 ± 20.2	22.3 ± 10.7	1.2 ± 0.7
$t\bar{t} + b\bar{b}$	8.6 ± 4.2	3.6 ± 1.8	48.9 ± 23.7	39.8 ± 18.8	13.4 ± 7.1
Single Top	93.2 ± 16.7	3.0 ± 1.0	87.6 ± 15.8	7.3 ± 2.5	0.4 ± 0.4
V+jets	14.5 ± 11.0	1.3 ± 0.8	16.0 ± 7.4	0.0 ± 0.0	0.0 ± 0.0
$t\bar{t}$+V	3.6 ± 0.9	0.3 ± 0.2	16.4 ± 3.2	3.2 ± 0.9	0.5 ± 0.2
Diboson	1.7 ± 0.9	0.0 ± 0.0	1.2 ± 1.0	0.1 ± 0.0	0.0 ± 0.0
Total bkg	2983.4 ± 590.4	85.6 ± 25.6	3089.2 ± 650.6	263.6 ± 79.9	22.5 ± 9.8
$t\bar{t}H$	1.4 ± 0.2	0.4 ± 0.1	8.1 ± 1.1	3.6 ± 0.6	1.0 ± 0.3
Data	3123	115	2943	319	27
S/B	0.00047	0.0051	0.0026	0.014	0.046
Data/B	1.0 ± 0.2	1.3 ± 0.4	1.0 ± 0.2	1.2 ± 0.3	1.2 ± 0.5

against all background subprocesses and are solely used to model the background. Hadronization and detector effects are taken into account via transfer functions derived from simulation, which map the measured four-momenta to the final-state particles in the matrix element. In each event, only the four jets that most likely originate from b quarks are considered explicitly when evaluating the probability densities; contributions from further jets are marginalized by integration. The four jets are selected using the likelihood ratio between the hypotheses that four or two jets in the event arose from b quarks and the rest from light quarks, based on the expected b-tagging probability densities from simulation.

The BDT and MEM discriminants perform slightly differently in terms of signal and background separation. While the BDT achieves in a slightly better separation against the inclusive $t\bar{t}$ background, the MEM is by construction especially powerful in separating against the challenging $t\bar{t} + b\bar{b}$ background. The correlation between the BDT and MEM discriminants have been studied in different control regions in data and found to be well-modeled by the simulation. In the lepton+jets channel, the two discrinants are combined with the scheme described below. This results in the best sensitivity, and it is robust against binning effects and overoptimization.

In the two and three b-tag categories, a BDT is used as final discriminant, which includes the MEM as input variable (in the three b-tags categories). These categories contain a relatively large number of events, which is favourable for training the BDT.

In the categories with four or more b-tags, events are further separated into two sub-categories, one with low and one with high BDT output, defined by the median of the BDT output distribution. In each sub-category, the MEM is used as final discriminant. The high BDT output subcategory is expected to be enhanced with signal events, and the MEM discriminant achieves additional separation against the residual $t\bar{t} + b\bar{b}$ background contributions. The choice of the median contributes to a robust result by ensuring a sufficient number of events in each subcategory.

In the boosted category, which contains the events in which both a boosted top quark and a boosted Higgs boson candidate are reconstructed and identified, a BDT is used as final discriminant. Important input variables are derived from the reconstructed $t\bar{t}$ and $H \to b\bar{b}$ system, for example the invariant mass of the two b jets assigned to the Higgs boson decay. The sensitivity in the boosted category benefits from reduced combinatorics when assigning reconstructed objects to the top quark and Higgs boson decay products.

The final discriminant outputs provide better discrimination between signal and background than any of the input variables individually. The output distributions of the background and signal processes are fit to the data simultaneously in all channels and categories to set limits on the Higgs boson production cross section, as described in Section 7.

6 Systematic Uncertainties

In Table 3, all sources of systematic uncertainties considered in the analysis are listed. They affect either the yields of the signal or background processes, or the discriminant shape, or both. In the last case, the yield and shape effects are treated as entirely correlated and are varied simultaneously. The uncertainties are taken into account via nuisance parameters in the final fit procedure described in Section 7.

The effect of the uncertainties is evaluated individually in each category of each analysis channel, where the effects from the same source are treated as fully correlated. The impact of the sys-

Table 3: Systematic uncertainties considered in the analysis.

Source	Type	Remarks
Luminosity	rate	Signal and all backgrounds
Lepton ID/trigger efficiency	shape	Signal and all backgrounds
Pileup	shape	Signal and all backgrounds
Jet energy scale	shape	Signal and all backgrounds
b-tag HF fraction	shape	Signal and all backgrounds
b-tag HF stats (linear)	shape	Signal and all backgrounds
b-tag HF stats (quadratic)	shape	Signal and all backgrounds
b-tag LF fraction	shape	Signal and all backgrounds
b-tag LF stats (linear)	shape	Signal and all backgrounds
b-tag LF stats (quadratic)	shape	Signal and all backgrounds
b-tag charm (linear)	shape	Signal and all backgrounds
b-tag charm (quadratic)	shape	Signal and all backgrounds
QCD scale ($t\bar{t}H$)	rate	Scale uncertainty of NLO $t\bar{t}H$ prediction
QCD scale ($t\bar{t}$)	rate	Scale uncertainty of NLO $t\bar{t}$ prediction
QCD scale ($t\bar{t}$+hf)	rate	Additional scale uncertainty of NLO $t\bar{t}$+hf predictions
QCD scale (t)	rate	Scale uncertainty of NLO single t prediction
QCD scale (V)	rate	Scale uncertainty of NNLO W and Z prediction
QCD scale (VV)	rate	Scale uncertainty of NLO diboson prediction
pdf (gg)	rate	Pdf uncertainty for gg initiated processes except $t\bar{t}H$
pdf (gg $t\bar{t}H$)	rate	Pdf uncertainty for $t\bar{t}H$
pdf ($q\bar{q}$)	rate	PDF uncertainty of $q\bar{q}$ initiated processes ($t\bar{t}$ W, W, Z)
pdf (qg)	rate	PDF uncertainty of qg initiated processes (single t)
Q^2 scale ($t\bar{t}$)	shape	Renormalization and factorization scale uncertainties of the $t\bar{t}$ ME generator, independent for additional jet flavors
PS Scale ($t\bar{t}$)	shape	Renormalization and factorization scale uncertainties of the parton shower (for $t\bar{t}$ events), independent for additional jet flavors

tematic variations differ between the categories. As an example, the change in background and signal event yield due to the different uncertainties is listed in Table 4 for the ≥ 6 jets, 3 b-tags category of the lepton+jets channel, which shows high sensitivity and at the same time contains a relatively large number of events such that the variations are statistically significant.

Table 4: Specific effect of systematic uncertainties that affect the discriminant shape on the predicted background and signal yields for events in the ≥ 6 jets, 3 b-tags category of the lepton+jets channel. Here, only the sum of the largest background processes, $t\bar{t}$+lf, $t\bar{t}$+b, $t\bar{t}$+2b, $t\bar{t}$+$b\bar{b}$, and $t\bar{t}$+$c\bar{c}$, are considered.

Process	$t\bar{t}$ rate up/down [%]	$t\bar{t}H$ rate up/down [%]
Jet energy Scale	$+11.3/-10.1$	$+7.7/-7.0$
Jet energy Resolution	$-0.1/+0.1$	$-0.1/+0.1$
Pile-Up	$-0.1/+0.0$	$+0.1/-0.2$
Electron Efficiency	$+1.6/-1.6$	$+1.6/-1.6$
Muon Efficiency	$+1.2/-1.2$	$+1.2/-1.2$
b-Tag HF contamination	$-3.5/+8.4$	$+0.2/+0.6$
b-Tag HF stats (linear)	$-6.4/+6.2$	$-5.3/+4.9$
b-Tag HF stats (quadratic)	$+4.2/-4.4$	$+3.3/-3.6$
b-Tag LF contamination	$+7.1/-5.1$	$+5.5/-4.2$
b-Tag LF stats (linear)	$-3.2/+6.5$	$-0.6/+1.1$
b-Tag LF stats (quadratic)	$+0.5/+1.2$	$-0.8/+1.1$
b-Tag charm Uncertainty (linear)	$-12.6/+16.9$	$-0.6/+0.7$
b-Tag charm Uncertainty (quadratic)	$+1.4/-1.4$	$+0.0/-0.0$
Q2 scale ($t\bar{t}$+lf)	$-1.9/+2.8$	–
Q2 scale ($t\bar{t}$+b)	$-0.6/+0.9$	–
Q2 scale ($t\bar{t}$+2b)	$-0.5/+0.8$	–
Q2 scale ($t\bar{t}$+$b\bar{b}$)	$-0.9/+1.3$	–
Q2 scale ($t\bar{t}$+$c\bar{c}$)	$-1.6/+2.4$	–
PS scale ($t\bar{t}$+lf)	$4.4/-8.7$	–
PS scale ($t\bar{t}$+b)	$-1.3/+0.8$	–
PS scale ($t\bar{t}$+2b)	$-1.0/+0.4$	–
PS scale ($t\bar{t}$+$b\bar{b}$)	$-2.0/+1.3$	–
PS scale ($t\bar{t}$+$c\bar{c}$)	$-4.3/+2.3$	–

