Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

Abstract: Results of a search for $H \rightarrow \tau\tau$ decays are presented, based on the full set of proton-proton collision data recorded by the ATLAS experiment at the LHC during 2011 and 2012. The data correspond to integrated luminosities of 4.5 fb$^{-1}$ and 20.3 fb$^{-1}$ at centre-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV respectively. All combinations of leptonic ($\tau \rightarrow \ell\nu\bar{\nu}$ with $\ell = e, \mu$) and hadronic ($\tau \rightarrow \text{hadrons } \nu$) tau decays are considered. An excess of events over the expected background from other Standard Model processes is found with an observed (expected) significance of 4.5 (3.4) standard deviations. This excess provides evidence for the direct coupling of the recently discovered Higgs boson to fermions. The measured signal strength, normalised to the Standard Model expectation, of $\mu = 1.43^{+0.43}_{-0.37}$ is consistent with the predicted Yukawa coupling strength in the Standard Model.

Keywords: Hadron-Hadron Scattering

ArXiv ePrint: 1501.04943
1 Introduction

The investigation of the origin of electroweak symmetry breaking and, related to this, the experimental confirmation of the Brout-Englert-Higgs mechanism \([1–6]\) is one of the prime goals of the physics programme at the Large Hadron Collider (LHC) \([7]\). With the discovery of a Higgs boson with a mass of approximately 125 GeV by the ATLAS \([8]\) and CMS \([9]\) collaborations, an important milestone has been reached. More precise measurements of the properties of the discovered particle \([10, 11]\) as well as tests of the spin-parity quantum numbers \([12–14]\) continue to be consistent with the predictions for the Standard Model (SM) Higgs boson.

These measurements rely predominantly on studies of the bosonic decay modes, \(H \to \gamma\gamma\), \(H \to ZZ^*\) and \(H \to WW^*\). To establish the mass generation mechanism for fermions as implemented in the SM, it is of prime importance to demonstrate the direct coupling of the Higgs boson to fermions and the proportionality of its strength to mass \([15]\).

The most promising candidate decay modes are the decays into tau leptons, \(H \to \tau\tau\), and bottom quarks (\(b\)-quarks), \(H \to b\bar{b}\). Due to the high background, the search for decays to \(b\bar{b}\) is restricted to Higgs bosons produced in modes which have a more distinct signature but a lower cross-section, such as \(H\) production with an associated vector boson. The smaller rate of these processes in the presence of still large background makes their detection challenging. More favourable signal-to-background conditions are expected for \(H \to \tau\tau\) decays. Recently, the CMS Collaboration published evidence for \(H \to \tau\tau\) decays at a significance in terms of standard deviations of 3.2\(\sigma\) \([16]\), and an excess corresponding to a significance of 2.1\(\sigma\) in the search for \(H \to b\bar{b}\) decays \([17]\). The combination of channels provides evidence for fermionic couplings with a significance of 3.8\(\sigma\) \([18]\). The yield of events in the search for \(H \to b\bar{b}\) decays observed by the ATLAS Collaboration has a signal significance of 1.4\(\sigma\) \([19]\). The Tevatron experiments have observed an excess corresponding to 2.8\(\sigma\) in the \(H \to b\bar{b}\) search \([20]\).

In this paper, the results of a search for \(H \to \tau\tau\) decays are presented, based on the full proton-proton dataset collected by the ATLAS experiment during the 2011 and 2012 data-taking periods, corresponding to integrated luminosities of 4.5 fb\(^{-1}\) at a centre-of-mass energy of \(\sqrt{s} = 7\) TeV and 20.3 fb\(^{-1}\) at \(\sqrt{s} = 8\) TeV. These results supersede the earlier upper limits on the cross section times the branching ratio obtained with the 7 TeV data \([21]\). All combinations of leptonic (\(\tau \to \ell\nu\bar{\nu}\) with \(\ell = e, \mu\)) and hadronic (\(\tau \to \text{hadrons} \nu\)) tau decays are considered.\(^1\) The corresponding three analysis channels are denoted by \(\tau\text{lep}\tau\text{lep}, \tau\text{lep}\tau\text{had},\) and \(\tau\text{had}\tau\text{had}\) in the following. The search is designed to be sensitive to the major production processes of a SM Higgs boson, i.e. production via gluon fusion (ggF) \([22]\), vector-boson fusion (VBF) \([23]\), and associated production (\(VH\)) with \(V = W\) or \(Z\). These production processes lead to different final-state signatures, which are exploited by defining an event categorisation. Two dedicated categories are considered to achieve both a good signal-to-background ratio and good resolution for the reconstructed \(\tau\tau\) invariant mass. The VBF category, enriched in events produced via vector-boson fusion,

\(^1\)Throughout this paper the inclusion of charge-conjugate decay modes is implied.
is defined by the presence of two jets with a large separation in pseudorapidity.\(^2\) The boosted category contains events where the reconstructed Higgs boson candidate has a large transverse momentum. It is dominated by events produced via gluon fusion with additional jets from gluon radiation. In view of the signal-to-background conditions, and in order to exploit correlations between final-state observables, a multivariate analysis technique, based on boosted decision trees (BDTs) \([24–26]\), is used to extract the final results. As a cross-check, a separate analysis where cuts on kinematic variables are applied is carried out.

2 The ATLAS detector and object reconstruction

The ATLAS detector \([27]\) is a multi-purpose detector with a cylindrical geometry. It comprises an inner detector (ID) surrounded by a thin superconducting solenoid, a calorimeter system and an extensive muon spectrometer in a toroidal magnetic field. The ID tracking system consists of a silicon pixel detector, a silicon microstrip detector (SCT), and a transition radiation tracker (TRT). It provides precise position and momentum measurements for charged particles and allows efficient identification of jets containing \(b\)-hadrons (\(b\)-jets) in the pseudorapidity range \(|\eta| < 2.5\). The ID is immersed in a 2 T axial magnetic field and is surrounded by high-granularity lead/liquid-argon (LAr) sampling electromagnetic calorimeters which cover the pseudorapidity range \(|\eta| < 3.2\). A steel/scintillator tile calorimeter provides hadronic energy measurements in the central pseudorapidity range \((|\eta| < 1.7)\). In the forward regions \((1.5 < |\eta| < 4.9)\), the system is complemented by two end-cap calorimeters using LAr as active material and copper or tungsten as absorbers. The muon spectrometer (MS) surrounds the calorimeters and consists of three large superconducting eight-coil toroids, a system of tracking chambers, and detectors for triggering. The deflection of muons is measured within \(|\eta| < 2.7\) by three layers of precision drift tubes, and cathode strip chambers in the innermost layer for \(|\eta| > 2.0\). The trigger chambers consist of resistive plate chambers in the barrel \((|\eta| < 0.95)\) and thin-gap chambers in the end-cap regions \((0.95 < |\eta| < 2.4)\).

A three-level trigger system \([28]\) is used to select events. A hardware-based Level-1 trigger uses a subset of detector information to reduce the event rate to a value to 75 kHz or less. The rate of accepted events is then reduced to about 400 Hz by two software-based trigger levels, Level-2 and the Event Filter.

The reconstruction of the basic physics objects used in this analysis is described in the following. The primary vertex referenced below is chosen as the proton-proton vertex candidate with the highest sum of the squared transverse momenta of all associated tracks.

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeters matched to a track in the ID. They are required to have a transverse en-
ergy, $E_T = E \sin \theta$, greater than 15 GeV, to be within the pseudorapidity range $|\eta| < 2.47$, and to satisfy the medium shower shape and track selection criteria defined in ref. [29]. Candidates found in the transition region between the end-cap and barrel calorimeters (1.37 $< |\eta| < 1.52$) are not considered. Typical reconstruction and identification efficiencies for electrons satisfying these selection criteria range between 80% and 90% depending on $E_T$ and $\eta$.

Muon candidates are reconstructed using an algorithm [30] that combines information from the ID and the MS. The distance between the $z$-position of the point of closest approach of the muon inner-detector track to the beam-line and the $z$-coordinate of the primary vertex is required to be less than 1 cm. This requirement reduces the contamination due to cosmic-ray muons and beam-induced backgrounds. Muon quality criteria such as inner detector hit requirements are applied to achieve a precise measurement of the muon momentum and reduce the misidentification rate. Muons are required to have a momentum in the transverse plane $p_T > 10$ GeV and to be within $|\eta| < 2.5$. Typical efficiencies for muons satisfying these selection criteria are above 95% [30].

Jets are reconstructed using the anti-$k_t$ jet clustering algorithm [31, 32] with a radius parameter $R = 0.4$, taking topological energy clusters [33] in the calorimeters as inputs. Jet energies are corrected for the contribution of multiple interactions using a technique based on jet area [34] and are calibrated using $p_T$- and $\eta$-dependent correction factors determined from data and simulation [35–37]. Jets are required to be reconstructed in the range $|\eta| < 4.5$ and to have $p_T > 30$ GeV. To reduce the contamination of jets by additional interactions in the same or neighbouring bunch crossings (pile-up), tracks originating from the primary vertex must contribute a large fraction of the $p_T$ when summing the scalar $p_T$ of all tracks in the jet. This jet vertex fraction (JVF) is required to be at least 75% (50%) for jets with $|\eta| < 2.4$ in the 7 TeV (8 TeV) dataset. Moreover, for the 8 TeV dataset, the JVF selection is applied only to jets with $p_T < 50$ GeV. Jets with no associated tracks are retained.

In the pseudorapidity range $|\eta| < 2.5$, $b$-jets are selected using a tagging algorithm [38]. The $b$-jet tagging algorithm has an efficiency of 60–70% for $b$-jets in simulated $t\bar{t}$ events. The corresponding light-quark jet misidentification probability is 0.1–1%, depending on the jet’s $p_T$ and $\eta$.

Hadronically decaying tau leptons are reconstructed starting from clusters of energy in the electromagnetic and hadronic calorimeters. The $\tau_{\text{had}}$ reconstruction is seeded by the anti-$k_t$ jet finding algorithm with a radius parameter $R = 0.4$. Jet-specific cleaning selection such as the JVF requirement is not needed for the tau candidate seeds, due to stricter vertex requirements. Tracks with $p_T > 1$ GeV within a cone of radius 0.2 around the cluster barycentre are matched to the $\tau_{\text{had}}$ candidate, and the $\tau_{\text{had}}$ charge is determined from the sum of the charges of its associated tracks. The rejection of jets is provided in a separate identification step using discriminating variables based on tracks with $p_T > 1$ GeV and the energy deposited in calorimeter cells found in the core region ($\Delta R < 0.2$) and in the region $0.2 < \Delta R < 0.4$ around the $\tau_{\text{had}}$ candidate’s direction. Such discriminating variables

---

In the following, the $\tau_{\text{had}}$ symbol always refers to the visible decay products of the $\tau$ hadronic decay.
are combined in a boosted decision tree and three working points, labelled tight, medium and loose [39], are defined, corresponding to different $\tau_{\text{had}}$ identification efficiency values.

In this analysis, $\tau_{\text{had}}$ candidates with $p_T > 20$ GeV and $|\eta| < 2.47$ are used. The $\tau_{\text{had}}$ candidates are required to have charge $\pm 1$, and must be 1- or 3-track (prong) candidates. In addition, a sample without the charge and track multiplicity requirements is retained for background modelling in the $\tau_{\text{had}}\tau_{\text{had}}$ channel, as described in section 6.2. The identification efficiency for $\tau_{\text{had}}$ candidates satisfying the medium criteria is of the order of 55–60%. Dedicated criteria [39] to separate $\tau_{\text{had}}$ candidates from misidentified electrons are also applied, with a selection efficiency for true $\tau_{\text{had}}$ decays of 95%. The probability to misidentify a jet with $p_T > 20$ GeV as a $\tau_{\text{had}}$ candidate is typically 1–2%.

Following their reconstruction, candidate leptons, hadronically decaying taus and jets may point to the same energy deposits in the calorimeters (within $\Delta R < 0.2$). Such overlaps are resolved by selecting objects in the following order of priority (from highest to lowest): muons, electrons, $\tau_{\text{had}}$, and jet candidates. For all channels, the leptons that are considered in overlap removal with $\tau_{\text{had}}$ candidates need to only satisfy looser criteria than those defined above, to reduce misidentified $\tau_{\text{had}}$ candidates from leptons. The $p_T$ threshold of muons considered in overlap removal is also lowered to 4 GeV.

The missing transverse momentum (with magnitude $E_T^{\text{miss}}$) is reconstructed using the energy deposits in calorimeter cells calibrated according to the reconstructed physics objects ($e$, $\gamma$, $\tau_{\text{had}}$, jets and $\mu$) with which they are associated [40]. The transverse momenta of reconstructed muons are included in the $E_T^{\text{miss}}$ calculation, with the energy deposited by these muons in the calorimeters taken into account. The energy from calorimeter cells not associated with any physics objects is scaled by a soft-term vertex fraction and also included in the $E_T^{\text{miss}}$ calculation. This fraction is the ratio of the summed scalar $p_T$ of tracks from the primary vertex not matched with objects to the summed scalar $p_T$ of all tracks in the event also not matched to objects. This method allows to achieve a more accurate reconstruction of the $E_T^{\text{miss}}$ in high pile-up conditions [41].

3 Data and simulated samples

After data quality requirements, the integrated luminosities of the samples used are 4.5 fb$^{-1}$ at $\sqrt{s} = 7$ TeV and 20.3 fb$^{-1}$ at $\sqrt{s} = 8$ TeV.

Samples of signal and background events were simulated using various Monte Carlo (MC) generators, as summarised in table 1. The generators used for the simulation of the hard-scattering process and the model used for the simulation of the parton shower, of the hadronisation and of the underlying-event activity are listed. In addition, the cross section values to which the simulation is normalised and the perturbative order in QCD of the respective calculations are provided.

The signal contributions considered include the three main processes for Higgs boson production at the LHC: ggF, VBF, and associated $VH$ production processes. The contributions from the associated $t\bar{t}H$ production process are found to be small and are neglected. The ggF and VBF production processes are simulated with POWHEG [42–45] interfaced to PYTHIA8 [46]. In the POWHEG event generator, the CT10 [47] parameter-
isation of the parton distribution functions (PDFs) is used. The overall normalisation of the ggF process is taken from a calculation at next-to-next-to-leading-order (NNLO) \cite{48-53} in QCD, including soft-gluon resummation up to next-to-next-to-leading logarithm terms (NNLL) \cite{54}. Next-to-leading order (NLO) electroweak (EW) corrections are also included \cite{55, 56}. Production by VBF is normalised to a cross section calculated with full NLO QCD and EW corrections \cite{57-59} with an approximate NNLO QCD correction applied \cite{60}. The associated VH production process is simulated with PYTHIA8. The CTEQ6L1 \cite{61} parameterisation of PDFs is used for the PYTHIA8 event generator. The predictions for VH production are normalised to cross sections calculated at NNLO in QCD \cite{62}, with NLO EW radiative corrections \cite{63} applied.

Additional corrections to the shape of the generated $p_T$ distribution of Higgs bosons produced via ggF are applied to match the distribution from a calculation at NNLO including the NNLL corrections provided by the HRes2.1 \cite{64} program. In this calculation, the effects of finite masses of the top and bottom quarks \cite{64, 65} are included and dynamical renormalisation and factorisation scales, $\mu_R, \mu_F = \sqrt{m_H^2 + p_T^2}$, are used. A reweighting is performed separately for events with no more than one jet at particle level and for events with two or more jets. In the latter case, the Higgs boson $p_T$ spectrum is reweighted to match the MinLO HJJ predictions \cite{66}. The reweighting is derived such that the inclusive Higgs boson $p_T$ spectrum and the $p_T$ spectrum of events with at least two jets match the HRes2.1 and MinLO HJJ predictions respectively, and that the jet multiplicities are in agreement with (N)NLO calculations from JetVHETO \cite{67-69}.

The NLO EW corrections for VBF production depend on the $p_T$ of the Higgs boson, varying from a few percent at low $p_T$ to $\sim 20\%$ at $p_T = 300$ GeV \cite{70}. The $p_T$ spectrum of the VBF-produced Higgs boson is therefore reweighted, based on the difference between the POWHEG+PYTHIA calculation and the HAWK \cite{57, 58} calculation which includes these corrections.

