Search for flavour-changing neutral current top quark decays $t \to Hq$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector


DOI
10.1007/JHEP12(2015)061

Publication date
2015

Document Version
Final published version

Published in
The Journal of High Energy Physics

Link to publication

Citation for published version (APA):
https://doi.org/10.1007/JHEP12(2015)061

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).
Search for flavour-changing neutral current top quark decays $t \rightarrow Hq$ 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 flavour-changing neutral current decays of a top quark to an up-type quark ($q = u, c$) and the Standard Model Higgs boson, where the Higgs boson decays to $b\bar{b}$, is presented. The analysis searches for top quark pair events in which one top quark decays to $Wb$, with the $W$ boson decaying leptonically, and the other top quark decays to $Hq$. The search is based on $pp$ collisions at $\sqrt{s} = 8$ TeV recorded in 2012 with the ATLAS detector at the CERN Large Hadron Collider and uses an integrated luminosity of 20.3 fb$^{-1}$. Data are analysed in the lepton-plus-jets final state, characterised by an isolated electron or muon and at least four jets. The search exploits the high multiplicity of $b$-quark jets characteristic of signal events, and employs a likelihood discriminant that uses the kinematic differences between the signal and the background, which is dominated by $t\bar{t} \rightarrow WbWb$ decays. No significant excess of events above the background expectation is found, and observed (expected) 95% CL upper limits of 0.56% (0.42%) and 0.61% (0.64%) are derived for the $t \rightarrow Hc$ and $t \rightarrow Hu$ branching ratios respectively. The combination of this search with other ATLAS searches in the $H \rightarrow \gamma\gamma$ and $H \rightarrow WW^*$, $\tau\tau$ decay modes significantly improves the sensitivity, yielding observed (expected) 95% CL upper limits on the $t \rightarrow Hc$ and $t \rightarrow Hu$ branching ratios of 0.46% (0.25%) and 0.45% (0.29%) respectively. The corresponding combined observed (expected) upper limits on the $|\lambda_{tcH}|$ and $|\lambda_{taH}|$ couplings are 0.13 (0.10) and 0.13 (0.10) respectively. These are the most restrictive direct bounds on $tqH$ interactions measured so far.

KEYWORDS: Hadron-Hadron Scattering, Top Physics, Higgs Physics, FCNC Interactions

ArXiv ePrint: 1509.06047
## Contents

1 Introduction .................................................. 2

2 ATLAS detector ............................................... 4

3 Object reconstruction ........................................ 4

4 Data sample and event preselection ......................... 6

5 Background and signal modelling ............................ 6

6 Analysis strategy ............................................... 10
   6.1 Event categorisation .................................... 10
   6.2 Discrimination of signal from background .......... 10
    6.2.1 Signal probability ................................ 12
    6.2.2 Background probability ............................ 13
    6.2.3 Final discriminant ................................ 15

7 Systematic uncertainties ..................................... 15
   7.1 Luminosity .............................................. 15
   7.2 Reconstructed objects .................................. 18
   7.3 Background modelling .................................. 19
   7.4 Signal modelling ....................................... 20

8 Statistical analysis .......................................... 21

9 Results ......................................................... 22
   9.1 $H \rightarrow b \bar{b}$ ....................................... 22
   9.2 $H \rightarrow \gamma \gamma$ .................................... 24
   9.3 $H \rightarrow W^+W^-, \tau^+\tau^-$ .......................... 29
   9.4 Combination of searches ................................ 30

10 Conclusion .................................................... 34

A Pre-fit and post-fit event yields in the $t\bar{t} \rightarrow WbHq, H \rightarrow b \bar{b}$ search 36

B Pre-fit event yields in the $t\bar{t} \rightarrow WbHq, H \rightarrow WW^*, \tau\tau$ search ........................................ 39

The ATLAS collaboration ........................................ 47
1 Introduction

Following the observation of a Higgs boson by the ATLAS and CMS collaborations \cite{1, 2}, a comprehensive programme of measurements of its properties is underway looking for deviations from the Standard Model (SM) predictions. An interesting possibility is the presence of flavour-changing neutral current (FCNC) interactions between the Higgs boson, the top quark, and a $u$- or $c$-quark, $tqH$ ($q = u, c$). Since the Higgs boson is lighter than the top quark, with a measured mass $m_H = 125.09 \pm 0.24$ GeV \cite{3}, such interactions would manifest themselves as FCNC top quark decays, $t \to Hq$. In the SM, such decays are extremely suppressed relative to the dominant $t \to Wb$ decay mode, since $tqH$ interactions are forbidden at the tree level and even suppressed at higher-orders in the perturbative expansion due to the Glashow-Iliopoulos-Maiani (GIM) mechanism \cite{4}. As a result, the SM predictions for the $t \to Hq$ branching ratios are exceedingly small: $\text{BR}(t \to Hu) \sim 10^{-17}$ and $\text{BR}(t \to Hc) \sim 10^{-15}$ \cite{5-8}. On the other hand, large enhancements in these branching ratios are possible in some beyond-SM scenarios, where the GIM suppression can be relaxed and/or new particles can contribute to the loops, yielding effective couplings orders of magnitude larger than those of the SM. Examples include quark-singlet models \cite{9}, two-Higgs-doublet models (2HDM) of type I, with explicit flavour conservation, and of type II, such as the minimal supersymmetric SM (MSSM) \cite{10-12}, or supersymmetric models with R-parity violation \cite{13}. In those scenarios, typical branching ratios can be as high as $\text{BR}(t \to Hq) \sim 10^{-5}$. An even larger branching ratio of $\text{BR}(t \to Hc) \sim 10^{-3}$ can be reached in 2HDM without explicit flavour conservation (type III), since a tree-level FCNC coupling is not forbidden by any symmetry \cite{14-16}. While other FCNC top couplings, $tq\gamma$, $tqZ$, $tqq$, are also enhanced relative to the SM prediction in those scenarios beyond the SM, the largest enhancements are typically for the $tqH$ couplings, and in particular the $tcH$ coupling. See ref. \cite{7} for a review.

Searches for $t \to Hq$ decays have been performed by the ATLAS and CMS collaborations, taking advantage of the large samples of $t\bar{t}$ events collected during Run 1 of the LHC. In these searches, one of the top quarks is required to decay into $Wb$, while the other top quark decays into $Hq$, yielding $t\bar{t} \to WbHq$.\footnote{In the following $WbHq$ is used to denote both $W^+bHq$ and its charge conjugate, $HqW^-\bar{b}$. Similarly, $WbWb$ is used to denote $W^+W^-\bar{b}$}. Assuming SM decays for the Higgs boson and $m_H = 125$ GeV, the most sensitive single-channel searches have been performed in the $H \to \gamma\gamma$ decay mode which, despite the tiny branching ratio of $\text{BR}(H \to \gamma\gamma) \simeq 0.2\%$, is characterised by very small background and excellent diphoton mass resolution. The resulting observed (expected) 95\% confidence level (CL) upper limits on $\text{BR}(t \to Hq)$ are 0.79\% (0.51\%) and 0.69\% (0.81\%), respectively from the ATLAS \cite{17} and CMS \cite{18} collaborations. These searches are insensitive to the difference between $t \to Hu$ and $t \to Hc$, and thus the above limits can be interpreted as applying to the sum $\text{BR}(t \to Hu)+\text{BR}(t \to Hc)$. The CMS Collaboration has also reinterpreted searches in multilepton (three or four leptons) final states \cite{18} in the context of $t\bar{t} \to WbHq$ with $H \to WW^*, \tau\tau$, resulting in an observed (expected) upper limit of $\text{BR}(t \to Hc) < 1.28\% (1.17\%)$ at the 95\% CL. Multi-lepton searches are able to exploit a significantly larger branching ratio for the Higgs boson...
decay compared to the $H \to \gamma\gamma$ decay mode, and are also characterised by relatively small backgrounds. However, in general they do not have good mass resolution, so any excess would be hard to interpret as originating from $t \to Hq$ decays. The combination of CMS searches in diphoton and multilepton (three or four leptons) final states yields an observed (expected) upper limit of $\text{BR}(t \to Hq) < 0.56\%$ (0.65\%) at the 95\% CL [18].

Upper limits on the branching ratios $\text{BR}(t \to Hq)$ ($q = u, c$) can be translated to upper limits on the non-flavour-diagonal Yukawa couplings $\lambda_{tqH}$ appearing in the following Lagrangian:

$$
\mathcal{L}_{\text{FCNC}} = \lambda_{tcH} iHc + \lambda_{tuaH} iHu + h.c.
$$

(1.1)

The branching ratio $\text{BR}(t \to Hq)$ is estimated as the ratio of its partial width [8] to the SM $t \to Wb$ partial width [19], which is assumed to be dominant. Both predicted partial widths include next-to-leading-order (NLO) QCD corrections. Using the expression derived in ref. [17], the coupling $|\lambda_{tqH}|$ can be extracted as $|\lambda_{tqH}| = (1.92 \pm 0.02) \sqrt{\text{BR}(t \to Hq)}$.

The results presented in this paper fill a gap in the current programme of searches for $t \to Hq$ decays at the LHC by considering the dominant decay mode $H \to b\bar{b}$, which has $\text{BR}(H \to b\bar{b}) \simeq 58\%$. This search is focused on the $t\bar{t} \to WbHq$ ($q = u, c$) process, with $W \to \ell\nu$ ($\ell = e, \mu, \tau$) and $H \to b\bar{b}$, resulting in a lepton-plus-jets final state with high $b$-jet multiplicity, which can be effectively exploited to suppress the overwhelming $t\bar{t}$ background. Early studies of the prospects for this search at the LHC were performed in ref. [20]. Only events with an electron or muon, including those produced via leptonically decaying taus, are considered. The lepton-plus-jets final state also allows the kinematic reconstruction of the final state and in particular the dijet invariant mass spectrum from the $H \to b\bar{b}$ decay, providing additional handles that would help in detecting $t\bar{t} \to WbHq$ events. Most of this paper is devoted to the discussion of this particular search, for which background estimation techniques, systematic uncertainties and statistical treatment closely follow those used in recent ATLAS searches using the same final-state signature [21, 22]. This paper also includes a reinterpretation of the ATLAS search for $t\bar{t}H$ associated production, with $H \to WW^*, ZZ^*, \tau\tau$, resulting in multilepton final states [23]. This reinterpretation only considers the final states with a significant expected contribution from $t\bar{t} \to WbHq$, $H \to WW^*, \tau\tau$ signal, namely two same-charge leptons with and without an identified hadronic tau lepton and three leptons. A combination of the three ATLAS searches for $t\bar{t} \to WbHq$, probing the $H \to b\bar{b}$, $H \to WW^*, \tau\tau$, and $H \to \gamma\gamma$ decay modes, is performed and bounds are set on $\text{BR}(t \to Hc)$ and $\text{BR}(t \to Hu)$, as well as on the corresponding non-flavour-diagonal Yukawa couplings.

This paper is organised as follows. A brief description of the ATLAS detector is provided in section 2. Subsequent sections are devoted to a detailed discussion of the $t\bar{t} \to WbHq, H \to b\bar{b}$ search, covering the object reconstruction (section 3), the data sample and event preselection (section 4), the modelling of the backgrounds and the signal (section 5), the analysis strategy (section 6), and the systematic uncertainties (section 7). Section 8 provides a discussion of the statistical methods used. Section 9 presents the

---

\[2\] An exception is the $H \to ZZ^* \to \ell^+\ell^-\ell'^+\ell'^-$ ($\ell, \ell' = e, \mu$) decay mode, which has a very small branching ratio and thus is not promising for this search.
results obtained by the three individual ATLAS searches as well as their combination. Finally, the conclusions are given in section 10.

2 ATLAS detector

The ATLAS detector [24] consists of the following main subsystems: an inner tracking system, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner detector provides tracking information from silicon pixel and microstrip detectors in the pseudorapidity\(^3\) range \(|\eta| < 2.5\) and from a straw-tube transition radiation tracker covering \(|\eta| < 2.0\), all immersed in a 2 T axial magnetic field provided by a superconducting solenoid. The electromagnetic (EM) sampling calorimeter uses lead as the absorber material and liquid-argon (LAr) as the active medium, and is divided into barrel (\(|\eta| < 1.475\)) and end-cap (1.375 < \(|\eta| < 3.2\)) regions. Hadron calorimetry is also based on the sampling technique, with either scintillator tiles or LAr as the active medium, and with steel, copper, or tungsten as the absorber material. The calorimeters cover \(|\eta| < 4.9\). The muon spectrometer measures the deflection of muons with \(|\eta| < 2.7\) using multiple layers of high-precision tracking chambers located in a toroidal field of approximately 0.5 T and 1 T in the central and end-cap regions of ATLAS, respectively. The muon spectrometer is also instrumented with separate trigger chambers covering \(|\eta| < 2.4\). A three-level trigger system [25] is used to select interesting events. The first-level trigger is implemented in custom electronics and uses a subset of detector information to reduce the event rate to at most 75 kHz. This is followed by two software-based trigger levels exploiting the full detector information and yielding a typical recorded event rate of 400 Hz during 2012.

3 Object reconstruction

Electron candidates [26] are reconstructed from energy clusters in the EM calorimeter that are matched to reconstructed tracks in the inner detector. Electron clusters are required to have a transverse energy \(E_T\) greater than 25 GeV and \(|\eta_{\text{cluster}}| < 2.47\), excluding the transition region 1.37 < \(|\eta_{\text{cluster}}| < 1.52\) between sections of the EM calorimeter. The longitudinal impact parameter of the electron track with respect to the event’s primary vertex (see section 4), \(z_0\), is required to be less than 2 mm. Electrons are required to satisfy “tight” quality requirements [26] based on calorimeter, tracking and combined variables that provide good separation between prompt electrons and jets. To reduce the background from non-prompt electrons resulting from semileptonic decays of \(b\)- or \(c\)-hadrons, and from jets with a high fraction of their energy deposited in the EM calorimeter, electron candidates must also satisfy calorimeter- and track-based isolation requirements. The calorimeter isolation variable is based on the energy sum of cells within a cone of size

\(^3\)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 \(\eta = -\ln\tan(\theta/2)\).
$\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2} = 0.2$ around the direction of each electron candidate, and an $\eta$-dependent requirement is made, giving an average efficiency of 90% across $\eta$ for prompt electrons from $Z$ boson decays. This energy sum excludes cells associated with the electron cluster and is corrected for leakage from the electron cluster itself as well as for energy deposits from additional $pp$ interactions within the same bunch crossing (“pileup”). A further 90%-efficient isolation requirement is made on the track transverse momentum ($p_T$) sum around the electron (excluding the electron track itself) in a cone of size $\Delta R = 0.3$.

Muon candidates [27, 28] are reconstructed from track segments in the various layers of the muon spectrometer that are matched with tracks found in the inner detector. The final candidates are refitted using the complete track information from both detector systems and are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. The longitudinal impact parameter of the muon track with respect to the primary vertex, $z_0$, is required to be less than 2 mm. Muons are required to satisfy a $p_T$-dependent track-based isolation requirement: the scalar sum of the $p_T$ of the tracks within a cone of variable size $\Delta R = 10$ GeV/$p_T^\mu$ around the muon (excluding the muon track itself) must be less than 5% of the muon $p_T$ ($p_T^\mu$). This requirement has good signal efficiency and background rejection even under high-pileup conditions, as well as in boosted configurations where the muon is close to a jet. For muons from $W$ boson decays in simulated $t\bar{t}$ events, the average efficiency of the isolation requirement is about 95%.

