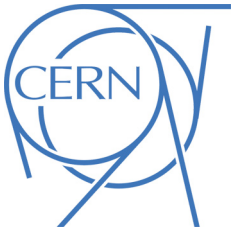
ATLAS NOTE

ATLAS-CONF-2014-061

October 7, 2014



Evidence for Higgs boson Yukawa couplings in the $H \to \tau\tau$ decay mode with the ATLAS detector

The ATLAS Collaboration

Abstract

Results of a search for $H \to \tau\tau$ 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 ($\tau \to \ell\nu\bar{\nu}$ with $\ell = e, \mu$) and hadronic ($\tau \to$ hadrons ν) 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 (3.5) standard deviations. This excess provides evidence for the direct coupling of the recently discovered Higgs boson with mass m_H=125 GeV to fermions. The measured signal strength, normalised to the Standard Model expectation, of $\mu = 1.42\ ^{+0.44}_{-0.38}$ is consistent with the predicted Yukawa coupling strength in the Standard Model.

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, 13] have strengthened the hypothesis of its consistency with 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 for fermions as implemented in the SM, it is of prime importance to demonstrate the direct coupling of the Higgs boson to fermions and its proportionality to mass [14]. The most prominent candidate decay modes are the decays into tau leptons, $H \to \tau\tau$, and bottom quarks (b-quarks), $H \to b\bar{b}$. The search for decays to $b\bar{b}$ requires the restriction to Higgs bosons produced in association with vector bosons or $t\bar{t}$ pairs, and by vector-boson fusion. 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 has published evidence for $H \to \tau\tau$ at a significance of three standard deviations (σ) [15] and an excess of events above the expected background corresponding to a significance of 2.1σ in the search for $H \to b\bar{b}$ decays [16] for a Higgs boson mass, m_H, of 125 GeV. The combination of the channels, based on a dataset corresponding to integrated luminosities of 5 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 7$ TeV and ~20 fb^{-1} at $\sqrt{s} = 8$ TeV, provides evidence for fermionic couplings of the newly discovered Higgs boson with a significance of 3.8σ [17]. Recently, the ATLAS Collaboration has observed an excess of events above the expected background in the search for $H \to b\bar{b}$ decays [18] corresponding to a significance of 1.4σ for m_H=125 GeV, based on the full dataset. In the search for $H \to \tau\tau$ decays, the ATLAS Collaboration has set upper limits on the cross section times the branching ratio, normalised to the SM prediction, between 2.9 and 11.7 in the mass range 100–150 GeV from 4.7 fb^{-1} of data collected at $\sqrt{s} = 7$ TeV [19].

In this note, 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. All combinations of leptonic ($\tau \to \ell\nu\bar{\nu}$ with $\ell = e, \mu$) and hadronic ($\tau \to$ hadrons ν) tau decays are considered.[1] The corresponding three analysis channels are denoted as $\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$, 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) [20], vector-boson fusion (VBF) [21], and the associated production (VH) with $V = W$ or Z. These production processes lead to different final state signatures, which have been exploited by defining an event categorisation. Two dedicated categories are considered to achieve both a good signal-to-background ratio and a good resolution for the reconstruction of the $\tau\tau$ invariant mass. The *VBF category*, enriched in events produced via vector-boson fusion, is defined by the presence of two jets with a large separation in pseudorapidity.[2] The *Boosted category* contains events with a large transverse momentum of the reconstructed Higgs boson candidate. 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,

[1] Throughout this paper the inclusion of charge-conjugate decay modes is implied.

[2] The ATLAS experiment uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam direction. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse (x, y) plane, ϕ being the azimuthal angle around the beam direction. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. The distance ΔR in the $\eta - \phi$ space is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

a multivariate analysis technique, based on boosted decision trees (BDTs) [22–24], 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 [25] 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 embedded 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 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 $|\eta| < 3.2$. An iron/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 ($|\eta| < 1.05$) and thin-gap chambers in the end-cap regions ($1.05 < |\eta| < 2.4$).

A three-level trigger system [26] 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 of at most 75 kHz. 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 is selected by choosing the vertex candidate with the highest sum of the squared transverse momentum of all tracks matched to the candidate.

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeters matched to a track in the ID. They are required to have an energy in the transverse plane $E_{\mathrm{T}} > 15$ GeV, be within the pseudorapidity range $|\eta| < 2.47$ and pass the *medium* shower shape and track selection criteria defined in Ref. [27]. Candidates found in the calorimeter transition region ($1.37 < |\eta| < 1.52$) are not considered. Typical reconstruction and identification efficiencies for electrons passing these selection cuts range between 80% and 90% depending on E_{T} and η.

Muon candidates are reconstructed using an algorithm [28] that combines information from the ID and the MS. They are required to have a momentum in the transverse plane $p_{\mathrm{T}} > 10$ GeV and to be within $|\eta| < 2.5$. Typical efficiencies for muons passing these selection criteria are above 95% [29].

Jets are reconstructed using the anti-k_t jet clustering algorithm [30, 31] with a radius parameter $R = 0.4$, taking topological energy clusters [32] in the calorimeters as inputs. Jet energies are corrected for the contribution of pile-up interactions using a jet-area based technique [33] and are calibrated using p_{T} and η dependent correction factors determined from simulation and data [34–36]. Jets are required to be reconstructed in the range $|\eta| < 4.5$ and to have $p_{\mathrm{T}} > 30$ GeV. To reduce the contamination of jets from multiple interactions in the same or neighbouring bunch crossings (pile-up), for jets with $|\eta| < 2.4$, the scalar sum of the p_{T} of tracks matched to jets and originating from the primary vertex is required to be at least 75% (50%) of the scalar sum of the transverse momenta of all tracks in the jet for the 7 TeV (8 TeV) dataset (jet vertex fraction, JVF). Moreover, for the 8 TeV dataset, the JVF selection is applied only to jets with $p_{\mathrm{T}} < 50$ GeV. Jets with no associated tracks are retained.

In the pseudorapidity range $|\eta| < 2.5$, b-jets are selected using a tagging algorithm [37]. The b-jet tagging algorithm used has an efficiency of 60–70% for b-jets in simulated $t\bar{t}$ events [38]. The

corresponding light-quark jet misidentification probability is 0.1–0.5%, depending on the jet p_{T} and η [39].

Hadronically decaying tau leptons are reconstructed starting from clusters of energy in the electromagnetic and hadronic calorimeters. The τ_{had} [3] reconstruction is seeded by the anti-k_t jet finding algorithm with a radius parameter $R = 0.4$. Tracks in a cone of radius $\Delta R < 0.2$ from the cluster barycentre are associated to the τ_{had} candidate, and the τ_{had} charge is determined from the sum of the charges of the tracks. The rejection against jets is provided in a separate identification step using discriminating variables based on tracks with $p_{\mathrm{T}} > 1$ GeV and calorimeter cells found in the core region ($\Delta R < 0.2$) and in the region $0.2 < \Delta R < 0.4$ around the τ_{had} candidate direction. Such discriminating variables are combined in a boosted decision tree and three working points, labelled *tight*, *medium* and *loose* [40], are defined, corresponding to different τ_{had} identification efficiency values.