The uncertainty in the luminosity estimate is 2.7% [68]. Electron and muon identification and trigger efficiency uncertainties were estimated by comparing variations in measured efficiency between data and MC simulation using a high-purity sample of Z-boson decays and are found to be 2–4%. Effects of the uncertainty in the distribution of the number of pileup interactions are evaluated by varying the cross section used to predict the number of pileup interactions in MC by $\pm 5\%$ from its nominal value. The uncertainty of the jet energy scale [58] (resolution) are evaluated by varying the energy scale (resolution) correction of all jets in the signal and background predictions by one standard deviation. The uncertainty of the CSV b-tagging scale factors is evaluated by applying alternative scale factors based on varying the following systematic effects by one standard deviation, separately for the different jet flavors: the contamination of background processes in the control samples, the jet energy scale uncertainty — which is correlated with the overall jet energy scale uncertainty — and the statistical uncertainty in the scale factor evaluation. The impact of the latter one is parameterized as the sum of two orthogonal contributions: a linear and a quadratic term, which allow an overall tilt and a shift of the center of CSV distribution, respectively. Both for the jet energy scale and for the b-tagging scale factor uncertainties, the event categorization and successive evaluation of the discriminant is re-evaluated after applying the systematic variations to account for migration effects.

Theoretical uncertainties of the cross sections used to predict the rates of various processes are propagated to the yield estimates. All rates are estimated using cross sections of at least NLO accuracy, which have uncertainties arising primarily from PDFs and the choice of factorization and renormalization scales (both in the matrix element and the parton shower). The cross section uncertainties are each separated into their PDF and scale components and correlated where appropriate between processes. For example, the PDF uncertainty for processes originating primarily from gluon-gluon initial states, such as $t\bar{t}$ and $t\bar{t}H$ production, are treated as 100% correlated. The $t\bar{t}$+$b\bar{b}$ process, and to lesser extent the $t\bar{t}$+2b, $t\bar{t}$+b, and $t\bar{t}$+$c\bar{c}$ production, represent important sources of irreducible background. Neither control region studies nor higher-order theoretical calculations can currently constrain the normalization of these contributions to better than 50% accuracy; therefore a conservative extra 50% rate uncertainty is assigned to the $t\bar{t}$ +heavy-flavor processes. This rate uncertainty has the largest impact on the median expected limit; omitting it in the computation improves the limit by approximately 10%. The effect of the scale uncertainties on the discriminant output shape is also taken into account for the $t\bar{t}$+jets production using event weights obtained directly from the MC generator in case of the matrix element and dedicated samples generated with different scale choices in case of the parton shower. The factorization and renormalization scales of the matrix element generator and also the scales of the initial-state radiation and the final-state radiation of the parton shower of the general-purpose MC event generator are varied simultaneously by a factor of 0.5 and 2. These scale variations are treated as uncorrelated between the matrix element generator and the general-purpose MC event generator. The rate and shape variation due to the PDF uncertainty has been evaluated by comparing the results to those obtained when using the sub-PDFs of the NNPDF set and taking into account that the sub-PDFs should have Gaussian-distributed deviations from the nominal PDF. Since the rate and shape uncertainties due to the used NNPDF PDF set have been found to be negligible for the final discriminants and output variables, these uncertainties have not been propagated into the final analysis.

The impact of statistical fluctuations in the signal and background prediction due to the limited number of simulated events is accounted for using the approach described in [69, 70].

7 Results and Statistical Interpretation

The signal strength modifier $\mu = \sigma/\sigma_{\mathrm{SM}}$ for the $t\bar{t}H$ production cross section is determined in a simultaneous binned maximum-likelihood fit to data in all analysis categories. The distributions of the final discriminant in all analysis categories before the fit to data are displayed in Figs. 2 to 4. The fit procedure takes into account sytematic uncertainties that modify the shape and normalization of the distribution, as described in Section 6. The distributions of the final discriminant after the fit are shown in Figs. 5 to 7. The value obtained for μ is both compatible with the SM expectation and no signal, therefore an upper limit at 95% confidence level (CL) is determined using a modified frequentist CL_s method [71, 72].

The expected and observed upper limits at the 95% CL on μ under the background-only hypothesis in each channel are listed in Table 5. A Higgs-boson mass of $m_H = 125\,\mathrm{GeV}$ has been assumed in the signal model. The limits in each individual category are listed in Appendices A and B for the lepton+jets and the dilepton channel, respectively.

When combining all categories and channels, we obtain an observed (expected) upper limit of $\mu < 2.6$ (3.6) at the 95% confidence level. The best-fit value of μ is $-2.0^{+1.8}_{-1.8}$, which is 1.7 standard deviations from the standard model expectation of $\mu = 1$. The upper limits and best-fit values are illustrated in Fig. 8.

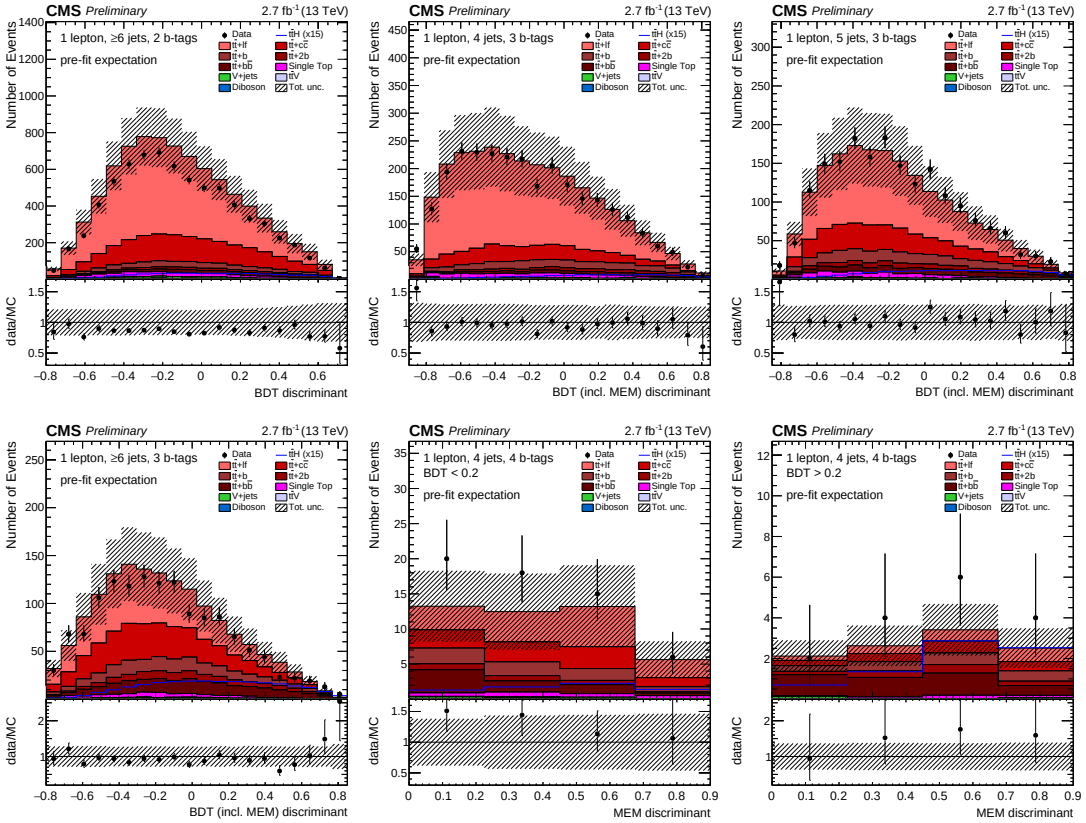

Figure 2: Final discriminant shapes in the different analysis categories in the lepton+jets channel before the fit to data. The expected background contributions (filled histograms) are stacked, and the expected signal distribution (line) for a Higgs-boson mass of $m_H = 125\,\text{GeV}$ is superimposed. Each contribution is normalized to an integrated luminosity of $2.7\,\text{fb}^{-1}$, and the signal contribution is additionally scaled by a factor of 15 for better readability. The distributions in data (markers) are also shown. In the top row the ≥ 6 jets 2 b-tag, the 4 jets 3 b-tag, and the 5 jets 3 b-tag category are shown. Below are the ≥ 6 jets 3 b-tag category, the 4 jets 4 b-tag category with low BDT output, and the 4 jets 4 b-tag category with high BDT output.