Jets are reconstructed with the anti-$k_t$ algorithm [29–31] with a radius parameter $R = 0.4$, using calibrated topological clusters [32, 33] built from energy deposits in the calorimeters. Prior to jet finding, a local cluster calibration scheme [34] is applied to correct the topological cluster energies for the non-compensating response of the calorimeter, as well as for the energy lost in dead material and via out-of-cluster leakage. The corrections are obtained from simulations of charged and neutral particles. After energy calibration [35], jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. To reduce the contamination due to jets originating from pileup interactions, a requirement on the absolute value of the jet vertex fraction (JVF) variable above 0.5 is applied to jets with $p_T < 50$ GeV and $|\eta| < 2.4$. This requirement ensures that at least 50% of the scalar sum of the $p_T$ of the tracks with $p_T > 1$ GeV associated with a jet comes from tracks originating from the primary vertex. During jet reconstruction, no distinction is made between identified electrons and jet energy deposits. Therefore, if any of the jets lie within $\Delta R = 0.2$ of a selected electron, the closest jet is discarded in order to avoid double-counting of electrons as jets. Finally, any electron or muon within $\Delta R = 0.4$ of a selected jet is discarded.

Jets containing $b$-hadrons are identified ($b$-tagged) via an algorithm [36] that uses multivariate techniques to combine information from the impact parameters of displaced tracks as well as topological properties of secondary and tertiary decay vertices reconstructed within the jet. For each jet, a value for the multivariate $b$-tagging discriminant is calculated. The jet is considered $b$-tagged if this value is above a given threshold. The threshold used in this search corresponds to 70% efficiency to tag a $b$-quark jet, with a light-jet rejection factor of $\sim 130$ and a charm-jet rejection factor of 5, as determined for jets with $p_T > 20$ GeV and $|\eta| < 2.5$ in simulated $t\bar{t}$ events.

$^4$Light-jet denotes a jet originating from the hadronisation of a light quark ($u$, $d$, $s$) or gluon.
The missing transverse momentum ($E_{\text{T}}^{\text{miss}}$) is constructed from the vector sum of all calorimeter energy deposits contained in topological clusters. All topological cluster energies are corrected using the local cluster calibration scheme discussed previously in the context of the jet energy calibration. Those topological clusters associated with a high-$p_T$ object (e.g. jet or electron) are further calibrated using their respective energy corrections. In addition, contributions from the $p_T$ of selected muons are included in the calculation of $E_{\text{T}}^{\text{miss}}$.

4 Data sample and event preselection

This search is based on $pp$ collision data at $\sqrt{s} = 8$ TeV collected by the ATLAS experiment between April and December 2012. Only events recorded with a single-electron or single-muon trigger under stable beam conditions and for which all detector subsystems were operational are considered. The corresponding integrated luminosity is $20.3 \pm 0.6 \text{ fb}^{-1}$ [38]. Single-lepton triggers with different $p_T$ thresholds are combined in a logical OR in order to increase the overall efficiency. The $p_T$ thresholds are 24 or 60 GeV for the electron triggers and 24 or 36 GeV for the muon triggers. The triggers with the lower $p_T$ threshold include isolation requirements on the candidate lepton, resulting in inefficiencies at high $p_T$ that are recovered by the triggers with higher $p_T$ threshold.

Events satisfying the trigger selection are required to have at least one reconstructed vertex with at least five associated tracks with $p_T > 400$ MeV, consistent with originating from the beam collision region in the $x$-$y$ plane. The average number of $pp$ interactions per bunch crossing is approximately 20, resulting in several vertices reconstructed per event. If more than one vertex is found, the hard-scatter primary vertex is taken to be the one which has the largest sum of the squared transverse momenta of its associated tracks. For the event topologies considered in this paper, this requirement leads to a probability to reconstruct and select the correct hard-scatter primary vertex larger than 99%.

Preselected events are required to have exactly one electron or muon, as defined in section 3, that matches, within $\Delta R = 0.15$, the lepton candidate reconstructed by the trigger. In addition, at least four jets are required, of which at least two must be $b$-tagged.

5 Background and signal modelling

After the event preselection, the main background is $t\bar{t} \rightarrow WbWb$ production, possibly in association with jets, denoted by $t\bar{t}$+jets in the following. Single top quark production and production of a $W$ boson in association with jets ($W$+jets) contribute to a lesser extent. Small contributions arise from multijet, $Z$+jets and diboson ($WW, WZ, ZZ$) production, as well as from the associated production of a vector boson $V$ ($V = W, Z$) or a Higgs boson and a $t\bar{t}$ pair ($t\bar{t}V$ and $t\bar{t}H$). Signal and all backgrounds are estimated from simulation and normalised to their theoretical cross sections, with the exception of the multijet background, which is estimated with data-driven methods [39].

Simulated samples of $t\bar{t}$ events are generated with the NLO generator POWHEG-BOX 2.0 [40–43] using the CT10 [44] set of parton distribution functions (PDF). The nominal
sample is interfaced to Pythia 6.425 [45] for parton showering and hadronisation with the CTEQ6L1 PDF set and the Perugia2011C [46] set of optimised parameters for the underlying event (UE) description, referred to as the “UE tune”. An alternative sample, used to study the uncertainty related to the hadronisation model, is interfaced to HERWIG v6.520 [47] with the CTEQ6L1 PDF set and JIMMY v4.31 [48] to simulate the UE. All samples are generated assuming a top quark mass of 172.5 GeV and top quark decays exclusively through $t \rightarrow Wb$. The $t\bar{t}$ process is normalised to a cross section of $253 \pm 15 \pm 16 \text{pb}$, computed using Top++ v2.0 [49] at next-to-next-to-leading order (NNLO) in QCD, including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [50–54], and using the MSTW 2008 NNLO [55, 56] PDF set. Theoretical uncertainties result from variations of the factorisation and renormalisation scales, as well as from uncertainties on the PDF and $\alpha_s$. The latter two represent the largest contribution to the overall theoretical uncertainty on the cross section and were calculated using the PDF4LHC prescription [57] with the MSTW 2008 68% CL NNLO, CT10 NNLO [44, 58] and NNPDF2.3 5f FFN [59] PDF sets. In the case where a non-zero BR($t \rightarrow Hq$) is assumed, an additional factor of $[1 – \text{BR}(t \rightarrow Hq)]^2$ is applied to the sample normalisation. It is not possible to generate the $t\bar{t}Hq$ signal with Powheg-Box, and a different event generator is used instead, as discussed below.

The $t\bar{t}$ samples are generated inclusively, but events are categorised depending on the flavour content of additional particle jets not originating from the decay of the $t\bar{t}$ system.\(^5\) Details about this categorisation scheme can be found in ref. [21]. In this way, a distinction is made between $t\bar{t}+b\bar{b}$, $t\bar{t}+c\bar{c}$ and $t\bar{t}$-light-jets events. The first two categories are generically referred to as $t\bar{t}$+HF events (with HF standing for “heavy flavour”), while the latter category also includes events with no additional jets. The modelling of $t\bar{t}$+HF in Powheg-Box+Pythia is via the parton-shower evolution. To study uncertainties related to this simplified description, an alternative $t\bar{t}$+jets sample is generated with Madgraph5 1.5.11 [60] using the CT10 PDF set. It includes tree-level diagrams with up to three additional partons (including $b$- and $c$-quarks) and is interfaced to Pythia 6.425.

Since the best possible modelling of the $t\bar{t}$+jets background is a key aspect of this search, a correction is applied to simulated $t\bar{t}$ events in Powheg-Box+Pythia based on the ratio of the differential cross sections measured in data and simulation at $\sqrt{s} = 7 \text{TeV}$ as a function of top quark $p_T$ and $t\bar{t}$ system $p_T$ [61]. This correction significantly improves agreement between simulation and data at $\sqrt{s} = 8 \text{TeV}$ in distributions such as the jet multiplicity and the $p_T$ of decay products of the $t\bar{t}$ system [21], and is applied only to $t\bar{t}$+light-jets and $t\bar{t}+c\bar{c}$ events. The modelling of the $t\bar{t}+b\bar{b}$ background is improved by reweighting the Powheg-Box+Pythia prediction to an NLO prediction of $t\bar{t}+b\bar{b}$ with massive $b$ quarks and including parton showering [62], based on Sherpa+OpenLoops [63, 64] using the CT10 PDF set. Such treatment is not possible for the $t\bar{t}+c\bar{c}$ background since a corresponding NLO prediction is not currently available. More details about the modelling of the $t\bar{t}$+jets background can be found in ref. [21].

\(^5\)Particle jets are reconstructed by clustering stable particles excluding muons and neutrinos using the anti-$k_t$ algorithm with a radius parameter $R = 0.4$. 
Samples of single-top-quark backgrounds corresponding to the $t$-channel, $s$-channel, and $Wt$ production mechanisms are generated with Powheg-Box 2.0 [65, 66] using the CT10 PDF set and interfaced to Pythia 6.425 with the CTEQ6L1 PDF set in combination with the Perugia2011C UE tune. Overlaps between the $t\bar{t}$ and $Wt$ final states are avoided using the “diagram removal” scheme [67]. The single-top-quark samples are normalised to the approximate NNLO theoretical cross sections [68–70], calculated using the MSTW 2008 NNLO PDF set.

Samples of $W/Z+$jets events are generated with up to five additional partons using the Alpgen v2.14 [71] LO generator with the CTEQ6L1 PDF set and interfaced to Pythia 6.426. To avoid double-counting of partonic configurations generated by both the matrix-element calculation and the parton shower, a parton-jet matching scheme (“MLM matching”) [72] is employed. The $W+$jets samples are generated separately for $W+$light-jets, $Wb\bar{b}$+jets, $Wc\bar{c}$+jets, and $Wc+$jets. The $Z+$jets samples are generated separately for $Z+$light-jets, $Zb\bar{b}$+jets, and $Zc\bar{c}$+jets. Overlap between $VQ\bar{Q}$+jets ($V = W, Z$ and $Q = b, c$) events generated from the matrix-element calculation and those generated from parton-shower evolution in the $W/Z+$light-jets samples is avoided via an algorithm based on the angular separation between the extra heavy quarks: if $\Delta R(Q, \bar{Q}) > 0.4$, the matrix-element prediction is used, otherwise the parton-shower prediction is used. Both the $W+$jets and $Z+$jets background contributions are normalised to their inclusive NNLO theoretical cross sections [73]. Further corrections are applied to $W/Z+$jets events in order to better describe data in the preselected sample. Normalisation factors for each of the $W+$jets categories ($Wb\bar{b}$+jets, $Wc\bar{c}$+jets, $Wc+$jets and $W+$light-jets) are derived for events with one lepton and at least four jets by simultaneously analysing six different event categories, defined by the $b$-tag multiplicity (0, 1 and $\geq 2$) and the sign of the lepton charge [74]. The $b$-tag multiplicity provides information about the heavy-flavour composition of the $W+$jets background, while the lepton charge is used to determine the normalisation of each component, exploiting the expected charge asymmetry for $W+$jets production in $pp$ collisions as predicted by Alpgen. In the case of $Z+$jets events, a correction to the heavy-flavour fraction is derived to reproduce the relative rates of $Z+$2-jets events with zero and one $b$-tagged jet observed in data. In addition, the $Z$ boson $p_T$ spectrum is compared between data and the simulation in $Z+$2-jets events, and a reweighting function is derived in order to improve the modelling. This reweighting function is also applied to the $W+$jets simulated sample and it was verified that this correction further improves the agreement between data and simulation for $W+$jets events. In any case, $W/Z+$jets events constitute a very small background in this analysis after final event selection.

The $WW=WZ=ZZ+$jets samples are generated with up to three additional partons using Alpgen v2.13 and the CTEQ6L1 PDF set, interfaced to Herwig v6.520 and Jimmy v4.31 for parton showering, hadronisation and UE modelling. The MLM parton-jet matching scheme is used. The $WW+$jets samples require at least one of the $W$ bosons to decay leptonically, while the $WZ=ZZ+$jets samples require one $Z$ boson to decay leptonically and the other boson decays inclusively. Additionally, $WZ+$jets samples requiring the $W$ boson to decay leptonically and the $Z$ boson to decay hadronically, are generated with up to three additional partons (including massive $b$- and $c$-quarks) using Sherpa v1.4.1 and the CT10 PDF set. All diboson samples are normalised to their NLO theoretical cross sections [75].
Samples of $t\bar{t}V$ events, including $t\bar{t}WW$, are generated with up to two additional partons using MADGRAPH5 1.3.28 with the CTEQ6L1 PDF set, and interfaced to PYTHIA 6.425 with the AUET2B UE tune [76]. A sample of $t\bar{t}H$ events is generated with the POWHEG framework [77], which combines the POWHEG-BOX generator and NLO matrix elements obtained from the HELAC-Oneloop package [78]. The sample is generated using the CT10nlo PDF set [44]. Showering is performed with PYTHIA 8.1 [79] using the CTEQ6L1 PDF set and the AU2 UE tune [76, 80]. Inclusive decays of the Higgs boson are assumed in the generation of the $t\bar{t}H$ sample. The $t\bar{t}V$ samples are normalised to the NLO cross-section predictions [81]. The $t\bar{t}H$ sample is normalised using the NLO cross section [82–84] and the Higgs decay branching ratios [85–88] collected in ref. [89].

The multijet background contributes to the selected data sample via several production and misreconstruction mechanisms. In the electron channel, it consists of non-prompt electrons (from semileptonic $b$- or $c$-hadron decays) as well as misidentified photons (e.g. from a conversion of a photon into an $e^+e^-$ pair) or jets with a high fraction of their energy deposited in the EM calorimeter. In the muon channel, the multijet background is predominantly from non-prompt muons. Its normalisation and shape are estimated directly from data by using the “matrix method” technique [39], which exploits differences in lepton-identification-related properties between prompt and isolated leptons and leptons that are either non-isolated or result from the misidentification of photons or jets. Further details can be found in ref. [22].

The $t\bar{t} \rightarrow WbHq$ signal process is modelled using the PROTOS v2.2 [90, 91] LO generator with the CTEQ6L1 PDF set, and interfaced to PYTHIA 6.426 and the Perugia2011C UE tune. Two separate samples are generated corresponding to $t\bar{t} \rightarrow WbHc$ and $t\bar{t} \rightarrow WbHu$, with the $W$ boson forced to decay leptonically, $W \rightarrow \ell\nu \ (\ell = e, \mu, \tau)$. The top quark and Higgs boson masses are set to 172.5 GeV and 125 GeV, respectively. The Higgs boson is allowed to decay to all SM particles with branching ratios as given in ref. [89]. The signal sample is normalised to the same NNLO cross section as used for the $t\bar{t} \rightarrow WbWb$ sample, and the corresponding branching ratios: $\sigma(t\bar{t} \rightarrow W(\rightarrow \ell\nu)bHq)[1 - BR(t \rightarrow Hq)]BR(W \rightarrow \ell\nu)\sigma_{t\bar{t}}$, with $BR(W \rightarrow \ell\nu) = 0.324$ and $BR(t \rightarrow Hq)$ depending on the branching ratio being tested. Typically a reference branching ratio of $BR(t \rightarrow Hq) = 1\%$ is used. The case of both top quarks decaying into $Hq$ is neglected in the analysis given existing upper limits on $BR(t \rightarrow Hq)$ (see section 1). In order to improve the modelling of the signal kinematics, a two-step reweighting procedure is applied: the first step is designed to correct the spectrum of top quark $p_T$ and $t\bar{t}$ system $p_T$ to match that of the uncorrected $t\bar{t} \rightarrow WbWb$ POWHEG-BOX+PYTHIA sample; the second step involves the same correction to the top quark $p_T$ and $t\bar{t}$ system $p_T$ applied to the $t\bar{t}$+jets background (see discussion above).

Finally, all generated samples are processed through a simulation [92] of the detector geometry and response using GEANT4 [93]. Additional minimum-bias $pp$ interactions are simulated with the PYTHIA 8.1 generator with the MSTW 2008 LO PDF set and the A2 UE tune [94]. They are overlaid on the simulated signal and background events according to the luminosity profile of the recorded data. The contributions from these pileup interactions are modelled both within the same bunch crossing as the hard-scattering process and
in neighbouring bunch crossings. All simulated samples are processed through the same reconstruction software as the data. Simulated events are corrected so that the object identification efficiencies, energy scales, and energy resolutions match those determined from data control samples.