In this analysis, τ_{had} candidates with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.47$ are used. The τ_{had} candidates are required to have charge ± 1, and must be 1- or 3-track (prong) candidates. In addition, a two-track sample (where the charge requirement is dropped) is retained for background studies, as described in Section 6.2. The identification efficiency for τ_{had} candidates passing the medium identification criteria is of the order of 55–60%. Dedicated criteria [40] to separate τ_{had} candidates from misidentified electrons are also applied, with a selection efficiency for true τ_{had} decays of 95%. The probability to misidentify a jet with $p_{\mathrm{T}} > 20$ GeV as a τ_{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 in the order of priority muons, electrons, τ_{had}, and jet candidates. For all channels, the leptons that are considered for overlap removal with τ_{had} candidates need only to satisfy looser criteria than those defined above, to reduce misidentified τ_{had} candidates from leptons. The p_{T} threshold of muons considered for overlap removal is also lowered to 4 GeV.

The missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) is reconstructed using the energy deposits in calorimeter cells calibrated according to the reconstructed physics objects (e, γ, τ_{had}, jets and μ) to which they are associated [41]. The transverse momenta of reconstructed muons are included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, with the energy deposited by these muons in the calorimeters taken into account. The energy from calorimeter cells not associated with any other objects is scaled by the soft-term vertex fraction and also included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. This fraction is the ratio of the scalar sum of the p_{T} of tracks from the primary vertex unmatched to objects to the scalar sum p_{T} of all tracks in the event also unmatched to objects. This method allows a better reconstruction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in high pile-up conditions [42].

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 given.

The signal contributions considered include the three main processes for Higgs boson production at the LHC: the gluon fusion (ggF), the vector-boson fusion (VBF), and the associated VH production processes. The contributions from the associated $t\bar{t}H$ production have been found to be small and are neglected. The gluon fusion and the VBF production are simulated with Powheg [43–46] interfaced to

[3] In the following, the τ_{had} symbol always refers to the visible decay product of the τ hadronic decay.

Pythia8 [47]. In the Powheg event generator the CT10 [48] parametrisation of the parton density functions (PDFs) is used. The overall normalisation of the *ggF* process is taken from a calculation at next-to-next-to-leading order (NNLO) [49–54] in QCD, including soft-gluon resummation up to the order of next-to-next-to-leading logarithm (NNLL) [55]. Next-to-leading order (NLO) electroweak (EW) corrections are also included [56, 57]. The *VBF* production is normalised to a cross section calculated with full NLO QCD and EW corrections [58–60] with an approximate NNLO QCD correction applied [61]. The associated VH production process is simulated with Pythia8. The Cteq6L1 [62] parametrisation of PDFs is used for the Pythia8 event generator. The predictions for VH production are normalised to cross sections calculated at NNLO in QCD [63], with NLO EW radiative corrections [64] applied.

Additional corrections to the shape of the generated p_{T} distribution of Higgs bosons produced via gluon fusion are applied to match the distribution from a calculation at NNLO including the NNLL corrections provided by the HRes2.1 [65] program. In this calculation, the effects of finite masses of the top and bottom quarks [65, 66] are included and dynamical renormalisation and factorisation scales, $\mu_{\mathrm{R}}, \mu_{\mathrm{F}} = \sqrt{m_H^2 + p_{\mathrm{T}}^2}$, are used. A reweighting is performed separately for events with less than or equal to 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 [67]. 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 matches the HRes2.1 and MinLo HJJ predictions respectively, and that the jet multiplicities are in agreement with (N)NLO calculations from JetVHeto [68–70].

The NLO EW corrections for the *VBF* production depend on the p_{T} of the Higgs boson, varying from a few percent at low p_{T} to $\sim 20\%$ at $p_{\mathrm{T}} = 300$ GeV [71]. The *VBF*-produced Higgs boson p_{T} spectrum is therefore reweighted, based on the difference between the Powheg+Pythia and the Hawk [58, 59] calculation, which includes these corrections.

The main and largely irreducible $Z/\gamma^* \to \tau\tau$ background is modelled using $Z/\gamma^* \to \mu\mu$ events from data,[4] 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 τ decays and the detector response of 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, $Z \to \mu\mu$ events can be selected from the data with high efficiency and purity. In order to replace the muons in the selected events, all tracks associated to the muons are removed and calorimeter cell energies associated to the muons are corrected by subtracting the corresponding energy depositions for a single simulated $Z \to \mu\mu$ event with the same kinematics. Finally, both the track information and the calorimeter cell energies of a simulated $Z \to \tau\tau$ decay are added to the data event. The decays of the tau leptons are simulated by Tauola [72], matched to the kinematics of the muons in data they replace, including polarisation and spin correlations [73], and accounting for the mass difference between the muons and the tau leptons. This hybrid sample is referred to as *embedded data* in the following.

Other background processes are simulated using different generators, each interfaced to Pythia [47, 74] or Herwig [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 simulated using Tauola [72]. Photos [76] provides photon radiation from charged leptons for all samples. 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 jets) and the parton shower. For WW production the loop-induced $gg \to WW$ process is also generated using the gg2WW [79]

[4] These processes are hereafter for simplicity denoted as $Z \to \tau\tau$ and $Z \to \mu\mu$ respectively, even though the whole continuum above and below the Z peak is considered.

Signal (m_H = 125 GeV)	MC generator	$\sigma \times \mathcal{B}$ [pb] $\sqrt{s}$ = 8 TeV		
ggF, $H \to \tau\tau$	Powheg [43–46] + Pythia8 [47]	1.22	NNLO+NNLL	[49–54, 84]
VBF, $H \to \tau\tau$	Powheg + Pythia8	0.100	(N)NLO	[58–60, 84]
WH, $H \to \tau\tau$	Pythia8	0.0445	NNLO	[63, 84]
ZH, $H \to \tau\tau$	Pythia8	0.0262	NNLO	[63, 84]
Background	MC generator	$\sigma \times \mathcal{B}$ [pb] $\sqrt{s}$ = 8 TeV		
$W(\to \ell\nu)$, $(\ell = e, \mu, \tau)$	Alpgen [77]+Pythia8	36800	NNLO	[85, 86]
$Z/\gamma^*(\to \ell\ell)$, 60 GeV< $m_{\ell\ell}$ < 2 TeV	Alpgen+Pythia8	3910	NNLO	[85, 86]
$Z/\gamma^*(\to \ell\ell)$, 10 GeV< $m_{\ell\ell}$ < 60 GeV	Alpgen+Herwig [75]	13000	NNLO	[85, 86]
VBF $Z/\gamma^*(\to \ell\ell)$	Sherpa [87]	1.1	LO	[87]
$t\bar{t}$	Powheg + Pythia8	253†	NNLO+NNLL	[88–93]
Single top : Wt	Powheg + Pythia8	22†	NNLO	[94]
Single top : s-channel	Powheg + Pythia8	5.6†	NNLO	[95]
Single top : t-channel	AcerMC [80]+Pythia6 [74]	87.8†	NNLO	[96]
$q\bar{q} \to WW$	Alpgen+Herwig	54†	NLO	[97]
$gg \to WW$	gg2WW [79]+Herwig	1.4†	NLO	[79]
WZ, ZZ	Herwig	30†	NLO	[97]
$H \to WW$	same as for $H \to \tau\tau$ signal	4.7†		

