Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in $pp$ collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: A search for Higgs boson production in association with a pair of top quarks ($t\bar{t}H$) is performed, where the Higgs boson decays to $b\bar{b}$, and both top quarks decay hadronically. The data used correspond to an integrated luminosity of 20.3 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 8$ TeV collected with the ATLAS detector at the Large Hadron Collider. The search selects events with at least six energetic jets and uses a boosted decision tree algorithm to discriminate between signal and Standard Model background. The dominant multijet background is estimated using a dedicated data-driven technique. For a Higgs boson mass of 125 GeV, an upper limit of 6.4 (5.4) times the Standard Model cross section is observed (expected) at 95% confidence level. The best-fit value for the signal strength is $\mu = 1.6 \pm 2.6$ times the Standard Model expectation for $m_H = 125$ GeV. Combining all $t\bar{t}H$ searches carried out by ATLAS at $\sqrt{s} = 8$ and 7 TeV, an observed (expected) upper limit of 3.1 (1.4) times the Standard Model expectation is obtained at 95% confidence level, with a signal strength $\mu = 1.7 \pm 0.8$.

KEYWORDS: Hadron-Hadron scattering (experiments)

ArXiv ePrint: 1604.03812
Contents

1 Introduction 2

2 The ATLAS detector 3

3 Object reconstruction 3

4 Event selection 4

5 Signal and background modelling 4
   5.1 Signal model 4
   5.2 Simulated backgrounds 5
   5.3 Common treatment of MC samples 6
   5.4 Multijet background estimation using data: the TRF\textsubscript{MJ} method 7
   5.5 Validation of the TRF\textsubscript{MJ} method in data and simulation 8

6 Multijet trigger efficiency 10

7 Event classification 10

8 Analysis method 10

9 Systematic uncertainties 14

10 Statistical methods 19

11 Results 19

12 Combination of $t\bar{t}H$ results at $\sqrt{s} = 7$ and 8 TeV 24
   12.1 Individual $t\bar{t}H$ measurements and results 24
      12.1.1 $H \rightarrow b\bar{b}$ (single lepton and dilepton $t\bar{t}$ decays) 25
      12.1.2 $H \rightarrow (WW^{(*)}, \tau\tau, ZZ^{(*)}) \rightarrow$ leptons 25
      12.1.3 $H \rightarrow \gamma\gamma$ 25
   12.2 Correlations 26
   12.3 Results of the combination 26
      12.3.1 Signal strength 26
      12.3.2 Couplings 26

13 Conclusion 28

The ATLAS collaboration 35
1 Introduction

After the discovery of a new boson with a mass of around 125 GeV in July 2012 by the ATLAS [1] and CMS [2] collaborations, the focus has now shifted to confirming whether this particle is the Standard Model (SM) Higgs boson [3–6] or another boson. While any deviation from SM predictions would indicate the presence of new physics, all measurements of the properties of this new boson thus far performed at the Large Hadron Collider (LHC), including spin, parity, total width, and coupling to SM particles, are consistent with the SM prediction [7–12].

Because of its large mass, the top quark is the fermion with the largest Yukawa coupling ($y_t$) to the Higgs field in the SM, with a value close to unity. The coupling $y_t$ is experimentally accessible by measuring the gluon fusion (ggF) production process or the $H \to \gamma\gamma$ decay, where a sizeable contribution derives from a top-quark loop. This case requires the assumption that no new physics contributes with additional induced loops in order to measure $y_t$. Currently, the only process where $y_t$ can be accessed directly is the production of a top-quark pair in association with a Higgs boson ($t\bar{t}H$).

The results of searches for the Higgs boson are usually expressed in terms of the signal-strength parameter $\mu$, which is defined as the ratio of the observed to the expected number of signal events. The latter is calculated using the SM cross section times branching ratio [13]. The combined $t\bar{t}H$ signal strength measured by the CMS Collaboration [14], obtained by merging searches in several final states, is $\mu = 2.8 \pm 1.0$. The ATLAS Collaboration has searched for a $t\bar{t}H$ signal in events enriched in Higgs boson decays to two massive vector bosons or $\tau$ leptons in the multilepton channel [15], finding $\mu = 2.1^{+1.4}_{-1.2}$ for $t\bar{t}H(H \to b\bar{b})$ [16] in final states with at least one lepton obtaining $\mu = 1.5 \pm 1.1$, and for $t\bar{t}H(H \to \gamma\gamma)$ [17] measuring $\mu = 1.3^{+2.6}_{-1.7}$.

Among all $t\bar{t}H$ final states, the one where both $W$ bosons from $t \to Wb$ decay hadronically and the Higgs boson decays into a $b\bar{b}$ pair has the largest branching ratio, but also the least signal purity. This paper describes a search for this all-hadronic $t\bar{t}H(H \to b\bar{b})$ decay mode. The analysis uses proton-proton collision data corresponding to an integrated luminosity of $20.3$ fb$^{-1}$ at center-of-mass energy $\sqrt{s} = 8$ TeV recorded with the ATLAS detector at the LHC.

At Born level, the signal signature is eight jets, four of which are $b$-quark jets. The dominant background is the non-resonant production of multijet events. For this analysis, a data-driven method is applied to estimate the multijet background by extrapolating its contribution from a control region with the same jet multiplicity, but a lower multiplicity of jets containing $b$-hadrons than the signal process. The parameters used for the extrapolation are measured from a control region and checked using Monte Carlo (MC) simulations. Other subdominant background processes are estimated using MC simulations. To maximise the signal sensitivity, the events are categorised according to their number of jets and jets identified as containing $b$-hadrons ($b$-tagged). A boosted decision tree (BDT) algorithm, based on event shape and kinematic variables, is used to discriminate the signal from the background. The extraction of $\mu$ is performed through a fit to the BDT discriminant distribution. After the fit the dominant uncertainty is the $t\bar{t} + b\bar{b}$ production cross...
section. The sensitivity is also limited by systematic uncertainties from the data-driven method used for the modelling of the large non-resonant multijet production.

2 The ATLAS detector

The ATLAS detector [18] consists of an inner tracking detector surrounded by a thin superconducting solenoid magnet providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer incorporating three large superconducting toroid magnets. The inner detector (ID) comprises the high-granularity silicon pixel detector and the silicon microstrip tracker covering the pseudorapidity\(^1\) range \(|\eta| < 2.5\), and the straw-tube transition radiation tracker covering \(|\eta| < 2.0\). The electromagnetic calorimeter covers \(|\eta| < 3.2\) and consists of a barrel and two endcap high-granularity lead/liquid-argon (LAr) calorimeters. An additional thin LAr presampler covers \(|\eta| < 1.8\). Hadron calorimetry is provided by a steel/scintillator-tile calorimeter, which covers the region \(|\eta| < 1.7\), and two copper/LAr hadron endcap calorimeters. To complete the pseudorapidity coverage, copper/LAr and tungsten/LAr forward calorimeters cover up to \(|\eta| = 4.9\). Muon tracking chambers precisely measure the deflection of muons in the magnetic field generated by superconducting air-core toroids in the region \(|\eta| < 2.7\). A three-level trigger system selects events for offline analysis [19]. The hardware-based Level-1 trigger is used to reduce the event rate to a maximum of 75 kHz, while the two software-based trigger levels, Level-2 and Event Filter (EF), reduce the event rate to about 400 Hz.

3 Object reconstruction

The all-hadronic \(t\bar{t}H\) final state is composed of jets originating from \((u, d, s)\)-quarks or gluons (light jets) and jets from \(c\)- or \(b\)-quarks (heavy-flavour jets). Electrons and muons, selected in the same way as in ref. [16], are used only to veto events that would overlap with the \(t\bar{t}H\) searches in final states with leptons.

At least one reconstructed primary vertex is required, with at least five associated tracks with \(p_T \geq 400\) MeV, and a position consistent with the luminous region of the beams in the transverse plane. If more than one vertex is found, the primary vertex is taken to be the one which has the largest sum of the squared transverse momenta of its associated tracks.

Jets are reconstructed with the anti-\(k_t\) algorithm [20–22], with a radius parameter \(R = 0.4\) in the \((\eta, \phi)\) plane. They are built from calibrated topological clusters of energy deposits in the calorimeters [18]. Prior to jet finding, a local cluster calibration scheme [23, 24] is applied to correct the topological cluster energies for the effects of non-compensating calorimeter response, dead material, and out-of-cluster leakage. After energy calibration

\(^1\)ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis coinciding with the axis of the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \(\phi\) being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle \(\theta\) as \(\eta = -\ln \tan(\theta/2)\). Transverse momentum and energy are defined as \(p_T = p \sin \theta\) and \(E_T = E \sin \theta\) respectively.
based on in-situ measurements [25], jets are required to have transverse momentum $p_T > 25$ GeV and $|\eta| < 2.5$. During jet reconstruction, no distinction is made between identified electrons and jet energy deposits. To avoid double counting electrons as jets, any jet within a cone of size $\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2} = 0.2$ around a reconstructed electron is discarded. After this, electrons within a $\Delta R = 0.4$ of a remaining jet are removed.

To avoid selecting jets from additional $pp$ interactions in the same event (pile-up), a loose selection is applied to the jet vertex fraction (JVF), defined as the ratio of the scalar sum of the $p_T$ of tracks matched to the jet and originating from the primary vertex to that of all tracks matched to the jet. This criterion, JVF $\geq 0.5$, is only applied to jets with $p_T < 50$ GeV and $|\eta| < 2.4$.

Jets are $b$-tagged by means of the MV1 algorithm [26]. It combines information from track impact parameters and topological properties of secondary and tertiary decay vertices which are reconstructed within the jet. The working point used for this search corresponds to a 60% efficiency to tag a $b$-quark jet, a light-jet rejection factor of approximately 700 and a charm-jet rejection factor of 8, as determined for jets with $p_T > 25$ GeV and $|\eta| < 2.5$ in simulated $t\bar{t}$ events [26]. The tagging efficiencies obtained in simulation are adjusted to match the results of the calibrations performed in data [26].

4 Event selection

This search is based on data collected using a multijet trigger, which requires at least five jets passing the EF stage, each having $p_T > 55$ GeV and $|\eta| < 2.5$. Events are discarded if any jet with $p_T > 20$ GeV is identified as out-of-time activity from a previous $pp$ collision or as calorimeter noise [27].

The five leading jets in $p_T$ are required to have $p_T > 55$ GeV with $|\eta| < 2.5$ and all other jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. Events are required to have at least six jets, of which at least two must be $b$-tagged. Events with well-identified isolated muons or electrons with $p_T > 25$ GeV are discarded in order to avoid overlap with other $t\bar{t}H$ analyses.