6 Analysis strategy

This section presents an overview of the analysis strategy followed by the \( t\bar{t} \rightarrow WbHq, H \rightarrow b\bar{b} \) search.

6.1 Event categorisation

Given the focus on the \( W \rightarrow \ell\nu \) and \( H \rightarrow b\bar{b} \) decay modes, the \( t\bar{t} \rightarrow WbHq \) signal is expected to have typically four jets, of which three or four are \( b \)-tagged. The latter case corresponds to the \( t\bar{t} \rightarrow WbHc \) signal where the charm quark, as well as the three \( b \)-quark jets, are \( b \)-tagged. Additional jets can also be present because of initial- or final-state radiation. In order to optimise the sensitivity of the search, the selected events are categorised into different channels depending on the number of jets (4, 5 and \( \geq 6 \)) and on the number of \( b \)-tagged jets (2, 3 and \( \geq 4 \)). Therefore, the total number of analysis channels considered in this search is nine: \((4 \ j, 2 \ b), (4 \ j, 3 \ b), (4 \ j, 4 \ b), (5 \ j, 2 \ b), (5 \ j, 3 \ b), (5 \ j, \geq 4 \ b), (\geq 6 \ j, 2 \ b), (\geq 6 \ j, 3 \ b), (\geq 6 \ j, \geq 4 \ b), \) where \((n \ j, m \ b)\) indicates \( n \) selected jets and \( m \) \( b \)-tagged jets.

The overall rate and composition of the \( t\bar{t}+\text{jets} \) background strongly depends on the jet and \( b \)-tag multiplicities, as illustrated in figure 1. The \( t\bar{t}+\text{light-jets} \) background is dominant in events with exactly two or three \( b \)-tagged jets, with the two \( b \)-quarks from the top quark decays being tagged in both cases, and a charm quark from the hadronic \( W \) boson decay also being tagged in the latter case. Contributions from \( t\bar{t}+c\bar{c} \) and \( t\bar{t}+b\bar{b} \) become significant as the jet and \( b \)-tag multiplicities increase, with the \( t\bar{t}+b\bar{b} \) background being dominant for events with \( \geq 6 \) jets and \( \geq 4 \) \( b \)-tags.

In the channels with four or five jets and three or at least four \( b \)-tags, which dominate the sensitivity of this search, selected signal events have a \( H \rightarrow b\bar{b} \) decay in more than 95% of the events. The channels most sensitive to the \( t\bar{t} \rightarrow WbHu \) and \( t\bar{t} \rightarrow WbHc \) signals are \((4 \ j, 3 \ b)\) and \((4 \ j, 4 \ b)\) respectively. Because of the better signal-to-background ratio in the \((4 \ j, 4 \ b)\) channel, this analysis is expected to have better sensitivity for \( t\bar{t} \rightarrow WbHc \) than for \( t\bar{t} \rightarrow WbHu \) signal. The rest of the channels have significantly lower signal-to-background ratios, but they are useful for calibrating the \( t\bar{t}+\text{jets} \) background prediction and constraining the related systematic uncertainties (see section 7) through a likelihood fit to data (see section 8). This strategy was first used in the ATLAS search for \( t\bar{t}H \) associated production, with \( H \rightarrow b\bar{b} \) [21], and is adopted in this analysis. A table summarising the observed and expected yields before the fit to data in each of the analysis channels can be found in appendix A.

6.2 Discrimination of signal from background

After event categorisation, the signal-to-background ratio is very low even in the most sensitive analysis channels, and a suitable discriminating variable between signal and back-
Figure 1. Comparison between the data and background prediction for the yields in each of the analysis channels considered before the fit to data (pre-fit). Backgrounds are normalised to their nominal cross sections discussed in section 5. The expected $t\bar{t} \to WbHc$ and $t\bar{t} \to WbHu$ signals (dashed histograms) are shown separately normalised to BR($t\to Hq$) = 1%. The $t\bar{t} \to WbWb$ background is normalised to the SM prediction. The small contributions from $W/Z+$jets, single top, diboson and multijet backgrounds are combined into a single background source referred to as “Non-$t\bar{t}$”. The bottom panel displays the ratio of data to the SM background (“Bkg”) prediction. The hashed area represents the total uncertainty on the background.

A powerful discriminant between signal and background can be defined as:

$$D(x) = \frac{P^{\text{sig}}(x)}{P^{\text{sig}}(x) + P^{\text{bkg}}(x)},$$

(6.1)

where $P^{\text{sig}}(x)$ and $P^{\text{bkg}}(x)$ represent the probability density functions (pdf) of a given event under the signal hypothesis ($t\bar{t} \to WbHq$) and under the background hypothesis ($t\bar{t} \to WbWb$) respectively. Both pdfs are functions of $x$, representing the four-momentum vectors of all final-state particles at the reconstruction level: the lepton ($\ell$), the neutrino ($\nu$; reconstructed as discussed below), and the $N_{\text{jets}}$ selected jets in a given analysis channel.

Since both signal and background result from the $t\bar{t}$ decay, there are few experimental handles available to discriminate between them. The most prominent features are the different resonances present in the decay (i.e. the Higgs boson in the case of $t\bar{t} \to WbHq$ and a hadronically decaying $W$ boson in the case of $t\bar{t} \to WbWb$), and the different flavour content of the jets forming those resonances. This is the main information exploited in the construction of $P^{\text{sig}}(x)$ and $P^{\text{bkg}}(x)$ in this analysis, so that $x$ is extended to include not only the four-momenta of jets $p_{\text{jet}}$, but also the value of their multivariate $b$-tagging discriminant $w_{\text{jet}}$, i.e., $x \equiv \{p_{\ell}, p_{\nu}, (p_{\text{jet}_i}, w_{\text{jet}_i})\}$ ($i = 1, \ldots, N_{\text{jets}}$). There is also some angular informa-
tion from the different spins of the daughter resonances (Higgs and W boson) that could be exploited, but it is expected to be subleading in importance and is neglected in this analysis.

The calculation of \( P_{\text{sig}}(x) \) and \( P_{\text{bkg}}(x) \) is discussed in detail in sections 6.2.1 and 6.2.2 respectively. In the following, \( b_1 \) denotes the b-quark jet from the semileptonic top quark decay, \( q_1 \) and \( b_2 \) denote the light-quark jet \((q_2 = u \text{ or } c)\) and b-quark jet from the hadronic top quark decay in background and signal events respectively, \( q_1 \) and \( q_2 \) denote the up-type-quark jet \((u \text{ or } c)\) and down-type-quark jet \((d \text{ or } s)\) from the W boson decay respectively, and \( b_1 \) and \( b_2 \) denote the two b-quark jets from the Higgs boson decay. The level of separation achieved between signal and background with the resulting discriminant \( D \) is illustrated in section 6.2.3.

### 6.2.1 Signal probability

The construction of \( P_{\text{sig}}(x) \) will now be described step by step to illustrate the method. If the partonic origin of each jet were known [see figure 2(a)], \( P_{\text{sig}}(x) \) would be defined in this analysis as the product of the normalised pdfs for each of the reconstructed invariant masses in the event: the semileptonic top quark mass \((M_{t\nu b_1})\), the hadronic top quark mass \((M_{b_1 b_2})\) and the Higgs boson mass \((M_{b_1 b_2})\). Since \( M_{b_1 b_2} \) and \( M_{b_1 b_2} \) are correlated, their difference in quadrature, \( X_{b_1 b_2} = M_{b_1 b_2} - M_{b_1 b_2,\text{MC}} \), is used instead of \( M_{b_1 b_2} \). Therefore the expression for \( P_{\text{sig}} \) just making use of the above kinematic information, denoted by \( P_{\text{kin}} \), is:

\[
P_{\text{kin}}(x) = P_{\text{sig}}(M_{t\nu b_1})P_{\text{sig}}(X_{b_1 b_2})P_{\text{sig}}(M_{b_1 b_2}). \tag{6.2}
\]

The distributions of these invariant masses are obtained from simulated signal events using the reconstructed lepton and/or jets corresponding to the correct parton-jet assignment, determined by matching a given quark (before final-state radiation) to the closest jet with \( \Delta R < 0.3 \). The corresponding pdfs are constructed as unit-normalised one-dimensional histograms. To compute \( M_{t\nu b_1} \), the neutrino four-momentum is needed, which is reconstructed as follows. Initially, the \( x \) and \( y \) components of the neutrino momentum, \( p_x,\nu \) and \( p_y,\nu \), are identified with those of the reconstructed \( E_T^{\text{miss}} \) vector. The \( z \) component of the neutrino momentum, \( p_z,\nu \), is inferred by solving \( M_W^2 = (p_t + p_\nu)^2 \), with \( M_W = 80.4 \text{ GeV} \) being the W boson mass. If two real solutions (“2sol”) exist, they are sorted according to their absolute value of \( |p_z,\nu| \), i.e., \( |p_z,\nu_1| < |p_z,\nu_2| \). It is found that in 62% of the cases \( p_{z,\nu,1} \) is closer than \( p_{z,\nu,2} \) to the generator-level neutrino \( p_{z,\nu} \). In this case, two different pdfs are constructed, one for each solution, and \( P_{2\text{sol}}(M_{t\nu b_1}) \) is defined as the average of the two pdfs weighted by their fractions (0.62 for \( p_{z,\nu,1} \) and 0.38 for \( p_{z,\nu,2} \)). If no real solution (“nosol”) exists, which happens in about 30% of the cases, the \( p_{z,\nu} \) and \( p_{y,\nu} \) components are scaled by a common factor until the discriminant of the quadratic equation is exactly zero, yielding only one solution for \( p_{z,\nu} \). This solution for \( p_{z,\nu} \) is used to compute \( M_{t\nu b_1} \) from which the corresponding \( P_{\text{nosol}}(M_{t\nu b_1}) \) is constructed. In the calculation of \( P_{\text{kin}}(x) \) from equation (6.2), \( P_{\text{sig}}(M_{t\nu b_1}) \) is identified with \( P_{2\text{sol}}(M_{t\nu b_1}) \) or with \( P_{\text{nosol}}(M_{t\nu b_1}) \), depending on how many neutrino solutions can be found for the event.

In practice, the partonic origin of the jets is not known, so it is necessary to evaluate \( P_{\text{sig}}(x) \) by averaging over the \( N_p \) possible parton-jet assignments, which dilutes the
kinematic information. At this point \( b \)-tagging information can be used to suppress the impact from parton-jet assignments that are inconsistent with the correct parton flavours as follows:

\[
P^{\text{sig}}(x) = \frac{\sum_{k=1}^{N_p} P^{\text{sig}}_{\text{btag}}(x^k) P^{\text{sig}}_{\text{kin}}(x^k)}{\sum_{k=1}^{N_p} P^{\text{sig}}_{\text{btag}}(x^k)},
\]

(6.3)

where \( P^{\text{sig}}_{\text{kin}}(x) \) is given by equation (6.2) and \( P^{\text{sig}}_{\text{btag}}(x) \) is defined as:

\[
P^{\text{sig}}_{\text{btag}}(x) = P_b(\text{jet}_1)P_b(\text{jet}_2)P_b(\text{jet}_3)P_{\text{q}}h(\text{jet}_4),
\]

(6.4)

with \( \text{jet}_i \) \((i = 1, \ldots, 4)\) representing the parton-jet assignment being evaluated, and \( P_f(\text{jet}_i) \) denoting the probability that \( \text{jet}_i \), characterised by its four-momentum \( p_{\text{jet}_i} \) and \( b \)-tagging weight value \( w_{\text{jet}_i} \), originates from a parton with flavour \( f \) (\( b, c, \) or \( l; l \) for light parton). The calibration of the \( b \)-tagging algorithm is performed for fixed thresholds on the multivariate \( b \)-tagging discriminant variable, corresponding to different average \( b \)-tagging efficiencies in \( t\bar{t} \) events of 60\%, 70\%, and 80\%, also referred to as “operating points” (OP). The corresponding thresholds are denoted by \( w_{\text{OP}}^{\text{cut}} \), with OP = 60\%, 70\%, or 80\%. Parameterisations of the \( b \)-tagging efficiencies for different jet flavours as functions of jet \( p_T \) and \( \eta \) are available for each of these operating points, \( \epsilon_f^{\text{OP}}(p_T, \eta) \), which can be used to compute \( P_f \) as follows: if the jet \( b \)-tagging weight falls between the thresholds for operating points \( \text{OP}_1 \) and \( \text{OP}_2 \), \( w_{\text{cut}}^{\text{OP}_1} < w_{\text{jet}} \leq w_{\text{cut}}^{\text{OP}_2} \), then \( P_f = \epsilon_f^{\text{OP}_1} - \epsilon_f^{\text{OP}_2} \); alternatively, if the jet \( b \)-tagging weight is below (above) the threshold corresponding to the 80\% (60\%) operating point, then \( P_f = 1 - \epsilon_f^{0.80} \) (\( P_f = \epsilon_f^{0.60} \)).

### 6.2.2 Background probability

The calculation of \( P^{\text{bkg}} \) follows a similar approach to that discussed in section 6.2.1, although it is slightly more complicated to account for the varying fraction and different kinematic features of the \( t\bar{t} + \text{light-jets} \), \( t\bar{t} + c\bar{c} \) and \( t\bar{t} + b\bar{b} \) backgrounds as a function of the analysis channel. This is particularly relevant in the (4 \( j, 3 b \)) and (4 \( j, 4 b \)) channels, which dominate the sensitivity of the search. While \( t\bar{t} + \text{light-jets} \) events often have both jets from the hadronic \( W \) boson decay among the four selected jets [see figure 2(b)], this is seldom the case for \( t\bar{t} + b\bar{b} \) and \( t\bar{t} + c\bar{c} \) events, especially in the (4 \( j, 4 b \)) channel. In this case the four \( b \)-tagged jets typically originate from the two \( b \)-quarks from the top quark decays, the charm quark from the \( W \) boson decay, and an extra heavy-flavour quark (\( b \) or \( c \)) produced in association with the \( t\bar{t} \) system, while the jet associated with the down-type quark from the \( W \) boson decay is not reconstructed [see figure 2(c)].

To account for this, the following kinematic variables are considered: \( M_{\text{bt}\text{bt}} \), \( X_{q_4j}b_h \) and \( M_{q_1j} \), with \( X_{q_4j}b_h \equiv M_{q_4j}b_h \otimes M_{q_1j} \), were \( j \) denotes an extra quark-jet which can either originate from the \( W \) boson decay (\( q_2 \)) or from an extra heavy-flavour quark (\( b \) or \( c \)) produced in association with the \( t\bar{t} \) system. For each of these possibilities, occurring in a fraction \( f_j \) of the cases, corresponding pdfs are constructed. As a generalisation of equation (6.3), the
expression for $P_{\text{bkg}}(x)$ becomes:

$$P_{\text{bkg}}(x) = \sum_{k=1}^{N_k} \sum_{j \in \{b,c,q_2\}} f_j P_{\text{bkg},j}^{btag}(x^k) P_{\text{kin}}^{btag}(x^k),$$

with

$$P_{\text{kin}}^{btag}(x) = P_{\text{bkg}}(M_{t\bar{t}b\bar{b}c}) P_{\text{bkg}}(X_{q_1j_b}) P_{\text{bkg}}(M_{q_1j}),$$

and

$$P_{\text{bkg},j}^{btag}(x) = P_b(\text{jet}_1) P_{q_1}(\text{jet}_2) P_j(\text{jet}_3) P_b(\text{jet}_4).$$

where $P_f(\text{jet}_i)$ are computed as discussed in section 6.2.1. In the above expression, $P_j = P_l$ for $j = q_2$, the down-type quark in the $W$ boson decay, and $P_{q_1} = f_c P_c + (1 - f_c) P_l$, where $f_c$ is the fraction of events where the up-type quark from the $W$ boson decay assigned to the jet is a charm quark. This fraction is different in each analysis channel, primarily depending on the $b$-tag multiplicity requirements. It varies from $\sim 50\%$ for events in the $(4j, 2b)$ channel to $\sim 90\%$ for events in the $(4j, 4b)$ channel.
6.2.3 Final discriminant

The final discriminant $D$ is computed for each event as given in equation (6.1), using the definitions for $P^{\text{sig}}$ and $P^{\text{bkg}}$ given in equations (6.3) and (6.5), respectively. Since this analysis has higher expected sensitivity to a $t\bar{t} \rightarrow WbHc$ signal than to a $t\bar{t} \rightarrow WbHu$ signal and, in order to allow probing of the BR($t \rightarrow Hc$) versus BR($t \rightarrow Hu$) plane, the discriminant optimised for $t\bar{t} \rightarrow WbHc$ is used for both the $Hc$ and $Hu$ decay modes. It was verified that using the $t\bar{t} \rightarrow WbHc$ discriminant for the $t\bar{t} \rightarrow WbHu$ search does not result in a significant sensitivity loss. Figure 3 compares the shape of the $D$ distribution between the $t\bar{t} \rightarrow WbHc$ and $t\bar{t} \rightarrow WbHu$ signals and the $t\bar{t} \rightarrow WbWb$ background in each of the channels considered in this analysis.