Table 5: Best-fit value of the signal strength modifier μ and expected and observed 95% CL upper limits (UL) in the lepton+jets and dilepton channels as well as the combined results. The one standard deviation ($\pm 1\sigma$) confidence intervals of the best-fit value and the expected limit is also quoted. Expected limits are calculated with the asymptotic method [73].

Channel	Best-fit μ	Observed UL	Expected UL
Lepton+jets	$-0.4^{+2.1}_{-2.1}$	4.0	$4.1^{+1.8}_{-1.2}$
Dilepton	$-4.7^{+3.7}_{-3.8}$	5.2	$7.7^{+3.6}_{-2.3}$
Combined	$-2.0^{+1.8}_{-1.8}$	2.6	$3.6^{+1.6}_{-1.1}$

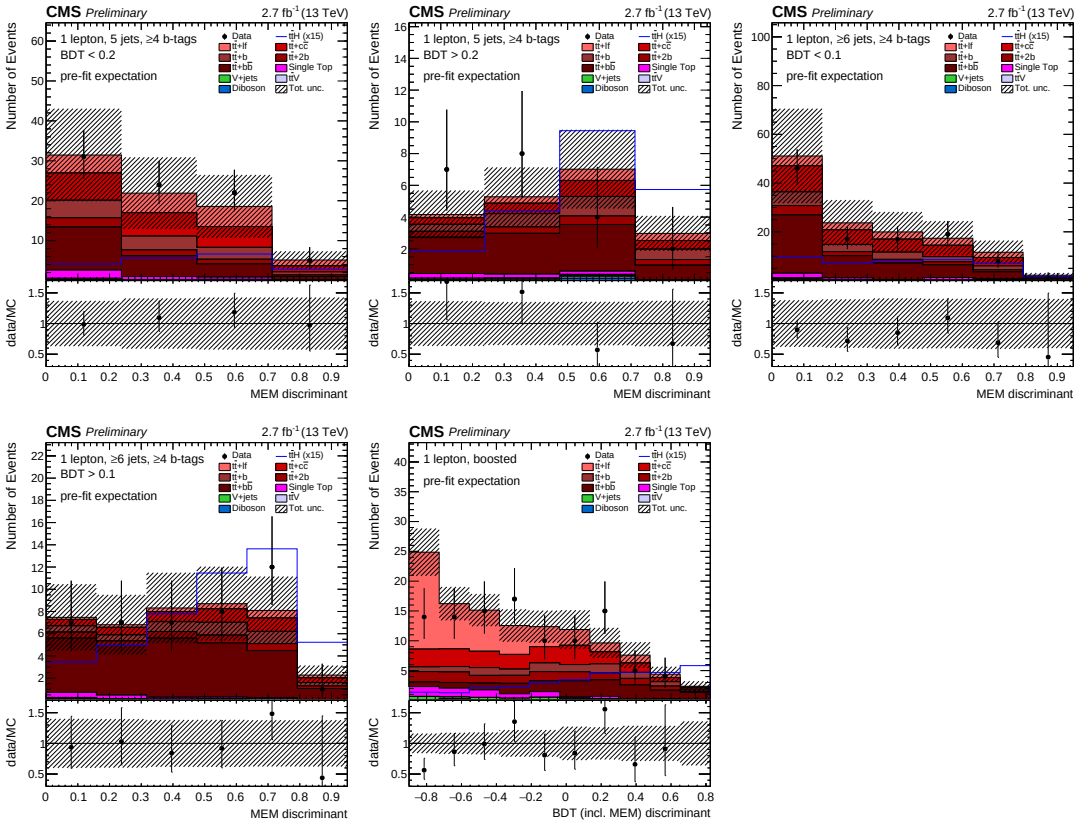

Figure 3: Final discriminant shapes in the different analysis categories in the lepton+jets channel before the fit to data. The expected background contributions (filled histograms) are stacked, and the expected signal distribution (line) for a Higgs-boson mass of $m_H = 125\,\text{GeV}$ is superimposed. Each contribution is normalized to an integrated luminosity of $2.7\,\text{fb}^{-1}$, and the signal contribution is additionally scaled by a factor of 15 for better readability. The distributions in data (markers) are also shown. In the top row the 5 jets 4 b-tag with low BDT output, the 5 jets 4 b-tag category with high BDT output, and the 6 jet category with low BDT output is shown. Below are the 6 jets 4 b-tag category with high BDT output, and the boosted category.

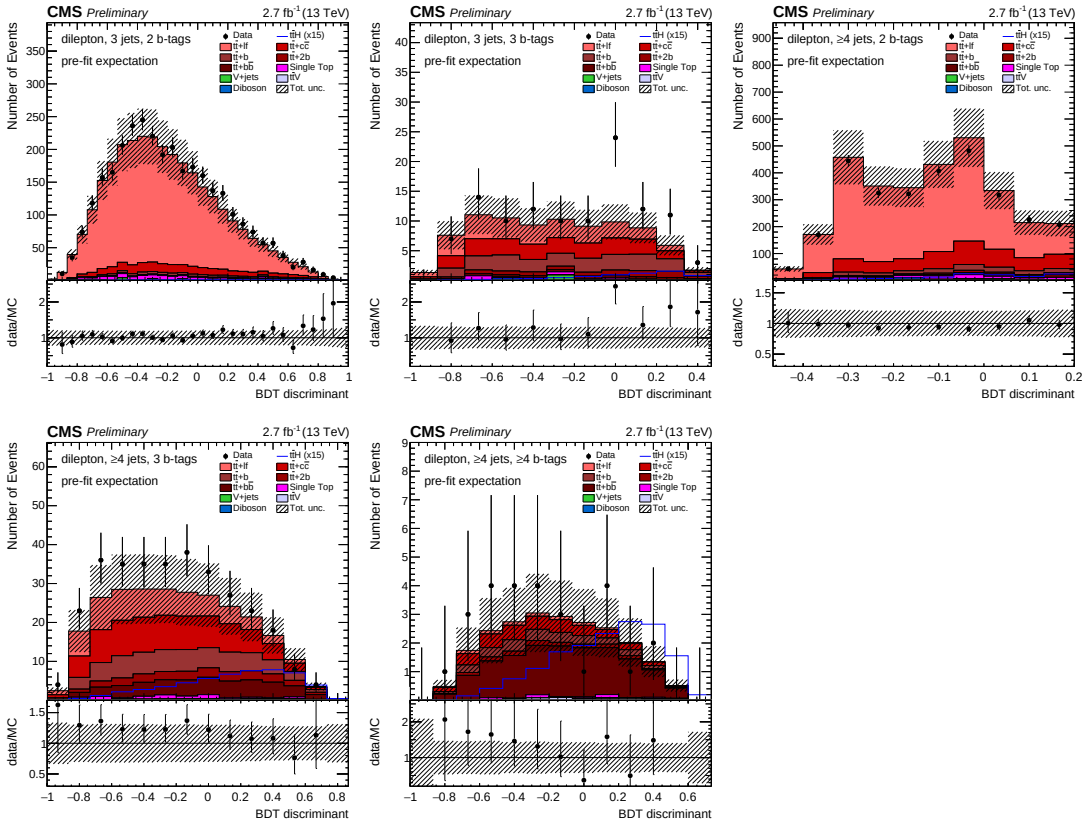

Figure 4: Final discriminant shapes in all analysis categories in the dilepton channel before the fit to data. In the top row the 3 jets, 2 b-tags, 3 jets, 3 b-tags, and $\geq$ 4 jets, 2 b-tags are shown. Below are the $\geq$ 4 jets, 3 b-tags and $\geq$ 4 jets, $\geq$ 4 b-tags categories. The expected background contributions (filled histograms) are stacked, and the expected signal distribution (line) for a Higgs-boson mass of $m_H = 125\,\text{GeV}$ is superimposed. Each contribution is normalized to an integrated luminosity of 2.7 fb^{-1}, and the signal contribution is additionally scaled by a factor of 15 for better readability. The distributions in data (markers) are also shown.