Table 1: Monte Carlo generators used to model the signal and background processes at $\sqrt{s}$ = 8 TeV. The cross sections times branching fractions ($\sigma \times \mathcal{B}$) used for the normalisation of some processes (many of these are subsequently normalised to data) are included in the last column together with the QCD perturbative order of the calculation. For the signal processes the $H \to \tau\tau$ branching ratio is included, and for the W and Z/γ^* 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 †).

program. In the AcerMC [80], Alpgen, and Herwig event generators the Cteq6L1 parametrisation of the PDFs is used, while the CT10 parametrisation 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 interactions were simulated using the AU2 [83] tuning of Pythia8. They are overlaid on the signal and background simulated 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.

Trigger	Trigger level thresholds, p_{T} [GeV]	Analysis level thresholds [GeV] $\sqrt{s}$ = 7 TeV					
		$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$		$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$		$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$	
Single electron	20–22	$e\mu$:	$p_{\mathrm{T}}(e) > 22-24$ $p_{\mathrm{T}}(\mu) > 10$	$e\tau$:	$p_{\mathrm{T}}(e) > 25$ $p_{\mathrm{T}}(\tau) > 20$	–	
Single muon	18	$\mu\mu$:	$p_{\mathrm{T}}(\mu_1) > 20$ $p_{\mathrm{T}}(\mu_2) > 10$	$\mu\tau$:	$p_{\mathrm{T}}(\mu) > 22$ $p_{\mathrm{T}}(\tau) > 20$	–	
Di-electron	12/12	ee:	$p_{\mathrm{T}}(e_1) > 15$ $p_{\mathrm{T}}(e_2) > 15$	–		–	
Di-τ_{had}	29/20	–		–		$\tau\tau$:	$p_{\mathrm{T}}(\tau_1) > 35$ $p_{\mathrm{T}}(\tau_2) > 25$
Trigger	Trigger level thresholds, p_{T} [GeV]	Analysis level thresholds [GeV] $\sqrt{s}$ = 8 TeV					
		$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$		$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$		$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$	
Single electron	24	$e\mu$: ee:	$p_{\mathrm{T}}(e) > 26$ $p_{\mathrm{T}}(\mu) > 10$ $p_{\mathrm{T}}(e_1) > 26$ $p_{\mathrm{T}}(e_2) > 15$	$e\tau$:	$p_{\mathrm{T}}(e) > 26$ $p_{\mathrm{T}}(\tau) > 20$	–	
Single muon	24	–		$\mu\tau$:	$p_{\mathrm{T}}(\mu) > 26$ $p_{\mathrm{T}}(\tau) > 20$	–	
Di-electron	12/12	ee:	$p_{\mathrm{T}}(e_1) > 15$ $p_{\mathrm{T}}(e_2) > 15$	–		–	
Di-muon	18/8	$\mu\mu$:	$p_{\mathrm{T}}(\mu_1) > 20$ $p_{\mathrm{T}}(\mu_2) > 10$	–		–	
Electron+muon	12/8	$e\mu$:	$p_{\mathrm{T}}(e) > 15$ $p_{\mathrm{T}}(\mu) > 10$	–		–	
Di-τ_{had}	29/20	–		–		$\tau\tau$:	$p_{\mathrm{T}}(\tau_1) > 35$ $p_{\mathrm{T}}(\tau_2) > 25$

Table 2: Summary of the triggers used to select events for the different analysis channels at the two centre-of-mass energies. Both the transverse momentum thresholds applied at trigger level as well as in the analysis are listed. When more than one trigger is used, a logical OR is taken and the trigger efficiencies are calculated accordingly.

4 Event selection and categorisation

4.1 Event selection

Single lepton, dilepton and di-hadronic tau 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 and more stringent identification requirements were applied for the data taking at $\sqrt{s} = 8$ TeV 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 primary vertex is required with at least four associated tracks and a position consistent with the beam spot.

With respect to the object identification requirements described in Section 2, tighter criteria are applied to address the different background contributions and compositions in the different analysis channels. Higher p_{T} thresholds are applied to electrons, muons, and τ_{had} candidates according to the trigger

		$\tau_{\text{lep}}\tau_{\text{lep}}$	$\tau_{\text{lep}}\tau_{\text{had}}$
Electrons	7 TeV	$I(p_T, 0.4) < 0.06$ $I(E_T, 0.2) < 0.08$	$I(p_T, 0.4) < 0.06$ $I(E_T, 0.2) < 0.06$
	8 TeV	$I(p_T, 0.4) < 0.17$ $I(E_T, 0.2) < 0.09$	$I(p_T, 0.4) < 0.06$ $I(E_T, 0.2) < 0.06$
Muons	7 TeV	$I(p_T, 0.4) < 0.06$ $I(E_T, 0.2) < 0.04$	$I(p_T, 0.4) < 0.06$ $I(E_T, 0.2) < 0.06$
	8 TeV	$I(p_T, 0.4) < 0.18$ $I(E_T, 0.2) < 0.09$	$I(p_T, 0.4) < 0.06$ $I(E_T, 0.2) < 0.06$

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.

conditions satisfied by the event, as listed in Table 2. For the channels involving leptonic tau decays, $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$, additional isolation criteria on electrons and muons, based on tracking and calorimeter information, are used to suppress the background from misidentified jets or from semileptonic decays of charm and bottom hadrons. The calorimeter isolation variable $I(E_{\text{T}}, \Delta R)$ is defined as the sum of the total transverse energy in the calorimeter in a given cone of size ΔR 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 $I(p_{\text{T}}, \Delta R)$ is defined as the sum of the transverse momenta of tracks within a cone of ΔR 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 defined, if no tracks with $p_{\text{T}} > 0.5$ GeV are found 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%) jet rejection is obtained 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.

$\tau_{\text{lep}}\tau_{\text{lep}}$ 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 τ_{had} candidate are vetoed. For the τ_{had} candidates considered the criteria used to reject electrons misidentified as τ_{had} candidates are tightened to a working-point of 85% signal efficiency [40].