To enhance the sensitivity, the selected events are categorised into various distinct regions, according to their jet and $b$-tag multiplicities: the region with $m$ jets, of which $n$ are $b$-jets, is referred to as “$(m, n)$”.

5 Signal and background modelling

5.1 Signal model

The $t\bar{t}H$ signal process is modelled using matrix elements calculations obtained from the HELAC-Oneloop package [28] with next-to-leading order (NLO) accuracy in $\alpha_s$. POWHEG-BOX [29–31] serves as an interface to the MC programs used to simulate the parton shower and hadronisation. The samples created using this approach are referred to as PowHel samples [32]. They include all SM Higgs boson and top-quark decays and use the CT10NLO [33] parton distribution function (PDF) sets with the factorisation ($\mu_F$) and renormalisation ($\mu_R$) scales set to $\mu_F = \mu_R = m_t + m_H/2$. The PowHel $t\bar{t}H$ samples
use Pythia 8.1 [34] to simulate the parton shower with the CTEQ6L1 [35] PDF and the AU2 underlying-event set of generator parameters (tune) [36], while HERWIG [37] is used to estimate systematic uncertainties due to the fragmentation modelling.

For these $t\bar{t}H$ samples the cross-section normalisations and the Higgs boson decay branching fractions are taken from the NLO QCD and from the NLO QCD + EW theoretical calculations [13] respectively. The masses of the Higgs boson and the top quark are set to $125$ GeV and to $172.5$ GeV respectively.

5.2 Simulated backgrounds

The dominant background to the all-hadronic $t\bar{t}H$ signal is multijet production, followed by $t\bar{t}$ + jets production. Small background contributions come from the production of a single top quark and from the associated production of a vector boson and a $t\bar{t}$ pair, $t\bar{t}V$ ($V = W, Z$). The multijet background is determined from data using a dedicated method described in section 5.4. The other background contributions are estimated using MC simulations.

The multijet events, which are used for jet trigger studies and for the validation of the data-driven multijet background estimation, are simulated with Pythia 8.1 using the NNPDF2.3 LO [38] PDFs.

The main $t\bar{t}$ sample is generated using the POWHEG NLO generator with the CT10NLO PDF set, assuming a value of the top-quark mass of 172.5 GeV. It is interfaced to Pythia 6.425 [39] with the CTEQ6L1 PDF set and the Perugia2011C [40] underlying-event tune; this combination of generator and showering programs is hereafter referred to as POWHEG+PYTHIA. The sample is normalised to the top++2.0 theoretical calculation performed at next-to-next-to leading order (NNLO) in QCD and includes resummation of next-to-next-to leading logarithmic (NNLL) soft gluon terms [41–46]. A second $t\bar{t}$ sample is generated using fully matched NLO predictions with massive $b$-quarks [47] within the SHERPA with OPENLOOPS framework [48, 49] henceforth referred to as SHERPA+OPENLOOPS. The SHERPA+OPENLOOPS NLO sample is generated following the four-flavour scheme using the SHERPA 2.0 pre-release and the CT10NLO PDF set. The renormalisation scale is set to $\mu_R = \prod_{i=t,\bar{t},b} E_{T,i}^{1/4}$, where $E_{T,i}$ is the transverse energy of parton $i$, and the factorisation and resummation scales are both set to $(E_{T,t} + E_{T,\bar{t}})/2$.

The prediction from SHERPA+OPENLOOPS is expected to model the $t\bar{t} + b\bar{b}$ contribution more accurately than POWHEG+PYTHIA, since the latter MC produces $t\bar{t} + b\bar{b}$ exclusively via the parton shower. The SHERPA+OPENLOOPS sample is not passed through full detector simulation. Thus, $t\bar{t}$ + jets events from POWHEG+PYTHIA are categorised into three non-overlapping samples, $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$, and $t\bar{t}$ + light-jets, hereafter called $t\bar{t}$ + light, using a labelling based on an algorithm that matches hadrons to particle jets. Then, $t\bar{t} + b\bar{b}$ events from POWHEG+PYTHIA are reweighted to reproduce the SHERPA+OPENLOOPS NLO $t\bar{t} + b\bar{b}$ prediction. The reweighting is done at generator level using a finer categorisation to distinguish events where one particle jet is matched to two $b$-hadrons, or where only one $b$-hadron is matched. The reweighting is applied using several kinematic variables such as the top-quark $p_T$, the $t\bar{t}$ system $p_T$, and, where this can be defined, $\Delta R$ and $p_T$ of the dijet system not originating from the top-quark decay [16].
Unlike $t\bar{t} + b\bar{b}$, no fully matched NLO predictions exist for $t\bar{t} + c\bar{c}$ and $t\bar{t} + \text{light}$ events. A dedicated reweighting is therefore applied to the top-quark $p_T$ spectra as well as to the $p_T$ spectra of the $t\bar{t}$ system of $t\bar{t} + \text{light}$ and $t\bar{t} + c\bar{c}$ events in POWHEG+PYTHIA, based on the ratio of data to simulation of the measured differential cross sections at $\sqrt{s} = 7$ TeV \cite{50}. No such reweighting is applied to the $t\bar{t} + b\bar{b}$ sample, which is already corrected to match the best available theory calculation.

Samples of single-top-quark events produced in the $s$- and $Wt$-channels are generated with POWHEG-BOX 2.0 using the CT10NLO PDF set. The samples are interfaced to PYTHIA 6.425 with the CTEQ6L1 set of parton distribution functions and Perugia2011C underlying-event tune. The $t$-channel production mode is generated with ACERMC \cite{51} interfaced to PYTHIA 6.425 with the CTEQ6L1 PDF set and the Perugia2011C underlying-event tune. Overlaps between the $t\bar{t}$ and $Wt$ final states are removed \cite{52}. The single-top-quark samples are normalised to the approximate NNLO theoretical cross sections \cite{53, 54} using the MSTW2008 NNLO PDF set \cite{55, 56}.

The samples of $tV$ ($V = W, Z$) events are generated with the MadGraph v5 LO generator \cite{57} and the CTEQ6L1 PDF set. PYTHIA 6.425 with the AUET2B tune is used to generate the parton shower. The $tV$ samples are normalised to NLO cross-sections \cite{58, 59}.

Finally, event samples for single top quark plus Higgs boson production, $t\bar{H}q\bar{b}$ and $t\bar{H}W$, are generated. The cross sections are computed using the MG5\_AMC@NLO generator \cite{60} at NLO in QCD. For $t\bar{H}q\bar{b}$, samples are generated with MadGraph in the four-flavour scheme and $\mu_F = \mu_R = 75$ GeV then showered with PYTHIA 8.1 with the CTEQ6L1 PDF and the AU2 underlying-event tune. For $t\bar{H}W$, computed with the five-flavour scheme, dynamic $\mu_F$ and $\mu_R$ scales are used and events are generated at NLO with MG5\_AMC@NLO+HERWIG++ \cite{61, 62}. These two processes together are referred to as $t\bar{H}$.

A summary of the cross-section values and their uncertainties for the signal as well as for the MC simulated background processes is given in table 1.

5.3 Common treatment of MC samples

All samples using HERWIG are also interfaced to JIMMY v4.31 \cite{63} to simulate the underlying event. With the exception of SHERPA, all MC samples use PHOTOS 2.15 \cite{64} to simulate photon radiation and TAUOLA 1.20 \cite{65} to simulate $\tau$ decays. The samples are then processed through a simulation \cite{66} of the detector geometry and response using GEANT4 \cite{67}. The single-top-quark sample produced in the $t$-channel is simulated with a parameterised calorimeter response \cite{68}.

All simulated events are processed through the same reconstruction software as the data. Simulated events are corrected so that the lepton and jet identification efficiencies, energy scales and energy resolutions match those in data.

When selecting based on the output value of the $b$-tagging algorithm, the number of selected simulated events is significantly reduced, leading to large statistical fluctuations in the resulting distributions for samples with a high $b$-tag multiplicity. Therefore, rather than tagging the jets individually, the normalisation and the shape of these distributions are predicted by calculating the probability that a jet with a given flavour, $p_T$, and $\eta$ is
Table 1. Production cross sections for signal $t\bar{t}H$, at $m_H = 125$ GeV, and various simulated background processes. The quoted errors arise from variations of the renormalisation and factorisation scales and uncertainties in the parton distribution functions.

$b$-tagged [69]. The method is validated by verifying that the predictions reproduce the normalisation and shape obtained for a given working point of the $b$-tagging algorithm. The method is applied to all simulated signal and background samples.

5.4 Multijet background estimation using data: the TRF$_{MJ}$ method

A data-driven technique, the tag rate function for multijet events (TRF$_{MJ}$) method, is used to estimate the multijet background. After measuring $\varepsilon_{MJ}$, the probability of $b$-tagging a third jet in a sample of events with at least two $b$-tagged jets, the TRF$_{MJ}$ method uses $\varepsilon_{MJ}$ to extrapolate the multijet background from the regions with lower $b$-tag multiplicity to the search regions with higher $b$-tag multiplicity but otherwise identical event selection.

In the first step, the $b$-tagging rate is measured in data samples selected with various single-jet triggers, which are enriched in multijet events and have limited ($\approx 10\%$) overlap with the search region. The events in this TRF$_{MJ}$ extraction region are required to have at least three jets with $p_T > 25$ GeV and $|\eta| < 2.5$, with at least two $b$-tagged jets. Excluding the two jets with the highest $b$-tagging weight in the event, $\varepsilon_{MJ}$ is defined as the rate of $b$-tagging any other jet in the event. It is parameterised as a function of the jet $p_T$ and $\eta$, and also of the average $\Delta R$ between this jet and the two jets in the event with highest $b$-tagging weight, $\langle \Delta R_{(j, hMV1)} \rangle$. The $p_T$ and $\eta$ dependence of $\varepsilon_{MJ}$ reflects the corresponding sensitivity of the $b$-tagging efficiency to these variables. In multijet events, the $\Delta R$ dependence of $\varepsilon_{MJ}$ is correlated with the multi-$b$-jet production mechanism. This affects $\varepsilon_{MJ}$, shown in figure 1, which decreases by up to a factor two as $\Delta R$ increases for fixed $p_T$ and $\eta$.