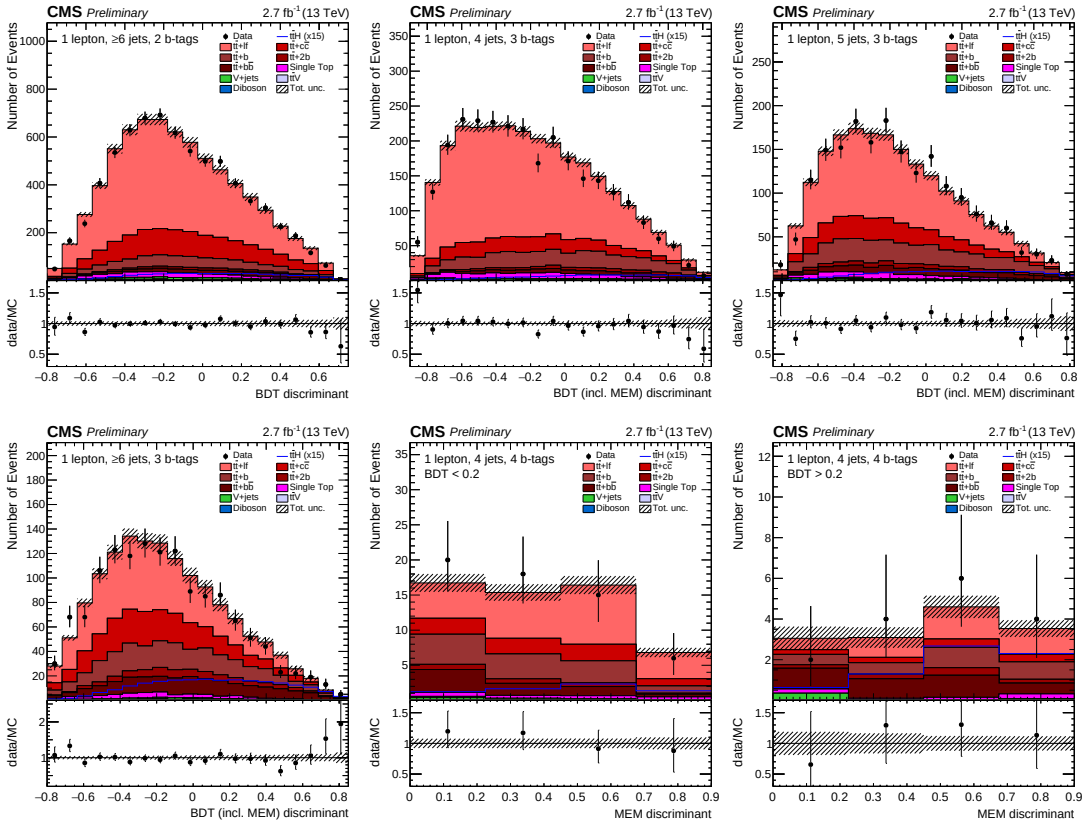

Figure 5: Final discriminant shapes in the different analysis categories in the lepton+jets channel after the fit to data. The expected background contributions (filled histograms) are stacked, and the expected signal distribution (line) for a Higgs-boson mass of $m_H = 125\,\text{GeV}$ is superimposed. Each contribution is normalized to an integrated luminosity of $2.7\,\text{fb}^{-1}$, and the signal contribution is additionally scaled by a factor of 15 for better readability. The distributions in data (markers) are also shown. In the top row the 5 jets 4 b-tag with low BDT output, the 5 jets 4 b-tag category with high BDT output, and the 6 jet category with low BDT output is shown. Below are the 6 jets 4 b-tag category with high BDT output, and the boosted category.

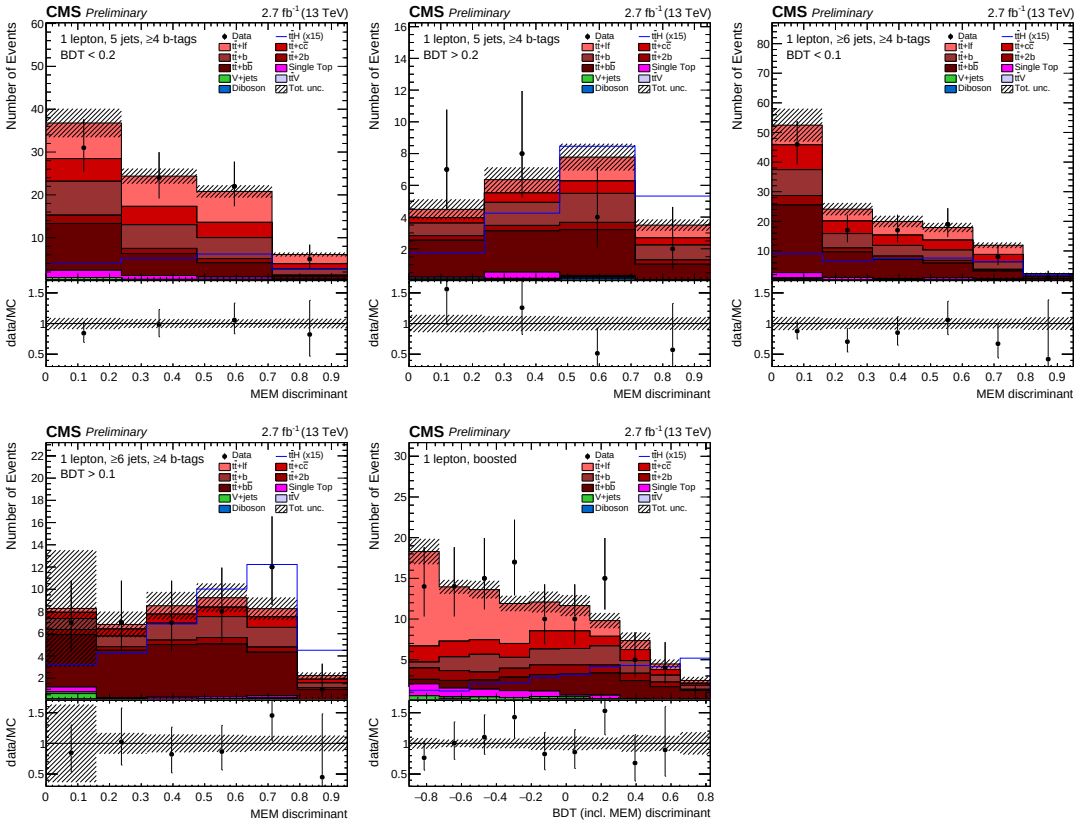

Figure 6: Final discriminant shapes in the different analysis categories in the lepton+jets channel after the fit to data. The expected background contributions (filled histograms) are stacked, and the expected signal distribution (line) for a Higgs-boson mass of $m_H = 125\,\text{GeV}$ is superimposed. Each contribution is normalized to an integrated luminosity of $2.7\,\text{fb}^{-1}$, and the signal contribution is additionally scaled by a factor of 15 for better readability. The distributions in data (markers) are also shown. In the top row the 5 jets 4 b-tag with low BDT output, the 5 jets 4 b-tag category with high BDT output, and the 6 jet category with low BDT output is shown. Below are the 6 jets 4 b-tag category with high BDT output, and the boosted category.

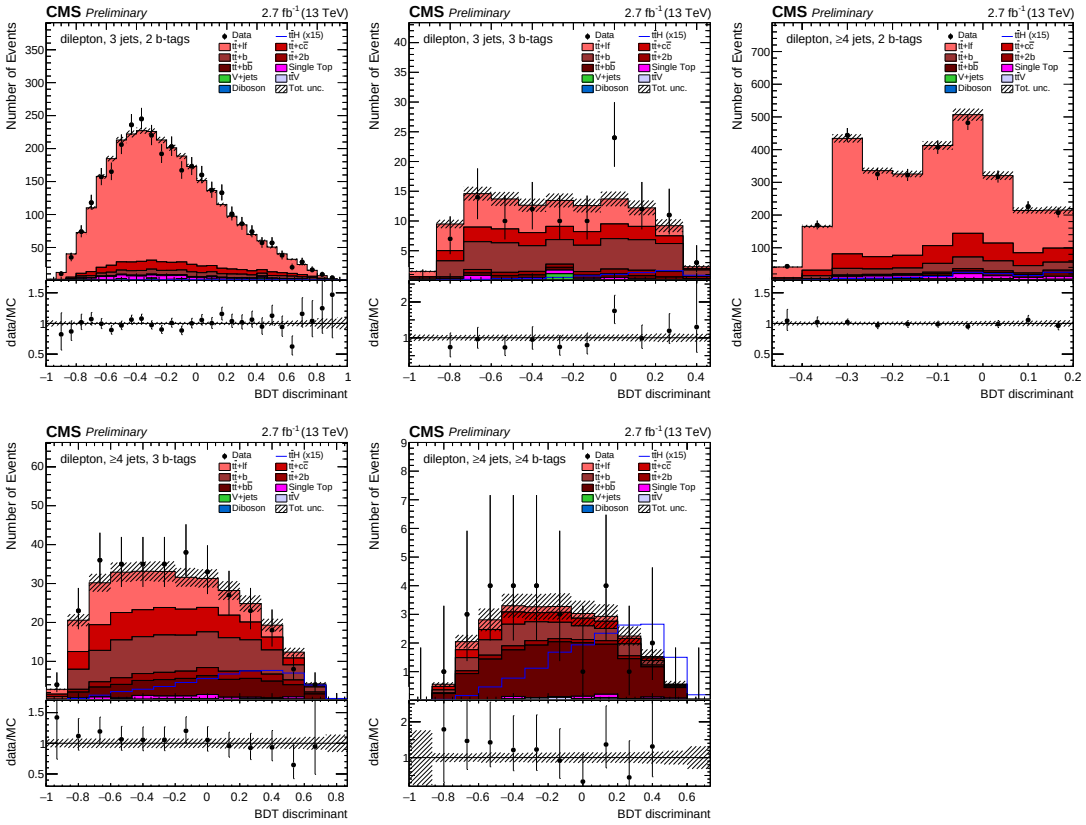

Figure 7: Final discriminant shapes in all analysis categories in the dilepton channel after the fit to data. In the top row the 3 jets, 2 b-tags, 3 jets, 3 b-tags, and $\geq$ 4 jets, 2 b-tags are shown. Below are the $\geq$ 4 jets, 3 b-tags and $\geq$ 4 jets, $\geq$ 4 b-tags categories. The expected background contributions (filled histograms) are stacked, and the expected signal distribution (line) for a Higgs-boson mass of $m_H = 125$ GeV is superimposed. Each contribution is normalized to an integrated luminosity of 2.7 fb^{-1}, and the signal contribution is additionally scaled by a factor of 15 for better readability. The distributions in data (markers) are also shown.