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 charmonium and bottomonium resonances, 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 the missing transverse momentum $E_{\text{T}}^{\text{miss}}$. In order 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 $E_{\text{T}}^{\text{miss}}$ 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* $E_{\text{T}}^{\text{miss}}$ ($E_{\text{T}}^{\text{miss,HPTO}}$) is also used to reject background from Z/γ^* production. It is calculated from the high p_{T} objects in the event, i.e. from the two leptons and jets with $p_{\text{T}} > 25$ GeV. Due to the presence of real neutrinos, the two $E_{\text{T}}^{\text{miss}}$ variables are strongly correlated for signal events but only loosely correlated for background from $Z \to ee$ and $Z \to \mu\mu$ decays.

To further suppress background contributions from misidentified leptons[5] a minimal value of the

[5]Misidentified leptons (τ_{had} candidates) are also referred to as "fake" leptons (τ_{had} candidates) in this paper.

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-tagged jet with $p_{\mathrm{T}} > 25$ GeV.

Within the colinear approximation [98], 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], $x_{\tau 1(2)} = p_{\mathrm{vis1(2)}}/(p_{\mathrm{vis1(2)}} + p_{\mathrm{mis1(2)}})$, are expected to lie in the interval $0 < x_{\tau 1(2)} < 1$, and hence corresponding requirements are applied to further reject non-tau background contributions.

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

$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel: Exactly one isolated lepton and one τ_{had} candidate with OS charges, passing the p_{T} thresholds listed in Table 2, are required. The criteria used to reject electrons misidentified as τ_{had} are also tightened in this channel to a working-point of 85% signal efficiency [40].

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 $E_{\mathrm{T}}^{\mathrm{miss}}$ is applied and events with $m_{\mathrm{T}} > 70$ GeV are rejected. Contributions from $t\bar{t}$ events are reduced by rejecting events with a b-tagged jet with $p_{\mathrm{T}} > 30$ GeV.

$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel: One isolated medium and one isolated tight τ_{had} candidate with OS charges are required. Events with electron or muon candidates are rejected. For all data, $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to exceed 20 GeV and its direction must either be between the two visible τ_{had} candidates in ϕ or within $\Delta\phi < \pi/4$ of the nearest τ_{had} candidate. In order to further reduce the background from multijet production, additional cuts on the ΔR and pseudorapidity separation $\Delta\eta$ between the two τ_{had} candidates are applied.

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

4.2 Analysis categories

In order to exploit signal-sensitive event topologies, two analysis categories are defined in an exclusive way:

- 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 using the two highest-p_{T} jets in the event. In the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel there is an additional requirement that $m_{\tau\tau}^{\mathrm{vis}} > 40$ GeV, , to eliminate low-mass Z/γ^* events.. Although this category is dominated by *VBF* events, it also includes smaller contributions from gluon-fusion and VH production.

- The *Boosted* category targets events with a boosted Higgs boson produced via gluon fusion. Higgs boson candidates are required to have a large transverse momentum, $p_{\mathrm{T}}^H > 100$ GeV. The p_{T}^H is reconstructed using the vector sum of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse momentum of the visible tau decay products. In the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel at least one jet with $p_{\mathrm{T}} > 40$ GeV is required. In order to define an orthogonal category, events passing the *VBF* categorisation are not considered. This category also includes small contributions from *VBF* and *VH* production.

[6] p_{vis} is defined as the total momentum of the visible decay products of the tau lepton, p_{mis} is defined as the momentum of the neutrino reconstructed using the colinear approximation.

[7] $m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}(\ell)\, E_{\mathrm{T}}^{\mathrm{miss}} \cdot (1 - \cos\Delta\phi)}$, where $\Delta\phi$ is the azimuthal separation between the directions of the lepton and the missing transverse momentum vector.

Channel	Pre-selection cuts
$\tau_{\text{lep}}\tau_{\text{lep}}$	Exactly two isolated opposite-sign leptons Events with τ_{had} candidates are rejected 30 GeV $< m_{\tau\tau}^{\text{vis}} <$ 100 (75) GeV for DF (SF) events $\Delta\phi_{\ell\ell} < 2.5$ $E_{\text{T}}^{\text{miss}} >$ 20 (40) GeV for DF (SF) events $E_{\text{T}}^{\text{miss,HPTO}} >$ 40 GeV for SF events $p_{\text{T}}(\ell_1) + p_{\text{T}}(\ell_2) >$ 35 GeV Events with a b-tagged jet with $p_{\text{T}} >$ 25 GeV are rejected $0.1 < x_{\tau_1}, x_{\tau_2} < 1$ $m_{\tau\tau}^{\text{coll}} > m_Z -$ 25 GeV
$\tau_{\text{lep}}\tau_{\text{had}}$	Exactly one isolated lepton and one medium τ_{had} candidate with opposite charges $m_{\text{T}} <$ 70 GeV Events with a b-tagged jet with $p_{\text{T}} >$ 30 GeV are rejected
$\tau_{\text{had}}\tau_{\text{had}}$	One isolated medium and one isolated tight opposite-sign τ_{had}-candidate Events with leptons are vetoed $E_{\text{T}}^{\text{miss}} >$ 20 GeV $E_{\text{T}}^{\text{miss}}$ points between the two visible taus in ϕ, or $\min[\Delta\phi(\tau, E_{\text{T}}^{\text{miss}})] < \pi/4$ $0.8 < \Delta R(\tau_{\text{had}_1}, \tau_{\text{had}_2}) < 2.4$ $\Delta\eta(\tau_{\text{had}_1}, \tau_{\text{had}_2}) < 1.5$
Channel	*VBF* category selection cuts
$\tau_{\text{lep}}\tau_{\text{lep}}$	At least two jets with $p_{\text{T}}(j_1) >$ 40 GeV and $p_{\text{T}}(j_2) >$ 30 GeV $\Delta\eta(j_1, j_2) > 2.2$
$\tau_{\text{lep}}\tau_{\text{had}}$	At least two jets with $p_{\text{T}}(j_1) >$ 50 GeV and $p_{\text{T}}(j_2) >$ 30 GeV $\Delta\eta(j_1, j_2) > 3.0$ $m_{\tau\tau}^{\text{vis}} >$ 40 GeV
$\tau_{\text{had}}\tau_{\text{had}}$	At least two jets with $p_{\text{T}}(j_1) >$ 50 GeV and $p_{\text{T}}(j_2) >$ 30 GeV $p_{\text{T}}(j_2) >$ 35 GeV for jets with $\lvert\eta\rvert > 2.4$ $\Delta\eta(j_1, j_2) > 2.0$
Channel	*Boosted* category selection cuts
$\tau_{\text{lep}}\tau_{\text{lep}}$	At least one jet with $p_{\text{T}}>$ 40 GeV
All	Failing the *VBF* selection $p_{\text{T}}^{H} >$ 100 GeV

Table 4: Summary of the event selection for the three analysis channels. The cuts used in both the pre-selection 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, τ_{had}, jets). The variables are defined in the text.

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 those passing the category requirements, are allowed.