In the search region the TRF$_{MJ}$ method starts from the data sample with exactly two $b$-tagged jets subtracting the contributions from all other backgrounds obtained from MC simulation. Multijet background samples containing $m$ jets ($m \geq 6$), out of which $n$ are
Figure 1. Dependence of $\varepsilon_{\text{MJ}}$ on the jet transverse momentum $p_T$, in regions of jet pseudorapidity $\eta$ and average $\Delta R$ between this jet and the two jets in the event with highest $b$-tagging weight, $\langle \Delta R_{(j,hMV1)} \rangle$. The $p_T$ bin boundaries are 25 (lowest), 40, 55, 70, 100, 200, 400, 600, 900 GeV (highest), chosen such as to have uniform number of events across bins of $\langle \Delta R_{(j,hMV1)} \rangle$.

$b$-tagged ($n \geq 3$) are then constructed, using an event weight $w(m_j,nb)$, which is calculated from $\varepsilon_{\text{MJ}}$ analogously to the method described in ref. [69], accounting for the fact that the starting sample contains two $b$-tagged jets. In each multijet event emulated using TRF$_{\text{MJ}}$ by means of $\varepsilon_{\text{MJ}}$, $(m-2)$ jets not originally $b$-tagged can be used for the emulation of the properties of additional $b$-tagged jets. This procedure allows to emulate observables that depend on the number of $b$-tagged jets.

5.5 Validation of the TRF$_{\text{MJ}}$ method in data and simulation

Validation of the TRF$_{\text{MJ}}$ method is performed by a ‘closure test’, separately in data and simulation. This is performed using the same data samples that were employed to estimate $\varepsilon_{\text{MJ}}$. In these low jet multiplicity samples, the TRF$_{\text{MJ}}$ method, which is applied to the events with exactly two $b$-tagged jets, is used to predict distributions in events with at least three $b$-tagged jets. Using $\varepsilon_{\text{MJ}}$ derived independently in data and simulation, the predicted distributions are compared to those resulting when directly applying $b$-tagging. This is done for a number of variables, such as $b$-tagged jet $p_T$, angular distance between $b$-tagged jets, and event shapes. As an example, for events with at least three jets and at least three $b$-tagged jets ($\geq3j$, $\geq3b$), figure 2 shows the closure test in data for the third-leading-jet $p_T$, $H_T$ (the scalar sum of the $p_T$ of all jets), and CentralityMass (defined as $H_T$ divided by the invariant mass of the jets). Figure 3 shows the results of the closure test in simulated multijet events for distributions of the leading-jet $p_T$, $m_{jj}^{\text{min}}$ (the minimum mass of all jet pairs in the event), and the third-leading $b$-tagged jet $p_T$. The definitions of these variables can be found in table 3. In both data and simulated multijet events with at least three $b$-tagged jets, the predicted and observed number of events agree within 5%. In events with a higher $b$-tagged jet multiplicity the numbers agree within the large statistical uncertainty. For this reason the systematic uncertainties related to the TRF$_{\text{MJ}}$ method are not estimated in the validation regions.
Figure 2. Comparison of the shapes predicted by the TRF$_{MJ}$ method (red histograms) and direct $b$-tagging (black circles) in data events with at least three jets and at least three $b$-tagged jets for (a) the third-leading $b$-tagged jet $p_T$, (b) $H_T$, and (c) Centrality$_{mass}$. The definitions of the variables are listed in table 3. Events were selected with various single-jet triggers. The TRF$_{MJ}$ prediction is normalised to the same number of events as the data. The uncertainty band for the TRF$_{MJ}$ predictions shown in the ratio plot represents statistical uncertainties only.

Figure 3. Comparison of the shapes predicted for the TRF$_{MJ}$ method (red histograms) and direct $b$-tagging (black circles) in Pythia 8.1 multijet events with at least three jets and at least three $b$-tagged jets for (a) leading-jet $p_T$, (b) $m_{jj}^{\text{min}}$ and (c) the third-leading $b$-tagged jet $p_T$ in the event. The definitions of the variables are listed in table 3. Distributions are normalised to the same area. The uncertainty band for the TRF$_{MJ}$ predictions shown in the ratio plot represents statistical uncertainties only.
6 Multijet trigger efficiency

Not all jets are reconstructed at the trigger level, mainly due to the Level-1 sliding window algorithm and the Level-1 resolution [70]. The multijet trigger efficiency with respect to the offline selection is derived in terms of the efficiency for a single jet to be associated with a complete jet trigger chain, i.e., a complete sequence of jets reconstructed at Level-1, Level-2 and EF satisfying the requirements described in section 4. This single-jet trigger efficiency, \( \epsilon_{\text{trig}} \), is evaluated in intervals of offline reconstructed \( p_T \) and \( \eta \):

\[
\epsilon_{\text{trig}}(p_T, \eta) = \frac{N_{\text{trig}}(p_T, \eta)}{N(p_T, \eta)},
\]

(6.1)

where \( N_{\text{trig}}(p_T, \eta) \) is the number of jets matched with a trigger chain and \( N(p_T, \eta) \) is the total number of jets within a given offline reconstructed \( p_T \) and \( \eta \) interval. Figure 4 shows that for large jet \( p_T \), \( \epsilon_{\text{trig}} \) reaches a plateau close to unity.

For both data and simulation, \( \epsilon_{\text{trig}}(p_T, \eta) \) is derived using events triggered by a single-jet trigger with a \( p_T \) threshold of 110 GeV, and only the offline jets which are in the hemisphere opposite to the trigger jet are used. To avoid additional trigger bias, events are discarded if more than one jet with \( p_T \geq 110 \) GeV is reconstructed. The ratio of \( \frac{\epsilon_{\text{data}}(p_T, \eta)}{\epsilon_{\text{MC,dijet}}(p_T, \eta)} \), where the latter is estimated in simulated dijet events, is referred to as SF\(_{\text{trig}}(p_T, \eta)\). In the analysis, for each MC sample considered, the final number of events passing the multijet trigger is estimated by weighting each jet by the product of \( \epsilon_{\text{trig}}\alpha(p_T, \eta) \) and SF\(_{\text{trig}}(p_T, \eta)\). The parameters \( \epsilon_{\text{trig}}(p_T, \eta) \) and SF\(_{\text{trig}}(p_T, \eta)\) are estimated for jet \( p_T \) up to 100 GeV. Figure 4 shows the \( p_T \) dependence of \( \epsilon_{\text{data}}(p_T, \eta), \epsilon_{\text{MC,tH}}(p_T, \eta), \epsilon_{\text{MC,dijet}}(p_T, \eta) \) and SF\(_{\text{trig}}(p_T, \eta)\) for jets within \( |\eta| < 2.5 \), together with the uncertainties from the difference between \( \epsilon_{\text{MC,tH}}(p_T, \eta) \) and \( \epsilon_{\text{MC,dijet}}(p_T, \eta) \), which is taken as the systematic uncertainty of the method.

7 Event classification

Six independent analysis regions are considered for the fit used in the analysis: two control regions (6j, 3b), (6j, \( \geq 4b \)) and four signal regions (7j, 3b), (7j, \( \geq 4b \)), (\( \geq 8j \), 3b) and (\( \geq 8j \), \( \geq 4b \)). In addition, the three regions with exactly two \( b \)-tagged jets, (6j, 2b), (7j, 2b) and (\( \geq 8j \), 2b), are used to predict the multijet contribution to higher \( b \)-tagging multiplicity regions, using the TRF\(_{\text{MJ}}\) method, as described above. The event yields in the different analysis regions prior to the fit are summarised in table 2.

The regions are analysed separately and combined statistically to maximise the overall sensitivity. The most sensitive regions, (\( \geq 8j \), 3b) and (\( \geq 8j \), \( \geq 4b \)), are expected to contribute more than 50% of the total significance.

8 Analysis method

The Toolkit for Multivariate Data Analysis (TMVA) [71] is used to train a BDT to separate the \( t\bar{t}H \) signal from the background. A dedicated BDT is defined and optimised in each of
Figure 4. Single-jet trigger efficiencies, $\epsilon_{\text{trig}}$, (top) for data, simulated dijet events, and $t\bar{t}H$ events, as a function of jet $p_T$ for jets with $|\eta| < 2.5$; (bottom) $S_{\text{trig}}(p_T, \eta) = \epsilon_{\text{trig}}^{\text{data}}(p_T, \eta)/\epsilon_{\text{trig}}^{\text{MC,dijet}}(p_T, \eta)$. The uncertainty on $S_{\text{trig}}$, shown as the green shaded area, is estimated from the difference between the efficiencies in dijet and $t\bar{t}H$ simulated events in the denominator of $S_{\text{trig}}$.

Table 2. Event yields from simulated backgrounds and the signal as well as data in each of the analysis regions prior to the fit (pre-fit). The quoted uncertainties are the sum in quadrature of the statistical and systematic uncertainties in the yields for all samples but the multijet background. The multijet normalisation and its systematic uncertainty are determined by the fit, so only its statistical uncertainty is quoted here. Since the numbers are rounded, the sum of all contributions may not equal the total value. The signal-to-background ratio, $S/B$, and the significance, $S/\sqrt{B}$, are also given. The $tH$ background is not shown as it amounts to fewer than 1.5 events in each region.
the six analysis regions. The variables entering the BDT and their definitions are listed in table 3.

The input variables include event-shape variables such as Centrality\textsubscript{Mass} and aplanarity, global event variables, such as $S_T$ (the modulus of the vector sum of the jet $p_T$), $H_{T5}$ (the scalar sum of the jet $p_T$ starting from the fifth jet in $p_T$ order), $m_{jj}^{\text{min}}$ (the smallest invariant mass of all dijet combinations), and the minimum $\Delta R$ between jets. The $p_T$ of the softest jet in the event is the only individual kinematic variable that enters the BDT directly. Other variables are calculated from pairs of objects: $\Delta R(b,b)_{\text{max}}$ (the $\Delta R$ between the two $b$-tagged jets with highest vector sum $p_T$), $m_{R(b,b)}$ (the invariant mass of the two $b$-tagged jets with the smallest $\Delta R$), $(E_{T1} + E_{T2})/\sum E_{T}^{\text{jets}}$ (the sum of the transverse energies of the two leading jets divided by the sum of the transverse energies of all jets), $m_{2\text{jets}}$ (the mass of the dijet pair, which, when combined with any $b$-tagged jet, maximises the magnitude of the vector sum of the $p_T$ of the three-jet system) and $m_{2b\text{jets}}$ (the invariant mass of the two $b$-tagged jets which are selected by requiring that the invariant mass of all the remaining jets is maximal). Two variables are calculated as the invariant mass of three jets: $m_{\text{top,1}}$ is computed from the three jets whose invariant mass is nearest to the top quark mass, taking into account the jet energy resolutions; the $m_{\text{top,2}}$ calculation uses the same algorithm but excludes the jets which enter $m_{\text{top,1}}$. Finally, a log-likelihood ratio variable, $\Lambda$, is used; it is related to the probability of an event to be a signal candidate, compared to the probability of being a background candidate.