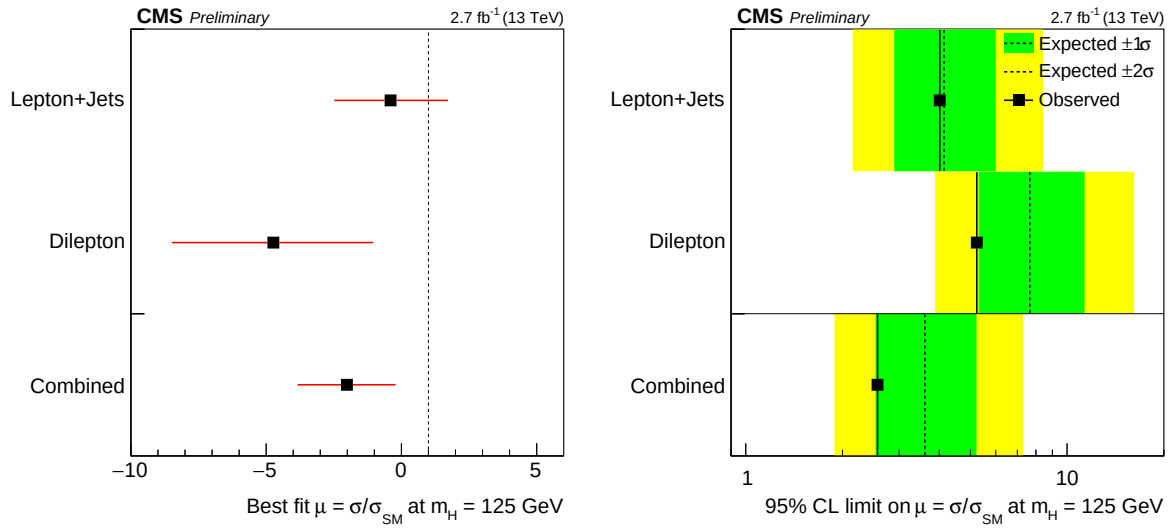

Figure 8: (left) Best-fit values of the signal strength modifiers μ with their $\pm1\sigma$ confidence intervals. (right) Median expected and observed 95% CL upper limits on μ. The expected limits are displayed together with $\pm1\sigma$ and $\pm2\sigma$ confidence intervals.

8 Conclusions

A search for the associated production of a Higgs boson and a top quark-antiquark pair ($t\bar{t}H$) is performed using the first data recorded with the CMS detector at a center-of-mass energy of 13 TeV in 2015. Candidate events are selected in final states compatible with the Higgs boson decay $H \to b\bar{b}$ and the lepton+jets or dilepton decay channel of the $t\bar{t}$ pair. Selected events are split into mutually exclusive categories according to their $t\bar{t}$ decay channel and jet content, including a category for "boosted" jets from hadronic decays of top quarks or $H \to b\bar{b}$ decays with large transverse momenta. In each category a powerful discriminant is constructed to separate the $t\bar{t}H$ signal from the $t\bar{t}$-dominated background, based on boosted decision trees and the matrix element method. Using $2.7\,\text{fb}^{-1}$ of pp collsion data an observed (expected) upper limit on the $t\bar{t}H$ production cross section relative to the SM expectations of $\mu = 2.6\,(3.6)$ at the 95% confidence level is obtained.

A Lepton+Jets Additional Material

In the following, the input variables used to train the BDTs in each category of the lepton+jets channel are presented. In Tables 6 and 7, all variables used in any of the categories are described, and in Table 8, the variables used per category are listed. The expected upper limits at 95% CL on the signal strength modifier μ under the background-only hypothesis in the lepton+jets channel are listed in Table 9 and displayed in Fig. 9 for the individual categories and for the combined fit in all categories.

Table 6: Variables used in the BDT training in the lepton+jets channel (continued in Table 7).

Event variable	Description
Object and event kinematics	
jet 1, 2, 3, 4 p_T	Jet transverse momenta, jets ordered in p_{T}
HT	Scalar sum of transverse momentum for all jets with $p_T > 30$ GeV/c
MET	Missing transverse energy
$\sum p_T$(jets,leptons,MET)	Sum of the p_T of all jets, leptons, and MET
mass(lepton,jet,MET)	Invariant mass of the 4-vector sum of all jets, leptons, and MET
avg ΔR(tag,tag)	Average ΔR between b-tagged jets
avg $\Delta\eta$(jet,jet)	Average $\Delta\eta$ between jets
max $\Delta\lvert\eta\rvert$ (jet, avg jet $\lvert\eta\rvert$)	max difference between jet $\lvert\eta\rvert$ and avg $\lvert\eta\rvert$ of jets
max $\Delta\lvert\eta\rvert$ (tag, avg jet $\lvert\eta\rvert$)	max difference between tag $\lvert\eta\rvert$ and avg $\lvert\eta\rvert$ of jets
max $\Delta\lvert\eta\rvert$ (tag, avg tag $\lvert\eta\rvert$)	max difference between tag $\lvert\eta\rvert$ and avg η of tags
min ΔR(tag,tag)	ΔR between the two closest b-tagged jets
M3	Invariant mass of the 3-jet system with largest transverse momentum
min ΔR(lepton,jet)	ΔR between the lepton and the closest jet (LJ channel)
mass(lepton,closest tag)	Invariant mass of the lepton and the closest b-tagged jet in ΔR (LJ channel)
closest tagged dijet mass	Invariant mass of the two b-tagged jets that are closest in ΔR
tagged dijet mass closest to 125	Invariant mass of the b-tagged pair closest to 125 GeV/c^2
best Higgs mass	A minimum-chi-squared fit to event kinematics is used to select two b-tagged jets as top-decay products. Of the remaining b-tagged jets, the invariant mass of the two with highest E_{T} is saved.
$\sqrt{\Delta\eta(t^{lep}, bb) \times \Delta\eta(t^{had}, bb)}$	Square root of the product of abs $\Delta\eta$ (leptonic top, bb) and abs $\Delta\eta$ (hadronic top, bb), where the bb-system and the candidates for the leptonic and hadronic tops are found with the best higgs mass algorithm
(Σ jet p_T)/(Σ jet E)	Ratio of the sum of the transverse momentum of all jets and the sum of the energy of all jets
CSVv2IVF b-tag	
first- to fifth-highest CSV	First- to fifth-highest highest CSVv2IVF discriminator value of all jets
avg CSV (tags/all)	Average b-tag discriminator value for b-tagged/all jets
dev from avg CSV (tags)	Squared difference between the CSVv2IVF discriminator value of a given b-tagged jet and the average CSVv2IVF discriminator value among b-tagged jets, summed over all b-tagged jets
sphericity	Sphericity: $3/2(\lambda_2 + \lambda_3)$ (λ_i: eigenvalues of momentum tensor)
aplanarity	Aplanarity: $3/2\lambda_1$ (λ_i: eigenvalues of momentum tensor)
H_1, H_2, H_3, H_4	Fox-Wolfram moments [74]

Table 7: continued from Table 6.

Event variable	Description
MEM discriminator	
MEM discriminator	MEM discriminator
Boosted object and event reconstruction	
τ_2/τ_1 Higgs cand.	2-subjettiness to 1-subjettiness ratio of Higgs candidate fat jet [75]
m(Higgs, di-filterjet)	Invariant mass of boosted Higgs candidate reconstructed from filtered subjets B1 and B2
$\Delta\eta$(top,Higgs)	Pseudo rapidity difference between boosted top candidate and boosted Higgs candidate
MEM discriminator (using subjets)	MEM discriminator using the subjets from the boosted top candidate

Table 8: BDT input variable assignment per category in the lepton+jets channel.