7 Systematic uncertainties

Several sources of systematic uncertainty are considered that can affect the normalisation of signal and background and/or the shape of their corresponding final discriminant distributions. Each source of systematic uncertainty is considered to be uncorrelated with the other sources. Correlations of a given systematic uncertainty are maintained across processes and channels. Table 1 presents a list of all systematic uncertainties considered in the analysis and indicates whether they are taken to be normalisation-only, or to affect both shape and normalisation.

The leading sources of systematic uncertainty vary depending on the analysis channel considered, but they typically originate from $t\bar{t}+\text{jets}$ modelling (including $t\bar{t}+\text{HF}$) and $b$-tagging. For example, the total systematic uncertainty in the background normalisation in the $(4\ j, 4\ b)$ channel, which dominates the sensitivity in the case of the $t\bar{t} \rightarrow WbHc$ search, is approximately 20%, with the largest contributions originating from $t\bar{t}+\text{HF}$ normalisation, $b$-tagging efficiency, $c$-tagging efficiency, light-jet tagging efficiency and $t\bar{t}$ cross section. However, as shown in section 9, the fit to data in the nine analysis channels allows the overall background uncertainty to be reduced significantly, to approximately 4.4%. The reduced uncertainty results from the significant constraints provided by the data on some systematic uncertainties, as well as the anti-correlations among sources of systematic uncertainty resulting from the fit to the data. The total systematic uncertainty on the $t\bar{t} \rightarrow WbHc$ signal normalisation in the $(4\ j, 4\ b)$ channel is approximately 17%, with similar contributions from uncertainties related to $b$-tagging and overall signal modelling. After the fit, this uncertainty is reduced to 7.8%. Table 2 presents a summary of the systematic uncertainties for the $t\bar{t} \rightarrow WbHc$ search and their impact on the normalisation of the signal and the main backgrounds in the $(4\ j, 4\ b)$ channel.

The following sections describe each of the systematic uncertainties considered in the analyses.

7.1 Luminosity

The uncertainty on the integrated luminosity is 2.8%, affecting the overall normalisation of all processes estimated from the simulation. It is estimated from a calibration of the lu-
Figure 3. Comparison of the shape of the $D$ discriminant distribution between the $t\bar{t} \rightarrow WbHc$ (red dashed) and $t\bar{t} \rightarrow WbHu$ (blue dotted) signals, and the $t\bar{t} \rightarrow WbWb$ background (black solid) in each of the channels considered in the analysis: (a) ($4 j$, 2 b), (b) ($4 j$, 3 b), (c) ($4 j$, 4 b), (d) ($5 j$, 2 b), (e) ($5 j$, 3 b), (f) ($5 j$, 4 b), (g) ($\geq 6 j$, 2 b), (h) ($\geq 6 j$, 3 b), and (i) ($\geq 6 j$, $\geq 4 b$).
<table>
<thead>
<tr>
<th>Systematic uncertainty</th>
<th>Type</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Luminosity</strong></td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td><strong>Reconstructed Objects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electron</td>
<td>SN</td>
<td>5</td>
</tr>
<tr>
<td>Muon</td>
<td>SN</td>
<td>6</td>
</tr>
<tr>
<td>Jet reconstruction</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Jet vertex fraction</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>SN</td>
<td>22</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Missing transverse momentum</td>
<td>SN</td>
<td>2</td>
</tr>
<tr>
<td>$b$-tagging efficiency</td>
<td>SN</td>
<td>6</td>
</tr>
<tr>
<td>$c$-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>High-$p_T$ tagging</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td><strong>Background Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>$t\bar{t}$ modelling: $p_T$ reweighting</td>
<td>SN</td>
<td>9</td>
</tr>
<tr>
<td>$t\bar{t}$ modelling: parton shower</td>
<td>SN</td>
<td>3</td>
</tr>
<tr>
<td>$t\bar{t}$+HF: normalisation</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td>$t\bar{t}$+cc: $p_T$ reweighting</td>
<td>SN</td>
<td>2</td>
</tr>
<tr>
<td>$t\bar{t}$+cc: generator</td>
<td>SN</td>
<td>4</td>
</tr>
<tr>
<td>$t\bar{t}$+bb: NLO shape</td>
<td>SN</td>
<td>8</td>
</tr>
<tr>
<td>$W$+jets normalisation</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>$W$ $p_T$ reweighting</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>$Z$+jets normalisation</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>$Z$ $p_T$ reweighting</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Single top normalisation</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>Single top model</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Diboson normalisation</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>$tW$ cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>$tW$ model</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>$tHW$ cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>$tHW$ model</td>
<td>SN</td>
<td>2</td>
</tr>
<tr>
<td>Multijet normalisation</td>
<td>N</td>
<td>4</td>
</tr>
<tr>
<td><strong>Signal Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>Higgs boson branching ratios</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>$t\bar{t}$ modelling: $p_T$ reweighting</td>
<td>SN</td>
<td>9</td>
</tr>
<tr>
<td>$t\bar{t}$ modelling: $p_T$ reweighting non-closure</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>$t\bar{t}$ modelling: parton shower</td>
<td>N</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1. List of systematic uncertainties considered. An “N” means that the uncertainty is taken as affecting only the normalisation for all relevant processes and channels, whereas “SN” means that the uncertainty is taken on both shape and normalisation. Some of the systematic uncertainties are split into several components for a more accurate treatment.
minosity scale derived from beam-separation scans performed in November 2012, following the same methodology as that detailed in ref. [38].

7.2 Reconstructed objects

Uncertainties associated with leptons arise from the reconstruction, identification and trigger. These efficiencies are measured using tag-and-probe techniques on $Z \rightarrow \ell^+\ell^-$ (where $\ell = e, \mu$) data and simulated samples. The small differences found are corrected for in the simulation. Negligible sources of uncertainty originate from the corrections applied to adjust the lepton momentum scale and resolution in the simulation to match those in data. The combined effect of all these uncertainties results in an overall normalisation uncertainty on the signal and background of approximately 1.5%.

Uncertainties associated with jets arise from the efficiency of jet reconstruction and identification based on the JVF variable, as well as the jet energy scale and resolution. The largest contribution results from the jet energy scale, whose uncertainty dependence on jet $p_T$ and $\eta$ is split into 22 uncorrelated sources that are treated independently in the analysis. It affects the normalisation of signal and backgrounds by approximately 3–4% in the most sensitive search channels, (4j, 3b) and (4j, 4b), and up to 12% in the channels with ≥6 jets.

Uncertainties associated with energy scales and resolutions of leptons and jets are propagated to $E_T^{\text{miss}}$. Additional uncertainties originating from the modelling of the underlying

<table>
<thead>
<tr>
<th></th>
<th>Pre-fit $WbHc$</th>
<th>$t\bar{t} + \text{LJ}$</th>
<th>$t\bar{t} + c\bar{c}$</th>
<th>$t\bar{t} + b\bar{b}$</th>
<th>Post-fit $WbHc$</th>
<th>$t\bar{t} + \text{LJ}$</th>
<th>$t\bar{t} + c\bar{c}$</th>
<th>$t\bar{t} + b\bar{b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>±2.8</td>
<td>±2.8</td>
<td>±2.8</td>
<td>±2.8</td>
<td>±2.6</td>
<td>±2.6</td>
<td>±2.6</td>
<td>±2.6</td>
</tr>
<tr>
<td>Lepton efficiencies</td>
<td>±1.5</td>
<td>±1.5</td>
<td>±1.5</td>
<td>±1.5</td>
<td>±1.5</td>
<td>±1.5</td>
<td>±1.5</td>
<td>±1.5</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>±3.3</td>
<td>±2.9</td>
<td>±2.3</td>
<td>±5.8</td>
<td>±1.4</td>
<td>±1.2</td>
<td>±1.8</td>
<td>±4.1</td>
</tr>
<tr>
<td>Jet efficiencies</td>
<td>±1.2</td>
<td>—</td>
<td>±1.9</td>
<td>±1.7</td>
<td>±0.9</td>
<td>—</td>
<td>±1.4</td>
<td>±1.2</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>—</td>
<td>±1.2</td>
<td>±2.8</td>
<td>±2.9</td>
<td>—</td>
<td>—</td>
<td>±1.0</td>
<td>±1.1</td>
</tr>
<tr>
<td>$b$-tagging eff.</td>
<td>±7.9</td>
<td>±5.5</td>
<td>±5.2</td>
<td>±10</td>
<td>±5.7</td>
<td>±3.9</td>
<td>±3.7</td>
<td>±6.6</td>
</tr>
<tr>
<td>$c$-tagging eff.</td>
<td>±7.0</td>
<td>±6.6</td>
<td>±13</td>
<td>±3.5</td>
<td>±6.3</td>
<td>±6.0</td>
<td>±11</td>
<td>±3.2</td>
</tr>
<tr>
<td>Light-jet tagging eff.</td>
<td>±0.8</td>
<td>±18</td>
<td>±3.2</td>
<td>±1.5</td>
<td>±0.6</td>
<td>±13</td>
<td>±2.3</td>
<td>±1.1</td>
</tr>
<tr>
<td>$t\bar{t}$: reweighting</td>
<td>±5.9</td>
<td>±2.7</td>
<td>±4.2</td>
<td>—</td>
<td>±3.8</td>
<td>±1.9</td>
<td>±2.3</td>
<td>—</td>
</tr>
<tr>
<td>$t\bar{t}$: parton shower</td>
<td>±5.4</td>
<td>±4.8</td>
<td>±10</td>
<td>±4.9</td>
<td>±1.7</td>
<td>±1.5</td>
<td>±6.5</td>
<td>±3.1</td>
</tr>
<tr>
<td>$t\bar{t}$+HF: normalisation</td>
<td>—</td>
<td>—</td>
<td>±50</td>
<td>±50</td>
<td>—</td>
<td>—</td>
<td>±32</td>
<td>±16</td>
</tr>
<tr>
<td>$t\bar{t}$+HF: modelling</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>±7.7</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>±7.4</td>
</tr>
<tr>
<td>Signal modelling</td>
<td>±6.9</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>±6.9</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Theor. cross sections</td>
<td>±6.2</td>
<td>±6.2</td>
<td>±6.2</td>
<td>±6.2</td>
<td>±3.9</td>
<td>±3.9</td>
<td>±3.9</td>
<td>±3.9</td>
</tr>
<tr>
<td>Total</td>
<td>±17</td>
<td>±22</td>
<td>±54</td>
<td>±53</td>
<td>±7.8</td>
<td>±14</td>
<td>±28</td>
<td>±15</td>
</tr>
</tbody>
</table>

Table 2. $t\bar{t} \rightarrow WbHc, H \rightarrow bb$ search: summary of the systematic uncertainties considered in the (4j, 4b) channel and their impact (in %) on the normalisation of the signal and the main backgrounds, before and after the fit to data. The $t\bar{t} \rightarrow WbHc$ signal and the $t\bar{t}$+light-jets background are denoted by "$WbHc$" and "$t\bar{t}$+LJ" respectively. Only sources of systematic uncertainty resulting in a normalisation change of at least 0.5% are displayed. The total post-fit uncertainty can differ from the sum in quadrature of individual sources due to the anti-correlations between them resulting from the fit to the data.
event, in particular its impact on the $p_T$ scale and resolution of unclustered energy, are negligibly small.

The leading uncertainties associated with reconstructed objects in this analysis originate from the modelling of the $b$-, $c$-, and light-jet-tagging efficiencies in the simulation, which is corrected to match the efficiencies measured in data control samples [36, 95] through dedicated scale factors. Uncertainties on these factors include a total of six independent sources affecting $b$-jets and four independent sources affecting $c$-jets. Each of these uncertainties has a different jet-$p_T$ dependence. Twelve sources of uncertainty affecting light jets are considered, which depend on jet $p_T$ and $\eta$. The above sources of systematic uncertainty are taken as uncorrelated between $b$-jets, $c$-jets, and light-jets. They have their largest impact in the $(4j, 4b)$ channel, resulting in 10%, 13%, and 18% normalisation uncertainties on the $t\bar{t}+b\bar{b}$, $tt+c\bar{c}$, and $t\bar{t}$+light-jets background associated with the uncertainties on the $b$-, $c$-, and light-jet-tagging scale factors, respectively. An additional uncertainty is included due to the extrapolation of these scale factors to jets with $p_T$ beyond the kinematic reach of the data calibration samples used ($p_T > 300$ GeV for $b$- and $c$-jets, and $p_T > 750$ GeV for light-jets), taken to be correlated among the three jet flavours. This uncertainty has a very small impact in this analysis (e.g. $< 0.2\%$ on the signal and background normalisations in the $(4j, 4b)$ channel).

7.3 Background modelling

A number of sources of systematic uncertainty affecting the modelling of $t\bar{t}$+jets are considered. A brief summary is provided below, with further details available in ref. [21], as the uncertainty treatment is identical. An uncertainty of $+6.1\%/-6.4\%$ is assumed for the inclusive $t\bar{t}$ production cross section [49], including contributions from varying the factorisation and renormalisation scales, and uncertainties arising from the PDF, $\alpha_S$, and the top quark mass. Uncertainties associated with the reweighting procedure applied to $t\bar{t}$+light-jets and $t\bar{t}+c\bar{c}$ processes include the nine leading sources of uncertainty in the differential cross-section measurement at $\sqrt{s} = 7$ TeV [61]. Additional uncertainties assigned to the modelling of the $t\bar{t}+c\bar{c}$ background include a 50% normalisation uncertainty, the full differences between applying and not applying the reweightings of the top quark and $t\bar{t}$ $p_T$ spectra, as well as smaller uncertainties associated with the choice of LO generator. Uncertainties affecting the modelling of $t\bar{t}+b\bar{b}$ production include a normalisation uncertainty of 50% (taken to be uncorrelated with the same uncertainty assigned to the $t\bar{t}+c\bar{c}$ background) and shape uncertainties (including inter-category migration effects) associated with the NLO prediction from SHERPA+OPENLOOPS, which is used for reweighting of the default POWHEG-BOX $t\bar{t}+b\bar{b}$ prediction. These include three different scale variations, a different shower-recoil model scheme, and two alternative PDF sets (MSTW and NNPDF). Additional uncertainties are assessed for the contributions to the $t\bar{t}+b\bar{b}$ background originating from multiple parton interactions or final-state radiation from top decay products, which are not part of the NLO prediction. Finally, an uncertainty due to the choice of parton shower and hadronisation model is derived by comparing events produced by POWHEG-BOX interfaced to PYTHIA or HERWIG. This uncertainty is taken to be uncorrelated between the $t\bar{t}$+light-jets, $t\bar{t}+c\bar{c}$ and $t\bar{t}+b\bar{b}$ processes.
Uncertainties affecting the modelling of the $W$+jets background include a 7% normalisation uncertainty for events with $\geq 4$ jets coming from the data-driven normalisation procedure. The corresponding normalisation uncertainty for $Z$+jets is 5% for events with $\geq 2$ jets. In addition, a 24% normalisation uncertainty is added in quadrature for each additional inclusive jet-multiplicity bin beyond the one where the background is normalised, based on a comparison among different algorithms for merging LO matrix elements and parton shower simulations [96]. For example, $W$+jets events with exactly 4 jets, exactly 5 jets and $\geq 6$ jets are assigned normalisation uncertainties of 7%, 7% + 24% = 31% and 7% + 24% + 24% = 35%. Finally, the full size of the $W$ and $Z$ boson $p_T$ correction, after symmetrisation, is taken as a systematic uncertainty. The above uncertainties are taken as uncorrelated between $W$+jets and $Z$+jets.