The $\Lambda$ variable is the sum of the logarithms of ratios of relative probability densities for $W$ boson, top quark and Higgs boson resonances to be reconstructed in the event. For a given resonance $X$ decaying to two jets, the $\Lambda$ component is built as $\Lambda_X(m_{jj}) = \ln \frac{P_{\text{sig}}(m_{jj})}{P_{\text{bkg}}(m_{jj})}$ within a mass window $w_X = \pm 30$ GeV around the given particle mass:

$$P_{\text{sig}}(m_{jj}) = \begin{cases} s \cdot G(m_{jj}|m_X, \sigma_X), & \text{for } |m_{jj} - m_X| \leq w_X, \\ 1 - s, & \text{for } |m_{jj} - m_X| > w_X. \end{cases} \tag{8.1}$$

$$P_{\text{bkg}}(m_{jj}) = \begin{cases} b \cdot \text{Rect}(m_X, w_X), & \text{for } |m_{jj} - m_X| \leq w_X, \\ 1 - b, & \text{for } |m_{jj} - m_X| > w_X. \end{cases} \tag{8.2}$$

Here $s$ and $b$ are the probabilities to find a jet pair with an invariant mass within $\pm w_X$ of $m_X$. They are calculated from the signal simulation and from the multijet background respectively. The signal mass distribution is modelled with a Gaussian $G(m_{jj}|m_X, \sigma_X)$, while the background is modelled with a uniform distribution $\text{Rect}(m_X, w_X)$ between $m_X - w_X$ and $m_X + w_X$. Both functions $P_{\text{sig}}(m_{jj})$ and $P_{\text{bkg}}(m_{jj})$ are normalised to unity. For the top quark resonance the three-particle mass, $m_{jjb}$, is used. The width of the Gaussian is set to $\sigma_X = 18$ GeV for all resonances; this value corresponds to the expected experimental width of a Higgs boson with no combinatoric background.

The expression for the complete event $\Lambda$ is:

$$\Lambda(m_{jj}, m_{jjb}, m_{bb}) = \Lambda_W(m_{jj}|m_W, \sigma_X) + \Lambda_{\text{top}}(p_T,jj, m_{jjb}|m_{\text{top}}, \sigma_X) + \Lambda_H(p_T,bb, m_{bb}|m_H, \sigma_X). \tag{8.3}$$
<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>BDT rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centrality(J_{\text{Max}})</td>
<td>Scalar sum of the jet (p_T) divided by the invariant mass of the jets</td>
<td>1</td>
</tr>
<tr>
<td>Aplanarity</td>
<td>1.5(\lambda_2), where (\lambda_2) is the second eigenvalue of the momentum tensor built with all jets</td>
<td></td>
</tr>
<tr>
<td>(S_T)</td>
<td>The modulus of the vector sum of jet (p_T)</td>
<td></td>
</tr>
<tr>
<td>(H_{T,5})</td>
<td>Scalar sum of jet (p_T) starting from the fifth jet</td>
<td></td>
</tr>
<tr>
<td>(m_{\text{min}}^{\text{m}})</td>
<td>Smallest invariant mass of any combination of two jets</td>
<td>9</td>
</tr>
<tr>
<td>(\Delta R_{\text{min}})</td>
<td>Minimum (\Delta R) between two jets</td>
<td>6</td>
</tr>
<tr>
<td>(p_T^{\text{softest jet}})</td>
<td>(p_T) of the softest jet</td>
<td></td>
</tr>
<tr>
<td>(\Delta R(b,b)^{p_T^{\text{max}}})</td>
<td>(\Delta R) between two (b)-tagged jets with the largest vector sum (p_T)</td>
<td>11</td>
</tr>
<tr>
<td>(m_{\text{top}}^{\Delta R(b,b)^{p_T^{\text{max}}}})</td>
<td>Invariant mass of the combination of two (b)-tagged jets with the smallest (\Delta R)</td>
<td>3</td>
</tr>
<tr>
<td>(m_{\text{top}}^{E_T+\sum E_T})</td>
<td>Sum of the (E_T) of the two jets with leading (E_T) divided by the sum of the (E_T) of all jets</td>
<td>5</td>
</tr>
<tr>
<td>(m_{2\text{b-jets}})</td>
<td>The mass of the dijet pair, which, when combined with any (b)-tagged jet, maximises the magnitude of the vector sum of the (p_T) of the three-jet system</td>
<td>10</td>
</tr>
<tr>
<td>(m_{2\text{b-jets}})</td>
<td>The invariant mass of the two (b)-tagged jets which are selected by requiring that the invariant mass of all the remaining jets is maximal</td>
<td>12</td>
</tr>
<tr>
<td>(m_{\text{top},1})</td>
<td>Mass of the reconstructed top quark</td>
<td>13</td>
</tr>
<tr>
<td>(m_{\text{top},2})</td>
<td>Mass of the reconstructed top quark calculated from the jets not entering (m_{\text{top},1})</td>
<td>7</td>
</tr>
<tr>
<td>(\Lambda)</td>
<td>The logarithm of the ratio of event probabilities under the signal and background hypotheses</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3. List of variables used in the BDT in the six analysis regions. The numbers indicate the ranking of the corresponding variables, ordered by decreasing discriminating power. Variables not used in the BDT of a specific region are marked by a dash.
The three terms refer to $W$, top, and Higgs resonances respectively. For the top quark and Higgs boson resonances the masses, $m_{jjb}$ and $m_{bb}$, as well as the $p_T$, defined as the magnitude of the vector sum of the $p_T$ of the jets used to reconstruct the top quark, $p_{T,jj}$, and to reconstruct the Higgs boson, $p_{T,bb}$, are used. The value of $\Lambda$ is calculated for all possible jet combinations and the maximum $\Lambda$ of the event is chosen.

The variables entering the BDT are selected and ranked according to their separation power with an iterative procedure, which stops when adding more variables does not significantly improve the separation between signal and background. The cut-off corresponds to the point when adding a variable increases the significance, defined as $\sqrt{\sum_i S_i^2/B_i^2}$ where $S_i$ and $B_i$ are the expected signal and background yields in the $i^{th}$ bin of the BDT discriminant, by less than 1%.

Signal and background samples are classified as described in section 7, and then each subsample is further subdivided randomly into two subsamples of equal size for training and for testing.

The ranking of the input variables in terms of separation power for each analysis region is shown in table 3. The distributions of the BDT outputs for simulated signal and background events are shown in figure 5 for each analysis region. The figure shows a better separation between signal and background for low jet multiplicities than for high jet multiplicities. This is explained by the number of possible jet permutations. The number of jet permutations increases giving the background more configurations to mimic the signal.

9 Systematic uncertainties

The sources of systematic uncertainty considered in this analysis can be grouped into six main categories as summarised in table 4. Each systematic uncertainty is represented by an independent parameter, referred to as a nuisance parameter, and is parameterised with a Gaussian function for the shape uncertainties and a log-normal distribution for the normalisations [72]. They are centred around zero and one, respectively, with a width that corresponds to the given uncertainty. The uncertainties in the integrated luminosity, reconstruction of the physics objects, and the signal and background MC models are treated as in ref. [16]. The uncertainties related to the jet trigger as well as those related to the data-driven method to estimate the multijet background are discussed below. In total, 99 fit parameters are considered. The determination and treatment of the systematic uncertainties are detailed in this section. Their impact on the fitted signal strength is summarised in table 8 in section 11.

The systematic uncertainty in the luminosity for the data sample is 2.8%. It is derived following the same methodology as that detailed in ref. [73]. The trigger uncertainty is determined from the difference between $\epsilon_{\text{trig}}$, estimated using $t\bar{t}H$ and dijet MC events. Each jet in the event is weighted according to SF$_{\text{trig}}(p_T, \eta)$, the uncertainty of which is propagated to the shape and normalisation of the BDT output distribution, as shown in figure 6(a).

The uncertainties in physics objects are related to the reconstruction and $b$-tagging of jets. The jet energy resolution (JER) and the jet energy scale (JES) uncertainties are
<table>
<thead>
<tr>
<th>Systematic uncertainty source</th>
<th>Type</th>
<th>Number of components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>Trigger</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td><strong>Physics Objects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>SN</td>
<td>21</td>
</tr>
<tr>
<td>Jet vertex fraction</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td><em>b</em>-tagging efficiency</td>
<td>SN</td>
<td>7</td>
</tr>
<tr>
<td><em>c</em>-tagging efficiency</td>
<td>SN</td>
<td>4</td>
</tr>
<tr>
<td>Light-jet tagging efficiency</td>
<td>SN</td>
<td>12</td>
</tr>
<tr>
<td><strong>Background MC Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>t</em>+<em>t</em> cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td><em>t</em>+<em>t</em> modelling: <em>p_T</em> reweighting</td>
<td>SN</td>
<td>9</td>
</tr>
<tr>
<td><em>t</em>+<em>t</em> modelling: parton shower</td>
<td>SN</td>
<td>3</td>
</tr>
<tr>
<td><em>t</em>+<em>c</em>+<em>c</em>: heavy-flavour: normalisation</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td><em>t</em>+<em>c</em>+<em>c</em>: heavy-flavour reweighting</td>
<td>SN</td>
<td>2</td>
</tr>
<tr>
<td><em>t</em>+<em>c</em>+<em>c</em>: generator</td>
<td>SN</td>
<td>4</td>
</tr>
<tr>
<td><em>t</em>+<em>b</em>+<em>b</em>: NLO Shape</td>
<td>SN</td>
<td>8</td>
</tr>
<tr>
<td><em>t</em>+<em>V</em> cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td><em>t</em>+<em>V</em> modelling</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Single top cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td><strong>Data driven background</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multijet normalisation</td>
<td>N</td>
<td>6</td>
</tr>
<tr>
<td>Multijet TRF$_{MJ}$ parameterisation</td>
<td>S</td>
<td>6</td>
</tr>
<tr>
<td>Multijet $H_T$ correction</td>
<td>S</td>
<td>1</td>
</tr>
<tr>
<td>Multijet $S_T$ correction</td>
<td>S</td>
<td>1</td>
</tr>
<tr>
<td><strong>Signal Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>t</em>+$H$ scale</td>
<td>SN</td>
<td>2</td>
</tr>
<tr>
<td><em>t</em>+$H$ generator</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td><em>t</em>+$H$ hadronisation</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td><em>t</em>+$H$ parton shower</td>
<td>SN</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4. Sources of systematic uncertainty considered in the analysis grouped in six categories. “N” denotes uncertainties affecting only the normalisation for the relevant processes and channels, whereas “S” denotes uncertainties which are considered to affect only the shape of normalised distributions. “SN” denotes uncertainties affecting both shape and normalisation. Some sources of systematic uncertainty are split into several components. The number of components is also reported.
derived combining the information from test-beam data and simulation [25]. The JES uncertainties are split into 21 uncorrelated components. The largest of these uncertainties is due to the jet flavour composition. The JVF uncertainty is derived from $Z(\rightarrow \ell^+ \ell^-)+$ 1-jet events in data and simulation by varying the nominal cut value by 0.1 up and down.