≥ 4 jets, ≥ 2 b-tags boosted	4 jets, 3 b-tags	4 jets, ≥ 4 b-tags
avg ΔR(tag,tag)	H_1	closest tagged dijet mass
τ_2/τ_1 of Higgs cand.	b-tagging likelihood ratio	b-tagging likelihood ratio
third-highest CSV	$\sum p_T$(jets,leptons,MET)	$\sum p_T$(jets,lepton,MET)
fourth-highest CSV	MEM discriminator	avg ΔR(tag,tag)
$\Delta\eta$(top,Higgs)	avg CSV (tags)	H_3
aplanarity	avg CSV (all)	jet 1 p_T
m(Higgs, di-filterjet)	jet 2 p_T	
min ΔR(tag,tag)	jet 4 p_T	
avg CSV (all)		
MEM discriminator (using subjets)		
b-tagging likelihood ratio		
	5 jets, 3 b-tags	5 jets, ≥ 4 b-tags
	MEM discriminator	b-tagging likelihood ratio
	avg ΔR(tag,tag)	jet 3 p_T
	min ΔR(lepton,jet)	tagged dijet mass closest to 125
	b-tagging likelihood ratio	avg $\Delta\eta$(jet,jet)
	fourth-highest CSV	avg ΔR(tag,tag)
	H_1	H_1
	dev from avg CSV (tags)	fifth-highest CSV
	avg $\Delta\eta$(jet,jet)	(Σ jet p_T)/(Σ jet E)
	avg CSV (tags)	
	avg CSV (all)	
	max $\Delta\lvert\eta\rvert$ (tag, avg jet $\lvert\eta\rvert$)	
≥ 6 jets, 2 b-tags	≥ 6 jets, 3 b-tags	≥ 6 jets, ≥ 4 b-tags
avg $\Delta\eta$(tag,tag)	b-tagging likelihood ratio	$\sum p_T$(jets,leptons,MET)
avg ΔR(tag,tag)	$\sqrt{\Delta\eta(t^{lep}, bb) \times \Delta\eta(t^{had}, bb)}$	H_3
ΔR(jet1, jet2)	HT	best Higgs mass
b-tagging likelihood ratio	MEM discriminator	b-tagging likelihood ratio
max $\Delta\lvert\eta\rvert$ (tag, avg tags $\lvert\eta\rvert$)	$\sum p_T$(jets,lepton,MET)	tagged dijet mass closest to 125
third-highest CSV	H_1	fifth-highest CSV
sphericity	fourth-highest CSV	(Σ jet p_T)/(Σ jet E)
fourth-highest CSV	avg CSV (tags)	jet 4 p_T
max $\Delta\lvert\eta\rvert$ (tag, avg jet $\lvert\eta\rvert$)	max $\Delta\lvert\eta\rvert$ (tag, avg jet $\lvert\eta\rvert$)	sphericity
min ΔR(tag,tag)		max $\Delta\lvert\eta\rvert$ (tag, avg tag $\lvert\eta\rvert$)
		second-highest CSV

Table 9: Median expected 95% CLs upper limits on μ in the lepton+jets channel, calculated with the asymptotic method. The upper and lower range of one standard deviation is also quoted.

Category	Observed	Expected
4 jets, 3 b-tags	14.5	$18.6^{+8.2}_{-5.5}$
4 jets, $\geq$ 4 b-tags high BDT output	35.7	$25.6^{+13.4}_{-8.1}$
4 jets, $\geq$ 4 b-tags low BDT output	86.6	$84.2^{+41.3}_{-25.8}$
5 jets, 3 b-tags	16.0	$12.3^{+5.5}_{-3.6}$
5 jets, $\geq$ 4 b-tags high BDT output	7.5	$10.3^{+5.6}_{-3.4}$
5 jets, $\geq$ 4 b-tags low BDT output	35.2	$31.9^{+16.1}_{-9.9}$
$\geq$ 6 jets, 2 b-tags	25.4	$41.1^{+21.1}_{-13.1}$
$\geq$ 6 jets, 3 b-tags	9.6	$7.6^{+3.3}_{-2.2}$
$\geq$ 6 jets, $\geq$ 4 b-tags high BDT output	9.2	$8.3^{+4.4}_{-2.7}$
$\geq$ 6 jets, $\geq$ 4 b-tags low BDT output	15.4	$18.3^{+9.6}_{-5.8}$
$\geq$ 4 jets, $\geq$ 2 b-tags, boosted	7.5	$10.7^{+5.9}_{-3.5}$
lepton+jets combined	4.0	$4.1^{+1.8}_{-1.2}$

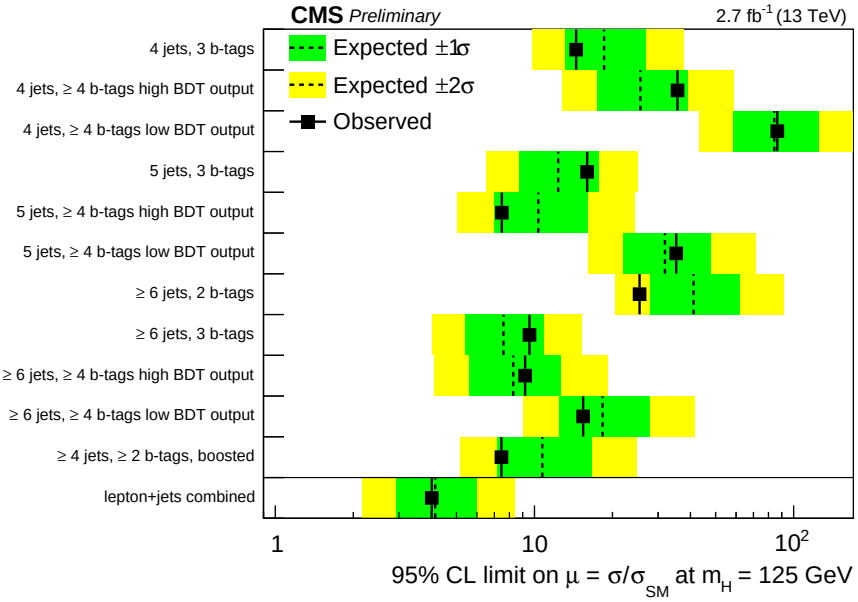

Figure 9: Visualisation of 95% CL upper limits on μ in the lepton+jets channel. The limits are calculated with the asymptotic method and displayed as the median and the $\pm1\sigma$ and $\pm2\sigma$ confidence intervals.

B Dilepton Additional Material

In the following, the input variables used to train the BDTs in each category of the dilepton channel are presented. In Table 10, all variables used in any of the categories are described, and in Table 11, the variables used per category are listed. The expected upper limits at 95% CL on the signal strength modifier μ under the background-only hypothesis in the dilepton channel are listed in Table 12 and displayed in Fig. 10 for the individual categories and for the combined fit in all categories.

Table 10: Variables used in the BDT training in the dilepton channel.

Event variable	Description
Object and event kinematics	
$\langle \Delta R_{tag,tag} \rangle$	Average ΔR between b-tagged jets
$\sum p_{T\,jets,leptons}$	Sum of the p_T of all jets and leptons
$\tau_{jet,jet}^{\text{max mass}}$	Twist angle between jet pair
min $\Delta R_{tag,tag}$	ΔR between the two closest b-tagged jets
max $\Delta\eta_{tag,tag}$	$\Delta\eta$ between the two furthest b-tagged jets
$m_{jet,jet}^{min\Delta R}$	Invariant mass of jet pair ΔR
$M_{higgs-like}^{jj}$	Invariant mass of a jet pair ordered in closeness to a Higgs mass
$m_{tag,tag}^{min\ \Delta R}$	Invariant mass of b-tag jet pair with minimum ΔR
$p_{T\ tag,tag}^{min\ \Delta R}$	Sum p_T of b-tag jet pair with minimum ΔR
Centrality (tags)	Ratio of the sum of the transverse momentum of all b-tagged jets and the sum of the energy of all b-tagged jets
H_T	Scalar sum of transverse momentum for all jets
min $\Delta R_{jet,jet}$	ΔR between the two closest jets
median $m_{jet,jet}$	Median invariant mass of all combinations of jet pairs
$m_{tag,tag}^{\text{max mass}}$	Invariant mass of b-tagged jet pair with maximum invariant mass combination
$\langle \Delta R_{jet,tag} \rangle$	Average ΔR between jets (with at least one b-tagged)
$p_{T\ jet,tag}^{min\ \Delta R}$	Sum p_T of jet pair with minimum ΔR between them (with at least one b-tag jet)
$\tau_{jet,tag}^{\text{max mass}}$	Twist angle between jet pair (with at least one b-tagged)
$m_{jet,tag,tag}^{\text{max } p_T}$	Invariant mass of the 3-jet system with the largest transverse momentum where at least two jets are b-tagged.
$M_{higgs-like}^{bj}$	Invariant mass of a jet pair (with at least one b-tagged) ordered in closeness to a Higgs mass.
CSVv2IVF b-tag	
$\langle d \rangle_{tagged/untagged}$	Average CSV b-tag discriminant value for b-tagged/un-b-tagged jets
Event shape	
H_0, H_1, H_2, H_3, H_4	Fox-Wolfram moments [74]

Table 11: BDT input variable assignment per category in the dilepton channel.