The main and largely irreducible $Z/\gamma^* \rightarrow \tau\tau$ background is modelled using $Z/\gamma^* \rightarrow \mu\mu$ events from data,\footnote{These processes are hereafter for simplicity denoted by $Z \rightarrow \tau\tau$ and $Z \rightarrow \mu\mu$ respectively, even though the whole continuum above and below the $Z$ peak is considered.} where the muon tracks and associated energy depositions in the calorimeters are replaced by the corresponding simulated signatures of the final-state particles of the tau decay. In this approach, essential features such as the modelling of the kinematics of the produced boson, the modelling of the hadronic activity of the event (jets and underlying event) as well as contributions from pile-up are taken from data. Thereby the dependence on the simulation is minimised and only the $\tau$ decays and the detector response to the tau-lepton decay products are based on simulation. By requiring two isolated, high-energy muons with opposite charge and a dimuon invariant mass $m_{\mu\mu} > 40$ GeV, $Z \rightarrow \mu\mu$ events can be selected from the data with high efficiency and purity. To replace the muons in the selected events, all tracks associated with the muons are removed and calorimeter cell energies associated with the muons are corrected by subtracting the corresponding energy depositions in a single simulated $Z \rightarrow \mu\mu$ event with the same kinematics. Finally, both the track information and the calorimeter cell energies from a simulated $Z \rightarrow \tau\tau$ decay are
added to the data event. The decays of the tau leptons are simulated by Tauola [71]. The
tau lepton kinematics are matched to the kinematics of the muons they are replacing, in-
cluding polarisation and spin correlations [72], and the mass difference between the muons
and the tau leptons is accounted for. This hybrid sample is referred to as embedded data
in the following.

Other background processes are simulated using different generators, each interfaced
to Pythia [46, 73] or Herwig [74, 75] to provide the parton shower, hadronisation and the
modelling of the underlying event, as indicated in table 1. For the Herwig samples, the
decays of tau leptons are also simulated using Tauola [71]. Photon radiation from charged
leptons for all samples is provided by Photos [76]. The samples for W/Z+jets production
are generated with Alpgen [77], employing the MLM matching scheme [78] between the
hard process (calculated with LO matrix elements for up to five partons) and the parton
shower. For WW production, the loop-induced gg → WW process is also generated, using
the gg2WW [79] program. In the AcerMC [80], Alpgen, and Herwig event generators,
the Cteq6L1 parameterisation of the PDFs is used, while the CT10 parameterisation is
used for the generation of events with gg2WW. The normalisation of these background
contributions is either estimated from control regions using data, as described in section 6,
or the cross sections quoted in table 1 are used.

For all samples, a full simulation of the ATLAS detector response [81] using the
Geant4 program [82] was performed. In addition, events from minimum-bias interac-
tions were simulated using the AU2 [83] parameter tuning of Pythia8. The AU2 tune
includes the set of optimized parameters for the parton shower, hadronisation, and multi-
ple parton interactions. They are overlaid on the simulated signal and background events
according to the luminosity profile of the recorded data. The contributions from these pile-
up interactions are simulated both within the same bunch crossing as the hard-scattering
process and in neighbouring bunch crossings. Finally, the resulting simulated events are
processed through the same reconstruction programs as the data.

4 Event selection and categorisation

4.1 Event selection

Single lepton, dilepton and di-τ had triggers were used to select the events for the analysis. A
summary of the triggers used by each channel at the two centre-of-mass energies is reported
in table 2. Due to the increasing luminosity and the different pile-up conditions, the online
p_{T} thresholds increased during data-taking in 2011 and again for 2012, and more stringent
identification requirements were applied for the data-taking in 2012. The p_{T} requirements
on the objects in the analysis are usually 2 GeV higher than the trigger requirements, to
ensure that the trigger is fully efficient.

In addition to applying criteria to ensure that the detector was functioning properly,
requirements to increase the purity and quality of the data sample are applied by rejecting
non-collision events such as cosmic rays and beam-halo events. At least one reconstructed
vertex is required with at least four associated tracks with p_{T} > 400 MeV and a posi-
Table 1. Monte Carlo generators used to model the signal and the background processes at $\sqrt{s} = 8$ TeV. The cross sections times branching fractions ($\sigma \times B$) used for the normalisation of some processes (many of these are subsequently normalised to data) are included in the last column together with the perturbative order of the QCD calculation. For the signal processes the $H \to \tau\tau$ SM branching ratio is included, and for the $W$ and $Z/\gamma^*$ background processes the branching ratios for leptonic decays ($\ell = e, \mu, \tau$) of the bosons are included. For all other background processes, inclusive cross sections are quoted (marked with a †).
\[
\sqrt{s} = 7 \text{ TeV}
\]

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Analysis level thresholds [GeV]</th>
<th>Trigger level thresholds, (p_T) [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single electron</td>
<td>(e\mu): (p_T^e &gt; 22 - 24) (p_T^\mu &gt; 10) (p_T^\tau &gt; 25) (p_T^\tau &gt; 20) (\tau\tau): (p_T^{\tau 1} &gt; 25) (p_T^{\tau 2} &gt; 25) (p_T^{\mu} &gt; 15)</td>
<td>(e\mu): (p_T^e &gt; 22 - 24) (p_T^\mu &gt; 10) (\mu\tau): (p_T^\mu &gt; 15) (\tau\tau): (p_T^{\tau 1} &gt; 25) (p_T^{\tau 2} &gt; 25)</td>
</tr>
<tr>
<td>Single muon</td>
<td>(\mu\mu): (p_T^\mu &gt; 26) (p_T^\mu &gt; 10) (p_T^\tau &gt; 26) (p_T^\tau &gt; 20) (p_T^\mu &gt; 15)</td>
<td>(\mu\mu): (p_T^\mu &gt; 26) (p_T^\mu &gt; 15) (p_T^\mu &gt; 15)</td>
</tr>
<tr>
<td>Di-electron</td>
<td>(ee): (p_T^{e 1} &gt; 15) (p_T^{e 2} &gt; 15)</td>
<td>(ee): (p_T^{e 1} &gt; 15) (p_T^{e 2} &gt; 15)</td>
</tr>
<tr>
<td>Di-(\tau_{\text{had}})</td>
<td>(29/20)</td>
<td>(29/20)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Analysis level thresholds [GeV]</th>
<th>Trigger level thresholds, (p_T) [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single electron</td>
<td>(e\mu): (p_T^e &gt; 24) (p_T^\mu &gt; 10) (p_T^\tau &gt; 26) (p_T^\tau &gt; 20) (\tau\tau): (p_T^{\tau 1} &gt; 35) (p_T^{\tau 2} &gt; 25)</td>
<td>(e\mu): (p_T^e &gt; 24) (p_T^\mu &gt; 10) (\mu\tau): (p_T^\mu &gt; 26) (\tau\tau): (p_T^{\tau 1} &gt; 35) (p_T^{\tau 2} &gt; 25)</td>
</tr>
<tr>
<td>Single muon</td>
<td>(\mu\mu): (p_T^\mu &gt; 24) (p_T^\mu &gt; 10)</td>
<td>(\mu\mu): (p_T^\mu &gt; 24) (p_T^\mu &gt; 10)</td>
</tr>
<tr>
<td>Di-electron</td>
<td>(ee): (p_T^{e 1} &gt; 15) (p_T^{e 2} &gt; 15)</td>
<td>(ee): (p_T^{e 1} &gt; 15) (p_T^{e 2} &gt; 15)</td>
</tr>
<tr>
<td>Di-muon</td>
<td>(\mu\mu): (p_T^\mu &gt; 24) (p_T^\mu &gt; 15)</td>
<td>(\mu\mu): (p_T^\mu &gt; 24) (p_T^\mu &gt; 15)</td>
</tr>
<tr>
<td>Electron+muon</td>
<td>(12/8)</td>
<td>(12/8)</td>
</tr>
<tr>
<td>Di-(\tau_{\text{had}})</td>
<td>(29/20)</td>
<td>(29/20)</td>
</tr>
</tbody>
</table>

**Table 2.** Summary of the triggers used to select events for the different analysis channels at the two centre-of-mass energies. The transverse momentum thresholds applied at trigger level and in the analysis are listed. When more than one trigger is used, a logical OR is taken and the trigger efficiencies are calculated accordingly.
as the sum of the total transverse energy in the calorimeter in a cone of size $\Delta R$ around the electron cluster or the muon track, divided by the $E_T$ of the electron cluster or the $p_T$ of the muon respectively. The track-based isolation $I(p_T, \Delta R)$ is defined as the sum of the transverse momenta of tracks within a cone of $\Delta R$ around the electron or muon track, divided by the $E_T$ of the electron cluster or the muon $p_T$ respectively. The isolation requirements applied are slightly different for the two centre-of-mass energies and are listed in table 3.

In the $\tau^{\text{had}}\tau^{\text{had}}$ channel, isolated taus are selected by requiring that there are no tracks with $p_T > 0.5 \text{ GeV}$ in an isolation region of $0.2 < \Delta R < 0.6$ around the tau direction. This requirement leads to a 12% (4%) efficiency loss for hadronic taus, while 30% (10%) of contamination from jets is rejected in 8 (7) TeV data.

After the basic lepton selection, further channel-dependent cuts are applied, as detailed in the following. The full event selection is summarised in table 4.

**The $\tau^{\text{lep}}\tau^{\text{lep}}$ channel:** exactly two isolated leptons with opposite-sign (OS) electric charges, passing the $p_T$ threshold listed in table 2, are required. Events containing a $\tau^{\text{had}}$ candidate are vetoed. For the $\tau^{\text{had}}$ candidates considered, the criteria used to reject electrons misidentified as $\tau^{\text{had}}$ candidates are tightened to a working-point of 85% signal efficiency [39].

In addition to the irreducible $Z \to \tau\tau$ background, sizeable background contributions from $Z \to \ell\ell$ and from $t\bar{t}$ production are expected in this channel. Background contributions from $Z$ decays, but also from low mass resonances (charmonium and bottomonium), are rejected by requirements on the invariant mass $m_{\tau\tau}^{\text{vis}}$ of the visible tau decay products, on the angle $\Delta\phi_{\ell\ell}$ between the two leptons in the transverse plane and on $E_T^{\text{miss}}$. To reject the large $Z \to \ell\ell$ contribution in events with same-flavour (SF) leptons ($ee, \mu\mu$), more stringent cuts on the visible mass and on $E_T^{\text{miss}}$ are applied for these events than for events with different-flavour (DF) leptons ($e\mu$). For SF final states, an additional variable named *high-$p_T$ objects* $E_T^{\text{miss}}$ ($E_T^{\text{miss,HPTO}}$) is also used to reject background from $Z/\gamma^*$ production. It is calculated from the high-$p_T$ objects in the event, i.e. from the two leptons and from jets with $p_T > 25$ GeV. Due to the presence of real neutrinos, the two $E_T^{\text{miss}}$ variables

<table>
<thead>
<tr>
<th>Channel</th>
<th>$7 \text{ TeV}$</th>
<th>$8 \text{ TeV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrons</td>
<td>$I(p_T, 0.4) &lt; 0.06$</td>
<td>$I(p_T, 0.4) &lt; 0.06$</td>
</tr>
<tr>
<td></td>
<td>$I(E_T, 0.2) &lt; 0.08$</td>
<td>$I(E_T, 0.2) &lt; 0.06$</td>
</tr>
<tr>
<td>Muons</td>
<td>$I(p_T, 0.4) &lt; 0.17$</td>
<td>$I(p_T, 0.4) &lt; 0.06$</td>
</tr>
<tr>
<td></td>
<td>$I(E_T, 0.2) &lt; 0.09$</td>
<td>$I(E_T, 0.2) &lt; 0.06$</td>
</tr>
<tr>
<td></td>
<td>$I(p_T, 0.4) &lt; 0.18$</td>
<td>$I(p_T, 0.4) &lt; 0.06$</td>
</tr>
<tr>
<td></td>
<td>$I(E_T, 0.2) &lt; 0.09$</td>
<td>$I(E_T, 0.2) &lt; 0.06$</td>
</tr>
</tbody>
</table>

**Table 3.** Summary of isolation requirements applied for the selection of isolated electrons and muons at the two centre-of-mass energies. The isolation variables are defined in the text.
are strongly correlated for signal events but only loosely correlated for background from $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ decays.

To further suppress background contributions from misidentified leptons\(^5\) a minimum value of the scalar sum of the transverse momenta of the two leptons is required. Contributions from $t\bar{t}$ events are further reduced by rejecting events with a $b$-jet with $p_T > 25$ GeV.

Within the collinear approximation\(^9\), i.e. assuming that the tau directions are given by the directions of the visible tau decay products and that the momenta of the neutrinos constitute the missing transverse momentum, the tau momenta can be reconstructed. For tau decays, the fractions of the tau momenta carried by the visible decay products,\(^6\) $x_{\tau,i} = p_{\text{vis},i} / (p_{\text{vis},i} + p_{\text{mis},i})$, with $i = 1,2$, are expected to lie in the interval $0 < x_{\tau,i} < 1$, and hence corresponding requirements are applied to further reject non-tau background contributions.

Finally, to avoid overlap between this analysis and the search for $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ decays, the $\tau\tau$ mass in the collinear approximation is required to satisfy the condition $m_{\tau\tau}^{\text{coll}} > m_Z - 25$ GeV.

The $\tau_{\text{lep}}\tau_{\text{had}}$ channel: exactly one isolated lepton and one $\tau_{\text{had}}$ candidate with OS charges, passing the $p_T$ thresholds listed in table 2, are required. The criteria used to reject electrons misidentified as $\tau_{\text{had}}$ are also tightened in this channel to a working-point of 85% signal efficiency\(^3\).

The production of $W$+jets and of top quarks constitute the dominant reducible background in this channel. To substantially reduce the $W$+jets contribution, a cut on the transverse mass\(^7\) constructed from the lepton and the missing transverse momentum is applied and events with $m_T > 70$ GeV are rejected. Contributions from $t\bar{t}$ events are reduced by rejecting events with a $b$-jet with $p_T > 30$ GeV.

The $\tau_{\text{had}}\tau_{\text{had}}$ channel: one isolated medium $\tau_{\text{had}}$ candidate and one isolated tight $\tau_{\text{had}}$ candidate with OS charges are required. Events with electron or muon candidates are rejected. For all data, the missing transverse momentum must satisfy $E_{\text{T}}^{\text{miss}} > 20$ GeV and its direction must either be between the two visible $\tau_{\text{had}}$ candidates in $\phi$ or within $\Delta \phi < \pi/4$ of the nearest $\tau_{\text{had}}$ candidate. To further reduce the background from multijet production, additional cuts on the $\Delta R$ and pseudorapidity separation $\Delta \eta$ between the two $\tau_{\text{had}}$ candidates are applied.

With these selections, there is no overlap between the individual channels.

### 4.2 Analysis categories

To exploit signal-sensitive event topologies, two analysis categories are defined in an exclusive way.

---

\(^5\)Misidentified leptons and $\tau_{\text{had}}$ candidates are also referred to as “fake” in this paper.

\(^6\)The variable $p_{\text{vis}}$ is defined as the total momentum of the visible decay products of the tau lepton while $p_{\text{mis}}$ is defined as the momentum of the neutrino system reconstructed using the collinear approximation.