Uncertainties affecting the modelling of the single-top-quark background include a $+5%/-4%$ uncertainty on the total cross section, which is estimated as a weighted average of the theoretical uncertainties on $t$-, $Wt$- and $s$-channel production [68–70]. Similarly to the case of $W/Z$+jets, an additional 24% normalisation uncertainty is added in quadrature for each additional inclusive jet-multiplicity bin above $\geq 3$ jets. An additional systematic uncertainty on $Wt$-channel production concerning the separation between $t\bar{t}$ and $Wt$ at NLO [97] is assessed by comparing the nominal sample, which uses the so-called “diagram subtraction” scheme, with an alternative sample using the “diagram removal” scheme.

Uncertainties on the diboson background normalisation include 5% from the NLO theoretical cross sections [75] and additional 24% normalisation uncertainties added in quadrature for each additional inclusive jet-multiplicity bin above $\geq 2$ jets. Uncertainties on the $t\bar{t}V$ and $t\bar{t}H$ normalisations are 15% and $+9%/-12%$ respectively, from the uncertainties on their respective NLO theoretical cross sections [81, 89, 98]. Additional small uncertainties arising from scale variations, which change the amount of initial-state radiation and thus the event kinematics, are also considered.

Uncertainties on the data-driven multijet background estimate receive contributions from the limited sample size in data, particularly at high jet and $b$-tag multiplicities, as well as from the uncertainty on the rate of fake leptons, estimated in different control regions (e.g. selected with a requirement on either the maximum $E_T^{\text{miss}}$ or $m_T^W$). A combined normalisation uncertainty of 50% due to all these effects is assigned, which is taken as correlated across jet and $b$-tag multiplicity bins, but uncorrelated between electron and muon channels. No explicit shape uncertainty is assigned since the large statistical uncertainties associated with the multijet background prediction, which are uncorrelated between bins in the final discriminant distribution, effectively cover all possible shape uncertainties.

### 7.4 Signal modelling

Several normalisation and shape uncertainties are taken into account for the $t\bar{t} \rightarrow WbHq$ signal. The uncertainty on the $t\bar{t}$ cross section (see above) also applies to the $t\bar{t} \rightarrow WbHq$ signal and is taken to be the same as, and fully correlated with, that assigned to the $t\bar{t} \rightarrow WbWb$ background. Uncertainties on the $H \rightarrow b\bar{b}$ branching ratio are taken into account following the recommendation in ref. [89]: $\pm1.1\%$ ($\Delta\alpha_S$), $\pm1.4\%$ ($\Delta m_b$) and $\pm0.8\%$ (theory). Additional modelling uncertainties originate from non-closure of the reweighting
procedure applied to correct the distributions of top quark $p_T$ and $t\bar{t}$ system $p_T$ from Protos to match those from the uncorrected Powheg-Box+Pythia simulation, and the uncertainties associated with the further reweighting in the same variables to match the differential cross-section measurements at $\sqrt{s} = 7$ TeV, taken to be fully correlated with those assigned to the $t\bar{t}$+light-jets background. Finally, an uncertainty from the choice of parton shower and hadronisation model is estimated by comparing the $t\bar{t} \rightarrow WbWb$ yields between Powheg-Box+Pythia and Powheg-Box+Herwig in the channels with two $b$-tags, which are enriched in $t\bar{t}$+light-jets, and thus taken to be representative of what would be the signal acceptance uncertainty due to differences in extra jet radiation and $b$-quark fragmentation between the two parton shower/hadronisation models.

8 Statistical analysis

The distributions of the final discriminants in each of the analysis channels considered are combined to test for the presence of a signal. The statistical analysis is based on a binned likelihood function $L(\mu, \theta)$ constructed as a product of Poisson probability terms over all bins considered in the search. In the case of several searches being combined, the product of Poisson probability terms is extended over all bins considered in all searches. The function $L(\mu, \theta)$ depends on the signal-strength parameter $\mu$, defined as a multiplicative factor to the yield for $t\bar{t} \rightarrow WbHq$ signal events normalised to a reference branching ratio $BR_{\text{ref}}(t \rightarrow Hq) = 1\%$, and $\theta$, a set of nuisance parameters that encode the effect of systematic uncertainties on the signal and background expectations and are implemented in the likelihood function as Gaussian or log-normal priors with their width parameters corresponding to the size of the respective uncertainties. The relationship between $\mu$ and the corresponding $BR(t \rightarrow Hq)$ is:

$$\mu = \frac{BR(t \rightarrow Hq)[1 - BR(t \rightarrow Hq)]}{BR_{\text{ref}}(t \rightarrow Hq)[1 - BR_{\text{ref}}(t \rightarrow Hq)].}$$

Equation (8.1)

For a given $\mu$ value, the SM $t\bar{t} \rightarrow WbWb$ background contribution is scaled accordingly in order to preserve the inclusive $t\bar{t}$ cross section. The corresponding multiplicative factor would be $[1 - BR(t \rightarrow Hq)]^2$, with $BR(t \rightarrow Hq)$ being a function of $\mu$ as can be derived from equation (8.1):

$$BR(t \rightarrow Hq) = \frac{1 - \sqrt{1 - 4BR_{\text{ref}}(t \rightarrow Hq)(1 - BR_{\text{ref}}(t \rightarrow Hq))\mu}}{2}.$$  

Equation (8.2)

Therefore, the total number of signal and background events in a given bin depends on $\mu$ and $\theta$. The best-fit $BR(t \rightarrow Hq)$ is obtained by performing a binned likelihood fit to the data under the signal-plus-background hypothesis, i.e. maximising the likelihood function $L(\mu, \theta)$ over $\mu$ and $\theta$. The nuisance parameters $\theta$ allow variations of the expectations for signal and background according to the corresponding systematic uncertainties, and their fitted values correspond to the deviations from the nominal expectations that globally provide the best fit to the data. This procedure allows a reduction of the impact of systematic uncertainties on the search sensitivity by taking advantage of the highly populated background-dominated channels included in the likelihood fit.
The test statistic \( q_{\mu} \) is defined as the profile likelihood ratio: 
\[
q_{\mu} = -2 \ln \left( \frac{L(\mu, \hat{\theta})}{L(\hat{\mu}, \hat{\theta})} \right),
\]
where \( \hat{\mu} \) and \( \hat{\theta} \) are the values of the parameters that maximise the likelihood function (with the constraint \( 0 \leq \hat{\mu} \leq \mu \)), and \( \hat{\theta}_{\mu} \) are the values of the nuisance parameters that maximise the likelihood function for a given value of \( \mu \). Statistical uncertainties in each bin of the discriminant distributions are also taken into account via dedicated nuisance parameters in the fit. The test statistic \( q_{\mu} \) is implemented in the RooFit package [99, 100] and is used to measure the compatibility of the observed data with the background-only hypothesis by setting \( \mu = 0 \) in the profile likelihood ratio: 
\[
q_{0} = -2 \ln \left( \frac{L(0, \hat{\theta}_{0})}{L(\hat{\mu}, \hat{\theta})} \right).
\]
The \( p \)-value (referred to as \( p_{0} \)) representing the compatibility of the data with the background-only hypothesis is estimated by integrating the distribution of \( q_{0} \) from background-only pseudo-experiments, approximated using the asymptotic formulae given in ref. [101], above the observed value of \( q_{0} \). The observed \( p_{0} \)-value is checked for each explored signal scenario. In the absence of any significant excess above the background expectation, upper limits on \( \mu \), and thus on \( \text{BR}(t \rightarrow Hq) \) via equation (8.2), are derived by using \( q_{\mu} \) in the CLs method [102, 103]. Values of \( \text{BR}(t \rightarrow Hq) \) yielding \( \text{CL}_{s} < 0.05 \), where \( \text{CL}_{s} \) is computed using the asymptotic approximation [101], are excluded at \( \geq 95\% \) CL.

9 Results

This section presents the results obtained from the individual searches for \( t\bar{t} \rightarrow WbHq \), as well as their combination.

9.1 \( H \rightarrow b\bar{b} \)

Following the statistical analysis discussed in section 8, a binned likelihood fit under the signal-plus-background hypothesis is performed on the distributions of the final discriminant in the nine analysis channels considered. Figures 4–6 show a comparison of the data and prediction in the final discriminant in each of the analysis channels, both pre- and post-fit to data, in the case of the \( t\bar{t} \rightarrow WbHc \) search. The post-fit yields can be found in appendix A. The best-fit branching ratio obtained is \( \text{BR}(t \rightarrow Hc) = [0.17 \pm 0.12 \text{ (stat.)} \pm 0.17 \text{ (syst.)}]\% \), assuming that \( \text{BR}(t \rightarrow Hu) = 0 \). A similar fit is performed for the \( t\bar{t} \rightarrow WbHu \) search, yielding \( \text{BR}(t \rightarrow Hu) = [-0.07 \pm 0.17 \text{ (stat.)} \pm 0.28 \text{ (syst.)}]\% \), assuming that \( \text{BR}(t \rightarrow Hc) = 0 \). The different measured values for the two branching ratios is the result of the different sensitivities of the \( t\bar{t} \rightarrow WbHc \) and \( t\bar{t} \rightarrow WbHu \) searches, as discussed in section 6.1.

The large number of events in the analysis channels considered, together with their different background compositions, allows the fit to place constraints on the combined effect of several sources of systematic uncertainty. As a result, an improved background prediction is obtained with significantly reduced uncertainty, not only in the signal-depleted channels, but also in the most sensitive analysis channels for this search, (4 j, 3 b) and (4 j, 4 b). The channels with two \( b \)-tags are used to constrain the leading uncertainties affecting the \( t\bar{t} \)+light-jets background prediction, while the channels with \( \geq 5 \) jets and \( \geq 3 \) \( b \)-tags are
sensitive to the uncertainties affecting the $t\bar{t}+\text{HF}$ background prediction. In particular, one of the main corrections applied by the fit is an increase of the $t\bar{t} + b\bar{b}$ normalisation by approximately 20% relative to the nominal prediction by adjusting the corresponding nuisance parameter. This results in an improved agreement between data and prediction in the (5 jet, $\geq 4$ b) and (6 jet, $\geq 4$ b) channels, where the $t\bar{t} + b\bar{b}$ process provides a significant, or, in the latter case, dominant background contribution. In addition, the corresponding uncertainty is reduced from the initial 50% down to about 16%. This correction is in agreement with that found in ref. [21]. However, in contrast with ref. [21], the $t\bar{t} + b\bar{b}$ normalisation is not one of the leading uncertainties affecting this search, since the $t\bar{t} + b\bar{b}$ background is subdominant in the (4 jet, 3 b) channel, and it has a very different final discriminant shape from that of the signal in the (4 jet, 4 b) channel.

As an illustration, figure 7 provides a summary of the leading 15 systematic uncertainties affecting the $t\bar{t} \rightarrow WbHc$ search, quantifying their impact on the signal strength $\mu$, both before and after the fit, and displaying the constraints provided by the data on the associated nuisance parameters. The pre-fit impact on $\mu$ is estimated by fixing the corresponding nuisance parameter at $\theta_0 \pm \Delta \theta$, where $\theta_0$ is the nominal value of the nuisance parameter and $\Delta \theta$ is its pre-fit uncertainty, and performing the fit again. The difference between the default and modified $\mu$, $\Delta \mu$, represents the effect of the systematic uncertainty in question on $\mu$. The same procedure is followed to estimate the post-fit impact on $\mu$, but the corresponding nuisance parameter is instead fixed at $\hat{\theta} \pm \sigma_{\theta}$, where $\hat{\theta}$ is the fitted value of the nuisance parameter and $\sigma_{\theta}$ is its post-fit uncertainty. For reference, $\Delta \mu = 0.05$ corresponds to $\Delta \text{BR}(t \rightarrow Hc) \simeq 0.05\%$.

Prior to the fit, the systematic uncertainties with the largest impact on $\mu$ are the leading uncertainty for light-jet tagging and the uncertainty on the $t\bar{t}$ background associated with the choice of parton shower and hadronisation models. The significant impact from light-jet tagging results from the large fraction of $t\bar{t}$+light-jets background present in the (4 jet, 4 b) channel, which peaks at high values of the final discriminant, like the signal, and thus cannot be strongly constrained by the fit. Because of this, this uncertainty remains the leading one after the fit. In contrast, the uncertainty related to $t\bar{t}$ modelling is significantly constrained by the fit since it has a large impact ($\sim 5$–16%) on the $t\bar{t}$+light-jets background normalisation in the highly populated channels with two $b$-tags. As a result, this uncertainty is ranked only fourth in importance after the fit, becoming comparable to uncertainties such as the choice of renormalisation scale for $t\bar{t} + b\bar{b}$, the leading uncertainty for $c$-jet tagging and the $t\bar{t} + c\bar{c}$ normalisation. Of these, the nuisance parameter associated with the choice of the renormalisation scale for $t\bar{t} + b\bar{b}$ is slightly pulled (by half of the prior uncertainty) to improve agreement with the data in the (4 jet, 4 b) and (5 jet, 4 b) channels. In these channels, this uncertainty causes variations of up to $\sim 5\%$ in the bin contents in some regions of the final discriminant, i.e. distorting its shape compared to that of the nominal prediction, but the sensitivity is not sufficient to constrain it significantly. The leading uncertainty from $c$-tagging causes small (few percent) distortions in the shape of

---

6The overall change in $t\bar{t} + b\bar{b}$ normalisation can be different across channels due to the different impact of other nuisance parameters affecting $t\bar{t} + b\bar{b}$ modelling, such as that related to parton shower and hadronisation, which is changed by the fit by half of the prior uncertainty.
the background, and also cannot be constrained by the fit. In contrast, the fit is sensitive to the $t\bar{t} + c\bar{c}$ normalisation and the second-leading uncertainty for $c$-tagging, through the comparison of data and predictions across channels with different $b$-tag multiplicity, yielding results in agreement with the nominal predictions but with half the initial uncertainties.

Other nuisance parameters have a smaller impact on the signal extraction and typically have small pulls or constraints. One exception is the nuisance parameter associated with the $t\bar{t}$ cross section, which affects the signal extraction indirectly through the existing small fraction of non-$t\bar{t}$ background, and after the fit is found to be consistent with the nominal prediction but is constrained owing to the large number of $t\bar{t}$ events. On the other hand, a slight pull is obtained for the nuisance parameter associated with one of the uncertainties for the top quark $p_T$ and $t\bar{t}$ system $p_T$ reweightings, which is used by the fit to improve agreement between data and prediction in the channels with two $b$-tags but which has very small effect on the background prediction in the signal region.

In the absence of a significant excess in data above the background expectation, 95% CL limits are set on $\text{BR}(t \rightarrow Hc)$ and $\text{BR}(t \rightarrow Hu)$. The observed (expected) 95% CL upper limits on the branching ratios are $\text{BR}(t \rightarrow Hc) < 0.56\% (0.42\%)$ and $\text{BR}(t \rightarrow Hu) < 0.61\% (0.64\%)$. These upper limits can be translated into corresponding observed (expected) limits on the couplings of $|\lambda_{tcH}| < 0.14 (0.12)$ and $|\lambda_{tH}| < 0.15 (0.15)$.