3 jets, 2 b-tags	3 jets, 3 b-tags	$\geq$ 4 jets, 2 b-tags	$\geq$ 4 jets, 3 b-tags	$\geq$ 4 jets, $\geq$ 4 b-tags
$\langle d \rangle_{untagged}$	$\langle d \rangle_{tagged}$	median $m_{jet,jet}$	min $\Delta R_{tag,tag}$	min $\Delta R_{tag,tag}$
$\sum p_{T\,jets,leptons}$	$m^{min\Delta R}_{tag,tag}$	H_1/H_0(tags)	$\langle d \rangle_{untagged}$	median $m_{jet,jet}$
$\tau^{\text{max mass}}_{jet,jet}$	$m^{\text{max mass}}_{\text{tag,tag}}$	$m^{min\Delta R}_{jet,jet}$	$\langle d \rangle_{tagged}$	max $\Delta\eta_{tag,tag}$
min $\Delta R_{tag,tag}$	max $\Delta\eta_{jet,jet}$	$\langle d \rangle_{untagged}$	$m^{min\Delta R}_{tag,tag}$	$M^{jj}_{higgs-like}$
max $\Delta\eta_{tag,tag}$	H_4/H_0(tags)	$H_2(jets)$	$M^{jj}_{higgs-like}$	H^{tags}_T
$m^{min\Delta R}_{jet,jet}$	$H_1(jets)$	$\sum p_{T\,jets,leptons}$	max $\Delta\eta_{tag,tag}$	$\langle d \rangle_{tagged}$
$M^{jj}_{higgs-like}$	$\tau^{\text{max mass}}_{jet,jet}$	$\langle \Delta R_{jet,tag} \rangle$	$\langle \Delta R_{jet,tag} \rangle$	$m^{min\Delta R}_{jet,tag}$
$m^{min\Delta R}_{tag,tag}$	$\sum p_{T\,jets,leptons}$	H^{jets}_T	$H_2(tags)$	$m^{min\Delta R}_{jet,jet}$
	min $\Delta R_{jet,jet}$	$m^{min\Delta R}_{tag,tag}$	$\sum p_{T\,jets,leptons}$	$m^{\text{max mass}}_{\text{tag,tag}}$
	$M^{bj}_{higgs-like}$	$p^{min\ \Delta R}_{T\ jet,tag}$	$\tau^{\text{max mass}}_{tag,tag}$	max $\Delta\eta_{jet,jet}$
			Centrality(jets & leptons)	Centrality(jets & leptons)
			$m^{\text{max } p_T}_{\text{jet,jet,jet}}$	Centrality(tags)

Table 12: Median expected 95% CLs upper limits on μ in the dilepton channel, calculated with the asymptotic method. The upper and lower range of one standard deviation is also quoted.

Category	Observed	Expected
3 jets, 2 b-tags	186.0	$114.8^{+52.6}_{-34.1}$
$\geq$ 3 jets, 3 b-tags	104.9	$48.6^{+26.2}_{-15.9}$
$\geq$ 4 jets, 2 b-tags	32.4	$40.1^{+16.8}_{-11.3}$
$\geq$ 4 jets, 3 b-tags	7.4	$10.8^{+5.2}_{-3.3}$
$\geq$ 4 jets, $\geq$ 4 b-tags	9.1	$12.2^{+7.5}_{-4.3}$
dilepton combined	5.2	$7.7^{+3.6}_{-2.3}$

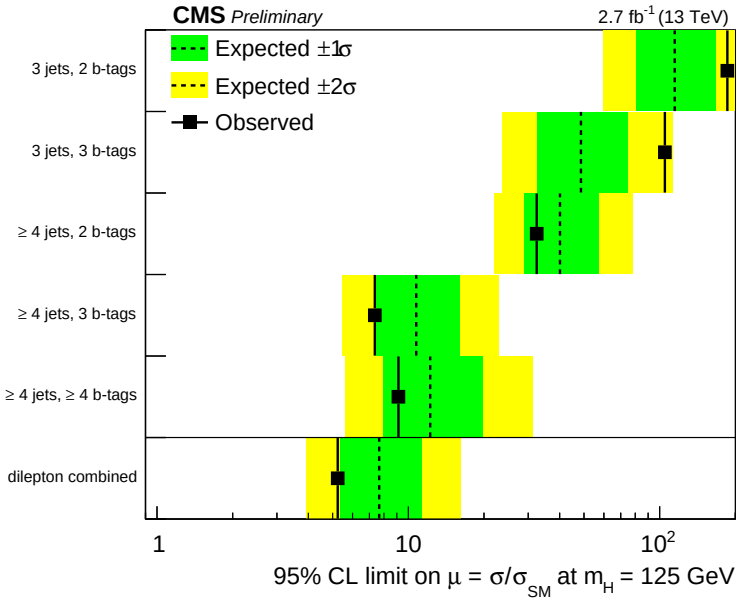

Figure 10: Visualisation of 95% CL upper limits on μ in the dilepton channel. The limits are calculated with the asymptotic method and displayed as the median and the $\pm 1\sigma$ and $\pm 2\sigma$ confidence intervals.

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), no. 1, 1–29, 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), no. 1, 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] CMS Collaboration, "Evidence for the direct decay of the 125 GeV Higgs boson to fermions", *Nature Phys.* **10** (2014), no. 5, 557–560, doi:10.1038/nphys3005, arXiv:1401.6527.

[4] ATLAS Collaboration, "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector", *JHEP* **04** (2015) 117, doi:10.1007/JHEP04(2015)117, arXiv:1501.04943.

[5] ATLAS Collaboration, "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC", *Phys. Lett. B* **726** (2013), no. 1-3, 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)].

[6] 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), no. 5, 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.

[7] ATLAS Collaboration, "Evidence for the spin-0 nature of the Higgs boson using ATLAS data", *Phys. Lett. B* **726** (2013), no. 1-3, 120–144, doi:10.1016/j.physletb.2013.08.026, arXiv:1307.1432.

[8] 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), no. 1, 012004, doi:10.1103/PhysRevD.92.012004, arXiv:1411.3441.

[9] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", doi:10.5170/CERN-2013-004, arXiv:1307.1347.

[10] G. Burdman, M. Perelstein, and A. Pierce, "Large Hadron Collider tests of a little Higgs model", *Phys. Rev. Lett.* **90** (2003), no. 24, 241802, doi:10.1103/PhysRevLett.90.241802, arXiv:hep-ph/0212228. [Erratum: Phys. Rev. Lett.92,049903(2004)].

[11] T. Han, H. E. Logan, B. McElrath, and L.-T. Wang, "Phenomenology of the little Higgs model", *Phys. Rev. D* **67** (2003), no. 9, 095004, doi:10.1103/PhysRevD.67.095004, arXiv:hep-ph/0301040.

[12] M. Perelstein, M. E. Peskin, and A. Pierce, "Top quarks and electroweak symmetry breaking in little Higgs models", *Phys. Rev. D* **69** (2004), no. 7, 075002, doi:10.1103/PhysRevD.69.075002, arXiv:hep-ph/0310039.

[13] H.-C. Cheng, I. Low, and L.-T. Wang, "Top partners in little Higgs theories with T-parity", *Phys. Rev. D* **74** (2006), no. 5, 055001, doi:10.1103/PhysRevD.74.055001, arXiv:hep-ph/0510225.

[14] H.-C. Cheng, B. A. Dobrescu, and C. T. Hill, "Electroweak symmetry breaking and extra dimensions", *Nucl. Phys. B.* **589** (2000), no. 1-3, 249–268, doi:10.1016/S0550-3213(00)00401-6, arXiv:hep-ph/9912343.

[15] M. Carena, E. Ponton, J. Santiago, and C. E. M. Wagner, "Light Kaluza Klein States in Randall-Sundrum Models with Custodial SU(2)", *Nucl. Phys. B.* **759** (2006), no. 1-2, 202–227, doi:10.1016/j.nuclphysb.2006.10.012, arXiv:hep-ph/0607106.

[16] R. Contino, L. Da Rold, and A. Pomarol, "Light custodians in natural composite Higgs models", *Phys. Rev. D* **75** (2007), no. 5, 055014, doi:10.1103/PhysRevD.75.055014, arXiv:hep-ph/0612048.