\(^7\) $m_T = \sqrt{2p_T E_{\text{T}}^{\text{miss}} \cdot (1 - \cos \Delta \phi)}$, where $\Delta \phi$ is the azimuthal separation between the directions of the lepton and the missing transverse momentum.
<table>
<thead>
<tr>
<th>Channel</th>
<th>Preselection cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$lep$\tau$lep</td>
<td>Exactly two isolated opposite-sign leptons</td>
</tr>
<tr>
<td></td>
<td>Events with $\tau_{\text{had}}$ candidates are rejected</td>
</tr>
<tr>
<td></td>
<td>$30 \text{ GeV} &lt; m_{\tau\tau}^{\text{vis}} &lt; 100 ,(75) \text{ GeV}$ for DF,(SF) events</td>
</tr>
<tr>
<td></td>
<td>$\Delta\phi_{ll} &lt; 2.5$</td>
</tr>
<tr>
<td></td>
<td>$E_{T}^{\text{miss}} &gt; 20,(40) \text{ GeV}$ for DF,(SF) events</td>
</tr>
<tr>
<td></td>
<td>$E_{T}^{\text{miss},\text{HPPTO}} &gt; 40\text{ GeV}$ for SF events</td>
</tr>
<tr>
<td></td>
<td>$p_{T}^{\ell_{1}} + p_{T}^{\ell_{2}} &gt; 35 \text{ GeV}$</td>
</tr>
<tr>
<td></td>
<td>Events with a $b$-tagged jet with $p_{T} &gt; 25 \text{ GeV}$ are rejected</td>
</tr>
<tr>
<td></td>
<td>$0.1 &lt; x_{\tau_{1}}, x_{\tau_{2}} &lt; 1$</td>
</tr>
<tr>
<td></td>
<td>$m_{\tau\tau}^{\text{coll}} &gt; m_{Z} - 25 \text{ GeV}$</td>
</tr>
<tr>
<td>$\tau$lep$\tau$had</td>
<td>Exactly one isolated lepton and one medium $\tau_{\text{had}}$ candidate with opposite charges</td>
</tr>
<tr>
<td>$\tau$had$\tau$had</td>
<td>One isolated medium and one isolated tight opposite-sign $\tau_{\text{had}}$-candidate</td>
</tr>
<tr>
<td></td>
<td>Events with leptons are vetoed</td>
</tr>
<tr>
<td></td>
<td>$E_{T}^{\text{miss}} &gt; 20 \text{ GeV}$</td>
</tr>
<tr>
<td></td>
<td>$E_{T}^{\text{miss}}$ points between the two visible taus in $\phi$, or min[$\Delta\phi(\tau, E_{T}^{\text{miss}})] &lt; \pi/4$</td>
</tr>
<tr>
<td></td>
<td>$0.8 &lt; \Delta R(\tau_{\text{had}1}, \tau_{\text{had}2}) &lt; 2.4$</td>
</tr>
<tr>
<td></td>
<td>$\Delta\eta(\tau_{\text{had}1}, \tau_{\text{had}2}) &lt; 1.5$</td>
</tr>
<tr>
<td>Channel</td>
<td>VBF category selection cuts</td>
</tr>
<tr>
<td>$\tau$lep$\tau$lep</td>
<td>At least two jets with $p_{T}^{j_{1}} &gt; 40 \text{ GeV}$ and $p_{T}^{j_{2}} &gt; 30 \text{ GeV}$</td>
</tr>
<tr>
<td></td>
<td>$\Delta\eta(j_{1}, j_{2}) &gt; 2.2$</td>
</tr>
<tr>
<td>$\tau$lep$\tau$had</td>
<td>At least two jets with $p_{T}^{j_{1}} &gt; 50 \text{ GeV}$ and $p_{T}^{j_{2}} &gt; 30 \text{ GeV}$</td>
</tr>
<tr>
<td></td>
<td>$\Delta\eta(j_{1}, j_{2}) &gt; 3.0$</td>
</tr>
<tr>
<td></td>
<td>$m_{\tau\tau}^{\text{vis}} &gt; 40 \text{ GeV}$</td>
</tr>
<tr>
<td>$\tau$had$\tau$had</td>
<td>At least two jets with $p_{T}^{j_{1}} &gt; 50 \text{ GeV}$ and $p_{T}^{j_{2}} &gt; 30 \text{ GeV}$</td>
</tr>
<tr>
<td></td>
<td>$p_{T}^{j_{2}} &gt; 35 \text{ GeV}$ for jets with $</td>
</tr>
<tr>
<td></td>
<td>$\Delta\eta(j_{1}, j_{2}) &gt; 2.0$</td>
</tr>
<tr>
<td>Channel</td>
<td>Boosted category selection cuts</td>
</tr>
<tr>
<td>$\tau$lep$\tau$lep</td>
<td>At least one jet with $p_{T} &gt; 40 \text{ GeV}$</td>
</tr>
<tr>
<td>All</td>
<td>Failing the VBF selection</td>
</tr>
<tr>
<td></td>
<td>$p_{T}^{H} &gt; 100 \text{ GeV}$</td>
</tr>
</tbody>
</table>

Table 4. Summary of the event selection for the three analysis channels. The requirements used in both the preselection and for the definition of the analysis categories are given. The labels (1) and (2) refer to the leading (highest $p_{T}$) and subleading final-state objects (leptons, $\tau_{\text{had}}$, jets). The variables are defined in the text.

- The VBF category targets events with a Higgs boson produced via vector boson fusion and is characterised by the presence of two high-$p_{T}$ jets with a large pseudorapidity separation (see table 4). The $\Delta\eta(j_{1}, j_{2})$ requirement is applied to the two highest-$p_{T}$ jets in the event. In the $\tau$lep$\tau$had channel, there is an additional requirement that $m_{\tau\tau}^{\text{vis}} > 40 \text{ GeV}$, to eliminate low-mass $Z/\gamma^{*}$ events. Although this category is dominated by VBF events, it also includes smaller contributions from ggF and $VH$ production.

- The boosted category targets events with a boosted Higgs boson produced via ggF.
Higgs boson candidates are therefore required to have large transverse momentum, $p_T^{H} > 100$ GeV. The $p_T^{H}$ is reconstructed using the vector sum of the missing transverse momentum and the transverse momentum of the visible tau decay products.

In the $\tau_{lep}\tau_{lep}$ channel, at least one jet with $p_T > 40$ GeV is required. The jet requirement selects a region of the phase space where the $E_T^{miss}$ of same-flavour events is well modelled by simulation. In order to define an orthogonal category, events passing the VBF category selection are not considered. This category also includes small contributions from VBF and $VH$ production.

While these categories are conceptually identical across the three channels, differences in the dominant background contributions require different selection criteria. For both categories, the requirement on jets is inclusive and additional jets, apart from those passing the category requirements, are allowed.

For the $\tau_{had}\tau_{had}$ channel, the so-called rest category is used as a control region. In this category, events passing the preselection requirements but not passing the VBF or boosted category selections are considered. This category is used to constrain the $Z \rightarrow \tau\tau$ and multijet background contributions. The signal contamination in this category is negligible.

### 4.3 Higgs boson candidate mass reconstruction

The di-tau invariant mass ($m_{\tau\tau}^{MMC}$) is reconstructed using the missing mass calculator (MMC) [100]. This requires solving an underconstrained system of equations for six to eight unknowns, depending on the number of neutrinos in the $\tau\tau$ final state. These unknowns include the $x$-, $y$-, and $z$-components of the momentum carried by the neutrinos for each of the two tau leptons in the event, and the invariant mass of the two neutrinos from any leptonic tau decays. The calculation uses the constraints from the measured $x$- and $y$-components of the missing transverse momentum, and the visible masses of both tau candidates. A scan is performed over the two components of the missing transverse momentum vector and the yet undetermined variables. Each scan point is weighted by its probability according to the $E_T^{miss}$ resolution and the tau decay topologies. The estimator for the $\tau\tau$ mass is defined as the most probable value of the scan points.

The MMC algorithm provides a solution for $\sim 99\%$ of the $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ events. This is a distinct advantage compared to the mass calculation using the collinear approximation where the failure rate is higher due to the implicit collinearity assumptions. The small loss rate of about $1\%$ for signal events is due to large fluctuations of the $E_T^{miss}$ measurement or other scan variables.

Figure 1 shows reconstructed $m_{\tau\tau}^{MMC}$ mass distributions for $H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ events in the $\tau_{lep}\tau_{had}$ VBF and boosted categories. The mass resolution, defined as the ratio between the full width at half maximum (FWHM) and the peak value of the mass distribution ($m_{peak}$), is found to be $\approx 30\%$ for all categories and channels.

### 5 Boosted decision trees

Boosted decision trees are used in each category to extract the Higgs boson signal from the large number of background events. Decision trees [24] recursively partition the parameter
Figure 1. The reconstructed invariant $\tau\tau$ mass, $m_{\tau\tau}^{\text{MMC}}$ for $H \rightarrow \tau\tau$ ($m_H = 125 \text{ GeV}$) and $Z \rightarrow \tau\tau$ events in MC simulation and embedding respectively, for events passing (a) the VBF category selection and (b) the boosted category selection in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel.

Separate BDTs are trained for each analysis category and channel with signal and background samples, described in section 6, at $\sqrt{s} = 8 \text{ TeV}$. They are then applied to the analysis of the data at both centre-of-mass energies. The separate training naturally exploits differences in event kinematics between different Higgs boson production modes. It also allows different discriminating variables to be used to address the different background compositions in each channel. A large set of potential variables was investigated, in each channel separately, and only those variables which led to an improved discrimination performance of the BDT were kept. For the training in the VBF category, only a VBF Higgs production signal sample is used, while training in the boosted category uses ggF, VBF, and VH signal samples. The Higgs boson mass is chosen to be $m_H = 125 \text{ GeV}$ for all signal samples. The BDT input variables used at both centre-of-mass energies are listed in table 5. Most of these variables have straightforward definitions, and the more complex ones are defined in the following.

- $\Delta R(\tau_1, \tau_2)$: the distance $\Delta R$ between the two leptons, between the lepton and $\tau_{\text{had}}$, or between the two $\tau_{\text{had}}$ candidates, depending on the decay mode.

- $p_T^{\text{Total}}$: magnitude of the vector sum of the transverse momenta of the visible tau decay products, the two leading jets, and $E_T^{\text{miss}}$. 


• Sum $p_T$: scalar sum of the $p_T$ of the visible components of the tau decay products and of the jets.

• $E_T^{\text{miss}}\phi$ centrality: a variable that quantifies the relative angular position of the missing transverse momentum with respect to the visible tau decay products in the transverse plane. The transverse plane is transformed such that the direction of the tau decay products are orthogonal, and that the smaller $\phi$ angle between the tau decay products defines the positive quadrant of the transformed plane. The $E_T^{\text{miss}}\phi$ centrality is defined as the sum of the $x$- and $y$-components of the $E_T^{\text{miss}}$ unit vector in this transformed plane.

• Sphericity: a variable that describes the isotropy of the energy flow in the event [101]. It is based on the quadratic momentum tensor

$$S^{\alpha\beta} = \frac{\sum_i p_i^{\alpha} p_i^{\beta}}{\sum_i |p_i|^2}.$$  \hspace{1cm} (5.1)

In this equation, $\alpha$ and $\beta$ are the indices of the tensor. The summation is performed over the momenta of the selected leptons and jets in the event. The sphericity of the event ($S$) is then defined in terms of the two smallest eigenvalues of this tensor, $\lambda_2$ and $\lambda_3$,

$$S = \frac{3}{2}(\lambda_2 + \lambda_3).$$  \hspace{1cm} (5.2)

• min$(\Delta \eta_{\ell_1,\ell_2,jets})$: the minimum $\Delta \eta$ between the dilepton system and either of the two jets.

• Object $\eta$ centrality: a variable that quantifies the $\eta$ position of an object (an isolated lepton, a $\tau_{\text{had}}$ candidate or a jet) with respect to the two leading jets in the event. It is defined as

$$C_{\eta,\eta_1,\eta_2}(\eta) = \exp \left[ -\frac{4}{(\eta_1 - \eta_2)^2} \left( \eta - \frac{\eta_1 + \eta_2}{2} \right)^2 \right],$$  \hspace{1cm} (5.3)

where $\eta$, $\eta_1$ and $\eta_2$ are the pseudorapidities of the object and the two leading jets respectively. This variable has a value of 1 when the object is halfway in $\eta$ between the two jets, $1/e$ when the object is aligned with one of the jets, and $< 1/e$ when the object is not between the jets in $\eta$. In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel the $\eta$ centrality of a third jet in the event, $C_{\eta,\eta_1}(\eta_{j_3})$, and the product of the $\eta$ centralities of the two leptons are used as BDT input variables, while in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel the $\eta$ centrality of the lepton, $C_{\eta,\eta_1}(\eta_{\ell})$, is used, and in the $\tau_{\text{had}}\tau_{\text{had}}$ channel the $\eta$ centrality of each $\tau$, $C_{\eta,\eta_1}(\eta_{\tau_1})$ and $C_{\eta,\eta_2}(\eta_{\tau_2})$, is used. Events with only two jets are assigned a dummy value of $-0.5$ for $C_{\eta,\eta_1}(\eta_{j_3})$.

Among these variables the most discriminating ones include $m_{\tau\tau}^{\text{MMC}}$, $\Delta R(\tau_1, \tau_2)$ and $\Delta \eta(j_1, j_2)$. Figures 2 and 3 show the distributions of selected BDT input variables. For the VBF category, the distributions of $\Delta \eta(j_1, j_2)$ and centrality variables are shown for all three channels. For the boosted category, the distributions of $\Delta R(\tau_1, \tau_2)$ and $E_T^{\text{miss}}\phi$
<table>
<thead>
<tr>
<th>Variable</th>
<th>VBF</th>
<th>Boosted</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{\text{MMC}}^{\tau\tau}$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\Delta R(\tau_1, \tau_2)$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\Delta \eta(p_{j1}, p_{j2})$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$m_{j1, j2}$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\eta_{j1} \times \eta_{j2}$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$p_T^{\text{Total}}$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sum $p_T$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$p_T^{j1} / p_T^{j2}$</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$E_T^{\text{miss}} \phi$ centrality</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$m_{\ell, \ell, j1}$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$m_{\ell_1, \ell_2}$</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$\Delta \phi(\ell_1, \ell_2)$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Sphericity</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$p_T^{\ell_1}$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$p_T^{\ell_2}$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$E_T^{\text{miss}} / p_T^{\ell_2}$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$m_T$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$\min(\Delta \eta_{\ell_1, \ell_2, \text{jet}es})$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$C_{\eta_1, \eta_2}(\eta_{\ell_1}) \cdot C_{\eta_1, \eta_2}(\eta_{\ell_2})$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$C_{\eta_1, \eta_2}(\eta_{\ell_i})$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$C_{\eta_1, \eta_2}(\eta_{\tau_1})$</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$C_{\eta_1, \eta_2}(\eta_{\tau_2})$</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 5. Discriminating variables used in the training of the BDT for each channel and category at $\sqrt{s} = 8$ TeV. The more complex variables are described in the text. The filled circles indicate which variables are used in each case.

centrality are shown for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels, and the distribution of the $p_T$ of the leading jet and the sphericity are shown for the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel. For all distributions, the data are compared to the predicted SM backgrounds at $\sqrt{s} = 8$ TeV. The corresponding uncertainties are indicated by the shaded bands. All input distributions are well described, giving confidence that the background models (from simulation and data) describe well the relevant input variables of the BDT. Similarly, good agreement is found for the distributions at $\sqrt{s} = 7$ TeV.
Figure 2. Distributions of important BDT input variables for the three channels and the two categories (VBF, left) and (boosted, right) for data collected at $\sqrt{s} = 8$ TeV. The distributions are shown for (a) the separation in pseudorapidity of the jets, $\Delta \eta(j_1, j_2)$, and (b) the transverse momentum of the leading jet $p_T^{j_1}$ in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, for (c) $\Delta \eta(j_1, j_2)$ and (d) $\Delta R(\tau_1, \tau_2)$, the distance $\Delta R$ between the lepton and $\tau_{\text{had}}$, in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel and for (e) $\Delta \eta(j_1, j_2)$ and (f) $\Delta R(\tau_1, \tau_2)$, the distance $\Delta R$ between the two $\tau_{\text{had}}$ candidates, in the $\tau_{\text{had}}\tau_{\text{had}}$ channel. The contributions from a Standard Model Higgs boson with $m_H = 125$ GeV are superimposed, multiplied by a factor of 50. These figures use background predictions made without the global fit defined in section 8. The error band includes statistical and pre-fit systematic uncertainties.
Figure 3. Distributions of important BDT input variables for the three channels and the two categories (VBF, left) and (boosted, right) for data collected at $\sqrt{s} = 8$ TeV. The distributions are shown for (a) the product of the lepton centralities, $C_{\eta_1, \eta_2}(\eta_l_1) \cdot C_{\eta_1, \eta_2}(\eta_l_2)$, and (b) the sphericity in the $\tau_{lep}\tau_{lep}$ channel, for (c) the centrality of the lepton, $C_{\eta_1, \eta_2}(\eta_l)$, and (d) the $E_{T}^{miss} \phi$ centrality in the $\tau_{lep}\tau_{had}$ channel, and for (e) the centrality of the subleading tau, $C_{\eta_1, \eta_2}(\eta_{\tau_2})$, and (f) the $E_{T}^{miss} \phi$ centrality in the $\tau_{had}\tau_{had}$ channel. The contributions from a Standard Model Higgs boson with $m_H = 125$ GeV are superimposed, multiplied by a factor of 50. These figures use background predictions made without the global fit defined in section 8. The error band includes statistical and pre-fit systematic uncertainties.
6 Background estimation

The different final-state topologies of the three analysis channels have different background compositions which necessitate different strategies for the background estimation. In general, the number of expected background events and the associated kinematic distributions are derived from a mixture of data-driven methods and simulation. The normalisation of several important background contributions is performed by comparing the simulated samples of individual background sources to data in regions which only have a small or negligible contamination from signal or other background events. The control regions used in the analysis are summarised in table 6.