The uncertainty related to the $b$-tagging is modelled with six independent parameters, while four parameters model the $c$-tagging uncertainty [26]. These are eigenvalues obtained by diagonalising the matrix which parameterises the tagging efficiency as a function of $p_T$, taking into account bin-to-bin correlations. Twelve parameters, which depend on $p_T$ and $\eta$, are used to parameterise the light-jet-tagging systematic uncertainties [74]. The per-jet $b$-tagging uncertainties are 3%–5%, about 10% for $c$-tagging and 20% for light jet tagging. An additional uncertainty is assigned to the $b$-tagging efficiency for jets with $p_T > 300$ GeV, which lacks statistics for an accurate calibration from data.

A combined uncertainty of ±6.0% is assigned to the $t\bar{t}$+jets production cross section, including modelling components due to the value of $\alpha_s$, the PDF used, the process energy scale, and the top quark mass. Other systematic uncertainties related to $t\bar{t}$+jets production are due to the modelling of parton showers and hadronisation.

Figure 5. Response of the BDT algorithm for simulated signal (dashed red), $t\bar{t}$+jets background (solid blue) and multijet background (dotted green) events in the (top) regions with 3 $b$-tags ((a) 6, (b) 7 and (c) $\geq 8$ jets) and in the (bottom) regions with $\geq 4$ $b$-tags ((d) 6, (e) 7 and (f) $\geq 8$ jets). The binning is the same as that used in the fit.
Figure 6. (a) Per event trigger scale factor $SF_{\text{trig}}$ (black dots) versus the BDT output of $t\bar{t}H$ events, shown with its corresponding systematic uncertainty (green band) for the $(\geq 8j, \geq 4b)$ region. (b) Comparison of the BDT output of the multijet background predicted with different sets of TRF$_{MJ}$. The nominal TRF$_{MJ}$ is represented by the red points. The bottom panel shows the ratios of the alternative TRF$_{MJ}$ predictions to the nominal set.

The systematic uncertainties arising from the reweighting procedure to improve $t\bar{t}$ background description by simulation (section 5.2), have been extensively studied in ref. [16] and adopted in this analysis. The largest uncertainties in the $t\bar{t}$ background description arise from radiation modelling, the choice of generator to simulate $t\bar{t}$ production, the JES, JER, and flavour modelling. These systematic uncertainties are applied to the $t\bar{t}$+light and $t\bar{t} + c\bar{c}$ components. Two additional systematic uncertainties, the full difference between applying and not applying the reweightings of the $t\bar{t}$ system $p_T$ and top quark $p_T$, are assigned to the $t\bar{t} + c\bar{c}$ component.

Four additional systematic uncertainties in the $t\bar{t} + c\bar{c}$ estimate are derived from the simultaneous variation of factorisation and renormalisation scales in Madgraph+Pythia. For the $t\bar{t}+bb$ background, three scale uncertainties are evaluated by varying the renormalisation and resummation scales. The shower recoil model uncertainty and two uncertainties due to the PDF choice in the Sherpa+OpenLoops NLO calculation are also taken into account.

The $t\bar{t}$+jets background is parameterised to allow a varying percentage of heavy flavours $c$ and $b$ in the additional jets not originating from the top quark decay products. An uncertainty of $\pm 50\%$ is assigned to the $t\bar{t} + bb$ and $t\bar{t} + c\bar{c}$ components of the $t\bar{t}$+jets cross section, which are treated as uncorrelated and are derived by comparing
Parameterisation variables in the TRF\textsubscript{MJ} method

<table>
<thead>
<tr>
<th>TRF\textsubscript{MJ} predictions</th>
<th>Parameterisation variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal set</td>
<td>$p_T,</td>
</tr>
<tr>
<td>Multijet set 1</td>
<td>$p_T, \Delta R_{MV1}, \Delta R_{(j,hMV1)}^{\text{min}}$</td>
</tr>
<tr>
<td>Multijet set 2</td>
<td>$p_T, \Delta R_{MV1}, \Delta R_{(j,j)}^{\text{min}}$</td>
</tr>
<tr>
<td>Multijet set 3</td>
<td>$p_T,</td>
</tr>
<tr>
<td>Multijet set 4</td>
<td>$p_T,</td>
</tr>
<tr>
<td>Multijet set 5</td>
<td>$p_T, \Delta R_{MV1}, \langle \Delta R_{(j,hMV1)} \rangle$</td>
</tr>
<tr>
<td>Multijet lowest MV1</td>
<td>Nominal set removing the two lowest MV1 jets from computation</td>
</tr>
<tr>
<td>Multijet random MV1</td>
<td>Nominal set removing randomly two MV1 jets from computation</td>
</tr>
<tr>
<td>Multijet HT RW</td>
<td>Nominal set with $H_T$ reweighting</td>
</tr>
<tr>
<td>Multijet ST RW</td>
<td>Nominal set with $S_T$ reweighting</td>
</tr>
</tbody>
</table>

**Table 5.** Alternative predictions of the multijet background with the TRF\textsubscript{MJ} method. Multijet sets 1 to 5 correspond to variations of the nominal set of variables describing $\varepsilon_{\text{MJ}}$. The next two sets specify the variation in the nominal set based on the two $b$-tagged jets which are used to compute $\varepsilon_{\text{MJ}}$. The last two refer to changes due to the residual mismodellings of $H_T$ and $S_T$. Each of these variations of the multijet background shape is quantified by one nuisance parameter in the fit.

$\text{POWHEG+PYTHIA}$ with a NLO result based on $\text{SHERPA}+\text{OPENLOOPS}$. The uncertainty in the $t\bar{t}+b\bar{b}$ contribution represents the dominant systematic effect in this analysis. An uncertainty of $\pm 30\%$ in the total cross section is assumed for $t\bar{t}+V$ [58, 59].

The multijet background is estimated using data in regions with exactly two $b$-tagged jets after subtraction of contributions from other events using MC simulation. All systematic uncertainties mentioned above are fully propagated to the data-driven multijet background estimation and treated in a correlated manner.

To estimate the uncertainties associated with the multijet background, the values of $\varepsilon_{\text{MJ}}$ are determined as a function of different sets of variables, listed in the first part of table 5, which are sensitive to the amount and the mechanism of heavy-flavour production. Alternative variables used are $\Delta R_{(j,j)}^{\text{min}}$, the minimum $\Delta R$ between the probed jet and any other jet in the event, $\Delta R_{(j,hMV1)}^{\text{min}}$, the minimum $\Delta R$ between the probed jet and the two jets with highest $b$-tag probability or $\langle \Delta R_{(j,hMV1)} \rangle$, its average value, and $\Delta R_{MV1}$, the $\Delta R$ between the two jets with the highest $b$-tag probability. In addition, different choices of methods to exclude $b$-tagged jets when determining $\varepsilon_{\text{MJ}}$ in the TRF\textsubscript{MJ} method are considered: the two $b$-tagged jets with the lowest MV1 weight or a random choice of two jets among all $b$-tagged jets in the event are chosen. The different sets of variables used to define $\varepsilon_{\text{MJ}}$ affect the shape of the BDT distribution for the multijet background, as shown in figure 6(b). Each of these shape variations is taken into account by a nuisance parameter in the fit. These parameterisations also affect the overall normalisation, with a maximum variation of 18\% in the 3-$b$-tag regions and 38\% in the $\geq 4$-$b$-tag regions. Residual mismodelling of $H_T$ and $S_T$ from the extraction region are also taken into account as systematic uncertainties. The normalisation of the multijet background is evaluated independently in each of the six analysis regions.
For the signal MC modelling, the PowHeL factorisation and renormalisation scales are varied independently by a factor two and 0.5. The kinematics of the MC simulated samples are then reweighted to reproduce the effects of these variations. The uncertainties related to the choice of PDFs are evaluated using the recommendations of PDF4LHC [75]. The systematic uncertainties from the parton shower and fragmentation models are evaluated using PowHeL+Herwig samples. The uncertainty due to the choice of generator is evaluated by comparing PowHeL+Pythia8 with MadGraph5_aMC@NLO+Herwig++.

10 Statistical methods

The binned distributions of the BDT output discriminants for each of the six analysis regions are combined as inputs to a test statistic to search for the presence of a signal. The analysis uses a maximum-likelihood fit [72] to measure the compatibility of the observed data with the background-only hypothesis, i.e., $\mu = 0$, and to make statistical inferences about $\mu$, such as upper limits, using the CLs method [76, 77] as implemented in the RooFit package [78].

A fit is performed under the signal-plus-background hypothesis to obtain the value of the signal strength, assuming a SM Higgs boson mass of $m_H = 125$ GeV. The value of $\mu$ is a free parameter in the fit. The normalisation of each component of the background and $\mu$ are determined simultaneously from the fit. Contributions from $t\bar{t}$+jets, $t\bar{t} + V$ and single-top-quark backgrounds are constrained by the uncertainties of the respective theoretical calculations, the uncertainty in the luminosity, and experimental data. The multijet background normalisations are free parameters in the fit and are independent in each region. The performance of the fit is validated using simulated events by injecting a signal with variable strength and comparing the known strength to the fitted value.

11 Results

The yields in the different analysis regions considered in the analysis after the fit (post-fit) are summarised in table 6. In each region, the variation of background and signal events with respect to the pre-fit values (cf. table 2) are modest and, in particular, the fitted multijet background component is well constrained by the fit within an uncertainty of 8%.

Figures 7 and 8 show the BDT output distributions for data and the predictions in each analysis region, both before (left panels) and after (right panels) the fit to data. The relative uncertainties decrease significantly in all regions due to the constraints provided by the data, exploiting the correlations between the uncertainties in the different analysis regions.