9.2 $H \rightarrow \gamma\gamma$

A search for $t\bar{t} \rightarrow WbHq, H \rightarrow \gamma\gamma$ published by the ATLAS Collaboration uses a data set corresponding to an integrated luminosity of 4.5 fb$^{-1}$ at $\sqrt{s} = 7$ TeV and 20.3 fb$^{-1}$ at $\sqrt{s} = 8$ TeV [17]. The event selection requires at least two reconstructed photon candidates and additional requirements to select $t \rightarrow Wb$ decays. Events are categorised into two channels, leptonic and hadronic, depending on the $W$ boson decay modes. The leptonic channel selects events with exactly one lepton ($e$ or $\mu$), at least two jets, and at least one $b$-tagged jet. The hadronic channel selects events with no reconstructed lepton, at least four jets, and at least one $b$-tagged jet. In both channels, additional requirements are made to select events compatible with $t\bar{t} \rightarrow WbHq$ production by exploiting the invariant masses of the reconstructed top quark candidates. Finally, the diphoton mass ($m_{\gamma\gamma}$) distribution of the selected events is analysed using a sideband technique in order to estimate the background in the signal region, defined to be $122 \text{ GeV} \leq m_{\gamma\gamma} \leq 129 \text{ GeV}$.

Based on the above strategy, this search has essentially no discrimination power between $t\bar{t} \rightarrow WbHc$ and $t\bar{t} \rightarrow WbHu$ signals, because their selection acceptances are very close, although not identical. To facilitate combining it with the other searches discussed in this paper, minor modifications to the inputs were made with respect to the published result, all having a negligible impact on the result. They include updates to the $t\bar{t}$ cross-section uncertainty and the uncertainty model for Higgs branching ratios, as well as the separate treatment of $t\bar{t} \rightarrow WbHc$ and $t\bar{t} \rightarrow WbHu$ signals taking into account their slightly different acceptances. The best-fit branching ratios obtained are $\text{BR}(t \rightarrow Hc) =$

7The main effect of this uncertainty is a change in normalisation for the background with almost no effect on its shape in the final discriminant.
The signal (solid red) is normalised to BR(\(t \rightarrow Hc\)) = 1% and the \(t \rightarrow WbWb\) background is normalised to the SM prediction, while in the post-fit distributions both signal and \(t \rightarrow WbWb\) background are normalised using the best-fit BR(\(t \rightarrow Hc\)). The small contributions from W/Z+jets, single top, diboson and multijet backgrounds are combined into a single background source referred to as “Non-t\(\bar{t}\)”. The bottom panels display the ratios of data to either the SM background prediction before the fit (“Bkg”) or the total signal-plus-background prediction after the fit (“Pred”). The hashed area represents the total uncertainty on the background.
Figure 5. $t\bar{t} \rightarrow WbHc, H \rightarrow b\bar{b}$ search: comparison between the data and prediction for the distribution of the $D$ discriminant used in the (4 j, 3 b) channel (a) before the fit and (b) after the fit, in the (5 j, 3 b) channel (c) before the fit and (d) after the fit, and in the ($\geq 6$ j, 3 b) channel (e) before the fit and (f) after the fit. The fit is performed on data under the signal-plus-background hypothesis. In the pre-fit distributions the $t\bar{t} \rightarrow WbHc$ signal (solid red) is normalised to $\text{BR}(t \rightarrow Hc) = 1\%$ and the $t\bar{t} \rightarrow WbWb$ background is normalised to the SM prediction, while in the post-fit distributions both signal and $t\bar{t} \rightarrow WbWb$ background are normalised using the best-fit $\text{BR}(t \rightarrow Hc)$. The small contributions from $W/Z + \text{jets}$, single top, diboson and multijet backgrounds are combined into a single background source referred to as “Non-$t\bar{t}$”. The bottom panels display the ratios of data to either the SM background prediction before the fit (“Bkg”) or the total signal-plus-background prediction after the fit (“Pred”). The hashed area represents the total uncertainty on the background.
Figure 6. $t\bar{t} \rightarrow WbHc, H \rightarrow bb$ search: comparison between the data and prediction for the distribution of the $D$ discriminant used in the (4 j, 4 b) channel (a) before the fit and (b) after the fit, in the (5 j, $\geq 4$ b) channel (c) before the fit and (d) after the fit, and in the ($\geq 6$ j, $\geq 4$ b) channel (e) before the fit and (f) after the fit. The fit is performed on data under the signal-plus-background hypothesis. In the pre-fit distributions the $t\bar{t} \rightarrow WbHc$ signal (solid red) is normalised to BR($t \rightarrow Hc$) = 1% and the $t\bar{t} \rightarrow WbWb$ background is normalised to the SM prediction, while in the post-fit distributions both signal and $t\bar{t} \rightarrow WbWb$ background are normalised using the best-fit BR($t \rightarrow Hc$). The small contributions from $W/Z+$jets, single top, diboson and multijet backgrounds are combined into a single background source referred to as “Non-$t\bar{t}$”. The bottom panels display the ratios of data to either the SM background prediction before the fit (“Bkg”) or the total signal-plus-background prediction after the fit (“Pred”). The hashed area represents the total uncertainty on the background.
Figure 7. $t\bar{t} \rightarrow WbHc$, $H \rightarrow b\bar{b}$ search: the fitted values of the nuisance parameters for the most important sources of systematic uncertainty and their impact on the measured signal strength. The points, which are drawn conforming to the scale of the bottom axis, show the deviation of each of the fitted nuisance parameters, $\hat{\theta}$, from $\theta_0$, which is the nominal value of that nuisance parameter, in units of the pre-fit standard deviation $\Delta \theta$. The error bars show the post-fit uncertainties, $\sigma_\theta$, which are close to 1 if the data do not provide any further constraint on that uncertainty. Conversely, a value of $\sigma_\theta$ much smaller than 1 indicates a significant reduction with respect to the original uncertainty. The nuisance parameters are sorted according to their post-fit effect on $\mu$ (hashed blue area), conforming to the scale of the top axis, with those with the largest impact at the top.

$[0.22 \pm 0.26 \text{ (stat.)} \pm 0.10 \text{ (syst.)}]\%$ and $\text{BR}(t \rightarrow Hu) = [0.23 \pm 0.27 \text{ (stat.)} \pm 0.10 \text{ (syst.)}]\%$ under the assumptions that $\text{BR}(t \rightarrow Hu) = 0$ and $\text{BR}(t \rightarrow Hc) = 0$ respectively. The observed (expected) 95% CL upper limits on the branching ratios from ref. [17] remain the best estimates, $\text{BR}(t \rightarrow Hq) < 0.79\% (0.51\%)$. The corresponding limits on the couplings are $|\lambda_{tqH}| < 0.17(0.14)$. These limits can be understood as applying to the sum of the $t \rightarrow Hc$ and $t \rightarrow Hu$ decay modes, or only to one of them, if the other decay mode is assumed to have a branching ratio equal to zero. In the former case, $\text{BR}(t \rightarrow Hq) \equiv \text{BR}(t \rightarrow Hc) + \text{BR}(t \rightarrow Hu)$ and $|\lambda_{tqH}| \equiv \sqrt{|\lambda_{tcH}|^2 + |\lambda_{tuH}|^2}$.
9.3 $H \to W^+W^-, \tau^+\tau^-$

The $H \to WW^*$ and $H \to \tau\tau$ decay modes are predicted to have significant branching ratios, of 21.5% and 6.3% respectively. The resulting signatures for signal events, $t\bar{t} \to WbHq \to W^\pm W^\mp bq$ and $t\bar{t} \to WbHq \to W^\pm\tau^\mp\tau^\pm bq$, can be effectively exploited to suppress backgrounds in the case of multilepton final states.

Recently, the ATLAS Collaboration has published a search for $t\bar{t}H$ production with $H \to WW^*$, $\tau\tau$ and $ZZ^*$ [23] that exploits several multilepton signatures resulting from the leptonic decays of $W$ and $Z$ bosons and/or the presence of $\tau$ leptons. That search considered five separate event categories depending on the number of reconstructed electrons or muons and hadronic $\tau$ candidates, of which the following three are considered for a reinterpretation in the context of the $t\bar{t} \to WbHq$ search:

- $2\ell0\tau_{\text{had}}$: two same-charge light leptons (e or $\mu$) with no hadronic $\tau$ candidates ($\tau_{\text{had}}$) and $\geq4$ jets with $\geq1$ b-tagged jets. This channel is sensitive to the process $t\bar{t} \to WbHq \to \ell^\pm\ell^\mp qqb2\nu$. To further improve the sensitivity, this category is further subdivided into six subcategories depending on the flavour of the leptons and the number of jets: $(ee, \mu\mu, e\mu) \times (4$ jets, $\geq5$ jets).

- $3\ell$: three light leptons with either $\geq3$ jets of which $\geq2$ are b-tagged, or $\geq4$ jets of which $\geq1$ are b-tagged. This channel is sensitive to the process $t\bar{t} \to WbHq \to \ell^\pm\ell^\mp\ell^\mp\tau^\pm qqb3\nu$.

- $2\ell1\tau_{\text{had}}$: two same-charge light leptons (e or $\mu$), one $\tau_{\text{had}}$ candidate, and $\geq4$ jets with $\geq1$ b-tagged jets. This channel is sensitive to the process $t\bar{t} \to WbHq \to \ell^\pm\ell^\mp\tau^\pm qqb3\nu$.

The two other categories considered in the $t\bar{t}H$ search in multilepton final states but not in this reinterpretation are: 4$\ell$ (four light leptons with $\geq2$ jets and $\geq1$ b-tagged jets) and 1$\ell2\tau_{\text{had}}$ (one light lepton and two opposite-charge $\tau_{\text{had}}$ candidates, with $\geq3$ jets and $\geq1$ b-tagged jets), which have very small signal acceptance and/or poor signal-to-background ratio due to the large number of jets misidentified as $\tau_{\text{had}}$ candidates. The minimum number of jets required in the $2\ell0\tau_{\text{had}}$ category is well matched to the number of partons expected from $t\bar{t} \to WbHq \to W^\pm W^\mp W^\pm bq$ signal events, while the 3$\ell$ and 2$\ell1\tau_{\text{had}}$ categories effectively require one or two jets beyond leading order. Because of this, and the gain in branching ratio resulting from only requiring two, as opposed to three, W bosons decaying leptonically, the $2\ell0\tau_{\text{had}}$ category dominates the sensitivity.

The largest background in the most sensitive category, 2$\ell4j$, is $t\bar{t}$ or single-top production with one of the leptons originating from a decay of a heavy-flavour hadron (“non-prompt lepton”), followed by $t\bar{t}W$ production. Smaller contributions arise from $t\bar{t}(Z/\gamma^*)$, $t\bar{t}H$, and diboson (primarily $WZ$) production, and dilepton events from $t\bar{t}$ and $Z/\gamma^*$ with the wrong charge sign measured for one electron. In the remaining categories the non-prompt lepton background decreases in importance, relative to the prompt-lepton contributions. The prompt contributions are estimated using the simulation, as typically the relevant processes (e.g., $t\bar{t}V$, $V = W, Z$) have not been measured at an accuracy exceeding that of theoretical predictions. The non-prompt lepton background predictions are computed or validated using data control regions with similar lepton kinematic selections but
with fewer jets (two or three). Further details are provided in ref. [23]. In the search for \(ttH\), these regions have almost no contamination from signal, but the same is not necessarily true for the \(tt \rightarrow WbHq\) process, which is characterised by a lower jet multiplicity than the \(ttH\) process. This could potentially lead to an overestimate of the non-prompt background in this reinterpretation, particularly in the \(2\ell0\text{had}\) categories, where the effect is more pronounced. The \(3\ell\) category estimates the non-prompt lepton background in the signal region by extrapolating, in terms of lepton quality, from a data sideband region that is sufficiently depleted in signal, using normalisation factors obtained from the simulation.

The non-prompt lepton background estimate for the \(2\ell0\text{had}\) categories is validated using simulation and control regions with two or three jets in which the lower \(p_T\) (subleading) lepton satisfies \(10\text{ GeV} < p_T < 30\text{ GeV}\). Because the number of non-prompt leptons increases significantly as the lepton \(p_T\) requirement is lowered, this region is dominated by the non-prompt contribution for any non-excluded value of BR(\(t \rightarrow Hq\)). To extract the non-prompt background yields in these regions, binned likelihood fits to the subleading lepton \(p_T\) distribution are performed, accounting for non-prompt lepton backgrounds, sources of prompt leptons (normalised to the SM expectation), and a potential \(tt \rightarrow WbHq\) contribution. The fitted \(tt \rightarrow WbHq\) contribution is in all cases found to be compatible with zero. The obtained best-fit non-prompt background yields in these low-\(p_T\) control regions are used to normalise the simulation and obtain estimates for the background in the \(2\ell0\text{had}\) signal regions. These are found to be compatible with the nominal data-driven predictions within the stated uncertainties. The nominal predictions of the non-prompt backgrounds and their associated uncertainties are therefore used without modification in this reinterpretation.

Figure 8 shows the event yields in each of the categories, which are used as input to the statistical analysis discussed in section 8. A table summarising the expected and observed yields can be found in appendix B. The best-fit branching ratio obtained is \(\text{BR}(t \rightarrow Hc) = [0.27 \pm 0.18\text{ (stat.)} \pm 0.21\text{ (syst.)}]\%\) assuming that \(\text{BR}(t \rightarrow Hu) = 0\). A similar fit is performed for the \(tt \rightarrow WbHu\) search, yielding \(\text{BR}(t \rightarrow Hu) = [0.23 \pm 0.18\text{ (stat.)} \pm 0.21\text{ (syst.)}]\%\), assuming that \(\text{BR}(t \rightarrow Hc) = 0\). The observed (expected) 95\% CL upper limits on the branching ratios are \(\text{BR}(t \rightarrow Hc) < 0.79\% (0.54\%)\) and \(\text{BR}(t \rightarrow Hu) < 0.78\% (0.57\%)\). The corresponding observed (expected) limits on the couplings are \(|\lambda_{tcH}| < 0.17(0.14)\) and \(|\lambda_{tstuH}| < 0.17(0.14)\) at 95\% CL.

9.4 Combination of searches

The three searches discussed in the sections 9.1–9.3 are combined using the statistical analysis discussed in section 8. In this combination, the only systematic uncertainties taken to be fully correlated among the three searches are the \(tt\) cross section and the integrated luminosity for \(\sqrt{s} = 8\text{ TeV}\) data. The dominant uncertainties on the Higgs boson branching ratios primarily affect the \(tt \rightarrow WbHq, H \rightarrow \gamma\gamma\) and \(tt \rightarrow WbHq, H \rightarrow b\bar{b}\) searches. Other uncertainties such as those associated with leptons, jet energy scale and \(b\)-tagging should be partially correlated between the \(tt \rightarrow WbHq, H \rightarrow \gamma\gamma\) search and the other two searches, but the differences in treatment between analyses (different lepton selection and \(p_T\) cuts, no uncertainty breakdown for jet energy scale and \(b\)-tagging in the
Figure 8. \(t\bar{t} \rightarrow WbHq, H \rightarrow WW^*, \tau\tau\) search: comparison between the data and background prediction for the yields in each of the analysis channels considered before the fit to data. The expected \(t\bar{t} \rightarrow WbHc\) and \(t\bar{t} \rightarrow WbHu\) signals (dashed histograms) are shown separately normalised to \(\text{BR}(t \rightarrow Hq) = 1\%\). The sum of instrumental backgrounds originating from non-prompt leptons and lepton charge misidentification is denoted by “Non-prompt/q-misid”. The small contribution from rare processes such as \(tZ, tWW,\) triboson, \(t\bar{t}t\) and \(tH\) production are combined into a single background source denoted by “Rare”. The bottom panel displays the ratio of data to the SM background (“Bkg”) prediction. The hashed area represents the total uncertainty on the background.