Common to all channels is the dominant $Z \rightarrow \tau\tau$ background, for which the kinematic distributions are taken from data by employing the embedding technique, as described in section 3. Background contributions from jets that are misidentified as hadronically decaying taus (fake backgrounds) are estimated by using either a fake-factor method or samples of non-isolated $\tau_{\text{had}}$ candidates. Likewise, samples of non-isolated leptons are used to estimate fake-lepton contributions from jets or hadronically decaying taus and leptons from other sources, such as heavy-quark decays.\(^8\)

Contributions from various other physics processes with leptons and/or $\tau_{\text{had}}$ candidates in the final state are estimated using the simulation, normalised to the theoretical cross sections, as given in table 1. A more detailed discussion of the estimation of the various background components in the different channels is given in the following.

6.1 Background from $Z \rightarrow \tau\tau$ production

A reliable modelling of the irreducible $Z \rightarrow \tau\tau$ background is an important ingredient of the analysis. It has been shown in other ATLAS analyses that existing $Z + \text{jets}$ Monte Carlo simulation needs to be reweighted to model data correctly [102–104]. Additionally, it is not possible to select a sufficiently pure and signal-free $Z \rightarrow \tau\tau$ control sample from data to model the background in the signal region. Therefore this background is estimated using embedded data, as described in section 3. This procedure was extensively validated using both data and simulation. To validate the subtraction procedure of the muon cell energies and tracks from data and the subsequent embedding of the corresponding information from simulation, the muons in $Z \rightarrow \mu\mu$ events are replaced by simulated muons. The calorimeter isolation energy in a cone of $\Delta R = 0.3$ around the muons from data before and after embedding is compared in figure 4(a). Good agreement is found, which indicates that no deterioration (e.g. possible energy biases) in the muon environment is introduced. Another important test validates the embedding of more complex $Z \rightarrow \tau\tau$ events, which can only be performed in the simulation. To achieve a meaningful validation, the same MC generator with identical settings was used to simulate both the $Z \rightarrow \mu\mu$ and $Z \rightarrow \tau\tau$ events. The sample of embedded events is corrected for the bias due to the trigger, reconstruction and acceptance of the original muons. These corrections are determined from data as a function of $p_T^\mu$ and $\eta(\mu)$, and allow the acceptance of the original selection to be corrected. The tau decay products are treated like any other objects obtained from the simulation, with one

\(^8\)For simplicity, leptons from heavy-quark decays are considered as fake leptons in the following.
Table 6. Summary and definition of the control regions used in the analysis. The requirements shown represent modifications to the signal region requirements. All other selections are applied as for the corresponding signal regions. The variable $m_{\tau\tau}^{\text{HPTO}}$ is the invariant mass of the $\tau\tau$-system obtained using the collinear approximation and the same objects as used in $E_{\text{T}}^{\text{miss, HPTO}}$. Each control region listed is actually two control regions, corresponding to the VBF or boosted categories, with the exception of the rest category. The first three control regions listed are used in the global fit defined in section 8.
important difference due to the absence of trigger simulation in this sample. Trigger effects are parameterised from the simulation as a function of the tau decay product $p_T$. After replacing the muons with simulated taus, kinematic distributions of the embedded sample can be directly compared to the fully simulated ones. As an example, the reconstructed invariant mass, $m_{\text{MMC}}^{\tau\tau}$, is shown in figure 4(b), for the $\tau_{\text{lep}}\tau_{\text{had}}$ final state. Good agreement is found and the observed differences are covered by the systematic uncertainties. Similarly, good agreement is found for other variables, such as the missing transverse momentum, the kinematic variables of the hadronically decaying tau lepton or of the associated jets in the event. A direct comparison of the $Z \rightarrow \tau\tau$ background in data and the modelling using the embedding technique also shows good agreement. This can be seen in several kinematic quantity distributions, which are dominated by $Z \rightarrow \tau\tau$ events, shown in figure 2.

The normalisation of this background process is taken from the final fit described in section 8. The normalisation is independent for the $\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$, and $\tau_{\text{had}}\tau_{\text{had}}$ analysis channels.

6.2 Background from misidentified leptons or hadronically decaying taus

For the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, all background sources resulting from misidentified leptons are treated together. In this approach, contributions from multijet and $W+\text{jets}$ production,
as well as the part of the $t\bar{t}$ background resulting from decays to leptons and hadrons ($t\bar{t} \to \ell\nu b q q$) are included. A control sample is defined in data by inverting the isolation requirements for one of the two leptons, while applying all other signal region requirements. The contributions from other background channels (dileptonic $t\bar{t}$ decays ($t\bar{t} \to \ell\nu b \ell\nu b$), $Z \to ee$, $Z \to \mu\mu$, and diboson production) are obtained from the simulation and are subtracted. From this control sample a template is created. The normalisation factor is obtained by fitting the $p_T$ distribution of the subleading lepton at an early stage of the preselection.

For the $\tau_{lep}\tau_{had}$ channel, the fake-factor method is used to derive estimates for the multijet, $W$+jets, $Z$+jets, and semileptonic $t\bar{t}$ background events that pass the $\tau_{lep}\tau_{had}$ selection due to a misidentified $\tau_{had}$ candidate. The fake factor is defined as the ratio of the number of jets identified as medium $\tau_{had}$ candidates to the number satisfying the loose, but not the medium, criteria. Since the fake factor depends on the type of parton initiating the jet and on the $p_T$ of the jet, it is determined as a function of $p_T$ separately for samples enriched in quark- and gluon-initiated jets. In addition, the fake factor is found to be different for 1-track and 3-track candidates. Three different, quark-jet dominated samples are used separately for the $W$+jets, $t\bar{t}$ and $Z$+jets background components. They are defined by selecting the high-$m_T$ region ($m_T > 70$ GeV), by inverting the $b$-jet veto and by requiring two leptons with an invariant mass consistent with $m_Z (80$ GeV $< m_\ell\ell < 100$ GeV) respectively. In addition, a multijet sample dominated by gluon-initiated jets is selected by relaxing the lepton identification and requiring it to satisfy the loose identification criteria. The derived fake factors are found to vary from 0.124 (0.082) for $p_T = 20$ GeV to 0.088 (0.038) for $p_T = 150$ GeV for 1-track (3-track) candidates in the VBF category. The corresponding values for the boosted category are 0.146 (0.084) for $p_T = 20$ GeV and 0.057 (0.033) for $p_T = 150$ GeV. To obtain the fake-background estimate for the VBF and boosted signal regions, these factors are then applied, weighted by the expected relative $W$+jets, $Z$+jets, multijet, and $t\bar{t}$ fractions, to the events in regions defined by applying the selections of the corresponding signal region, except that the $\tau_{had}$ candidate is required to pass the loose and to fail the medium $\tau_{had}$ identification. As an example, the good agreement between data and background estimates is shown in figure 5(a) for the reconstructed $\tau\tau$ mass for events in the high-$m_T$ region, which is dominated by $W$+jets production.

For the $\tau_{had}\tau_{had}$ channel, the multijet background is modelled using a template extracted from data that pass the VBF or boosted category selection, where, however, the taus fail the isolation and opposite-sign charge requirements (the number-of-tracks requirement is not enforced). The normalisation of the multijet background is first determined by performing a simultaneous fit of the multijet (modelled by the data sample just mentioned) and $Z \to \tau\tau$ (modelled by embedding) templates after the preselection cuts. The fit is performed for the distribution of the difference in pseudorapidity between the two hadronic tau candidates, $\Delta\eta(\tau_{had}, \tau_{had})$. The signal contribution is expected to be small in this category. The agreement between data and the background estimate for this distribution is shown in figure 5(b) for the rest category defined in section 4. The preselection normalisation is used as a reference point and starting value for the global fit (see below) and is used for validation plots. The final normalisations of the two important background components,
from multijet and $Z \rightarrow \tau\tau$ events, are extracted from the final global fit, as described in section 8, in which the $\Delta \eta(\tau_{\text{had}}, \tau_{\text{had}})$ distribution for the rest category is included.

### 6.3 $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ background

The Drell-Yan $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ background channels are important contributions to the final states with two same-flavour leptons. They also contribute to the other channels. As described below, a simulation based on ALPGEN is used to estimate these background sources. Correction factors are applied to account for differences between data and simulation.

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, the ALPGEN simulation is normalised to the data in the $Z$-mass control region, $80 \text{ GeV} < m_{\ell\ell} < 100 \text{ GeV}$, for each category, and separately for $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ events. The normalisation factors are determined from the final fit described in section 8. The distribution of the reconstructed $\tau\tau$ mass for events in this control region is shown in figure 6(a).

In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ background estimates are also based on simulation. The corrections applied for a $\tau_{\text{had}}$ candidate depend on whether it originates from a lepton from the $Z$ boson decay or from a jet. In the first case, corrections from data, derived from dedicated tag-and-probe studies, are applied to account for the difference in the rate of misidentified $\tau_{\text{had}}$ candidates resulting from leptons [21, 105]. This is particularly important for $Z \rightarrow ee$ events with a misidentified $\tau_{\text{had}}$ candidate originating from a true electron. In the second case, the fake-factor method described in section 6.2 is applied.
Figure 6. (a) The distribution of the reconstructed invariant ττ mass, $m_{\text{MMC}}^{\tau\tau}$, for events in the $Z \rightarrow \ell\ell$ control region, for the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel. (b) The distribution of the separation in pseudorapidity of the two leading jets, $\Delta\eta(j_1, j_2)$, for events in the top control region, for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. This figure uses background predictions made without the global fit defined in section 8. The error band includes statistical and pre-fit systematic uncertainties.

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the contribution of this background is very small and is taken from simulation.

6.4 W+jets background

Events with $W$ bosons and jets constitute a background to all channels since leptonic $W$ decays can feed into all signatures when the true lepton is accompanied by a jet which is falsely identified as a $\tau_{\text{had}}$ or a lepton candidate. This process can also contribute via semileptonic heavy quark decays that provide identified leptons.

As stated in section 6.2, for the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, the $W$+jets contributions are determined with data-driven methods. For the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the $W \rightarrow \tau_{\text{had}}\nu$ background is estimated from simulation. A correction is applied to account for differences in the $\tau_{\text{had}}$ misidentification probability between data and simulation.

6.5 Background from top-quark production

Background contributions from $t\bar{t}$ and single top-quark production, where leptons or hadronically decaying taus appear in decays of top quarks, are estimated from simulation in the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels. The normalisation is obtained from data control regions defined by requiring a $b$-jet instead of a $b$-veto. In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, a large value of the transverse mass $m_T$ is also required, to enhance the background from top-quark production and to suppress the signal contribution. This background is also found to be small for the $\tau_{\text{had}}\tau_{\text{had}}$ channel and it is estimated using simulation. The distribution of $\Delta\eta(j_1, j_2)$ for events in the top control region, for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, is shown in figure 6(b).
### Table 7

<table>
<thead>
<tr>
<th>Channel</th>
<th>Background</th>
<th>Scale factors (CR)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>VBF</td>
<td>Boosted</td>
</tr>
<tr>
<td>$\tau_{\text{lep}}\tau_{\text{lep}}$</td>
<td>Top</td>
<td>$0.99 \pm 0.07$</td>
<td>$1.01 \pm 0.05$</td>
</tr>
<tr>
<td></td>
<td>$Z \rightarrow ee$</td>
<td>$0.91 \pm 0.16$</td>
<td>$0.98 \pm 0.10$</td>
</tr>
<tr>
<td></td>
<td>$Z \rightarrow \mu\mu$</td>
<td>$0.97 \pm 0.13$</td>
<td>$0.96 \pm 0.08$</td>
</tr>
<tr>
<td>$\tau_{\text{lep}}\tau_{\text{had}}$</td>
<td>Top</td>
<td>$0.84 \pm 0.08$</td>
<td>$0.96 \pm 0.04$</td>
</tr>
</tbody>
</table>

The scale factors calculated in control regions (CR) for background normalisation. Only the statistical uncertainties are given. The background contributions shown in validation plots use the normalisation predicted from simulation multiplied by the corresponding scale factor. Final normalisations are taken from the global fit, described in section 8.

### 6.6 Diboson background

The production of pairs of vector bosons ($W^+W^-$, $ZZ$ and $W^\pm Z$), with subsequent decays to leptons or jets, contributes especially to the background in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel. For all analysis channels, these contributions are estimated from simulation, normalised to the NLO cross sections indicated in table 1.

### 6.7 Contributions from other Higgs boson decays

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, a non-negligible contribution from $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ exists and this process is considered as background. Its contribution is estimated for $m_H = 125$ GeV using simulation. The corresponding signal cross section is assumed to be the SM value and is indicated in table 1.

### 6.8 Validation of background estimates

As described above, the normalisation for important background sources that are modelled with simulation are determined by fitting to data in control regions. These normalisations are compared in table 7 to predictions based on the theoretical cross sections for the 8 TeV analysis. In most cases, the values obtained are compatible with unity within the statistical uncertainties shown. For the top control region in the VBF category of the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the value is also in agreement with unity if the experimental and theoretical systematic uncertainties are included. The control-region normalisations are used for validation plots, and they are used as starting values in the final global fit described in section 8. The global fit does not change any of these normalisations by more than 2%.

It is important to verify that the BDT output distributions in data control regions are well described after the various background determinations. Figure 7 shows distributions from important control regions for the $\sqrt{s} = 8$ TeV dataset, i.e. the $Z$-enriched control regions for the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, and the reconstructed $\tau\tau$ invariant mass sideband control region (defined as $m_{\tau\tau}^{\text{MMC}} < 100$ GeV or $m_{\tau\tau}^{\text{MMC}} > 150$ GeV) for the $\tau_{\text{had}}\tau_{\text{had}}$ channel. The distributions are shown for both the VBF and the boosted categories. All distributions are found to be well described, within the systematic uncertainties.
Figure 7. Distributions of the BDT output for data collected at $\sqrt{s} = 8$ TeV, compared to the expected background contributions in important control regions. The distributions are shown for the VBF (left) and boosted (right) categories (a,b) for the $Z\rightarrow \tau\tau$-enriched control region in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, (c,d) for the $Z\rightarrow \tau\tau$-enriched control region in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel and (e,f) for the $\tau\tau$ invariant mass sideband control region in the $\tau_{\text{had}}\tau_{\text{had}}$ channel. The contributions from a Standard Model Higgs boson with $m_H = 125$ GeV are superimposed, multiplied by a factor of 50. These figures use background predictions made without the global fit defined in section 8. The error band includes statistical and pre-fit systematic uncertainties.
7 Systematic uncertainties

The numbers of expected signal and background events, the input variables to the BDT, and thereby the BDT output and the final discrimination between signal and background are affected by systematic uncertainties. They are discussed below, grouped into three categories: experimental uncertainties, background modelling uncertainties, and theoretical uncertainties. For all uncertainties, the effects on both the total signal and background yields and on the shape of the BDT output distribution are evaluated. Table 8 gives a summary of the systematic uncertainties and their impact on the number of expected events for the signal and the total background for the analysis of the data taken at $\sqrt{s} = 8$ TeV.

The dominant sources that affect the shape of the BDT output distribution are marked in the table. All uncertainties are treated either as fully correlated or uncorrelated across channels. The latter are also marked in table 8. The effects of the systematic uncertainties at $\sqrt{s} = 7$ TeV are found to be similar and are not discussed here. The inclusion of the uncertainties in the profile likelihood global fit is described in section 8 and the effect of the most significant systematic uncertainties is presented in table 13.

7.1 Experimental uncertainties

The major experimental systematic uncertainties result from uncertainties on efficiencies for triggering, object reconstruction and identification, as well as from uncertainties on the energy scale and resolution of jets, hadronically decaying taus and leptons. In general, the effects resulting from lepton-related uncertainties are smaller than those from jets and taus. They are not discussed in detail, however, their impact is included in table 8. In addition, uncertainties on the luminosity affect the number of signal and background events from simulation.

- Luminosity: the uncertainty on the integrated luminosity is $\pm 2.8\%$ for the 8 TeV dataset and $\pm 1.8\%$ for the 7 TeV dataset. It is determined from a calibration of the luminosity scale derived from beam-separation scans performed in 2011 and 2012 using the method described in ref. [106].

- Efficiencies: the efficiencies for triggering, reconstructing and identifying electrons, muons, and $\tau_{\text{had}}$ candidates are measured in data using tag-and-probe techniques. The uncertainties on the $\tau_{\text{had}}$ identification efficiency are $\pm (2\text{–}3)\%$ for 1-prong and $\pm (3\text{–}5)\%$ for 3-prong tau decays [39]. The $b$-jet tagging efficiency has been measured from data using $t\bar{t}$ events, where both top quarks decay to leptons, with a total uncertainty of about $\pm 2\%$ for jets with transverse momenta up to 100 GeV [38, 107]. The MC samples used are corrected for differences in these efficiencies between data and simulation and the associated uncertainties are propagated through the analysis.