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

4.3 Higgs boson candidate mass reconstruction

The $\tau\tau$ invariant mass ($m_{\tau\tau}^{\text{MMC}}$) is reconstructed using the missing mass calculator (MMC) [99]. 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 undetected neutrinos for each of the two tau leptons in the event, and the invariant mass of the two neutrinos from any leptonic tau decays. This is done by using the constraints from the measured x- and y-components of $E_{\text{T}}^{\text{miss}}$ and the visible masses of both tau candidates. A scan is performed over the two components of the $E_{\text{T}}^{\text{miss}}$ vector and the yet undetermined variables. Each scan point is weighted by its probability according to the $E_{\text{T}}^{\text{miss}}$ 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 ~99% of the $H \to \tau\tau$ and $Z \to \tau\tau$ events. This is a distinct advantage compared to the mass calculation using the colinear approximation where the failure rate is higher due to the implicit colinearity assumptions. The small loss rate of about 1% for signal events is due to large fluctuations of the $E_{\text{T}}^{\text{miss}}$ measurement or other scan variables. In Figure 1 reconstructed $m_{\tau\tau}^{\text{MMC}}$ mass distributions are shown for $\tau_{\text{lep}}\tau_{\text{had}}$ signal events with a mass of 125 GeV in the *VBF* and *Boosted* categories. The mass resolution, $\mathcal{R}$, is found to be 15% and 16% for the *VBF* and *Boosted* categories respectively. The resolutions in the other categories are: $\mathcal{R}_{\tau_{\text{lep}}\tau_{\text{lep}}}^{VBF} \approx 16\%$, $\mathcal{R}_{\tau_{\text{lep}}\tau_{\text{lep}}}^{Boosted} \approx 16\%$, $\mathcal{R}_{\tau_{\text{had}}\tau_{\text{had}}}^{VBF} \approx 14\%$, and $\mathcal{R}_{\tau_{\text{had}}\tau_{\text{had}}}^{Boosted} \approx 14\%$. The distributions of reconstructed $m_{\tau\tau}^{\text{MMC}}$ for $Z \to \tau\tau$ background events are also shown in Figure 1.

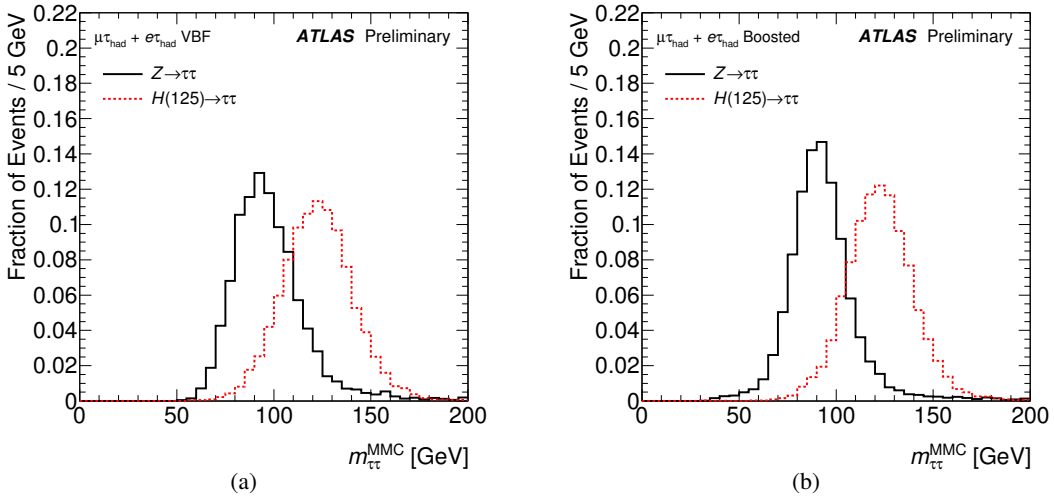


Figure 1: The reconstructed $m_{\tau\tau}^{\text{MMC}}$ mass distributions for $H \to \tau\tau$ (m_H = 125 GeV) and $Z \to \tau\tau$ events in MC simulation and embedding, respectively, for events passing the *VBF* selection (a) and the *Boosted* (b) selection in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel.

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 [22] recursively partition the parameter space into multiple regions where signal or background purities are enhanced. Boosting is a method which improves the performance and stability of decision trees and involves the combination of many trees into a single final discriminant [23, 24]. After boosting, the final score undergoes a transformation to map the scores on the interval -1 to 1. The most signal-like events have scores near 1 while the most background-like events have scores near -1.

Separate BDTs are trained for each analysis category and channel with signal and background samples, described in Section 6, at $\sqrt{s} = 8$ TeV. They are then applied to the analysis of the data of 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. For the training in the *VBF* category only a *VBF* signal sample is used, while in the *Boosted* category gluon fusion, *VBF*, and *VH* signal samples are included. The Higgs boson mass has been chosen to be $m_H = 125$ 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 in ΔR between the two leptons, between the lepton and τ_{had}, or between the two τ_{had} candidates, depending on the decay mode.
- $p_{\text{T}}^{\text{Total}}$: magnitude of the vector sum of the visible components of the tau decay products, the two leading jets, and $E_{\text{T}}^{\text{miss}}$.
- Sum p_{T}: scalar sum of the p_{T} of the visible components of the tau decay products and of the jets.
- $E_{\text{T}}^{\text{miss}}\phi$ centrality: a variable that quantifies the relative angular position of the missing transverse momentum with respect to the 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 ϕ angle between the tau decay products defines the positive quadrant of the transformed plane. $E_{\text{T}}^{\text{miss}}\phi$ centrality is defined as the sum of the x and y components of the $E_{\text{T}}^{\text{miss}}$ unit vector in this transformed plane.
- Sphericity: a variable that describes the isotropy of the energy flow in the event [100]. It is based on the quadratic momentum tensor

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i^{\,2}|}. \tag{1}$$

  In this equation, α and β 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, λ_2 and λ_3:

$$S = \frac{3}{2}(\lambda_2 + \lambda_3). \tag{2}$$

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

$$C_{\eta_1,\eta_2}(\eta) = \exp\left[\frac{-4}{(\eta_1 - \eta_2)^2}\left(\eta - \frac{\eta_1 + \eta_2}{2}\right)^2\right], \tag{3}$$

where η, η_1 and η_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 η between the two jets, $1/e$ when the object is aligned with one of the jets, and $< 1/e$ when the object is outside the jets. In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel the η centrality of a third jet in the event, $C_{\eta_1,\eta_2}(\eta_{j_3})$, and the product of the η centralities of the two leptons are used as BDT input variables, while in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel the η centrality of the lepton, $C_{\eta_1,\eta_2}(\eta_\ell)$, is used, and in the $\tau_{\text{had}}\tau_{\text{had}}$ channel the η centrality of each τ, $C_{\eta_1,\eta_2}(\eta_{\tau_1})$ and $C_{\eta_1,\eta_2}(\eta_{\tau_2})$, is used. Events with only two jets are assigned a dummy value of -0.5 for $C_{\eta_1,\eta_2}(\eta_{j_3})$.