The signal strength in the all-hadronic $t\bar{t}H$ decay mode, for $m_H = 125$ GeV, is measured to be:

$$\mu(m_H=125 \text{ GeV}) = 1.6 \pm 2.6.$$  \hspace{1cm} (11.1)

The expected uncertainty in the signal strength ($\mu = 1$) is $\pm 2.8$. The observed (expected) significance of the signal is 0.6 (0.4) standard deviations, corresponding to an observed (expected) $p$-value of 27% (34%), where the $p$-value is the probability to obtain a result at least as signal-like as observed if no signal were present.
Figure 7. Comparison between data and prediction for the BDT discriminant in the, from top to bottom, (6-8j, 3b) regions before (left) and after (right) the fit. The fit is performed under the signal-plus-background hypothesis. Pre-fit plots show an overlay of the multijet distribution normalised to data for illustration purposes only. The bottom panels display the ratios of data to the total prediction. The hashed areas represent the total uncertainty in the background predictions. The $ttH$ signal yield (solid red) is scaled by a fixed factor before the fit.
Figure 8. Comparison between data and prediction for the BDT discriminant in the, from top to bottom, (6-8j, ≥4b) regions before (left) and after (right) the fit. The fit is performed under the signal-plus-background hypothesis. Pre-fit plots show an overlay of the multijet distribution normalised to data for illustration purposes only. The bottom panels display the ratios of data to the total prediction. The hashed areas represent the total uncertainty in the background predictions. The $ttH$ signal yield (solid red) is scaled by a fixed factor before the fit.
Table 6. Event yields from simulated backgrounds and the signal as well as measured events in each of the analysis regions after the fit. The quoted uncertainties include statistical and systematic effects. The sum of all contributions may slightly differ from the total value due to rounding. The $tH$ background is not shown as fewer than 1.5 events in each region are predicted.

<table>
<thead>
<tr>
<th>Region</th>
<th>$6j, 3b$</th>
<th>$6j, \geq 4b$</th>
<th>$7j, 3b$</th>
<th>$7j, \geq 4b$</th>
<th>$\geq 8j, 3b$</th>
<th>$\geq 8j, \geq 4b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>$15940 \pm 320$</td>
<td>$1423 \pm 66$</td>
<td>$12060 \pm 350$</td>
<td>$1233 \pm 78$</td>
<td>$10020 \pm 490$</td>
<td>$1280 \pm 100$</td>
</tr>
<tr>
<td>$t\bar{t}$+light</td>
<td>$1750 \pm 270$</td>
<td>$55 \pm 13$</td>
<td>$1650 \pm 340$</td>
<td>$54 \pm 15$</td>
<td>$1550 \pm 450$</td>
<td>$54 \pm 21$</td>
</tr>
<tr>
<td>$t\bar{t}$+c</td>
<td>$350 \pm 170$</td>
<td>$22 \pm 11$</td>
<td>$490 \pm 240$</td>
<td>$28 \pm 14$</td>
<td>$750 \pm 360$</td>
<td>$66 \pm 33$</td>
</tr>
<tr>
<td>$t\bar{t}$+b</td>
<td>$230 \pm 120$</td>
<td>$31 \pm 17$</td>
<td>$350 \pm 190$</td>
<td>$63 \pm 34$</td>
<td>$560 \pm 320$</td>
<td>$139 \pm 75$</td>
</tr>
<tr>
<td>$t\bar{t}$+V</td>
<td>$15.0 \pm 6.2$</td>
<td>$1.9 \pm 1.5$</td>
<td>$23.3 \pm 8.9$</td>
<td>$3.6 \pm 2.2$</td>
<td>$43 \pm 15$</td>
<td>$8.7 \pm 4.2$</td>
</tr>
<tr>
<td>Single top</td>
<td>$184 \pm 59$</td>
<td>$6.7 \pm 3.6$</td>
<td>$153 \pm 52$</td>
<td>$9.4 \pm 4.4$</td>
<td>$123 \pm 48$</td>
<td>$11.8 \pm 5.8$</td>
</tr>
<tr>
<td>Total background</td>
<td>$18470 \pm 320$</td>
<td>$1539 \pm 58$</td>
<td>$14720 \pm 320$</td>
<td>$1391 \pm 69$</td>
<td>$13030 \pm 340$</td>
<td>$1561 \pm 63$</td>
</tr>
<tr>
<td>$t\bar{t}H$ ($m_H=125$ GeV)</td>
<td>$23.4 \pm 6.3$</td>
<td>$5.6 \pm 2.8$</td>
<td>$39.1 \pm 8.9$</td>
<td>$11.9 \pm 4.5$</td>
<td>$71 \pm 15$</td>
<td>$28.8 \pm 8.5$</td>
</tr>
</tbody>
</table>

Table 7. Observed and expected upper limits at 95% CL on $\sigma(t\bar{t}H)$ relative to the SM prediction assuming $m_H = 125$ GeV, for the background-only hypothesis. Confidence intervals around the expected limits under the background-only hypothesis are also provided, denoted by $\pm1\sigma$ and $\pm2\sigma$, respectively. The expected (median) upper limit at 95% CL assuming the SM prediction for $\sigma(t\bar{t}H)$ is shown in the last column.

<table>
<thead>
<tr>
<th>Region</th>
<th>Observed</th>
<th>Expected if $\mu = 0$</th>
<th>Expected if $\mu = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper limit on $\mu$ at 95%</td>
<td>6.4</td>
<td>$-2\sigma$</td>
<td>$-1\sigma$</td>
</tr>
</tbody>
</table>

The observed and expected limits are summarised in table 7. A $t\bar{t}H$ signal 6.4 times larger than predicted by the SM is excluded at 95% CL. A signal 5.4 times larger than the signal of a SM Higgs boson is expected to be excluded for the background-only hypothesis.

Figure 9 summarises the post-fit event yields for data, total background and signal expectations as a function of $\log_{10}(S/B)$. The signal is normalised to the fitted value of the signal strength ($\mu = 1.6$). A signal strength 6.4 times larger than predicted by the SM is also shown in figure 9.

Figure 10 shows the effect of the major systematic uncertainties on the fitted value of $\mu$ and the constraints provided by the data. The ranking, from top to bottom, is determined by the post-fit impact on $\mu$. This effect is calculated by fixing the corresponding nuisance parameter at $\hat{\theta} \pm 3\sigma$ and performing the fit again. Here $\hat{\theta}$ is the fitted value of the nuisance parameter and $\sigma$ is its post-fit uncertainty. The difference between the default and the modified $\mu$, $\Delta\mu$, represents the effect on $\mu$ of this particular systematic uncertainty. This is also shown in table 8.

The largest effect arises from the uncertainty in the normalisation of the irreducible $t\bar{t} + b\bar{b}$ background. The $t\bar{t} + b\bar{b}$ background normalisation is smaller by 30% in the fit than the prediction, resulting in a decrease of the observed $t\bar{t} + b\bar{b}$ yield with respect to the POWHEG+PYTHIA prediction. The second largest effect comes from the multijet background normalisation. The data-driven method focuses on modelling the shape of the
Sources of systematic uncertainty | ±1σ post-fit impact on μ
---|---
tt normalisation | 108%
Multijet normalisation | 71%
Multijet shape | 60%
Main contributions from t\(\bar{t}\) modelling | 34%–41%
Flavour tagging | 31%
Jet energy scale | 27%
Signal modelling | 22%
Luminosity+trigger+JVF+JER | 18%

Table 8. Effect of the different sources of systematic uncertainties on μ expressed in terms of percentage of the fitted value of μ sorted according to their post-fit effect.

Figure 9. Event yields as a function of \(\log_{10}(S/B)\), where \(S\) (expected signal yield) and \(B\) (expected background yield) are taken from the corresponding BDT discriminant bin. Events from all fitted regions are included. The predicted background is obtained from the global signal-plus-background fit. The \(t\bar{t}H\) signal is shown both for the best-fit value (μ = 1.6) and for the upper limit at 95% CL (μ = 6.4).

multijet background while the normalisation is constrained by the regions dominated by multijet background. The uncertainty in the normalisation parameters amounts to few percent and the values from each region are consistent with the variations applied to these parameters to account for systematic uncertainties. Two of the multijet background shape uncertainties are ranked fourth and fifth, and their pulls are slightly positive.

Other important uncertainties include \(b\)-tagging and JES. Uncertainties arising from jet energy resolution, jet vertex fraction, jet reconstruction and JES that affect primarily low-\(p_T\) jets, as well as the \(t\bar{t}+\text{light-jet}\) background modelling uncertainties, do not have a significant impact on the result.
Figure 10. The fitted values of the 20 nuisance parameters corresponding to the sources of systematic uncertainty with the largest impact on the fitted signal strength $\mu$. The points, which are drawn conforming to the scale of the bottom axis, show the deviation of each of the fitted nuisance parameters $\theta$ from $\theta_0$, which is the nominal value of that nuisance parameter, in units of the pre-fit standard deviation $\Delta \theta$. The plain yellow area represents the pre-fit impact on $\sigma_1$ and the hashed blue area its post-fit impact. The error bars show the post-fit uncertainties $\sigma_\theta$, which have size close to one if the data do not provide any further constraint on that uncertainty. Conversely, an error bar for $\sigma_\theta$ smaller than one indicates a reduction with respect to the original uncertainty. The nuisance parameters are sorted according to their post-fit impact $\Delta \theta$ (top horizontal scale). Multijet scale factors (SF) show the fitted values and uncertainties of the normalisation parameters that are freely floating in the fit. These normalisation parameters have a pre-fit value of unity.

12 Combination of $t\bar{t}H$ results at $\sqrt{s} = 7$ and 8 TeV

The sensitivity of the search for $t\bar{t}H$ production can be increased by statistically combining different Higgs boson decay channels. This combination is described in the following.

12.1 Individual $t\bar{t}H$ measurements and results

The $t\bar{t}H$ searches that are combined are:

- $t\bar{t}H(H \rightarrow b\bar{b})$ in the single-lepton and opposite-charge dilepton $t\bar{t}$ decay channels using data at $\sqrt{s} = 8$ TeV [16],
• $t\bar{t}H(H \to b\bar{b})$ in the all-hadronic $t\bar{t}$ decay channel using data at $\sqrt{s} = 8$ TeV as presented in this paper,

• $t\bar{t}H(H \to (WW^{(*)}, \tau\tau, ZZ^{(*)}) \to$ leptons) with two same-charge leptons ($e$ or $\mu$), three leptons, four leptons, two hadronically decaying $\tau$ leptons plus one lepton and one hadronically decaying $\tau$ lepton plus two leptons in the final state using data at $\sqrt{s} = 8$ TeV [15],

• $t\bar{t}H(H \to \gamma\gamma)$ at $\sqrt{s} = 7$ and 8 TeV in both the hadronic and leptonic ($e$ or $\mu$) $t\bar{t}$ pair decay channels [17].