- Energy scales: the uncertainties on the jet energy scale (JES) arise from several sources. These include, among others, varied response due to the jet flavour composition (quark- versus gluon-initiated jets), pile-up, $\eta$ intercalibration, and detector
response and modelling of in situ jet calibration [35, 36]. The impact of the JES uncertainty in this analysis is reduced because many of the background components are estimated using data. The tau energy scale is obtained by fitting the reconstructed visible mass for $Z \rightarrow \tau \tau$ events in data, which can be selected with a satisfactory purity. It is measured with a precision of $\pm (2-4)\%$ [108]. Since systematic uncertainties on the energy scales of all objects affect the reconstructed missing transverse momentum, it is recalculated after each variation is applied. The scale uncertainty on $E_T^{\text{miss}}$ due to the energy in calorimeter cells not associated with physics objects is also taken into account.

- Energy resolutions: systematic uncertainties on the energy resolution of taus, electrons, muons, jets, and $E_T^{\text{miss}}$ affect the final discriminant. The effects resulting from uncertainties on the tau energy resolution are small. The impact of changes in the amount of material (inactive material in the detector, e.g. support structures), in the hadronic shower model and in the underlying-event tune were studied in the simulation. They result in systematic uncertainties below 1% on the tau energy resolution. The jet energy resolution is determined by in situ measurements, as described in ref. [109], and affects signal modelling and background components modelled by the simulation. The uncertainty of the resolution on $E_T^{\text{miss}}$ is estimated by evaluating the energy resolution of each of the $E_T^{\text{miss}}$ terms. The largest impact results from the soft term (see section 2), arising both from the MC modelling and the effects of pile-up. It is evaluated using simulated $Z(\rightarrow \mu\mu)+\text{jets}$ events.

### 7.2 Background modelling uncertainties

The most significant systematic uncertainties on the background estimation techniques, as described in section 6, are detailed in the following for the three decay modes considered.

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, systematic uncertainties on the shape and normalisation of fake-lepton background sources are estimated by comparing samples of same-sign lepton events that pass and fail the lepton isolation criteria. These uncertainties amount to $\pm 33\%$ ($\pm 20\%$) at 8 TeV and $\pm 10.5\%$ ($\pm 13\%$) at 7 TeV for the boosted (VBF) category. The extrapolation uncertainty for the $Z \rightarrow \ell\ell$ background is obtained by varying the $m_{\ell\ell}$ window that defines the control region for this background, and amounts to about $\pm 6\%$. The corresponding extrapolation uncertainty for top-quark background sources is $\pm (3-6)\%$, obtained from the differences in event yields in the top-quark control regions when using different MC generators. Neither of these extrapolation uncertainties is significant for the final result. The dominant uncertainties on the normalisation of the $t\bar{t}$ background, obtained from the global fit, are the systematic uncertainties on the $b$-jet tagging efficiency and the jet energy scale.

In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, an important systematic uncertainty on the background determination comes from the estimated fake background, for which several sources of systematic uncertainty are considered. The statistical uncertainty on the effective fake factor is $\pm 4.3\%$ ($\pm 2.3\%$) in the 8 TeV VBF (boosted) category, and about $\pm 22\%$ ($\pm 11\%$) in the 7 TeV VBF (boosted) category. The dominant systematic uncertainty on the methodology itself arises
from the composition of the combined fake background ($W$+jets, $Z$+jets, multijet, and $t\bar{t}$ fractions), which is largely estimated based on simulated event samples as explained in section 6.2. The uncertainty is estimated by varying each fractional contribution by ±50%, which affects the effective fake factor by ±3% (±6%) and by ±10% (±15%) in the 8 TeV and 7 TeV boosted (VBF) categories respectively. As a closure test, the method was also applied in a region of data where the lepton and $\tau_{\text{had}}$ candidate have the same charge, rich in fake $\tau_{\text{had}}$ candidates. Very good agreement was observed between data and the method’s prediction, so that no additional in situ uncertainty was deemed necessary. In addition, the uncertainties on the normalisation of the $t\bar{t}$ background are important. As in the case of the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, the dominant contribution obtained from the global fit originates from systematic uncertainties on the $b$-jet tagging efficiency and the jet energy scale, along with statistical uncertainties on the observed data in the respective control regions.

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the major background from multijet production is determined using a data-driven template method. The default multijet template, derived from a sample in data where the $\tau_{\text{had}}$ candidates fail the isolation and opposite-sign charge requirements, is compared with an alternative template derived from a sample where the $\tau_{\text{had}}$ candidates fail just the opposite-sign charge requirement. The normalisation of the alternative template is fixed to that of the default template at preselection; the alternative multijet template is propagated into the various categories and gives a different set of yields from the default template. This difference, along with the difference in shape between the two templates, constitutes the systematic uncertainty on the background estimate. This leads to an overall multijet yield variation of 10% (3%) in the VBF (boosted) category at $\sqrt{s} = 8$ TeV and of 10% (30%) in the VBF (boosted) category $\sqrt{s} = 7$ TeV. However, there is a very strong shape dependence, such that the uncertainties on the BDT output are much larger at higher output values.

For the embedding method used in all channels, the major systematic uncertainties are related to the selection of $Z \rightarrow \mu\mu$ events in data and to the subtraction of the muon energy deposits in the calorimeters. The selection uncertainties are estimated by varying the muon isolation criteria in the selection from the nominal value of $I(p_T,0.2) < 0.2$ (see section 4) to tighter ($I(p_T,0.4) < 0.06$ and $I(E_T,0.2) < 0.04$) and looser (no isolation requirements) values. The muon-related cell energies to be subtracted are varied within ±20% (±30%) for the 8 TeV (7 TeV) data. In addition, systematic uncertainties on the corrections for trigger and reconstruction efficiencies are taken into account. Due to the combination of single-lepton and dilepton triggers used, the uncertainties are largest for the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel. All experimental systematic uncertainties relating to the embedded $\tau$ decay products (such as tau energy scale or identification uncertainties) are applied normally. The combined effect of all uncertainties on the signal and background yields is included in table 8. Because the $Z \rightarrow \tau\tau$ normalisation is determined in the final fit, the impact on the final result is much smaller.

### 7.3 Theoretical uncertainties

Theoretical uncertainties are estimated for the signal and for all background contributions modelled with the simulation. Since the major background contributions, from $Z \rightarrow \tau\tau$
Relative signal and background variations [%]

<table>
<thead>
<tr>
<th>Source</th>
<th>( \tau\ell\ell ) VBF</th>
<th>( \tau\ell\ell ) Boosted</th>
<th>( \tau\ell\text{had} ) VBF</th>
<th>( \tau\ell\text{had} ) Boosted</th>
<th>( \ell\ell\text{had} ) VBF</th>
<th>( \ell\ell\text{had} ) Boosted</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luminosity</td>
<td>(+2.8) ± 0.1</td>
<td>(+2.8) ± 0.1</td>
<td>(+2.8) ± 0.1</td>
<td>(+2.8) ± 0.1</td>
<td>(+2.8) ± 0.1</td>
<td>(+2.8) ± 0.1</td>
</tr>
<tr>
<td>Tau trigger*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Tau identification</td>
<td>–</td>
<td>–</td>
<td>(+3.3) ± 1.2</td>
<td>(+3.3) ± 1.8</td>
<td>(+6.6) ± 3.8</td>
<td>(+6.6) ± 5.1</td>
</tr>
<tr>
<td>Lepton ident. and trigger*</td>
<td>(+1.4) ± 1.3</td>
<td>(+1.4) ± 1.1</td>
<td>(+1.8) ± 0.5</td>
<td>(+1.8) ± 0.8</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>b-tagging</td>
<td>(+1.3) ± 1.6</td>
<td>(+1.6) ± 1.6</td>
<td>(&lt;0.1) ± 0.2</td>
<td>(+0.4) ± 0.2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(\tau) energy scale†</td>
<td>–</td>
<td>–</td>
<td>(+2.4) ± 1.3</td>
<td>(+2.4) ± 0.9</td>
<td>(+2.9) ± 2.5</td>
<td>(+2.9) ± 2.5</td>
</tr>
<tr>
<td>Jet energy scale and resolution†</td>
<td>(+9.5) ± 4.7</td>
<td>(+4.7) ± 3.7</td>
<td>(+5.1) ± 1.0</td>
<td>(+3.9) ± 0.4</td>
<td>(+10.1) ± 5.1</td>
<td>(+5.1) ± 0.2</td>
</tr>
<tr>
<td>(E_T^{\text{miss}}) soft scale &amp; resolution†</td>
<td>(+0.0) ± 0.0</td>
<td>(+0.0) ± 0.0</td>
<td>(+0.8) ± 0.2</td>
<td>(+0.4) ± 0.2</td>
<td>(+0.5) ± 0.2</td>
<td>(+0.1) ± 0.1</td>
</tr>
<tr>
<td><strong>Background Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modelling of fake backgrounds*</td>
<td>–</td>
<td>–</td>
<td>(+1.2) ± 0.2</td>
<td>(+2.6) ± 0.3</td>
<td>(+5.2) ± 0.6</td>
<td>(+0.6) ± 0.6</td>
</tr>
<tr>
<td>Embedding†</td>
<td>(+3.8) ± 4.3</td>
<td>(+3.6) ± 6.5</td>
<td>(+1.5) ± 1.2</td>
<td>(+2.2) ± 0.3</td>
<td>(+3.3) ± 0.3</td>
<td>(+3.3) ± 0.3</td>
</tr>
<tr>
<td>(Z\to\ell\ell) normalisation*</td>
<td>(+2.1) ± 0.7</td>
<td>(+0.7) ± 0.7</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>Theoretical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher-order QCD corrections†</td>
<td>(+11.3) ± 15.3</td>
<td>(+19.8) ± 7.6</td>
<td>(+23.7) ± 2.7</td>
<td>(+19.0) ± 14.7</td>
<td>(+10.7) ± 8.2</td>
<td>(+20.3) ± 15.4</td>
</tr>
<tr>
<td>UE/PS</td>
<td>(+1.8) ± 0.1</td>
<td>(+5.9) ± 0.1</td>
<td>(+3.8) ± 0.1</td>
<td>(+2.9) ± 0.1</td>
<td>(+4.6) ± 0.1</td>
<td>(+3.8) ± 0.1</td>
</tr>
<tr>
<td>Generator modelling</td>
<td>(+2.3) ± 0.1</td>
<td>(+1.2) ± 0.1</td>
<td>(+2.7) ± 0.1</td>
<td>(+1.3) ± 0.1</td>
<td>(+2.4) ± 0.1</td>
<td>(+1.2) ± 0.1</td>
</tr>
<tr>
<td>EW corrections</td>
<td>(+1.1) ± 0.4</td>
<td>(+1.3) ± 0.1</td>
<td>(+0.4) ± 0.1</td>
<td>(+1.1) ± 0.1</td>
<td>(+0.4) ± 0.1</td>
<td>(+0.4) ± 0.1</td>
</tr>
<tr>
<td>PDF†</td>
<td>(+4.5) ± 0.3</td>
<td>(+6.2) ± 0.2</td>
<td>(+3.9) ± 0.2</td>
<td>(+6.6) ± 0.2</td>
<td>(+4.3) ± 0.2</td>
<td>(+6.3) ± 0.1</td>
</tr>
<tr>
<td>BR ((H\to\tau\tau))</td>
<td>(+5.7) ± 5.7</td>
<td>(+5.7) ± 5.7</td>
<td>(+5.7) ± 5.7</td>
<td>(+5.7) ± 5.7</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 8. Impact of systematic uncertainties on the total signal, \(S\), (sum of all production modes) and on the sum of all background estimates, \(B\), for each of the three channels and the two signal categories for the analysis of the data taken at \(\sqrt{s} = 8\) TeV. Each systematic uncertainty is assumed to be correlated across the analysis channels, except those marked with a *. Uncertainties that affect the shape of the BDT-output distribution in a non-negligible way are marked with a †. All values are given before the global fit. The notation UE/PS refers to the underlying event and parton shower modelling.
and misidentification of hadronically decaying \( \tau \) leptons, are estimated using data-driven methods, they are not affected by these uncertainties. Uncertainties on the signal cross sections are assigned from missing higher-order corrections, from uncertainties in the PDFs, and from uncertainties in the modelling of the underlying event.

For VBF and VH Higgs boson production cross sections, the uncertainties due to missing higher-order QCD corrections are estimated by varying the factorisation and renormalisation scales by factors of two around the nominal scale \( m_W \), as prescribed by the LHC Higgs Cross section Working Group [110]. The resulting uncertainties range from \( \pm 2\% \) to \( \pm 4\% \), depending on the process and the category-specific selection considered. In addition, a \( 2\% \) uncertainty related to the inclusion of the NLO EWK corrections (see section 3) is assigned.

For Higgs boson production via ggF, the uncertainties on the cross sections due to missing higher order QCD corrections are estimated by varying the renormalisation and factorisation scales around the central values \( \mu_R = \mu_F = \sqrt{m_H^2 + p_T^2} \) in the NLO cross section calculations of \( H + 1\)-jet and \( H + 2\)-jet production. In the calculation of the uncertainties, appropriate cuts on the Higgs \( p_T \) (\( p_T^H > 100 \text{ GeV} \)) and on the jet kinematics (\( \Delta\eta, p_T \)) are applied at parton level for the boosted and VBF categories respectively. The resulting uncertainties on the ggF contributions are found to be about \( \pm 24\% \) in the boosted category and \( \pm 23\% \) in the VBF category. The ggF contribution is dominant in the boosted category, whereas it is only about \( 20\% \) of the signal in the VBF category. Since the two categories are exclusive, their anti-correlation is taken into account following the prescription in ref. [111].

In the present analysis, no explicit veto on jets is applied in the VBF selection, but enough kinematical information is provided as input to the BDT so that the high BDT-output region corresponds to a more exclusive region, where the probability of finding a third jet is reduced. Since the cross section for gluon-fusion events produced with a third jet is only known at LO, this could introduce a large uncertainty on the gluon-fusion contamination in the highest (and most sensitive) BDT-output bins. The uncertainty on the BDT shape of the ggF contribution is evaluated using the \( \text{Mcfm} \) Monte Carlo program [98], which calculates \( H + 3 \) jets at LO. Scale variations induce changes of the ggF contribution in the highest BDT bin of about \( \pm 30\% \). They are taken into account in the final fit.

Uncertainties related to the simulation of the underlying event and parton shower are estimated by comparing the acceptance from \( \text{Powheg}+\text{Pythia} \) to \( \text{Powheg}+\text{Herwig} \) for both VBF and ggF Higgs boson production modes. Differences in the signal yields range from \( \pm 1\% \) to \( \pm 8\% \) for the VBF and from \( \pm 1\% \) to \( \pm 9\% \) for ggF production, depending on the channel and category. The BDT-score distribution of the \( \text{Powheg}+\text{Pythia} \) and \( \text{Powheg}+\text{Herwig} \) samples are compatible with each other within statistical uncertainties.

The PDF uncertainties are estimated by studying the change in the acceptance when using different PDF sets or varying the CT10 PDF set within its uncertainties. The standard VBF \( \text{Powheg} \) sample and a \( \text{MC}@\text{NLO} \) [112] ggF sample, both generated with the CT10 PDFs, are reweighted to the MSTW2008NLO [113], NNPDF [114] and the
CT10 eigen-tunes parameterisation. The largest variation in acceptance for each category is used as a constant PDF uncertainty; it varies between approximately ±4.5% and ±6% for ggF production and between about ±0.8% and ±1.0% for VBF production. A shape uncertainty is also included to cover any difference between the BDT score in the default sample and the reweighted ones. The uncertainty on the total cross section for the VBF, VH and ggF production modes due to the PDFs is also considered.

Variations in the acceptance for different Monte Carlo generators are also included, comparing Powheg+Herwig samples to MC@NLO+Herwig samples for ggF, and to aMC@NLO+Herwig [115] samples for VBF. The generator modelling uncertainty is around ±2% for ggF and ±4% for VBF productions modes.

Finally, an uncertainty on the decay branching ratio, BR($H \rightarrow \tau\tau$), of ±5.7% [70] affects the signal rates.

The theoretical systematic uncertainties on the background predictions taken from the simulation are evaluated by applying the same procedures as used for the signal samples. Uncertainties resulting from the choice of QCD scales, PDF parameterisation and underlying-event model are estimated. The results are reported in table 8.