Among these variables, the most discriminating ones include: $m_{\tau\tau}^{\text{MMC}}$, $\Delta R(\tau_1, \tau_2)$ and $\Delta\eta(j_1, j_2)$. In Figure 2 the distributions of selected BDT input variables are shown. For the *VBF* category, the distributions of $\Delta\eta(j_1, j_2)$ are shown for all three channels. For the *Boosted* category the distributions of $\Delta R(\tau_1, \tau_2)$ 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 is shown for the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel. For all distributions the data are compared to the predictions from SM background processes 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.

Variable	*VBF*			*Boosted*		
	$\tau_{\text{lep}}\tau_{\text{lep}}$	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$	$\tau_{\text{lep}}\tau_{\text{lep}}$	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$
$m_{\tau\tau}^{\text{MMC}}$	•	•	•	•	•	•
$\Delta R(\tau_1, \tau_2)$	•	•	•		•	•
$\Delta\eta(j_1, j_2)$	•	•	•			
m_{j_1,j_2}	•	•	•			
$\eta_{j_1} \times \eta_{j_2}$		•	•			
$p_{\text{T}}^{\text{Total}}$		•	•			
Sum p_{T}					•	•
$p_{\text{T}}(\tau_1)/p_{\text{T}}(\tau_2)$					•	•
$E_{\text{T}}^{\text{miss}}\phi$ centrality		•	•	•	•	•
m_{ℓ,ℓ,j_1}				•		
m_{ℓ_1,ℓ_2}				•		
$\Delta\phi(\ell_1, \ell_2)$				•		
Sphericity				•		
$p_{\text{T}}^{\ell_1}$				•		
$p_{\text{T}}^{j_1}$				•		
$E_{\text{T}}^{\text{miss}}/p_{\text{T}}^{\ell_2}$				•		
m_{T}		•			•	
$\min(\Delta\eta_{\ell_1\ell_2,\text{jets}})$	•					
$C_{\eta_1,\eta_2}(\eta_{\ell_1}) \cdot C_{\eta_1,\eta_2}(\eta_{\ell_2})$	•					
$C_{\eta_1,\eta_2}(\eta_\ell)$		•				
$C_{\eta_1,\eta_2}(\eta_{j_3})$	•					
$C_{\eta_1,\eta_2}(\eta_{\tau_1})$			•			
$C_{\eta_1,\eta_2}(\eta_{\tau_2})$			•			

Table 5: Discriminating variables used in the training of the BDT for each channel and category at $\sqrt{s} = 8$ TeV. The filled circles indicate which variables are used in each case. Variables such as $\Delta R(\tau_1, \tau_2)$ are defined between the two leptons, between the lepton and τ_{had}, or between the two τ_{had} candidates, depending on the decay mode.

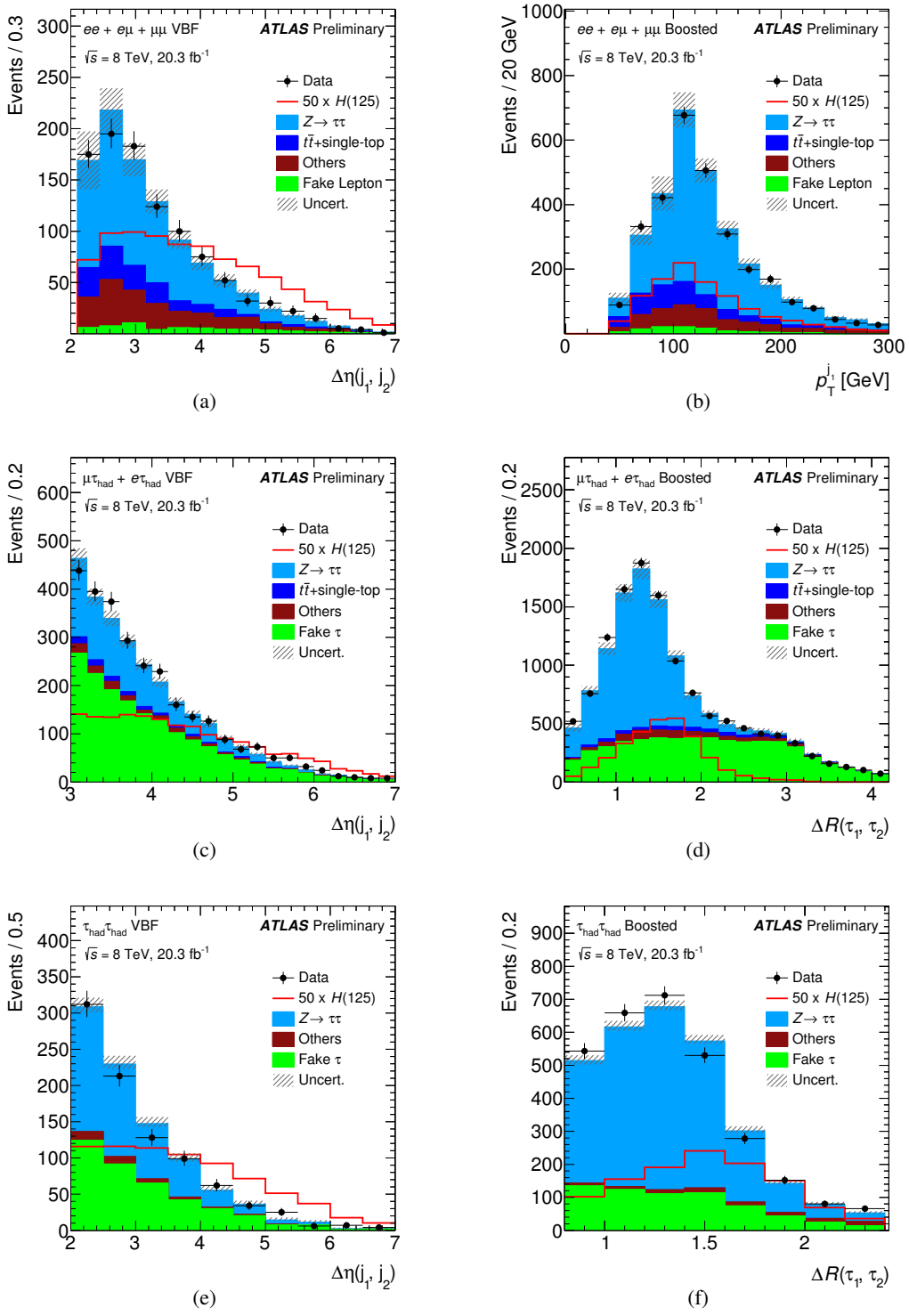


Figure 2: Distributions of important BDT input variables for the three channels and the two categories (*VBF*, left) and (*Boosted*, right) for data collected at $\sqrt{s} = 8$ TeV. The distributions are shown for (a) $\Delta\eta(j_1, j_2)$ and (b) $p_{\mathrm{T}}(j_1)$ in the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel, for (c) $\Delta\eta(j_1, j_2)$ and (d) $\Delta R(\tau_1, \tau_2)$ in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel and for (e) $\Delta\eta(j_1, j_2)$ and (f) $\Delta R(\tau_1, \tau_2)$ in the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ 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.