\(t\bar{t} \rightarrow WbHq, H \rightarrow \gamma \gamma\) search) make it difficult to account for correlations. However, given that the \(t\bar{t} \rightarrow WbHq, H \rightarrow \gamma \gamma\) search is completely dominated by the data statistics, the effect of this simplification is negligible. The uncertainties taken as correlated between the \(t\bar{t} \rightarrow WbHq, H \rightarrow b\bar{b}\) and \(t\bar{t} \rightarrow WbHq, H \rightarrow WW^*, \tau\tau\) searches include those associated with lepton isolation, the leading \(b\)-tagging and jet energy scale uncertainties, and the reweighting of top quark \(p_T\) and \(t\bar{t}\) system \(p_T\). The rest of the uncertainties are taken to be uncorrelated among the searches. This correlation scheme closely follows the procedure adopted in the combination of \(tH\) searches by ATLAS [104].

The first set of combined results is obtained for each branching ratio separately, setting the other branching ratio to zero. The best-fit combined branching ratios are \(\text{BR}(t \rightarrow Hc) = [0.22 \pm 0.10\; (\text{stat}) \pm 0.10\; (\text{syst})]\%\) and \(\text{BR}(t \rightarrow Hu) = [0.16 \pm 0.11\; (\text{stat}) \pm 0.12\; (\text{syst})]\%\). The difference between the central values of \(\text{BR}(t \rightarrow Hc)\) and \(\text{BR}(t \rightarrow Hu)\) originates from the ability of the \(H \rightarrow b\bar{b}\) search to probe both decay modes separately. A comparison of the best-fit branching ratios for the individual searches and their combination can be found in figure 9 for \(\text{BR}(t \rightarrow Hc)\) and figure 10 for \(\text{BR}(t \rightarrow Hu)\). Figure 11 shows the CLs versus branching ratio for the combination. The observed (expected) 95% CL combined upper limits on the branching ratios are \(\text{BR}(t \rightarrow Hc) < 0.46\% \ (0.25\%\) and \(\text{BR}(t \rightarrow Hu) < 0.45\% \ (0.29\%). The corresponding observed (expected) upper limits on the couplings are \(|\lambda_{tcH}| < 0.13\ (0.10)\) and \(|\lambda_{tuH}| < 0.13\ (0.10)\). A summary of the upper limits on the branching ratios obtained by the individual searches, as well as their combination, can be found in figure 12.
A similar set of results can be obtained by simultaneously probing both branching ratios. Although the $t\bar{t} \rightarrow WbHq$, $H \rightarrow \gamma\gamma$ and $t\bar{t} \rightarrow WbHq$, $H \rightarrow WW^*, \tau\tau$ searches are basically only sensitive to the sum of the two branching ratios, the $t\bar{t} \rightarrow WbHq$, $H \rightarrow b\bar{b}$ search has different sensitivity to each of them, and a simultaneous fit can be performed. The best-fit branching ratios obtained from the simultaneous fit are $\text{BR}(t \rightarrow Hc) = [0.34 \pm 0.22 \text{ (stat.)} \pm 0.15 \text{ (syst.)}]\%$ and $\text{BR}(t \rightarrow Hu) = [−0.17 \pm 0.25 \text{ (stat.)} \pm 0.17 \text{ (syst.)}]\%$, with a correlation coefficient of $−0.84$, as shown in figure 13. Figure 14(a) shows the 95% CL upper limits on the branching ratios in the BR($t \rightarrow Hu$) versus BR($t \rightarrow Hc$) plane. The corresponding upper limits on the couplings in the $|\lambda_{tuH}|$ versus $|\lambda_{tcH}|$ plane can be found in figure 14(b).
Figure 11. (a) CL$_s$ versus BR($t \to Hc$) and (b) CL$_s$ versus BR($t \to Hu$) for the combination of the searches, assuming that the other branching ratio is zero. The observed CL$_s$ values (solid black lines) are compared to the expected (median) CL$_s$ values under the background-only hypothesis (dotted black lines). The surrounding shaded bands correspond to the 68% and 95% CL intervals around the expected CL$_s$ values, denoted by $\pm 1\sigma$ and $\pm 2\sigma$, respectively. The solid red line at CL$_s$=0.05 denotes the value below which the hypothesis is excluded at 95% CL.

Figure 12. 95% CL upper limits on (a) BR($t \to Hc$) and (b) BR($t \to Hu$) for the individual searches as well as their combination, assuming that the other branching ratio is zero. The observed limits (solid lines) are compared to the expected (median) limits under the background-only hypothesis (dotted lines). The surrounding shaded bands correspond to the 68% and 95% CL intervals around the expected limits, denoted by $\pm 1\sigma$ and $\pm 2\sigma$, respectively. Because the asymptotic approximation is used in the calculation of CL$_s$, the results of the $H \to \gamma \gamma$ search reported in this figure differ slightly from those published in ref. [17], which remain the most accurate results.
Hc) [%] \rightarrow BR(t) Hc) [%] \rightarrow BR(t) Hu) [%] \rightarrow BR(t)

Figure 13. Best-fit BR(t \rightarrow Hc) and BR(t \rightarrow Hu) and the corresponding 68% CL (solid) and 95% CL (dotted) regions for the combination of the searches.

Figure 14. 95% CL upper limits (a) on the plane of BR(t \rightarrow Hu) versus BR(t \rightarrow Hc) and (b) on the plane of |\lambda_{tuH}| versus |\lambda_{tcH}| for the combination of the searches. The observed limits (solid lines) are compared to the expected (median) limits under the background-only hypothesis (dotted lines). The surrounding shaded bands correspond to the 68% and 95% CL intervals around the expected limits, denoted by ±1σ and ±2σ, respectively.

10 Conclusion

A search for flavour-changing neutral current decays of a top quark to an up-type quark (q = u, c) and the Standard Model Higgs boson, where the Higgs boson decays to bb, is presented. The analysis searches for top quark pair events in which one top quark decays to Wb, with the W boson decaying leptonically, and the other top quark decays to Hq. The search is based on pp collisions at \sqrt{s} = 8 TeV recorded in 2012 with the ATLAS detector at the CERN Large Hadron Collider and uses an integrated luminosity of 20.3 fb⁻¹. Data are analysed in the lepton-plus-jets final state, characterised by an isolated electron or muon with moderately high transverse momentum and at least four jets. In this search the dominant background process is t\bar{t} \rightarrow WbWb. To separate the signal from the background, the search exploits the high multiplicity of b-quark jets characteristic of signal events, and
employs a likelihood discriminant that combines information from invariant mass distributions and the flavour of the jets. No significant excess of events above the background expectation is found, and observed (expected) 95% CL upper limits of 0.56% (0.42%) and 0.61% (0.64%) are derived for the $t \rightarrow Hc$ and $t \rightarrow Hu$ branching ratios respectively.

Results from other ATLAS searches are also summarised, including a previous search probing $H \rightarrow \gamma \gamma$ decays, and a reinterpretation of a search for $tH$ production in multilepton final states, which exploits the $H \rightarrow WW^*, \tau\tau$ decay modes. All searches discussed in this paper have comparable sensitivity, and thus their combination represents a significant improvement over the individual results. The observed (expected) 95% CL combined upper limits on the $t \rightarrow Hc$ and $t \rightarrow Hu$ branching ratios are 0.46% (0.25%) and 0.45% (0.29%) respectively. The corresponding observed (expected) upper limits on the $|\lambda_{tH}|$ and $|\lambda_{taH}|$ couplings are 0.13 (0.10) and 0.13 (0.10) respectively. Upper limits in the $t \rightarrow Hc$ versus $t \rightarrow Hu$ branching ratio plane, as well as best-fit branching ratios, are also reported. These are the most restrictive direct bounds on $tgqH \ (g = u, c)$ interactions measured so far.

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, DNSRC and Lundbeck Foundation, 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; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNISW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC Ki, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, 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, Region 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; BSSF, 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.
A Pre-fit and post-fit event yields in the $t\bar{t} \rightarrow WbHq, H \rightarrow b\bar{b}$ search

Table 3 presents the observed and predicted yields in each of the analysis channels for the $t\bar{t} \rightarrow WbHq, H \rightarrow b\bar{b}$ search before the fit to data. Table 4 presents the observed and predicted yields in each of the analysis channels for the $t\bar{t} \rightarrow WbHc, H \rightarrow b\bar{b}$ search, after the fit to the data under the signal-plus-background hypothesis.

<table>
<thead>
<tr>
<th>Channel</th>
<th>4 j, 2 b</th>
<th>4 j, 3 b</th>
<th>4 j, 4 b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t} \rightarrow WbHc$</td>
<td>890 ± 100</td>
<td>394 ± 54</td>
<td>41.6 ± 7.2</td>
</tr>
<tr>
<td>$t\bar{t} \rightarrow WbHu$</td>
<td>851 ± 98</td>
<td>339 ± 49</td>
<td>3.81 ± 0.71</td>
</tr>
<tr>
<td>$t\bar{t}$ + light-jets</td>
<td>77400 ± 8100</td>
<td>6170 ± 860</td>
<td>53 ± 12</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>4900 ± 2600</td>
<td>680 ± 370</td>
<td>21 ± 11</td>
</tr>
<tr>
<td>$t\bar{t} + b\bar{b}$</td>
<td>1870 ± 990</td>
<td>680 ± 370</td>
<td>44 ± 23</td>
</tr>
<tr>
<td>$t\bar{t}V$</td>
<td>121 ± 21</td>
<td>15.5 ± 2.9</td>
<td>0.89 ± 0.19</td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>30.5 ± 4.2</td>
<td>12.7 ± 1.9</td>
<td>1.91 ± 0.34</td>
</tr>
<tr>
<td>$W$ + jets</td>
<td>4700 ± 1600</td>
<td>217 ± 78</td>
<td>5.4 ± 2.0</td>
</tr>
<tr>
<td>$Z$ + jets</td>
<td>1080 ± 450</td>
<td>50 ± 22</td>
<td>0.90 ± 0.50</td>
</tr>
<tr>
<td>Single top</td>
<td>4900 ± 1400</td>
<td>340 ± 100</td>
<td>6.8 ± 2.3</td>
</tr>
<tr>
<td>Diboson</td>
<td>212 ± 75</td>
<td>11.5 ± 4.1</td>
<td>0.24 ± 0.11</td>
</tr>
<tr>
<td>Multijet</td>
<td>1540 ± 550</td>
<td>100 ± 36</td>
<td>3.4 ± 1.2</td>
</tr>
<tr>
<td>Total background</td>
<td>96800 ± 9600</td>
<td>8300 ± 1100</td>
<td>138 ± 32</td>
</tr>
<tr>
<td>Data</td>
<td>98049</td>
<td>8752</td>
<td>161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel</th>
<th>5 j, 2 b</th>
<th>5 j, 3 b</th>
<th>5 j, 4 b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t} \rightarrow WbHc$</td>
<td>483 ± 96</td>
<td>242 ± 50</td>
<td>35.1 ± 7.7</td>
</tr>
<tr>
<td>$t\bar{t} \rightarrow WbHu$</td>
<td>473 ± 95</td>
<td>217 ± 46</td>
<td>8.4 ± 2.0</td>
</tr>
<tr>
<td>$t\bar{t}$ + light-jets</td>
<td>37600 ± 6600</td>
<td>3480 ± 750</td>
<td>61 ± 18</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>4300 ± 2300</td>
<td>810 ± 460</td>
<td>43 ± 28</td>
</tr>
<tr>
<td>$t\bar{t} + b\bar{b}$</td>
<td>1670 ± 860</td>
<td>890 ± 470</td>
<td>115 ± 61</td>
</tr>
<tr>
<td>$t\bar{t}V$</td>
<td>145 ± 24</td>
<td>26.5 ± 4.5</td>
<td>3.10 ± 0.60</td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>40.9 ± 4.8</td>
<td>22.3 ± 2.9</td>
<td>5.96 ± 0.98</td>
</tr>
<tr>
<td>$W$ + jets</td>
<td>1850 ± 790</td>
<td>131 ± 57</td>
<td>5.8 ± 2.7</td>
</tr>
<tr>
<td>$Z$ + jets</td>
<td>400 ± 200</td>
<td>29 ± 14</td>
<td>1.47 ± 0.76</td>
</tr>
<tr>
<td>Single top</td>
<td>1880 ± 740</td>
<td>195 ± 78</td>
<td>8.3 ± 3.1</td>
</tr>
<tr>
<td>Diboson</td>
<td>96 ± 41</td>
<td>8.0 ± 3.5</td>
<td>0.90 ± 0.19</td>
</tr>
<tr>
<td>Multijet</td>
<td>450 ± 160</td>
<td>68 ± 24</td>
<td>8.3 ± 3.0</td>
</tr>
<tr>
<td>Total background</td>
<td>48400 ± 7800</td>
<td>5700 ± 1100</td>
<td>252 ± 75</td>
</tr>
<tr>
<td>Data</td>
<td>49699</td>
<td>6199</td>
<td>286</td>
</tr>
</tbody>
</table>

Table 3.
Table 3 (cont’d). \( \bar{t}t \rightarrow WbHq, H \rightarrow b\bar{b} \) search: predicted and observed yields in each of the analysis channels considered. The prediction is shown before the fit to data. Also shown are the signal expectations for \( \bar{t}t \rightarrow WbHc \) and \( \bar{t}t \rightarrow WbHu \) assuming \( \text{BR}(t \rightarrow Hc) = 1\% \) and \( \text{BR}(t \rightarrow Hu) = 1\% \) respectively. The \( \bar{t}t \rightarrow WbWb \) background is normalised to the SM prediction. The quoted uncertainties are the sum in quadrature of statistical and systematic uncertainties on the yields.

<table>
<thead>
<tr>
<th>Channel</th>
<th>( \geq 6 \ j, 2 \ b )</th>
<th>( \geq 6 \ j, 3 \ b )</th>
<th>( \geq 6 \ j, 4 \ b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{t}t \rightarrow WbHc )</td>
<td>267 ± 68</td>
<td>145 ± 37</td>
<td>31.1 ± 8.3</td>
</tr>
<tr>
<td>( \bar{t}t \rightarrow WbHu )</td>
<td>259 ± 67</td>
<td>132 ± 34</td>
<td>10.3 ± 2.8</td>
</tr>
<tr>
<td>( \bar{t}t + \text{light-jets} )</td>
<td>18800 ± 4800</td>
<td>2000 ± 730</td>
<td>52 ± 40</td>
</tr>
<tr>
<td>( \bar{t}t + c\bar{c} )</td>
<td>3700 ± 2000</td>
<td>850 ± 500</td>
<td>79 ± 46</td>
</tr>
<tr>
<td>( \bar{t}t + b\bar{b} )</td>
<td>1430 ± 760</td>
<td>970 ± 520</td>
<td>240 ± 130</td>
</tr>
<tr>
<td>( \bar{t}tV )</td>
<td>182 ± 32</td>
<td>44.6 ± 8.1</td>
<td>8.4 ± 1.7</td>
</tr>
<tr>
<td>( \bar{t}tH )</td>
<td>64.2 ± 8.2</td>
<td>39.8 ± 5.4</td>
<td>16.1 ± 2.6</td>
</tr>
<tr>
<td>W+jets</td>
<td>880 ± 440</td>
<td>95 ± 47</td>
<td>8.5 ± 4.5</td>
</tr>
<tr>
<td>Z+jets</td>
<td>180 ± 100</td>
<td>19 ± 11</td>
<td>1.5 ± 0.9</td>
</tr>
<tr>
<td>Single top</td>
<td>840 ± 410</td>
<td>122 ± 62</td>
<td>11.9 ± 6.2</td>
</tr>
<tr>
<td>Diboson</td>
<td>50 ± 26</td>
<td>6.0 ± 3.0</td>
<td>0.54 ± 0.29</td>
</tr>
<tr>
<td>Multijet</td>
<td>176 ± 62</td>
<td>20.3 ± 7.2</td>
<td>0.93 ± 0.50</td>
</tr>
<tr>
<td>Total background</td>
<td>26400 ± 6100</td>
<td>4200 ± 1200</td>
<td>420 ± 160</td>
</tr>
<tr>
<td>Data</td>
<td>26185</td>
<td>4701</td>
<td>516</td>
</tr>
</tbody>
</table>

Table 4.
Table 4 (cont’d). $t\bar{t} \rightarrow WbHc, H \rightarrow b\bar{b}$ search: predicted and observed yields in each of the analysis channels considered. The background prediction is shown after the fit to data under the signal-plus-background hypothesis. The quoted uncertainties are the sum in quadrature of statistical and systematic uncertainties on the yields, computed taking into account correlations among nuisance parameters and among processes.
B Pre-fit event yields in the $t\bar{t} \to WbHq, H \to WW^*, \tau\tau$ search

Table 5 presents the observed and predicted yields in each of the analysis channels for the $t\bar{t} \to WbHq, H \to WW^*, \tau\tau$ search before the fit to data.