### 8 Signal extraction procedure

The BDT output in the six analysis categories provides the final discrimination between signal and background for both the 7 and 8 TeV datasets. A maximum-likelihood fit is performed on all categories simultaneously to extract the signal strength, $\mu$, defined as the ratio of the measured signal yield to the Standard Model expectation. The value $\mu = 0$ ($\mu = 1$) corresponds to the absence (presence) of a Higgs boson signal with the SM production cross section. The statistical analysis of the data employs a binned likelihood function $L(\mu, \hat{\theta})$, constructed as the product of Poisson probability terms, to estimate $\mu$.

The impact of systematic uncertainties on the signal and background expectations is described by nuisance parameters, $\hat{\theta}$, which are each parameterised by a Gaussian or log-normal constraint. The expected numbers of signal and background events in each bin are functions of $\hat{\theta}$. The test statistic $q_\mu$ is then constructed according to the profile likelihood ratio: $q_\mu = -2 \ln[L(\mu, \hat{\theta})/L(\hat{\mu}, \hat{\theta})]$, where $\hat{\mu}$ and $\hat{\theta}$ are the parameters that maximise the likelihood, and $\hat{\theta}$ are the nuisance parameter values that maximise the likelihood for a given $\mu$. This test statistic is used to measure the compatibility of the background-only hypothesis with the observed data.

The likelihood is maximised on the BDT distributions in the signal regions, with information from control regions included to constrain background normalisations. The fit includes the event yields from the $Z \rightarrow \ell\ell$ and top control regions in the $\tau_{lep}\tau_{lep}$ channel, and from the top control region of the $\tau_{lep}\tau_{had}$ channel; furthermore the $\Delta\eta(\tau_{had},\tau_{had})$ distribution in the rest control region of the $\tau_{had}\tau_{had}$ channel is also included.

The $Z \rightarrow \tau\tau$ background is constrained primarily in the signal regions, due to the difference between the BDT distributions for $Z \rightarrow \tau\tau$ events and the signal. For the $\tau_{had}\tau_{had}$ channel, the $Z \rightarrow \tau\tau$ and multijet background rates are constrained by the simultaneous
fit of the two signal regions and the $\Delta\eta(\tau_{\text{had}}\tau_{\text{had}})$ distribution in the rest category control region. The top and $Z \rightarrow \ell\ell$ background components for the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels are also allowed to float freely, but are primarily constrained by the inclusion of the respective control regions.

As described in section 7, a large number of systematic uncertainties, taken into account via nuisance parameters, affect the final results. It is important to investigate the behaviour of the global fit and in particular to investigate how far the nuisance parameters are pulled away from their nominal values and how well their uncertainties are constrained. Furthermore, it is important to understand which systematic uncertainties have the most impact on the final result. For this purpose a ranking of nuisance parameters is introduced. For each parameter, the fit is performed again with the parameter fixed to its fitted value shifted up or down by its fitted uncertainty, with all the other parameters allowed to vary. The ranking obtained for those nuisance parameters contributing most to the uncertainty on the signal strength is shown in figure 8 for the combined fit of the three channels at the two centre-of-mass energies. The parameters contributing most are those related to the jet energy scale, the normalisation uncertainties for $Z \rightarrow \tau\tau$ and top-quark events, and the tau energy scale. The uncertainties on the jet energy scale are decomposed into several uncorrelated components (among others: $\eta$ intercalibration of different calorimeter regions, jet energy response, and response to jets of different flavour). In addition, theoretical uncertainties on the branching ratio $\text{BR}(H \rightarrow \tau\tau)$ are found to have a significant impact. In general, good agreement is found between the pre-fit and post-fit values for these nuisance parameters and neither large pulls nor large constraints are observed.

The distributions of the BDT discriminants for all channels and categories for the data at 8 TeV are shown in figure 9, with background normalisations, signal normalisation, and nuisance parameters adjusted by the profile likelihood global fit.

The results for the numbers of fitted signal and background events, split into the various contributions, are summarised in tables 9, 10 and 11 for the three channels separately, for the dataset collected at 8 TeV centre-of-mass energy. In addition to the total number of events, the expected number of events in each of the two highest BDT output bins is given. The number of events observed in the data is also included. Within the uncertainties, good agreement is observed between the data and the model predictions for the sum of background components and a Standard Model Higgs boson with $m_H = 125$ GeV.
Figure 8. Impact of systematic uncertainties on the fitted signal-strength parameter $\hat{\mu}$ for the combined fit for all channels and both centre-of-mass energies. The systematic uncertainties are listed in decreasing order of their impact on $\hat{\mu}$ on the y-axis. The hatched blue and red boxes show the variations of $\hat{\mu}$ with respect to the total error on $\mu$, $\sigma_{\text{tot}}$, referring to the top x-axis, when fixing the corresponding individual nuisance parameter $\theta$ to its post-fit value $\hat{\theta}$ modified upwards or downwards by its post-fit uncertainty, and repeating the fit. The filled circles, referring to the bottom x-axis, show the pulls of the fitted nuisance parameters, i.e. the deviations of the fitted parameters $\hat{\theta}$ from their nominal values $\theta_0$, normalised to their nominal uncertainties $\Delta \theta$. The black lines show the post-fit uncertainties of the nuisance parameters, relative to their nominal uncertainties, which are indicated by the yellow band. The jet energy scale uncertainties are decomposed into uncorrelated components.
Figure 9. Distributions of the BDT discriminants for the data taken at $\sqrt{s} = 8$ TeV in the signal regions of the VBF (left) and boosted (right) categories for the $\tau_{lep} \tau_{lep}$ (top), $\tau_{lep} \tau_{had}$ (middle), and $\tau_{had} \tau_{had}$ (bottom) channels. The Higgs boson signal ($m_H = 125$ GeV) is shown stacked with a signal strength of $\mu = 1$ (dashed line) and $\mu = 1.4$ (solid line). The background predictions are determined in the global fit (that gives $\mu = 1.4$). The size of the statistical and systematic normalisation uncertainties is indicated by the hashed band. The ratios of the data to the model (background plus Higgs boson contributions with $\mu = 1.4$) are shown in the lower panels. The dashed red and the solid black lines represent the changes in the model when $\mu = 1.0$ or $\mu = 0$ are assumed respectively.
<table>
<thead>
<tr>
<th>Process/Category</th>
<th>VBF</th>
<th>Boosted</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDT output bin</td>
<td>All bins</td>
<td>Second to last bin</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>589 ± 24</td>
<td>9.7 ± 1.0</td>
</tr>
<tr>
<td>Fake background</td>
<td>57 ± 12</td>
<td>1.2 ± 0.6</td>
</tr>
<tr>
<td>Top</td>
<td>131 ± 19</td>
<td>0.9 ± 0.4</td>
</tr>
<tr>
<td>Others</td>
<td>196 ± 17</td>
<td>3.0 ± 0.4</td>
</tr>
<tr>
<td>ggF: $H \rightarrow WW$ ($m_H = 125\ GeV$)</td>
<td>2.9 ± 0.8</td>
<td>0.12 ± 0.04</td>
</tr>
<tr>
<td>VBF: $H \rightarrow WW$</td>
<td>3.4 ± 0.4</td>
<td>0.40 ± 0.06</td>
</tr>
<tr>
<td>WH: $H \rightarrow WW$</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>ZH: $H \rightarrow WW$</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>ggF: $H \rightarrow \tau\tau$ ($m_H = 125\ GeV$)</td>
<td>9.8 ± 3.4</td>
<td>0.73 ± 0.26</td>
</tr>
<tr>
<td>VBF: $H \rightarrow \tau\tau$</td>
<td>13.3 ± 4.0</td>
<td>2.7 ± 0.7</td>
</tr>
<tr>
<td>WH: $H \rightarrow \tau\tau$</td>
<td>0.25 ± 0.07</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>ZH: $H \rightarrow \tau\tau$</td>
<td>0.14 ± 0.04</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>Total background</td>
<td>980 ± 22</td>
<td>15.4 ± 1.8</td>
</tr>
<tr>
<td>Total signal</td>
<td>24 ± 6</td>
<td>3.5 ± 0.9</td>
</tr>
<tr>
<td>Data</td>
<td>1014</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 9. The predicted post-fit event yields in the $\tau_\text{lep}\tau_\text{lep}$ channel for $m_H = 125\ GeV$ for the total number of events and for the two highest bins of the BDT distributions for the data taken at $\sqrt{s} = 8\ TeV$. The background normalisations, signal normalisation, and their uncertainties represent the post-fit values (see text). The uncertainties on the total background and total signal represent the full statistical and systematic uncertainty, while the uncertainties on the individual background components represent the full systematic uncertainty only.
<table>
<thead>
<tr>
<th>Process/Category</th>
<th>VBF</th>
<th>Boosted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All bins</td>
<td>Second to last bin</td>
</tr>
<tr>
<td>Fake background</td>
<td>1680 ± 50</td>
<td>8.2 ± 0.9</td>
</tr>
<tr>
<td>$Z \rightarrow \tau \tau$</td>
<td>877 ± 29</td>
<td>7.6 ± 0.9</td>
</tr>
<tr>
<td>Top</td>
<td>82 ± 15</td>
<td>0.3 ± 0.4</td>
</tr>
<tr>
<td>$Z \rightarrow \ell\ell(\ell \rightarrow \tau_{\text{had}})$</td>
<td>54 ± 26</td>
<td>1.0 ± 0.7</td>
</tr>
<tr>
<td>Diboson</td>
<td>63 ± 11</td>
<td>1.0 ± 0.4</td>
</tr>
<tr>
<td>$\text{ggF: } H \rightarrow \tau \tau$ ($m_H = 125$ GeV)</td>
<td>16 ± 6</td>
<td>1.0 ± 0.4</td>
</tr>
<tr>
<td>$\text{VBF: } H \rightarrow \tau \tau$</td>
<td>31 ± 8</td>
<td>4.5 ± 1.1</td>
</tr>
<tr>
<td>$WH : H \rightarrow \tau \tau$</td>
<td>0.6 ± 0.4</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>$ZH : H \rightarrow \tau \tau$</td>
<td>0.16 ± 0.07</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>Total background</td>
<td>2760 ± 40</td>
<td>18.1 ± 2.3</td>
</tr>
<tr>
<td>Total signal</td>
<td>48 ± 12</td>
<td>5.5 ± 1.3</td>
</tr>
<tr>
<td>Data</td>
<td>2830</td>
<td>22</td>
</tr>
</tbody>
</table>

**Table 10.** The predicted post-fit event yields in the $\tau_{\text{lep}}\bar{\tau}_{\text{had}}$ channel for $m_H = 125$ GeV for the total number of events and for the two highest bins of the BDT distributions for the data taken at $\sqrt{s} = 8$ TeV. The background normalisations, signal normalisation, and their uncertainties represent the post-fit values (see text). The uncertainties on the total background and total signal represent the full statistical and systematic uncertainty, while the uncertainties on the individual background components represent the full systematic uncertainty only.
<table>
<thead>
<tr>
<th>Process/Category</th>
<th>VBF</th>
<th>Boosted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All bins</td>
<td>Second to last bin</td>
</tr>
<tr>
<td>Fake background</td>
<td>370 ± 18</td>
<td>2.3 ± 0.9</td>
</tr>
<tr>
<td>Others</td>
<td>37 ± 5</td>
<td>0.67 ± 0.22</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau$</td>
<td>475 ± 16</td>
<td>0.6 ± 0.7</td>
</tr>
<tr>
<td>ggF: $H \rightarrow \tau\tau$ ($m_H = 125$ GeV)</td>
<td>8.0 ± 2.7</td>
<td>0.67 ± 0.23</td>
</tr>
<tr>
<td>VBF: $H \rightarrow \tau\tau$</td>
<td>12.0 ± 3.1</td>
<td>1.8 ± 0.5</td>
</tr>
<tr>
<td>WH : $H \rightarrow \tau\tau$</td>
<td>0.25 ± 0.07</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>ZH : $H \rightarrow \tau\tau$</td>
<td>0.16 ± 0.04</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>Total background</td>
<td>883 ± 18</td>
<td>3.6 ± 1.3</td>
</tr>
<tr>
<td>Total signal</td>
<td>20 ± 5</td>
<td>2.5 ± 0.6</td>
</tr>
<tr>
<td>Data</td>
<td>892</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 11. The predicted post-fit event yields in the $\tau_{\text{had}}\tau_{\text{had}}$ channel for $m_H = 125$ GeV for the total number of events and for the two highest bins of the BDT distributions for the data taken at $\sqrt{s} = 8$ TeV. The background normalisations, signal normalisation, and their uncertainties represent the post-fit values (see text). The uncertainties on the total background and total signal represent the full statistical and systematic uncertainty, while the uncertainties on the individual background components represent the full systematic uncertainty only.
9 Results

As explained in the previous section, the observed signal strength is determined from a global maximum likelihood fit to the BDT output distributions in data, with nuisance parameters that are either free or constrained. The results are extracted for each channel and for each category individually as well as for combinations of categories and for the overall combination.

At the value of the Higgs boson mass obtained from the combination of the ATLAS $H \to \gamma\gamma$ and $H \to ZZ^*$ measurements [116], $m_H = 125.36$ GeV, the signal strength obtained from the combined $H \to \tau\tau$ analysis is:

$$\mu = 1.43^{+0.27}_{-0.26}(\text{stat.})^{+0.32}_{-0.25}(\text{syst.}) \pm 0.09(\text{theory syst.}).$$

The systematic uncertainties are split into two groups: systematic uncertainties (syst.) including all experimental effects as well as theoretical uncertainties on the signal region acceptance, such as those due to the QCD scales, the PDF choice, and the underlying event and parton shower; and, separately, theoretical uncertainties on the inclusive Higgs boson production cross section and $H \to \tau\tau$ branching ratio (theory syst.). The results for each individual channel and for each category as well as for their combination are shown in figure 10. They are based on the full dataset, however, separate combined results are given for the two centre-of-mass energies.

The probability $p_0$ of obtaining a result at least as signal-like as observed in the data if no signal were present is calculated using the test statistic $q_{\mu=0} = -2 \ln(\mathcal{L}(0, \hat{\theta})/\mathcal{L}(\hat{\mu}, \hat{\theta}))$ in the asymptotic approximation [117]. For $m_H = 125.36$ GeV, the observed $p_0$ value is $2.7 \times 10^{-6}$, which corresponds to a deviation from the background-only hypothesis of $4.5 \sigma$. This can be compared to an expected significance of $3.4 \sigma$. This provides evidence at the level of $4.5 \sigma$ for the decay of the Higgs boson into tau leptons. Table 12 shows the expected and observed significances for the signal strength measured in each channel separately.

Figure 11 shows the expected and observed number of events, in bins of $\log_{10}(S/B)$, for all signal region bins. Here, $S/B$ is the signal-to-background ratio calculated assuming $\mu = 1.4$ for each BDT bin in the signal regions. The expected signal yield for both $\mu = 1$ and the best-fit value $\mu = 1.4$ for $m_H = 125$ GeV is shown on top of the background prediction from the best-fit values. The background expectation where the signal-strength parameter is fixed to $\mu = 0$ is also shown for comparison.

To visualise the compatibility of this excess of events above background predictions with the SM Higgs boson at $m_H = 125$ GeV, a weighted distribution of events as a function of $m_{\tau\tau}^{\text{MMC}}$ is shown in figure 12. The events are weighted by a factor of $\ln(1 + S/B)$, which enhances the events compatible with the signal hypothesis. The excess of events in this mass distribution is consistent with the expectation for a Standard Model Higgs boson with $m_H = 125$ GeV. The distributions for the predicted excess in data over the background are also shown for alternative SM Higgs boson mass hypotheses of $m_H = 110$ GeV and $m_H = 150$ GeV. The data favour a Higgs boson mass of $m_H = 125$ GeV and are less consistent with the other masses considered.
Table 12. The expected and observed significances of the signal in each channel and category for the combined 7 and 8 TeV datasets.