Common to all channels is the dominant $Z \to \tau\tau$ 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 τ_{had} candidates. Likewise, samples of non-isolated leptons are used to estimate fake lepton contributions from either jets or hadronically decaying taus and leptons from other sources, such as heavy quark decay.[8]

Other non-fake contributions from various physics processes are estimated using the simulation, normalised to the theoretical cross sections, as given in Table 1. A more detailed discussion on the estimation of the various background components in the different channels is given in the following.

6.1 Backgrounds from $Z \to \tau\tau$ production

A reliable modelling of the irreducible $Z \to \tau\tau$ background is an important ingredient of the analysis. Since it is not possible to select a sufficiently pure and signal-free $Z \to \tau\tau$ control sample from data, the contribution of this background is estimated using embedded data. This procedure has been extensively validated by 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 $Z \to \mu\mu$ 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 are compared in Figure 3(a). Good agreement is found, which indicates that no deterioration in the muon environment is introduced. Another important test constitutes the validation of the embedding of more complex $Z \to \tau\tau$ 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 $Z \to \mu\mu$ and $Z \to \tau\tau$ 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 $p_{\text{T}}(\mu)$ and $\eta(\mu)$, and allow the acceptance of the original selection to be corrected. The tau decay product are treated as any other objects determined from the simulation, with one 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 by simulated taus, kinematic distributions of the embedded sample can be directly compared to the simulated ones. As an example, the reconstructed invariant mass, $m_{\tau\tau}^{\text{MMC}}$, is shown in Figure 3(b). Also in this case, 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 energy, the kinematic variables of the hadronically decaying tau lepton or of the associated jets in the event. A direct comparison of the $Z \to \tau\tau$ background in data and the modelling using the embedding technique also shows good agreement. This can be seen, e.g. from several distributions of kinematic quantities, which are dominated by $Z \to \tau\tau$ events, shown in Figure 2.

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

[8] Leptons from heavy quark decays are considered as fake leptons in the following.

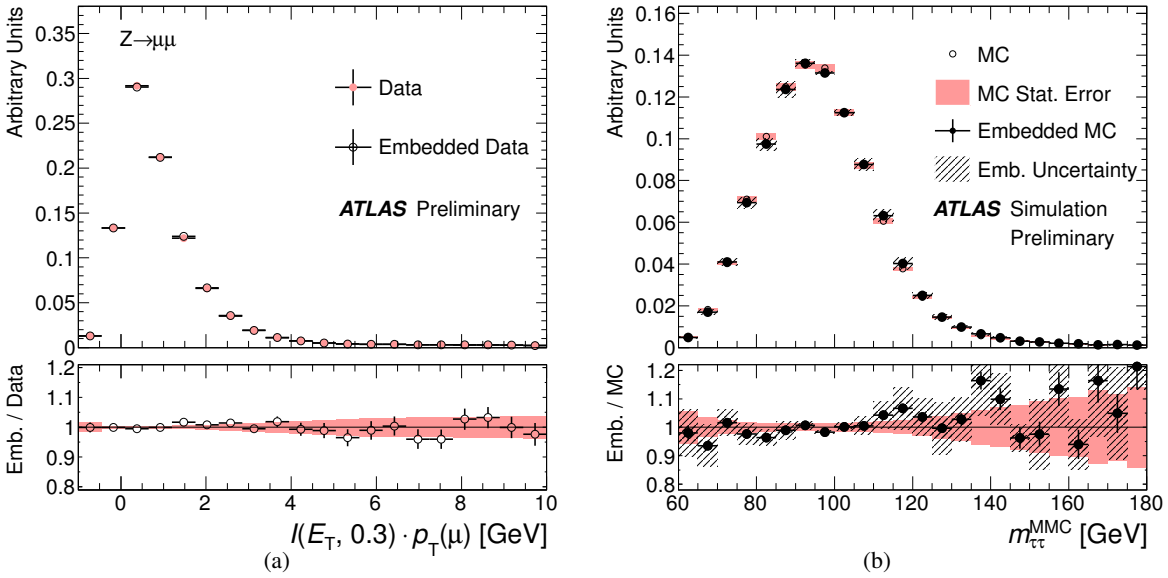


Figure 3: (a) The distribution of the calorimeter isolation energy $I(E_{\mathrm{T}}, 0.3) \cdot p_{\mathrm{T}}(\mu)$ within a cone of radius $\Delta R = 0.3$ around the muons in $Z \to \mu\mu$ events from data, before and after embedding of simulated muons. (b) The distribution of the reconstructed invariant $\tau\tau$ mass, $m_{\tau\tau}^{\mathrm{MMC}}$, for simulated $Z \to \tau\tau$ events, compared to the one obtained from simulated $Z \to \mu\mu$ events after tau embedding. The ratios of the values before and after embedding and between the embedded $Z \to \mu\mu$ and $Z \to \tau\tau$ events are given in (a) and (b) respectively. The errors in (a) and (b) on the ratios (points) represent the statistical uncertainties, while the systematic uncertainties are indicated by the hatched bands in (b). The shaded bands represent the statistical uncertainties from the $Z \to \mu\mu$ data event in (a) and from the $Z \to \tau\tau$ MC in (b).

6.2 Backgrounds 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+jets production, as well as the part of the $t\bar{t}$ background resulting from lepton-hadron decays ($t\bar{t} \to \ell\nu b\, qqb$) are included. A control sample is defined in data by inverting the isolation selections 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 sub-leading lepton at an early stage of the pre-selection.

For the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the fake-factor method is used to derive estimates for the multijet, W+jets, Z+jets, and semi-leptonic $t\bar{t}$ background events that pass the $\tau_{\text{lep}}\tau_{\text{had}}$ selection due to a misidentified τ_{had} candidate. The fake factor is defined as the ratio of the number of τ_{had} candidates identified as medium, to the number passing the loose, but not the medium criteria. Since the fake factor depends on the nature and on the p_{T} of the jet, it is determined as a function of p_{T} separately for quark- and gluon-enriched samples. 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 ($m_{\text{T}} > 70$ GeV), by inverting the b-tag veto and by requiring two leptons with an invariant mass consistent with m_Z ($80\ \text{GeV} < m_{\ell\ell} < 100\ \text{GeV}$) respectively. In addition, a gluon-jet dominated multijet sample is selected by relaxing the lepton identification and requiring the lepton to pass the loose identification criteria. The derived fake factors are found to vary from 0.124 (0.082) for $p_{\text{T}} = 20$ GeV to 0.088 (0.038) for $p_{\text{T}} = 150$ GeV for 1-track (3-track) candidates in the *VBF* category. The corresponding values for the *Boosted* category are 0.146 (0.084) for $p_{\text{T}} = 20$ GeV and 0.057 (0.033) for $p_{\text{T}} = 150$ 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 τ_{had} candidate is required to pass the loose and to fail the medium τ_{had} identification. As an example, the good agreement between data and background estimations is shown in Figure 4(a) for the reconstructed $\tau\tau$ mass for events in the high-m_{T} region, which is dominated by fakes from W+jets production.