<table>
<thead>
<tr>
<th></th>
<th>$ee4j$</th>
<th>$e\mu4j$</th>
<th>$\mu\mu4j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t} \to WbHc$</td>
<td>$5.4^{+0.8}_{-0.6}$</td>
<td>$15.0^{+1.7}_{-1.4}$</td>
<td>$11.1^{+1.1}_{-1.2}$</td>
</tr>
<tr>
<td>$t\bar{t} \to WbHu$</td>
<td>$5.5 \pm 0.7$</td>
<td>$15.9^{+1.7}_{-1.5}$</td>
<td>$9.7^{+1.0}_{-1.5}$</td>
</tr>
<tr>
<td>Non-prompt</td>
<td>$3.4 \pm 1.7$</td>
<td>$12 \pm 4$</td>
<td>$6.3 \pm 2.6$</td>
</tr>
<tr>
<td>$q$ mis-id</td>
<td>$1.8 \pm 0.7$</td>
<td>$1.4 \pm 0.6$</td>
<td>—</td>
</tr>
<tr>
<td>$t\bar{t}W$</td>
<td>$2.0 \pm 0.4$</td>
<td>$6.2 \pm 1.0$</td>
<td>$4.7 \pm 0.9$</td>
</tr>
<tr>
<td>$t\bar{t}(Z/\gamma^*)$</td>
<td>$0.75 \pm 0.20$</td>
<td>$1.5 \pm 0.3$</td>
<td>$0.80 \pm 0.22$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$0.7 \pm 0.4$</td>
<td>$1.9 \pm 1.0$</td>
<td>$0.53 \pm 0.30$</td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>$0.44 \pm 0.06$</td>
<td>$1.16 \pm 0.14$</td>
<td>$0.74 \pm 0.10$</td>
</tr>
<tr>
<td>Rare</td>
<td>$0.25^{+0.04}_{-0.02}$</td>
<td>$0.72 \pm 0.05$</td>
<td>$0.34^{+0.04}_{-0.03}$</td>
</tr>
<tr>
<td>Total background</td>
<td>$9.5 \pm 2.1$</td>
<td>$25 \pm 5$</td>
<td>$13.4 \pm 2.9$</td>
</tr>
<tr>
<td>Data</td>
<td>9</td>
<td>26</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$ee\geq5j$</th>
<th>$e\mu\geq5j$</th>
<th>$\mu\mu\geq5j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t} \to WbHc$</td>
<td>$2.9 \pm 0.6$</td>
<td>$10.2^{+2.0}_{-1.6}$</td>
<td>$6.6^{+1.5}_{-1.3}$</td>
</tr>
<tr>
<td>$t\bar{t} \to WbHu$</td>
<td>$2.7 \pm 0.8$</td>
<td>$8.2^{+1.6}_{-1.3}$</td>
<td>$7.2^{+1.2}_{-1.4}$</td>
</tr>
<tr>
<td>Non-prompt</td>
<td>$2.3 \pm 1.2$</td>
<td>$6.7 \pm 2.4$</td>
<td>$2.9 \pm 1.4$</td>
</tr>
<tr>
<td>$q$ mis-id</td>
<td>$1.1 \pm 0.5$</td>
<td>$0.85 \pm 0.35$</td>
<td>—</td>
</tr>
<tr>
<td>$t\bar{t}W$</td>
<td>$1.4 \pm 0.4$</td>
<td>$4.8 \pm 1.2$</td>
<td>$3.8 \pm 0.9$</td>
</tr>
<tr>
<td>$t\bar{t}(Z/\gamma^*)$</td>
<td>$0.98 \pm 0.26$</td>
<td>$2.1 \pm 0.5$</td>
<td>$0.95 \pm 0.25$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$0.47 \pm 0.29$</td>
<td>$0.38 \pm 0.30$</td>
<td>$0.7 \pm 0.4$</td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>$0.73 \pm 0.14$</td>
<td>$2.1 \pm 0.4$</td>
<td>$1.41 \pm 0.28$</td>
</tr>
<tr>
<td>Rare</td>
<td>$0.27 \pm 0.02$</td>
<td>$0.79^{+0.05}_{-0.04}$</td>
<td>$0.38 \pm 0.02$</td>
</tr>
<tr>
<td>Total background</td>
<td>$7.2 \pm 1.8$</td>
<td>$17 \pm 3$</td>
<td>$10.0 \pm 2.2$</td>
</tr>
<tr>
<td>Data</td>
<td>10</td>
<td>22</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 5.
Table 5 (cont’d). $t\bar{t}\rightarrow WbHq, H \rightarrow WW^{*}, \tau\tau$ search: predicted and observed yields in each of the event categories considered. The prediction is shown before the fit to data. Also shown are the signal expectations for $t\bar{t}\rightarrow WbHc$ and $t\bar{t}\rightarrow WbHu$ assuming $\text{BR}(t \rightarrow Hc) = 1\%$ and $\text{BR}(t \rightarrow Hu) = 1\%$ respectively. The quoted uncertainties are the sum in quadrature of statistical and systematic uncertainties on the yields.

<table>
<thead>
<tr>
<th>Event Category</th>
<th>3$\ell$</th>
<th>2$\ell$1$\gamma_{\text{had}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}\rightarrow WbHc$</td>
<td>6.4 ± 0.9</td>
<td>2.4 ± 0.4</td>
</tr>
<tr>
<td>$t\bar{t}\rightarrow WbHu$</td>
<td>5.2 ±1.0 -0.8</td>
<td>2.1 ±0.5 -0.4</td>
</tr>
<tr>
<td>Non-prompt</td>
<td>3.2 ± 0.7</td>
<td>0.4 ±0.6 -0.4</td>
</tr>
<tr>
<td>$q$ mis-id</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$t\bar{t}W$</td>
<td>2.3 ± 0.7</td>
<td>0.38 ± 0.12</td>
</tr>
<tr>
<td>$t\bar{t}(Z/\gamma^*)$</td>
<td>3.9 ± 0.8</td>
<td>0.37 ± 0.08</td>
</tr>
<tr>
<td>Diboson</td>
<td>0.86 ± 0.55</td>
<td>0.12 ± 0.11</td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>2.34 ± 0.35</td>
<td>0.47 ± 0.08</td>
</tr>
<tr>
<td>Rare</td>
<td>0.92 ±0.07 -0.06</td>
<td>0.10 ±0.02</td>
</tr>
<tr>
<td>Total background</td>
<td>13.7 ± 2.3</td>
<td>1.9 ± 0.6</td>
</tr>
<tr>
<td>Data</td>
<td>18</td>
<td>1</td>
</tr>
</tbody>
</table>

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


The ATLAS collaboration

A. Wildauer\textsuperscript{101}, H.G. Wilkens\textsuperscript{30}, H.H. Williams\textsuperscript{122}, S. Williams\textsuperscript{107}, C. Willis\textsuperscript{90}, S. Willocq\textsuperscript{86}, A. Wilson\textsuperscript{89}, J.A. Wilson\textsuperscript{18}, I. Wingertzer-Seez\textsuperscript{3}, F. Winklmeier\textsuperscript{110}, B.T. Winter\textsuperscript{21}, M. Wittgen\textsuperscript{143}, J. Wittkowski\textsuperscript{80}, S.J. Wollstadt\textsuperscript{83}, M.W. Wolter\textsuperscript{39}, H. Wolters\textsuperscript{126a,126c}, B.K. Wosiek\textsuperscript{39}, J. Wotschack\textsuperscript{30}, M.J. Woudstra\textsuperscript{84}, K.W. Woźniak\textsuperscript{39}, M. Wu\textsuperscript{55}, M. Wu\textsuperscript{31}, S.L. Wu\textsuperscript{173}, X. Wu\textsuperscript{49}, Y. Wu\textsuperscript{69}, T.R. Wyatt\textsuperscript{84}, B.M. Wynne\textsuperscript{46}, S. Xella\textsuperscript{36}, D. Xu\textsuperscript{3a}, L. Xu\textsuperscript{25}, B. Yabsley\textsuperscript{150}, S. Yacoobi\textsuperscript{145a}, R. Yaka\textsuperscript{67}, M. Yamada\textsuperscript{66}, D. Yamaguchi\textsuperscript{157}, Y. Yamaguchi\textsuperscript{118}, A. Yamamoto\textsuperscript{66}, S. Yamamoto\textsuperscript{155}, T. Yamanaka\textsuperscript{152}, K. Yamauchi\textsuperscript{103}, Y. Yamazaki\textsuperscript{67}, Z. Yan\textsuperscript{22}, H. Yang\textsuperscript{33e}, H. Yang\textsuperscript{173}, Y. Yang\textsuperscript{151}, W-M. Yao\textsuperscript{15}, Y.C. Yap\textsuperscript{80}, Y. Yasu\textsuperscript{56}, E. Yatsenko\textsuperscript{5}, K.H. Yau Wong\textsuperscript{21}, J. Ye\textsuperscript{40}, S. Ye\textsuperscript{25}, I. Yeletsikh\textsuperscript{65}, A.L. Yen\textsuperscript{57}, E. Yıldırım\textsuperscript{12}, K. Yorita\textsuperscript{171}, R. Yoshida\textsuperscript{6}, K. Yoshihara\textsuperscript{122}, C. Young\textsuperscript{143}, J.S. Young\textsuperscript{30}, S. Yousef\textsuperscript{22}, D.R. Yu\textsuperscript{19}, J. Yu\textsuperscript{8}, J.M. Yu\textsuperscript{89}, J. Yu\textsuperscript{114}, L. Yuan\textsuperscript{67}, S.P.Y. Yuen\textsuperscript{21}, A. Yurkewicz\textsuperscript{108}, I. Yusuf\textsuperscript{28,am}, B. Zabinski\textsuperscript{39}, R. Zaidan\textsuperscript{63}, A.M. Zaitsev\textsuperscript{130,ad}, J. Zaliecka\textsuperscript{14}, A. Zaman\textsuperscript{148}, S. Zambito\textsuperscript{57}, L. Zanello\textsuperscript{132a,132b}, D. Zanzi\textsuperscript{88}, C. Zeitnitz\textsuperscript{175}, M. Zeman\textsuperscript{128}, A. Zemla\textsuperscript{38a}, Q. Zeng\textsuperscript{143}, K. Zengel\textsuperscript{23}, O. Zenin\textsuperscript{130}, T. Ženiš\textsuperscript{144a}, D. Zerwas\textsuperscript{117}, D. Zhang\textsuperscript{89}, F. Zhang\textsuperscript{173}, G. Zhang\textsuperscript{33b}, H. Zhang\textsuperscript{33c}, J. Zhang\textsuperscript{6}, L. Zhang\textsuperscript{48}, R. Zhang\textsuperscript{33b,j}, X. Zhang\textsuperscript{33d}, Z. Zhang\textsuperscript{117}, X. Zhao\textsuperscript{40}, Y. Zhao\textsuperscript{33d,117}, Z. Zhao\textsuperscript{33b}, A. Zhemchugov\textsuperscript{65}, J. Zhong\textsuperscript{120}, B. Zhou\textsuperscript{89}, C. Zhou\textsuperscript{45}, L. Zhou\textsuperscript{35}, L. Zhou\textsuperscript{40}, M. Zhou\textsuperscript{148}, N. Zhou\textsuperscript{33f}, C.G. Zhu\textsuperscript{33d}, H. Zhu\textsuperscript{33a}, J. Zhu\textsuperscript{89}, Y. Zhu\textsuperscript{33b}, X. Zhuang\textsuperscript{33a}, K. Zhukov\textsuperscript{96}, A. Zibell\textsuperscript{174}, D. Ziems\textsuperscript{61}, N.I. Zimine\textsuperscript{65}, C. Zimmermann\textsuperscript{83}, S. Zimmermann\textsuperscript{48}, Z. Zinonos\textsuperscript{54}, M. Zinser\textsuperscript{83}, M. Ziolkowski\textsuperscript{141}, L. Živković\textsuperscript{13}, G. Zobernig\textsuperscript{173}, A. Zoccoli\textsuperscript{20a,20b}, M. zur Nedden\textsuperscript{16}, G. Zurzolo\textsuperscript{104a,104b}, and L. Zwailski\textsuperscript{30}

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 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
12 Institut de Física d’Altes Energies and Departament de Física de la Universitat Autònoma de Barcelona, Barcelona, Spain
13 Institute of Physics, University of Belgrade, Belgrade, Serbia
14 Department for Physics and Technology, University of Bergen, Bergen, Norway
15 Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley CA, United States of America
16 Department of Physics, Humboldt University, Berlin, Germany
17 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
18 School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
19 (a) Department of Physics, Bogazici University, Istanbul; (b) Department of Physics Engineering, Gaziantep University, Gaziantep; (c) Department of Physics, Dogus University, Istanbul, Turkey
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
(a) 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
Kyoto University of Education, 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, Jožef 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 Física 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
School of Physics, University of Melbourne, Victoria, Australia
Department of Physics, The University of Michigan, Ann Arbor MI, United States of America
Department of Physics and Astronomy, Michigan State University, East Lansing MI, United States of America
(a) INFN Sezione di Milano; (b) Dipartimento di Fisica, Università di Milano, Milano, Italy
B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Republic of Belarus
National Scientific and Educational Centre for Particle and High Energy Physics, Minsk, Republic of Belarus
Department of Physics, Massachusetts Institute of Technology, Cambridge MA, United States of America
Group of Particle Physics, University of Montreal, Montreal QC, Canada
P.N. Lebedev Institute of Physics, Academy of Sciences, Moscow, Russia
Institute for Theoretical and Experimental Physics (ITEP), Moscow, Russia
National Research Nuclear University MEPhI, Moscow, Russia
D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
Fakultät für Physik, Ludwig-Maximilians-Universität München, München, Germany
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany
Nagasaki Institute of Applied Science, Nagasaki, Japan
(a) INFN Sezione di Napoli; (b) Dipartimento di Fisica, Università di Napoli, Napoli, Italy
Department of Physics and Astronomy, University of New Mexico, Albuquerque NM, United States of America
Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, Netherlands
Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, Netherlands
Department of Physics, Northern Illinois University, DeKalb IL, United States of America
Budker Institute of Nuclear Physics, SB RAS, Novosibirsk, Russia
Department of Physics, New York University, New York NY, United States of America
Ohio State University, Columbus OH, United States of America
Faculty of Science, Okayama University, Okayama, Japan
Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman OK, United States of America
Department of Physics, Oklahoma State University, Stillwater OK, United States of America
\textsuperscript{ai} Also at Faculty of Physics, M.V.Lomonosov Moscow State University, Moscow, Russia
\textsuperscript{aj} Also at National Research Nuclear University MEPhI, Moscow, Russia
\textsuperscript{ak} Also at Department of Physics, Stanford University, Stanford CA, United States of America
\textsuperscript{al} Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
\textsuperscript{am} Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia
* Deceased