<table>
<thead>
<tr>
<th>Channel and Category</th>
<th>Expected Significance (σ)</th>
<th>Observed Significance (σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>τ\text{lep}τ\text{lep} VBF</td>
<td>1.15</td>
<td>1.88</td>
</tr>
<tr>
<td>τ\text{lep}τ\text{lep} Boosted</td>
<td>0.57</td>
<td>1.72</td>
</tr>
<tr>
<td>τ\text{lep}τ\text{lep} Total</td>
<td>1.25</td>
<td>2.40</td>
</tr>
<tr>
<td>τ\text{lep}τ\text{had} VBF</td>
<td>2.11</td>
<td>2.23</td>
</tr>
<tr>
<td>τ\text{lep}τ\text{had} Boosted</td>
<td>1.11</td>
<td>1.01</td>
</tr>
<tr>
<td>τ\text{lep}τ\text{had} Total</td>
<td>2.33</td>
<td>2.33</td>
</tr>
<tr>
<td>τ\text{had}τ\text{had} VBF</td>
<td>1.70</td>
<td>2.23</td>
</tr>
<tr>
<td>τ\text{had}τ\text{had} Boosted</td>
<td>0.82</td>
<td>2.56</td>
</tr>
<tr>
<td>τ\text{had}τ\text{had} Total</td>
<td>1.99</td>
<td>3.25</td>
</tr>
<tr>
<td>Combined</td>
<td>3.43</td>
<td>4.54</td>
</tr>
</tbody>
</table>

Table 13. Important sources of uncertainty on the measured signal-strength parameter \( \mu \). The contributions are given as absolute uncertainties on the best-fit value of \( \mu = 1.43 \). Various sub-components are combined assuming no correlations.

As discussed in section 8, the dominant uncertainties on the measurement of the signal-strength parameters include statistical uncertainties on the data from the signal regions, uncertainties on the jet and tau energy scales, uncertainties on the normalisation of the \( Z \to \tau \tau \) and \( t\bar{t} \) background components as well as theoretical uncertainties. The contributions of each of these significant sources to the uncertainty of the measured signal strength are summarised in table 13.
Figure 10. The best-fit value for the signal strength $\mu$ in the individual channels and their combination for the full ATLAS datasets at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The total $\pm 1\sigma$ uncertainty is indicated by the shaded green band, with the individual contributions from the statistical uncertainty (top, black), the experimental systematic uncertainty (middle, blue), and the theory uncertainty (bottom, red) on the signal cross section (from QCD scale, PDF, and branching ratios) shown by the error bars and printed in the central column.

The normalisation uncertainties on the $Z \to \tau\tau$ embedded sample are correlated across the categories in each respective channel. The global fit also constrains the normalisation for $Z \to \tau\tau$ more strongly than for the $Z \to \ell\ell$ and top-quark background components, as the low BDT-score region is dominated by $Z \to \tau\tau$ events.

The measurement of the overall signal strength discussed above does not give direct information on the relative contributions of the different production mechanisms. Therefore, the signal strengths of different production processes contributing to the $H \to \tau\tau$ decay mode are determined, exploiting the sensitivity offered by the use of the event categories in the analyses of the three channels. The data are fitted separating the vector-boson-mediated VBF and $VH$ processes from gluon-mediated ggF processes. Two signal strength parameters, $\mu_{\text{ggF}}$ and $\mu_{\text{VBF+VH}}$, which scale the SM-predicted rates to those ob-
Figure 11. Event yields as a function of $\log_{10}(S/B)$, where $S$ (signal yield) and $B$ (background yield) are taken from the BDT output bin of each event, assuming a signal strength $\mu = 1.4$. Events in all categories are included. The predicted background is obtained from the global fit (with $\mu = 1.4$), and signal yields are shown for $m_H = 125$ GeV at $\mu = 1$ and $\mu = 1.4$ (the best-fit value). The background-only distribution (dashed line) is obtained from the global fit, with $\mu$ fixed at zero.

served, are introduced. The two-dimensional 68% and 95% confidence level (CL) contours in the plane of $\mu_{ggF}^{\tau\tau}$ and $\mu_{VBF+VH}^{\tau\tau}$ [84] are shown in figure 13 for $m_H = 125.36$ GeV. The best-fit values are

$$\mu_{ggF}^{\tau\tau} = 2.0 \pm 0.8\text{(stat.)} \pm^{+1.2}_{-0.3}\text{(syst.)} \pm 0.3\text{(theory syst.)}$$

and

$$\mu_{VBF+VH}^{\tau\tau} = 1.24 \pm^{+0.49}_{-0.45}\text{(stat.)} \pm^{+0.31}_{-0.29}\text{(syst.)} \pm 0.08\text{(theory syst.),}$$

in agreement with the predictions from the Standard Model. The two results are strongly anti-correlated (correlation coefficient of $-48\%$). The observed (expected) significances of the $\mu_{ggF}^{\tau\tau}$ and $\mu_{VBF+VH}^{\tau\tau}$ signal strengths are $1.74\sigma$ ($0.95\sigma$) and $2.25\sigma$ ($1.72\sigma$) respectively.

A total cross section times branching ratio for $H \rightarrow \tau\tau$ with $m_H = 125$ GeV can also be measured. The central value is obtained from the product of the measured $\mu$ and the predicted cross section used to define it. The uncertainties are similarly obtained by scaling the uncertainties on $\mu$ by the predicted cross section, noting that theoretical uncertainties on the inclusive cross section cancel between $\mu$ and the predicted cross section and thus are not included for the production processes under consideration. These include the
Figure 12. Distributions of the reconstructed invariant $\tau\tau$ mass, $m_{\text{MMC}}$, where events are weighted by $\ln(1 + S/B)$ for all channels. These weights are determined by the signal ($S$) and background ($B$) predictions for each BDT bin. The bottom panel in each plot shows the difference between weighted data events and weighted background events (black points), compared to the weighted signal yields. The background predictions are obtained from the global fit with the $m_H = 125$ GeV signal hypothesis (signal strength $\mu = 1.4$). The $m_H = 125$ GeV signal is plotted with a solid red line, and, for comparison, signals for $m_H = 110$ GeV (blue) and $m_H = 150$ GeV (green) are also shown. The signal normalisations are taken from fits to data with the corresponding signal mass hypotheses, and the fitted $\mu$ values are given in the figure. The signal strengths are shown for the Standard Model expectations ($\mu = 1$) in (a), while in (b) the best-fit values are used.

uncertainties on the inclusive cross section due to the QCD scale and the PDF choice as well as the uncertainty on the branching ratio $H \rightarrow \tau\tau$; however, theoretical uncertainties on the acceptance of the signal regions from the QCD scale and PDF choice are retained, along with the uncertainties due to underlying event and parton shower, and the electroweak correction on VBF production. Table 14 gives the measured values for the total cross section at 7 and at 8 TeV, as well as the measured values at 8 TeV for gluon fusion production and for VBF and $VH$ production separately.

10 Cross-check with cut-based analysis

The search for the SM Higgs boson presented above is cross-checked for the dataset collected at $\sqrt{s} = 8$ TeV in an analysis where cuts on kinematic variables are applied. This search uses improved definitions of event categories and an improved fit model with respect to results previously published for the $\sqrt{s} = 7$ TeV dataset [21]. To allow a straightforward comparison of results, the multivariate and cut-based analyses have common components. The two analyses are performed for the same three channels, $\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$, they use the same preselection and share the same strategy for the estimation of background contributions and systematic uncertainties. As in the multivariate analysis, the irreducible
Figure 13. Likelihood contours for the combination of all channels in the ($\mu_{ggF}^{\tau\tau}$, $\mu_{VBF+VH}^{\tau\tau}$) plane. The signal strength $\mu$ is the ratio of the measured signal yield to the Standard Model expectation, for each production mode. The 68% and 95% CL contours are shown as dashed and solid lines respectively, for $m_H = 125.36$ GeV. The SM expectation is shown by a filled plus symbol, and the best fit to the data is shown as a star.

<table>
<thead>
<tr>
<th>Production Mode</th>
<th>Measured $\sigma \times BR$ [pb]</th>
<th>Predicted $\sigma \times BR$ [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 TeV</td>
<td>$1.0^{+0.9}<em>{-0.8}$ (stat.) $^{+0.9}</em>{-0.8}$ (syst.)</td>
<td>$1.09 \pm 0.11$</td>
</tr>
<tr>
<td>8 TeV</td>
<td>$2.1 \pm 0.4$ (stat.) $^{+0.5}_{-0.4}$ (syst.)</td>
<td>$1.39 \pm 0.14$</td>
</tr>
<tr>
<td>Gluon fusion, 8 TeV</td>
<td>$1.7 \pm 1.1$ (stat.) $^{+1.5}_{-1.1}$ (syst.)</td>
<td>$1.22 \pm 0.14$</td>
</tr>
<tr>
<td>VBF+VH, 8 TeV</td>
<td>$0.26 \pm 0.09$ (stat.) $^{+0.06}_{-0.05}$ (syst.)</td>
<td>$0.17 \pm 0.01$</td>
</tr>
</tbody>
</table>

Table 14. Measured and predicted total cross section times branching ratio for $H \rightarrow \tau\tau$ with $m_H = 125$ GeV, at 7 TeV and at 8 TeV for all production modes, as well as for gluon fusion and for VBF and $VH$ production separately, at 8 TeV only. The theoretical predictions are obtained as described in section 3.
\(Z \to \tau\tau\) background is estimated using the embedding procedure and the reducible ones are estimated using similar data-driven methods, as described in section 6. Finally, the same statistical methods are used to extract the results, although these are applied to different discriminating variables. While the multivariate analysis performs a fit to the BDT output distribution, the cut-based analysis relies on a fit to the \(\tau\tau\) invariant mass distribution. The \(\tau\tau\) invariant mass is calculated using the missing mass calculator, as described in section 4.3. The analysis is not designed to be sensitive to a specific value of the Higgs boson mass \(m_H\). The use of the mass as the discriminating variable is motivated not only by its power to separate the irreducible \(Z \to \tau\tau\) background from signal, but also by its sensitivity to the mass of the signal itself.

In the cut-based analysis, a categorisation is performed similar to that in the multivariate analysis, i.e. VBF and boosted categories are defined. To increase the separation power, subcategories are introduced for the \(\tau_{\text{lep}}\tau_{\text{had}}\) and \(\tau_{\text{had}}\tau_{\text{had}}\) channels. These subcategories target events produced via the same production mode, but select different phase-space regions with different signal-to-background ratios. With this strategy the most sensitive subcategories have a small number of events, but a high signal-to-background ratio. Although the combined sensitivity is dominated by the few highly sensitive subcategories, the others are important not just to increase the sensitivity but also to constrain the various background components.

An overview of the defined categories in the three channels is given in table 15. In all channels, the event categorisation is designed by splitting events first according to the production mode, either VBF-like or boosted ggF-like, and second, for the \(\tau_{\text{lep}}\tau_{\text{had}}\) and \(\tau_{\text{had}}\tau_{\text{had}}\) channels, by signal-to-background ratio. The events accepted in the VBF categories pass a common selection that requires the presence of the two forward jets distinctive of VBF production. In the \(\tau_{\text{lep}}\tau_{\text{had}}\) channel, tight and loose VBF subcategories are defined, via cuts on the mass of the dijet system, \(m_{jj}\), and \(p_T^H\), the transverse momentum of the Higgs boson candidate. In the \(\tau_{\text{had}}\tau_{\text{had}}\) channel, the variables used to select the most sensitive categories for both production modes are \(p_T^H\) and the separation \(\Delta R(\tau_1, \tau_2)\) between the two \(\tau_{\text{had}}\) candidates. In the VBF-like events, correlations between the invariant mass of the selected jets \(m_{jj}\) and \(\Delta \eta_{jj}\) of the jets characteristic of VBF production are also used. The subcategory with the highest purity is the VBF high-\(p_T^H\) subcategory, where tight cuts on \(p_T^H\) and \(\Delta R(\tau_1, \tau_2)\) reject almost all non-resonant background sources. The other two VBF-like subcategories are distinguished by a different signal-to-background ratio due to a tighter selection applied to the forward jets. For the \(\tau_{\text{had}}\tau_{\text{had}}\) channel, boosted subcategories are also defined. The division is based on the same cuts on \(p_T^H\) and \(\Delta R(\tau_1, \tau_2)\) as used in the VBF high-\(p_T^H\) category. Events with low transverse momentum are not used in any category because in such events the signal cannot be effectively distinguished from background channels. The proportion of the signal yield produced via VBF in the VBF-like subcategories is found to be \(80\%\) in the \(\tau_{\text{lep}}\tau_{\text{lep}}\) channel, between \(67\%\) and \(85\%\) in the \(\tau_{\text{lep}}\tau_{\text{had}}\) channel and between \(58\%\) and \(78\%\) in the \(\tau_{\text{had}}\tau_{\text{had}}\) channel.

The final results are derived from the combined fit of the \(m_{\tau\tau}\) distributions observed in the various subcategories. The number of fitted signal and background events in each channel and category is given in table 16. The combined mass distribution for the three
# VBF category selection criteria

### VBF tight
- $m_{j_{1,j_{2}}} > 500 \text{ GeV}$
- $p_T^{H} > 100 \text{ GeV}$
- $p_T^{had} > 30 \text{ GeV}$
- $m_{vis} > 40 \text{ GeV}$

### VBF loose
- Non tight VBF

### VBF high $p_T^H$
- $\Delta R(\tau_1, \tau_2) < 1.5$
- $p_T^{H} > 140 \text{ GeV}$

### VBF low $p_T^H$, tight
- $\Delta R(\tau_1, \tau_2) > 1.5$ or $p_T^{H} < 140 \text{ GeV}$
- $m_{j_{1,j_{2}}} [\text{GeV}] > (-250 \cdot |\Delta \eta(j_{1},j_{2})| + 1550)$

### VBF low $p_T^H$, loose
- $\Delta R(\tau_1, \tau_2) > 1.5$ or $p_T^{H} < 140 \text{ GeV}$
- $m_{j_{1,j_{2}}} [\text{GeV}] < (-250 \cdot |\Delta \eta(j_{1},j_{2})| + 1550)$

### Boosted category selection criteria

#### Boosted high $p_T^H$
- $\Delta R(\tau_1, \tau_2) < 1.5$
- $p_T^{H} > 100 \text{ GeV}$

#### Boosted low $p_T^H$
- $\Delta R(\tau_1, \tau_2) > 1.5$ or $p_T^{H} < 140 \text{ GeV}$

### Channel

<table>
<thead>
<tr>
<th>Channel</th>
<th>VBF category selection criteria</th>
</tr>
</thead>
</table>
| $\tau_{lep}$ | At least two jets with $p_T^{H} > 40 \text{ GeV}$ and $p_T^{H} > 30 \text{ GeV}$  
| $\tau_{lep}$ | $|\Delta \eta(j_{1,j_{2}})| > 3.0$
| $\tau_{lep}$ | $m_{j_{1,j_{2}}} > 400 \text{ GeV}$
| $\tau_{lep}$ | b-jet veto for jets with $p_T > 25 \text{ GeV}$
| $\tau_{lep}$ | Jet veto: no additional jet with $p_T > 25 \text{ GeV}$ within $|\eta| < 2.4$

### $\tau_{lep, had}$
- At least two jets with $p_T^{H} > 40 \text{ GeV}$ and $p_T^{H} > 30 \text{ GeV}$
- $E^{\text{miss}}_T > 20 \text{ GeV}$
- $|\Delta \eta(j_{1,j_{2}})| > 3.0$ and $\eta(j_{1}) \cdot \eta(j_{2}) < 0$, $m_{j_{1,j_{2}}} > 300 \text{ GeV}$
- $p_T^{\text{Total}} = |p_T^{j_{1}} + p_T^{j_{2}} + p_T^{H} + E^{\text{miss}}_T| < 30 \text{ GeV}$
- b-jet veto for jets with $p_T > 30 \text{ GeV}$
- min$(\eta(j_{1}),\eta(j_{2})) < \eta(\ell_{1}), \eta(\ell_{\text{had}}) < \max(\eta(j_{1}),\eta(j_{2}))$

### $\tau_{had}$
- At least two jets with $p_T^{H} > 50 \text{ GeV}$ and $p_T^{H} > 30 \text{ GeV}$
- $|\Delta \eta(\tau_{1},\tau_{2})| < 1.5$
- $|\Delta \eta(j_{1,j_{2}})| > 2.6$ and $m_{j_{1,j_{2}}} > 250 \text{ GeV}$
- min$(\eta(j_{1}),\eta(j_{2})) < \eta(\tau_{1}), \eta(\tau_{2}) < \max(\eta(j_{1}),\eta(j_{2}))$

### $\tau_{lep, had}$
- Exclude events passing the VBF selection
- b-jet veto for jets with $p_T > 25 \text{ GeV}$

### $\tau_{lep, had}$
- Exclude events passing the VBF selection
- $E^{\text{miss}}_T > 20 \text{ GeV}$
- $p_T^{H} > 100 \text{ GeV}$
- $p_T(\tau_{had}) > 30 \text{ GeV}$
- b-jet veto for jets with $p_T > 30 \text{ GeV}$

### $\tau_{had}$
- Exclude events passing the VBF selection
- $|\Delta \eta(\tau_{1},\tau_{2})| < 1.5$
- $p_T^{H} > 100 \text{ GeV}$

### Table 15.
Summary of the selection criteria used to define the VBF and boosted subcategories in the cut-based analysis for the three analysis channels. The labels (1) and (2) refer to the leading (highest $p_T$) and subleading final-state objects (leptons, $\tau_{had}$, jets). The variables are defined in the text.
channels is shown in figure 14, where events are weighted by ln(1 + S/B), based on the signal and background content of their channel and category. An excess of events above the expected SM background is visible in the mass region around 125 GeV.