First all $H \to b\bar{b}$ final states are combined, obtaining a signal strength for the $t\bar{t}H(H \to b\bar{b})$ combination, and then the outcome is combined with the remaining (non-$H \to b\bar{b}$) channels.

12.1.1 $H \to b\bar{b}$ (single lepton and dilepton $t\bar{t}$ decays)

The search for $t\bar{t}H$ production with $H \to b\bar{b}$ is performed in both the single-lepton and dilepton $t\bar{t}$ decay modes [16]. The single-lepton analysis requires one charged lepton with at least four jets, of which at least two need to be $b$-tagged, while the dilepton analysis requires two opposite-charge leptons with at least two jets, of which at least two must be $b$-tagged. The events are then categorised according to the jet and $b$-tagged jet multiplicity. The dominant background in the signal-enriched regions is from $t\bar{t} + b\bar{b}$ events. In these regions, neural networks [79] are built using kinematic information in order to separate the $t\bar{t}H$ signal from $t\bar{t}$ background. Furthermore, in the single-lepton channel, a matrix-element discriminant is built in the most signal-enriched regions and is used as an input to the neural network.

12.1.2 $H \to (WW^{(*)}, \tau\tau, ZZ^{(*)}) \to$ leptons

The $ttH$ search with $H \to (WW^{(*)}, \tau\tau, ZZ^{(*)}) \to$ leptons [15] exploits several multilepton signatures resulting from Higgs boson decays to vector bosons and/or $\tau$ leptons. Events are categorised based on the number of charged leptons and/or hadronically decaying $\tau$ leptons in the final state. The categorisation includes events with two same-charge leptons, three leptons, four leptons, one lepton and two hadronic $\tau$ leptons, as well as two same-charge leptons with one hadronically decaying $\tau$ lepton. Backgrounds include events with electron charge misidentification, which are estimated using data-driven techniques, non-prompt leptons arising from semileptonic $b$-hadron decays, mostly from $t\bar{t}$ events, again estimated from data-driven techniques, and production of $t\bar{t} + W$ and $t\bar{t} + Z$, which are estimated using MC simulations. Signal and background event yields are obtained from a simultaneous fit to all channels.

12.1.3 $H \to \gamma\gamma$

The $t\bar{t}H$ search in the $H \to \gamma\gamma$ channel [17] exploits the sharp peak in the diphoton mass distribution from the $H \to \gamma\gamma$ decay over the continuum background. The analysis is split according to the decay mode of the $t\bar{t}$ pair. A leptonic selection requires at least one
lepton and at least one $b$-tagged jet, and missing transverse momentum if there is only one $b$-tagged jet, whereas a hadronic selection requires a combination of jets and $b$-tagged jets. Contributions from peaking non-$t \bar{t}H$ Higgs boson production modes are estimated from MC simulations. The signal is extracted with a fit using the diphoton mass distribution as a discriminant.

12.2 Correlations

Nuisance parameters corresponding to the same source of uncertainty in different analyses are generally considered to be correlated with each other, except for the following sets:

- Nuisance parameters related to $b$-tagging (also $c$-tagging and light mis-tagging) are considered to be independent among the analyses as different $b$-tagging working points are employed.
- The electron identification uncertainty is considered to be uncorrelated between analyses due to different selections used.

12.3 Results of the combination

12.3.1 Signal strength

The result of the $t \bar{t}H(H \rightarrow b \bar{b})$ combination for the signal strength is $\mu = 1.4 \pm 1.0$. The observed signal strengths for the individual $t \bar{t}H(H \rightarrow b \bar{b})$ channels and for their combination are summarised in figure 11. The $t \bar{t} + b \bar{b}$ normalisation nuisance parameters obtained in the all-hadronic analysis ($-0.6 \pm 0.8$) and the lepton analysis ($+0.8 \pm 0.4$) The expected significance increases from 1.0$\sigma$ for the leptonic final state of $t \bar{t}H(H \rightarrow b \bar{b})$ to 1.1$\sigma$ for the combined $t \bar{t}H(H \rightarrow b \bar{b})$. Because the combined $t \bar{t}H(H \rightarrow b \bar{b})$ best-fit value of $\mu$ is lower than the leptonic-only value, the observed significance for the $t \bar{t}H(H \rightarrow b \bar{b})$ combination is reduced from 1.4$\sigma$ (leptonic [16]) to 1.35$\sigma$ (combined).

Figure 12 summarises the observed signal strength $\mu$ of the individual $t \bar{t}H$ channels ($H \rightarrow b \bar{b}$, $H \rightarrow \gamma \gamma$ and $H \rightarrow (WW^{(*)}, \tau \tau, ZZ^{(*)}) \rightarrow$ leptons) and the $t \bar{t}H$ combination. The observed (expected) significance of the combined $t \bar{t}H$ result is 2.33$\sigma$ (1.53$\sigma$).

The combination of all $t \bar{t}H$ analyses yields an observed (expected) 95% CL upper limit of 3.1 (1.4) times the SM cross section. The observed 95% CL limits for the individual $t \bar{t}H$ channels and for the combination are shown in figure 13 and in table 9.

The result for the best-fit value is $\mu = 1.7 \pm 0.8$.

12.3.2 Couplings

Sensitivity to $t - H$ and $W - H$ couplings stems from several sources: from the $t \bar{t}H$ production itself, from the Higgs boson decay branching fractions, from associated single top and Higgs boson production processes ($tHj b$ and $WtH$), where interference terms include both the $t \bar{t}H$ and $WWH$ vertices, and from the $H \rightarrow \gamma \gamma$ branching fraction, where again interferences between loop contributions from the top quark and the W boson are present. Different channels differ in their sensitivity to these components. A two-parameter
Figure 11. Summary of the measurements of the signal strength $\mu$ for $t\bar{t}H(H \rightarrow b\bar{b})$ production for the individual $H \rightarrow b\bar{b}$ channels and for their combination, assuming $m_H = 125$ GeV. The total (tot) and statistical (stat) uncertainties of $\mu$ are shown. The SM $\mu = 1$ expectation is shown as the grey line.

Figure 12. Summary of the measurements of the signal strength $\mu$ for the individual channels and for their combination, assuming $m_H = 125$ GeV. The total (tot) and statistical (stat) uncertainties of $\mu$ are shown. The SM $\mu = 1$ expectation is shown as the grey line.

fit is performed, assuming that all boson couplings scale with the same modifier $\kappa_V$, while all fermion couplings scale with the same modifier $\kappa_F$.

The parameterisation of the couplings for the $t\bar{t}H$ and $tH$ production modes and for the different Higgs boson decay modes is taken from refs. [7, 80]. Figure 14 shows the log-likelihood contours of $\kappa_F$ versus $\kappa_V$ for the combined $t\bar{t}H$ fit. The combination of all analysis channels slightly prefers positive $\kappa_F$. Additional studies, performed to determine the contribution of the individual analyses to the combined coupling measurement, indicate that the $t\bar{t}H$, $H \rightarrow \ell\ell$ $\rightarrow (WW^{(*)}, \tau\tau, ZZ^{(*)}) \rightarrow$ leptons analysis prefers somewhat enhanced $W$—
Figure 13. Upper limits on the signal strength $\mu$ for the individual channels as well as for their combination, at 95% CL. The observed limits (solid lines) are compared to the expected median limits under the background-only hypothesis (black dashed lines) and under the signal-plus-background hypothesis assuming the SM prediction for $\sigma(t\bar{t}H)$ (red dotted lines). The surrounding green and yellow bands correspond to the $\pm 1\sigma$ and $\pm 2\sigma$ ranges around the expected limits under the background-only hypothesis.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>95% CL upper limit</th>
<th>Signal strength</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>$-2\sigma$</td>
</tr>
<tr>
<td>$t\bar{t}H(H \rightarrow \gamma\gamma)$</td>
<td>6.7</td>
<td>2.6</td>
</tr>
<tr>
<td>$t\bar{t}H(H \rightarrow$ leptons)</td>
<td>4.7</td>
<td>1.3</td>
</tr>
<tr>
<td>$t\bar{t}H(H \rightarrow b\bar{b})$</td>
<td>3.3</td>
<td>1.3</td>
</tr>
<tr>
<td>$t\bar{t}H$ Combination</td>
<td>3.1</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 9. Observed and expected (median, for the background-only hypothesis) upper limits at 95% CL on $\sigma(t\bar{t}H)$ relative to the SM prediction, for the individual channels as well as for their combination. The $\pm 1\sigma$ and $\pm 2\sigma$ ranges around the expected limit are also given. The expected median upper limits at 95% CL assuming the SM prediction for $\sigma(t\bar{t}H)$ are shown in the last column.

The $H$ coupling, which can only be compatible with the $t\bar{t}H(H \rightarrow \gamma\gamma)$ rate if the interference between $t\bar{t}H$ and $WWH$ amplitudes is destructive, as expected in the SM.

13 Conclusion

A search for the SM Higgs boson produced in association with a pair of top quarks ($t\bar{t}H$) has been carried out with the ATLAS detector at the Large Hadron Collider. The search focuses on $H \rightarrow b\bar{b}$ decays with $t\bar{t}$ pairs decaying hadronically. The data used correspond to an integrated luminosity of 20.3 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 8$ TeV. The analysis is carried out in six different jet and $b$-tagged jet multiplicity regions. Discrimination between signal and background is obtained by employing a boosted decision tree multivariate classifier in all regions. No significant excess of events above the background expectation is found.
for the SM Higgs boson with a mass of 125 GeV. An observed (expected) 95% CL upper limit of 6.4 (5.4) times the SM cross section is obtained. By performing a fit under the signal-plus-background hypothesis, the ratio of the measured signal strength to the SM expectation is found to be $\mu = 1.6 \pm 2.6$.

The statistical combination of all $t\bar{t}H$ analyses performed at $\sqrt{s} = 7$ TeV and 8 TeV yields an observed (expected) upper limit of 3.1 (1.4) times the SM cross section at 95% CL. The combined measured signal strength is found to be $\mu = 1.7 \pm 0.8$.