[17] G. Burdman and L. Da Rold, "Electroweak Symmetry Breaking from a Holographic Fourth Generation", *JHEP* **12** (2007) 086, doi:10.1088/1126-6708/2007/12/086, arXiv:0710.0623.

[18] C. T. Hill, "Topcolor: Top quark condensation in a gauge extension of the standard model", *Phys. Lett. B* **266** (1991), no. 3, 419–424, doi:10.1016/0370-2693(91)91061-Y.

[19] A. Carmona, M. Chala, and J. Santiago, "New Higgs Production Mechanism in Composite Higgs Models", *JHEP* **07** (2012) 049, doi:10.1007/JHEP07(2012)049, arXiv:1205.2378.

[20] CMS Collaboration, "Search for the associated production of the Higgs boson with a top-quark pair", *JHEP* **09** (2014) 087, doi:10.1007/JHEP09(2014)087, 10.1007/JHEP10(2014)106, arXiv:1408.1682. [Erratum: JHEP10,106(2014)].

[21] ATLAS Collaboration, "Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector", *Phys. Lett. B* **749** (2015) 519–541, doi:10.1016/j.physletb.2015.07.079, arXiv:1506.05988.

[22] CMS Collaboration, "Search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method", *Eur. Phys. J. C.* **75** (2015), no. 6, 251, doi:10.1140/epjc/s10052-015-3454-1, arXiv:1502.02485.

[23] ATLAS Collaboration, "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector", *Eur. Phys. J. C.* **75** (2015), no. 7, 349, doi:10.1140/epjc/s10052-015-3543-1, arXiv:1503.05066.

[24] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", doi:10.5170/CERN-2011-002, arXiv:1101.0593.

[25] T. Plehn, G. P. Salam, and M. Spannowsky, "Fat Jets for a Light Higgs", *Phys. Rev. Lett.* **104** (2010), no. 11, 111801, doi:10.1103/PhysRevLett.104.111801, arXiv:0910.5472.

[26] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008), no. 8, S08004, doi:10.1088/1748-0221/3/08/S08004.

[27] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003), no. 3, 250, doi:10.1016/S0168-9002(03)01368-8.

[28] 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.

[29] E. Re, "Single-top Wt-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.

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

[31] 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.

[32] 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.

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

[34] CMS Collaboration, "Underlying event tunes and double parton scattering", CMS Physics Analysis Summary CMS-PAS-GEN-14-001, CMS Collaboration, 2014.

[35] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune", *Eur. Phys. J. C* **74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[36] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with W^- or H^-", *Phys. Rev. D* **82** (2010), no. 5, 054018, doi:10.1103/PhysRevD.82.054018, arXiv:hep-ph/1005.4451.

[37] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", *JHEP* **07** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[38] F. Maltoni, D. Pagani, and I. Tsinikos, "Associated production of a top-quark pair with vector bosons at NLO in QCD: impact on $t\bar{t}H$ searches at the LHC", arXiv:1507.05640.

[39] W. Beenakker et al., "Higgs radiation off top quarks at the Tevatron and the LHC", *Phys. Rev. Lett.* **87** (2001), no. 20, 201805, doi:10.1103/PhysRevLett.87.201805, arXiv:hep-ph/0107081.

[40] W. Beenakker et al., "NLO QCD corrections to $t\bar{t}H$ production in hadron collisions", *Nucl. Phys. B* **653** (2003), no. 1-2, 151, doi:10.1016/S0550-3213(03)00044-0, arXiv:hep-ph/0211352.

[41] S. Dawson, L. H. Orr, L. Reina, and D. Wackeroth, "Associated top quark Higgs boson production at the LHC", *Phys. Rev. D* **67** (2003), no. 7, 071503, doi:10.1103/PhysRevD.67.071503, arXiv:hep-ph/0211438.

[42] S. Dawson et al., "Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections", *Phys. Rev. D* **68** (2003), no. 3, 034022, doi:10.1103/PhysRevD.68.034022, arXiv:hep-ph/0305087.

[43] A. Djouadi, J. Kalinowski, and M. Spira, "HDECAY: A program for Higgs boson decays in the standard model and its supersymmetric extension", *Comput. Phys. Commun.* **108** (1998), no. 1, 56, doi:10.1016/S0010-4655(97)00123-9, arXiv:hep-ph/9704448.

[44] A. Djouadi, M. M. Mühlleitner, and M. Spira, "Decays of supersymmetric particles: The Program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface)", *Acta Phys. Polon. B* **38** (2007) 635, arXiv:hep-ph/0609292.

[45] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, "Precise predictions for the Higgs-boson decay $H \to WW/ZZ \to 4$ leptons", *Phys. Rev. D* **74** (2006), no. 1, 013004, doi:10.1103/PhysRevD.74.013004, arXiv:hep-ph/0604011.

[46] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, "Radiative corrections to the semileptonic and hadronic Higgs-boson decays $H \to WW/ZZ \to 4$ fermions", *JHEP* **02** (2007) 080, doi:10.1088/1126-6708/2007/02/080, arXiv:hep-ph/0611234.

[47] M. Cacciari et al., "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation", *Phys. Lett. B* **710** (2012), no. 4-5, 612, doi:10.1016/j.physletb.2012.03.013, arXiv:1111.5869.

[48] P. Baernreuther et al., "Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$", *Phys. Rev. Lett.* **109** (2012), no. 13, 132001, doi:10.1103/PhysRevLett.109.132001, arXiv:1204.5201.

[49] M. Czakon and A. Mitov, "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels", *JHEP* **12** (2012) 054, arXiv:1207.0236.

[50] M. Czakon and A. Mitov, "NNLO corrections to top-pair production at hadron colliders: the quark-gluon reaction", *JHEP* **01** (2013) 080, arXiv:1210.6832.

[51] M. Beneke et al., "Hadronic top-quark pair production with NNLL threshold resummation", *Nucl. Phys. B* **855** (2012), no. 3, 695, doi:10.1016/j.nuclphysb.2011.10.021, arXiv:1109.1536.

[52] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S{}^4)$", *Phys. Rev. Lett.* **110** (2013), no. 25, 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[53] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders", *Comput. Phys. Commun.* **185** (2014), no. 11, 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[54] CMS Collaboration, "Particle–flow event reconstruction in CMS and performance for jets, taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, CMS Collaboration, 2009.

[55] CMS Collaboration, "Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, CMS Collaboration, 2010.

[56] M. Cacciari, G. P. Salam, and G. Soyez, "The catchment area of jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.

[57] 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.

[58] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011), no. 11, P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[59] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013), no. 4, P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[60] 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, CMS Collaboration, 2016.

[61] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, "Stop Reconstruction with Tagged Tops", *JHEP* **10** (2010) 078, doi:10.1007/JHEP10(2010)078, arXiv:1006.2833.

[62] C. Anders et al., "Benchmarking an even better top tagger algorithm", *Phys. Rev.* **D89** (2014), no. 7, 074047, doi:10.1103/PhysRevD.89.074047, arXiv:1312.1504.

[63] G. Kasieczka et al., "Resonance Searches with an Updated Top Tagger", *JHEP* **06** (2015) 203, doi:10.1007/JHEP06(2015)203, arXiv:1503.05921.

[64] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC", *Phys. Rev. Lett.* **100** (2008), no. 24, 242001, doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470.

[65] G. P. Salam, "Towards Jetography", *Eur. Phys. J.* **C67** (2010) 637–686, doi:10.1140/epjc/s10052-010-1314-6, arXiv:0906.1833.

[66] A. Höcker et al., "TMVA: Toolkit for Multivariate Data Analysis", *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

[67] J. Kennedy and R. Eberhart, "Particle swarm optimization", in *Proceedings of the IEEE International Conference on neural networks*, volume 4, pp. 1942–1948. Nov, 1995. doi:10.1109/ICNN.1995.488968.

[68] CMS Collaboration, "CMS Luminosity Measurement for the 2015 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-15-001, CMS Collaboration, 2016.

[69] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput. Phys. Commun.* **77** (1993), no. 2, 219–228, doi:10.1016/0010-4655(93)90005-W.

[70] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra", in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*. 2011. arXiv:1103.0354.

[71] A. Read, "Modified frequentist analysis of search results (the CL_s method)", Technical Report CERN-OPEN-2000-005, CERN, 2000.

[72] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999), no. 2-3, 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[73] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J.* **C71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727. [Erratum: Eur. Phys. J.C73,2501(2013)].

[74] G. Fox and S. Wolfram, "Event shapes in e^+e^- annihilation", *Nuclear Physics B* **157** (1979), no. 3, 543–544, doi:10.1016/0550-3213(79)90120-2.

[75] J. Thaler and K. Tilburg, "Identifying boosted objects with N-subjettiness", *Journal of High Energy Physics* **2011** (2011), no. 3, 1–28, doi:10.1007/JHEP03(2011)015.