The signal strengths extracted in the three analysis channels and their combination are given in table 17. This cut-based analysis also provides evidence for $H \rightarrow \tau\tau$ decays, giving a combined signal strength of

$$\mu_{\text{Cut-based}} = 1.43 \pm 0.36^{+0.36}_{-0.35} \text{(stat.)} \pm 0.004^{+0.04}_{-0.03} \text{(syst.)} \pm 0.100 \text{(theory syst.)}$$

for $m_H = 125.36$ GeV. For comparison, the results obtained in the multivariate analysis for the dataset at $\sqrt{s} = 8$ TeV are also included in table 17. Good agreement between the results of the two analyses is found for the individual channels as well as for their combination. To further quantify the level of agreement, the correlation $\rho$ and the uncertainties on the differences between the $\mu$ values obtained, i.e. $\Delta \mu \pm \delta(\Delta \mu)$, were evaluated using the so-called jackknife technique [118, 119]. Using this method, the correlation between the $\mu$ values obtained in the two analyses is found to be between 0.55 and 0.75 for each of the three analysis channels. The results of the analyses are found to be fully compatible, with deviations $\Delta \mu/\delta(\Delta \mu)$ below 1 for all analysis channels as well as for the combined result.

The probability $p_0$ of obtaining a result at least as signal-like as observed if no signal were present is shown as a function of the mass in figure 15 for the cut-based analysis for the combined dataset at $\sqrt{s} = 8$ TeV. The observed $p_0$ values show a shallow minimum around 125 GeV, corresponding to a significance of 3.2$\sigma$. The expected significance for the cut-based analysis is superimposed on the figure and reaches a significance of 2.5$\sigma$ at $m_H = 125.36$ GeV. The corresponding significance values for the multivariate analysis for
Figure 14. Distribution of the reconstructed invariant $\tau\tau$ mass, $m_{\tau\tau}^{\text{MMC}}$, for the sum of all channels in the cut-based analysis for the data taken at $\sqrt{s} = 8$ TeV. The events are weighted by a factor of $\ln(1 + S/B)$ based on the signal ($S$) and background ($B$) yields in each category. The bottom panel shows the difference between weighted data events and weighted background events (black points), compared to the weighted signal yields. The background predictions are obtained from the global fit with the $m_H = 125$ GeV signal hypothesis (signal strength $\mu = 1.4$). The $m_H = 125$ GeV signal is plotted as a solid red line, and, for comparison, signals for $m_H = 110$ GeV (blue) and $m_H = 150$ GeV (green) are also shown. The signal normalisations are taken from fits to data with the corresponding signal mass hypotheses and the fitted $\mu$ values are given in the figure.

<table>
<thead>
<tr>
<th>Fitted $\mu$ values</th>
<th>$\sqrt{s}$</th>
<th>Multivariate analysis</th>
<th>Cut-based analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{\text{lep}}\tau_{\text{lep}}$</td>
<td>8 TeV</td>
<td>$1.9^{+1.0}_{-0.9}$</td>
<td>$3.2^{+1.4}_{-1.3}$</td>
</tr>
<tr>
<td>$\tau_{\text{lep}}\tau_{\text{had}}$</td>
<td>8 TeV</td>
<td>$1.1^{+0.6}_{-0.5}$</td>
<td>$0.7^{+0.7}_{-0.6}$</td>
</tr>
<tr>
<td>$\tau_{\text{had}}\tau_{\text{had}}$</td>
<td>8 TeV</td>
<td>$1.8^{+0.9}_{-0.7}$</td>
<td>$1.6^{+0.9}_{-0.7}$</td>
</tr>
<tr>
<td>All channels</td>
<td>8 TeV</td>
<td>$1.53^{+0.47}_{-0.41}$</td>
<td>$1.43^{+0.55}_{-0.49}$</td>
</tr>
</tbody>
</table>

Table 17. Fitted values of the signal strength for the different channels at $\sqrt{s} = 8$ TeV for the multivariate and cut-based analyses, measured at $m_H=125.36$ GeV. The results for the combinations of all channels are also given. The total uncertainties (statistical and systematic) are quoted.
Figure 15. Observed (solid red) and expected (dashed red) $p_0$ values as a function of $m_H$ for the combination of all channels in the cut-based analysis (CBA) for the data taken at $\sqrt{s} = 8$ TeV. The expected $p_0$ values are given for the background-only hypothesis. The corresponding observed and expected $p_0$ values for the multivariate analysis (MVA) are indicated for $m_H = 125$ GeV by a full and open star respectively. The axis labels on the right hand side and the dotted lines display the significance in units of Gaussian standard deviations.

The dataset at $\sqrt{s} = 8$ TeV are found to be 4.5$\sigma$ (observed) and 3.3$\sigma$ (expected). They are also indicated in the figure.

Given the mass sensitivity of the cut-based analysis, a two-dimensional likelihood fit for the signal strength $\mu$ and the mass $m_H$ is performed. The mass points are tested in steps of 5 GeV in the range between 100 GeV and 150 GeV. The best fit value is found at $\mu = 1.4$ and $m_H = 125$ GeV. The result is shown in the $(m_H, \mu)$ plane in figure 16 together with the 68% and 95% CL contours. This result indicates that the observation is compatible with the decay of a Standard Model Higgs boson with a mass of 125 GeV.

11 Conclusions

Evidence for decays of the recently discovered Higgs boson into pairs of tau leptons is presented. The analysis is based on the full set of proton-proton collision data recorded by the ATLAS experiment at the LHC during Run 1. The data correspond to integrated luminosities of 4.5 fb$^{-1}$ and 20.3 fb$^{-1}$ at centre-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV respectively. All combinations of leptonic and hadronic tau decay channels are included and event categories selecting both the vector boson fusion and highly boosted $\tau\tau$ signatures are considered in a multivariate analysis. An excess of events over the expected background from other Standard Model processes is found with an observed (expected) significance of 4.5 (3.4) standard deviations. This excess is consistent with resulting from
Figure 16. The results of the two-dimensional likelihood fit in the \((m_H, \mu)\) plane for the cut-based analysis for the data taken at \(\sqrt{s} = 8\) TeV. The signal strength \(\mu\) is the ratio of the measured signal yield to the Standard Model expectation. The 68\% and 95\% CL contours are shown as dashed and solid red lines respectively. The best-fit value is indicated as a red cross. The dashed and solid blue lines correspond to the expected 68\% and 95\% CL contours for \(m_H = 125.36\) GeV and \(\mu = 1.43\).

\(H \rightarrow \tau\tau\) decays with \(m_H = 125.36\) GeV. The measured signal strength, normalised to the Standard Model expectation, is

\[
\mu = 1.43 \pm 0.27\text{(stat.)} ^{+0.32}_{-0.25}\text{(syst.)} \pm 0.09\text{(theory syst.)}.
\]

This value is consistent with the predicted Yukawa coupling strength of the Higgs boson in the Standard Model.

The results of the multivariate analysis are cross-checked for the data collected at \(\sqrt{s} = 8\) TeV using a cut-based analysis. The results confirm the findings of the multivariate analysis, and an excess with a compatible signal strength is found. A two-dimensional fit of the signal strength \(\mu\) and \(m_H\) in the cut-based analysis indicates that the observed excess is compatible with the \(\tau\tau\) decay of the Higgs boson with a mass of 125 GeV.

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; STFC, 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; EPLANET, ERC and NSRF, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, DFG, HGF, MPG and AvH Foundation, Germany; GSRT and NSRF, Greece; RGC, Hong Kong SAR, China; ISF, MINERVA, GIF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; BRF and RCN, Norway; MNiSW and NCN, Poland; GRICES and FCT, Portugal; MNE/IFA, Romania; MES of Russia and ROSATOM, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America.

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 (USA) and in the Tier-2 facilities worldwide.

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

References


[21] ATLAS collaboration, Search for the standard model Higgs boson in the \( H \to \tau^+\tau^- \) decay mode in \( \sqrt{s} = 7 \) TeV pp collisions with ATLAS, JHEP 09 (2012) 070 [arXiv:1206.5971] [inSPIRE].


[27] ATLAS collaboration, The ATLAS experiment at the CERN Large Hadron Collider, 2008 JINST 3 S08003 [inSPIRE].


[65] D. de Florian, G. Ferrera, M. Grazzini and D. Tommasini, Higgs boson production at the LHC: transverse momentum resummation effects in the $H \rightarrow 2\gamma$, $H \rightarrow WW \rightarrow l\nu l\nu$ and $H \rightarrow ZZ \rightarrow 4l$ decay modes, JHEP 06 (2012) 132 [arXiv:1203.6321] [inSPIRE].


G. Corcella et al., \textit{HERWIG 6.5 release note}, \textcolor{blue}{hep-ph/0210213} [nSPIRE].


ATLAS collaboration, Determination of the $\tau$ energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC in 2012, ATLAS-CONF-2013-044 (2013).


ATLAS collaboration, Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb$^{-1}$ of pp collision data, Phys. Rev. D 90 (2014) 052004 [arXiv:1406.3827] [INSPIRE].

Fisica, Universidade do Minho, Braga; (f) Departamento de Fisica Teorica y del Cosmos and CAFEPE, Universidad de Granada, Granada (Spain); (g) Dep Fisica and CEFITEC of Faculdade de Ciencias e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

127 Institute of Physics, Academy of Sciences of the Czech Republic, Praha, Czech Republic
128 Czech Technical University in Prague, Praha, Czech Republic
129 Faculty of Mathematics and Physics, Charles University in Prague, Praha, Czech Republic
130 State Research Center Institute for High Energy Physics, Protvino, Russia
131 Particle Physics Department, Rutherford Appleton Laboratory, Didcot, United Kingdom
132 Ritsumeikan University, Kusatsu, Shiga, Japan
133 (a) INFN Sezione di Roma; (b) Dipartimento di Fisica, Sapienza Universit` a di Roma, Roma, Italy
134 (a) INFN Sezione di Roma Tor Vergata; (b) Dipartimento di Fisica, Universit` a di Roma Tor Vergata, Roma, Italy
135 (a) INFN Sezione di Roma Tre; (b) Dipartimento di Matematica e Fisica, Universit` a Roma Tre, Roma, Italy
136 (a) Facult´ e des Sciences Ain Chock, R´ eseau Universitaire de Physique des Hautes Energies - Universit´ e Hassan II, Casablanca; (b) Centre National de l’Energie des Sciences Techniques Nucleaires, Rabat; (c) Facult´ e des Sciences Semlalia, Universit´ e Cadi Ayyad, LPHEA-Marrakech; (d) Facult´ e des Sciences, Universit´ e Mohamed Premier and LPTPM, Oujda; (e) Facult´ e des sciences, Universit´ e Mohammed V-Agdal, Rabat, Morocco
137 DSM/IRFU (Institut de Recherches sur les Lois Fondamentales de l’Univers), CEA Saclay (Commissariat à l’Energie Atomique et aux Energies Alternatives), Gif-sur-Yvette, France
138 Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz CA, United States of America
139 Department of Physics, University of Washington, Seattle WA, United States of America
140 Department of Physics and Astronomy, University of Sheffield, Sheffield, United Kingdom
141 Department of Physics, Shinshu University, Nagano, Japan
142 Fachbereich Physik, Universit¨ at Siegen, Siegen, Germany
143 Department of Physics, Simon Fraser University, Burnaby BC, Canada
144 SLAC National Accelerator Laboratory, Stanford CA, United States of America
145 (a) Faculty of Mathematics, Physics & Informatics, Comenius University, Bratislava; (b) Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
146 (a) Department of Physics, University of Cape Town, Cape Town; (b) Department of Physics, University of Johannesburg, Johannesburg; (c) School of Physics, University of the Witwatersrand, Johannesburg, South Africa
147 (a) Department of Physics, Stockholm University; (b) The Oskar Klein Centre, Stockholm, Sweden
148 Physics Department, Royal Institute of Technology, Stockholm, Sweden
149 Departments of Physics & Astronomy and Chemistry, Stony Brook University, Stony Brook NY, United States of America
150 Department of Physics and Astronomy, University of Sussex, Brighton, United Kingdom
151 School of Physics, University of Sydney, Sydney, Australia
152 Institute of Physics, Academia Sinica, Taipei, Taiwan
153 Department of Physics, Technion: Israel Institute of Technology, Haifa, Israel
154 Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
155 Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece
156 International Center for Elementary Particle Physics and Department of Physics, The University of Tokyo, Tokyo, Japan
157 Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo, Japan
158 Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
159 Department of Physics, University of Toronto, Toronto ON, Canada
160 (a) TRIUMF, Vancouver BC; (b) Department of Physics and Astronomy, York University, Toronto ON, Canada
Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba, Japan
Department of Physics and Astronomy, Tufts University, Medford MA, United States of America
Centro de Investigaciones, Universidad Antonio Narino, Bogota, Colombia
Department of Physics and Astronomy, University of California Irvine, Irvine CA, United States of America

(a) INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine; (b) ICTP, Trieste; (c) Dipartimento di Chimica, Fisica e Ambiente, Università di Udine, Udine, Italy
Department of Physics, University of Illinois, Urbana IL, United States of America
Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
Instituto de Física Corpuscular (IFIC) and Departamento de Física Atómica, Molecular y Nuclear and Departamento de Ingeniería Electrónica and Instituto de Microelectrónica de Barcelona (IMB-CNM), University of Valencia and CSIC, Valencia, Spain
Department of Physics, University of British Columbia, Vancouver BC, Canada
Department of Physics and Astronomy, University of Victoria, Victoria BC, Canada
Department of Physics, University of Warwick, Coventry, United Kingdom
Waseda University, Tokyo, Japan
Department of Particle Physics, The Weizmann Institute of Science, Rehovot, Israel
Department of Physics, University of Wisconsin, Madison WI, United States of America
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität, Würzburg, Germany
Fachbereich C Physik, Bergische Universität Wuppertal, Wuppertal, Germany
Department of Physics, Yale University, New Haven CT, United States of America
Yerevan Physics Institute, Yerevan, Armenia
Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France

Also at Department of Physics, King’s College London, London, United Kingdom
Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
Also at Novosibirsk State University, Novosibirsk, Russia
Also at TRIUMF, Vancouver BC, Canada
Also at Department of Physics, California State University, Fresno CA, United States of America
Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
Also at Tomsk State University, Tomsk, Russia
Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
Also at Università di Napoli Parthenope, Napoli, Italy
Also at Institute of Particle Physics (IPP), Canada
Also at Particle Physics Department, Rutherford Appleton Laboratory, Didcot, United Kingdom
Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia
Also at Louisiana Tech University, Ruston LA, United States of America
Also at Institut Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain
Also at Department of Physics, National Tsing Hua University, Taiwan
Also at Department of Physics, The University of Texas at Austin, Austin TX, United States of America
Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia
Also at CERN, Geneva, Switzerland
Also at Ochadai Academic Production, Octoberoniz University, Tokyo, Japan
Also at Manhattan College, New York NY, United States of America
Also at Institute of Physics, Academia Sinica, Taipei, Taiwan
Also at LAL, Université Paris-Sud and CNRS/IN2P3, Orsay, France
Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan
Also at Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université Paris-Diderot and CNRS/IN2P3, Paris, France
Also at School of Physical Sciences, National Institute of Science Education and Research, Bhubaneswar, India

Also at Dipartimento di Fisica, Sapienza Università di Roma, Roma, Italy

Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia

Also at section de Physique, Université de Genève, Geneva, Switzerland

Also at International School for Advanced Studies (SISSA), Trieste, Italy

Also at Department of Physics and Astronomy, University of South Carolina, Columbia SC, United States of America

Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China

Also at Faculty of Physics, M.V.Lomonosov Moscow State University, Moscow, Russia

Also at National Research Nuclear University MEPhI, Moscow, Russia

Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary

Also at Department of Physics, Oxford University, Oxford, United Kingdom

Also at Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany

Also at Department of Physics, The University of Michigan, Ann Arbor MI, United States of America

Also at Discipline of Physics, University of KwaZulu-Natal, Durban, South Africa

Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia

* Deceased