**Acknowledgments**

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; NWO, Netherlands; RCN, Norway; MNIsW and NCN, Poland; FCT, Portugal; MINE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESSTD, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC,
United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; the Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.) and in the Tier-2 facilities worldwide.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


1 Department of Physics, University of Adelaide, Adelaide, Australia
2 Physics Department, SUNY Albany, Albany NY, United States of America
3 Department of Physics, University of Alberta, Edmonton AB, Canada
4 (a) Department of Physics, Ankara University, Ankara; (b) Istanbul Aydin University, Istanbul; (c) Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey
5 LAPP, CNRS/IN2P3 and Université Savoie Mont Blanc, Annecy-le-Vieux, France
6 High Energy Physics Division, Argonne National Laboratory, Argonne IL, United States of America
7 Department of Physics, University of Arizona, Tucson AZ, United States of America
8 Department of Physics, The University of Texas at Arlington, Arlington TX, United States of America
9 Physics Department, University of Athens, Athens, Greece
10 Physics Department, National Technical University of Athens, Zografou, Greece
11 Department of Physics, The University of Texas at Austin, Austin TX, United States of America
12 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
13 Institut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Barcelona, Spain
14 Institute of Physics, University of Belgrade, Belgrade, Serbia
15 Department for Physics and Technology, University of Bergen, Bergen, Norway
Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, CA, United States of America
17 Department of Physics, Humboldt University, Berlin, Germany
18 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
19 School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
20 (a) Department of Physics, Bogazici University, Istanbul; (b) Department of Physics Engineering, Gaziantep University, Gaziantep; (c) Istanbul Bilgi University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey; (d) Bahcesehir University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey, Turkey
21 Centro de Investigaciones, Universidad Antonio Narino, Bogota, Colombia
22 (a) INFN Sezione di Bologna; (b) Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy
23 Physikalisches Institut, University of Bonn, Bonn, Germany
24 Department of Physics, Boston University, Boston MA, United States of America
25 Department of Physics, Brandeis University, Waltham MA, United States of America
26 (a) Universidade Federal do Rio de Janeiro COPPE/EE/IF, Rio de Janeiro; (b) Electrical Circuits Department, Federal University of Juiz de Fora (UFJF), Juiz de Fora; (c) Federal University of Sao Joao del Rei (UFSJ), Sao Joao del Rei; (d) Instituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil
27 Physics Department, Brookhaven National Laboratory, Upton NY, United States of America
28 (a) Transilvania University of Brasov, Brasov, Romania; (b) National Institute of Physics and Nuclear Engineering, Bucharest; (c) National Institute for Research and Development of Isotopic and Molecular Technologies, Physics Department, Cluj Napoca; (d) University Politehnica Bucharest, Bucharest; (e) West University in Timisoara, Timisoara, Romania
29 Departamento de Fisica, Universidad de Buenos Aires, Buenos Aires, Argentina
30 Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
31 Department of Physics, Carleton University, Ottawa ON, Canada
32 CERN, Geneva, Switzerland
33 Enrico Fermi Institute, University of Chicago, Chicago IL, United States of America
34 (a) Departamento de Fisica, Pontificia Universidad Catolica de Chile, Santiago; (b) Departamento de Fisica, Universidad Técnica Federico Santa María, Valparaiso, Chile
35 (a) Institute of High Energy Physics, Chinese Academy of Sciences, Beijing; (b) Department of Modern Physics, University of Science and Technology of China, Anhui; (c) Department of Physics, Nanjing University, Jiangsu; (d) School of Physics, Shandong University, Shandong; (e) Department of Physics and Astronomy, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University, Shanghai; (also affiliated with PKU-CHEP); (f) Physics Department, Tsinghua University, Beijing 100084, China
36 Laboratoire de Physique Corpusculaire, Clermont Université and Université Blaise Pascal and CNRS/IN2P3, Clermont-Ferrand, France
37 Nevis Laboratory, Columbia University, Irvington NY, United States of America
38 Niels Bohr Institute, University of Copenhagen, Kobenhavn, Denmark
39 (a) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati; (b) Dipartimento di Fisica, Università della Calabria, Rende, Italy
40 (a) AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Krakow; (b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow, Poland
41 Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland
42 Physics Department, Southern Methodist University, Dallas TX, United States of America
43 Physics Department, University of Texas at Dallas, Richardson TX, United States of America
44 DESY, Hamburg and Zeuthen, Germany
45 Institut für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany
46 Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany
Department of Physics, Duke University, Durham NC, United States of America
SUPA - School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
INFN Laboratori Nazionali di Frascati, Frascati, Italy
Fakultät für Mathematik und Physik, Albert-Ludwigs-Universität, Freiburg, Germany
Section de Physique, Université de Genève, Geneva, Switzerland
(a) INFN Sezione di Genova; (b) Dipartimento di Fisica, Università di Genova, Genova, Italy
E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi; (b) High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia
II Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany
SUPA - School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
II Physikalisches Institut, Georg-August-Universität, Göttingen, Germany
Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, Grenoble, France
Department of Physics, Hampton University, Hampton VA, United States of America
Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge MA, United States of America
(a) Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg; (b) Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg; (c) ZITI Institut für technische Informatik, Ruprecht-Karls-Universität Heidelberg, Mannheim, Germany
Faculty of Applied Information Science, Hiroshima Institute of Technology, Hiroshima, Japan
(a) Department of Physics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong; (b) Department of Physics, The University of Hong Kong, Hong Kong; (c) Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
Department of Physics, Indiana University, Bloomington IN, United States of America
Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck, Austria
University of Iowa, Iowa City IA, United States of America
Department of Physics and Astronomy, Iowa State University, Ames IA, United States of America
Joint Institute for Nuclear Research, JINR Dubna, Dubna, Russia
KEK, High Energy Accelerator Research Organization, Tsukuba, Japan
Graduate School of Science, Kobe University, Kobe, Japan
Faculty of Science, Kyoto University, Kyoto, Japan
Department of Physics, Kyushu University, Fukuoka, Japan
Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina
Physics Department, Lancaster University, Lancaster, United Kingdom
(a) INFN Sezione di Lecce; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy
Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
Department of Physics, Jozef Stefan Institute and University of Ljubljana, Ljubljana, Slovenia
School of Physics and Astronomy, Queen Mary University of London, London, United Kingdom
Department of Physics, Royal Holloway University of London, Surrey, United Kingdom
Department of Physics and Astronomy, University College London, London, United Kingdom
Louisiana Tech University, Ruston LA, United States of America
Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université Paris-Diderot and CNRS/IN2P3, Paris, France
Fysiska institutionen, Lunds universitet, Lund, Sweden
Departamento de Fisica Teorica C-15, Universidad Autonoma de Madrid, Madrid, Spain
Institut für Physik, Universität Mainz, Mainz, Germany
School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
Department of Physics, University of Massachusetts, Amherst MA, United States of America
Department of Physics, McGill University, Montreal QC, Canada
90 School of Physics, University of Melbourne, Victoria, Australia
91 Department of Physics, The University of Michigan, Ann Arbor MI, United States of America
92 Department of Physics and Astronomy, Michigan State University, East Lansing MI, United States of America
93 (a) INFN Sezione di Milano; (b) Dipartimento di Fisica, Università di Milano, Milano, Italy
94 B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Republic of Belarus
95 National Scientific and Educational Centre for Particle and High Energy Physics, Minsk, Republic of Belarus
96 Group of Particle Physics, University of Montreal, Montreal QC, Canada
97 P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, Russia
98 Institute for Theoretical and Experimental Physics (ITEP), Moscow, Russia
99 National Research Nuclear University MEPhI, Moscow, Russia
100 D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
101 Fakultät für Physik, Ludwig-Maximilians-Universität München, München, Germany
102 Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany
103 Nagasaki Institute of Applied Science, Nagasaki, Japan
104 Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan
105 (a) INFN Sezione di Napoli; (b) Dipartimento di Fisica, Università di Napoli, Napoli, Italy
106 Department of Physics and Astronomy, University of New Mexico, Albuquerque NM, United States of America
107 Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, Netherlands
108 Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, Netherlands
109 Department of Physics, Northern Illinois University, DeKalb IL, United States of America
110 Budker Institute of Nuclear Physics, SB RAS, Novosibirsk, Russia
111 Department of Physics, New York University, New York NY, United States of America
112 Ohio State University, Columbus OH, United States of America
113 Faculty of Science, Okayama University, Okayama, Japan
114 Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman OK, United States of America
115 Department of Physics, Oklahoma State University, Stillwater OK, United States of America
116 Palacký University, RCPTM, Olomouc, Czech Republic
117 Center for High Energy Physics, University of Oregon, Eugene OR, United States of America
118 LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
119 Graduate School of Science, Osaka University, Osaka, Japan
120 Department of Physics, University of Oslo, Oslo, Norway
121 Department of Physics, Oxford University, Oxford, United Kingdom
122 (a) INFN Sezione di Pavia; (b) Dipartimento di Fisica, Università di Pavia, Pavia, Italy
123 Department of Physics, University of Pennsylvania, Philadelphia PA, United States of America
124 National Research Centre “Kurchatov Institute” B.P.Konstantinov Petersburg Nuclear Physics Institute, St. Petersburg, Russia
125 (a) INFN Sezione di Pisa; (b) Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa, Italy
126 Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh PA, United States of America
127 (a) Laboratório de Instrumentação e Física Experimental de Partículas - LIP, Lisboa; (b) Faculdade de Ciências, Universidade de Lisboa, Lisboa; (c) Department of Physics, University of Coimbra, Coimbra; (d) Centro de Física Nuclear da Universidade de Lisboa, Lisboa; (e) Departamento de Física, Universidade do Minho, Braga; (f) Departamento de Física Teórica y del Cosmos and CAFPE, Universidad de Granada, Granada (Spain); (g) Dep Fisica and CEFITEC of Faculdade de Ciencias e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia
 Also at CERN, Geneva, Switzerland
 Also at Georgian Technical University (GTU), Tbilisi, Georgia
 Also at Ochadai Academic Production, Ochanomizu University, Tokyo, Japan
 Also at Manhattan College, New York NY, United States of America
 Also at Hellenic Open University, Patras, Greece
 Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan
 Also at School of Physics, Shandong University, Shandong, China
 Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
 Also at section de Physique, Université de Genève, Geneva, Switzerland
 Also at Eotvos Lorand University, Budapest, Hungary
 Also at International School for Advanced Studies (SISSA), Trieste, Italy
 Also at Department of Physics and Astronomy, University of South Carolina, Columbia SC, United States of America
 Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China
 Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria
 Also at Faculty of Physics, M.V.Lomonosov Moscow State University, Moscow, Russia
 Also at Institute of Physics, Academia Sinica, Taipei, Taiwan
 Also at National Research Nuclear University MEPhI, Moscow, Russia
 Also at Department of Physics, Stanford University, Stanford CA, United States of America
 Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
 Also at Flensburg University of Applied Sciences, Flensburg, Germany
 Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia
 Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France

* Deceased