For the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the multijet background is modelled using a template extracted from data that pass the *VBF* or *Boosted* signal selection, where, however, the taus fail the isolation and opposite-sign charge requirements. 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 pre-selection cuts. The fit is performed for the distribution of the difference in pseudorapidity between the two hadronic tau candidates, $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}})$. The signal contribution is expected to be small in this category. The agreement between data and background estimation for this distribution is shown in Figure 4(b) for a control region using 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 $Z \to \tau\tau$ 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 $Z \to ee$ and $Z \to \mu\mu$ backgrounds

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

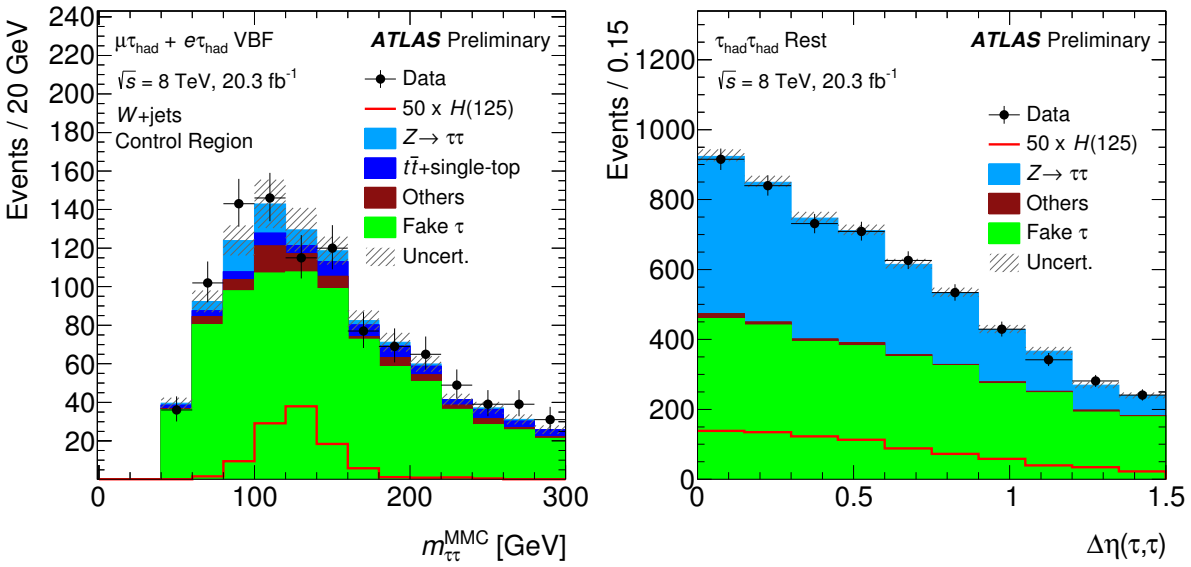


Figure 4: (a) The distribution of the reconstructed $\tau\tau$ mass for events in the W+jets control region, for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. (b) The distribution of $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}})$ for the $\tau_{\text{had}}\tau_{\text{had}}$ channel in the *Rest* control region. The expected SM Higgs boson signal contribution is superimposed, multiplied by a factor 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.

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, the ALPGEN simulation is normalised to the data in the Z-mass control region, 80 GeV $< m_{\ell\ell} <$ 100 GeV, for each category, and separately for $Z \to ee$ and $Z \to \mu\mu$ 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 5.

In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the $Z \to ee$ and $Z \to \mu\mu$ background estimates are also based on simulation. The corrections applied for a τ_{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 fake τ_{had} candidates resulting from leptons [19, 101]. This is particularly important for $Z \to ee$ events with a misidentified τ_{had} candidate originating from a true electron. In the second case, the fake-factor method described in Section 6.2 is applied.

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 τ_{had} or a lepton candidate. This process can also contribute via semi-leptonic 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 based on data-driven methods. For the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the $W \to \tau_{\text{had}}\nu$ background is estimated from simulation. A correction is applied to account for differences in the τ_{had} fake rate between data and simulation.

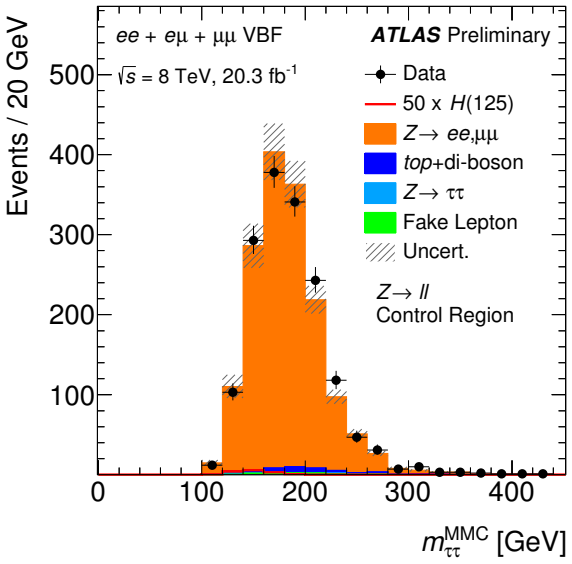


Figure 5: The distribution of the reconstructed mass $m_{\tau\tau}^{\mathrm{MMC}}$ for events in the $Z \to \ell\ell$ control region, for the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel. This figure uses background predictions made without the global fit defined in Section 8. The error band includes statistical and pre-fit systematic uncertainties.

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 $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ and $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channels. The normalisation is obtained from data control regions defined by requiring a b-tag instead of a b-veto. In the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel a large value of the transverse mass m_{T} is required in addition, to enhance the background from top production and to suppress the signal contribution. This background is also found to be small for the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel and it is estimated using simulation.

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 $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ 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 $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel a non-negligible contribution from $H \to WW \to \ell\nu\ell\nu$ 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 BDT output distributions in control regions

It is important to verify that the BDT output distributions in data control regions are well described after the various background determinations. The control regions used in the analysis are summarised in Table 6. In Figure 6, such distributions are shown for important control regions for the $\sqrt{s} = 8$ TeV dataset, i.e. the Z-enriched control regions for the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ and $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channels, and the reconstructed

$\tau\tau$ invariant mass side-band control region (defined as $m_{\tau\tau}^{\mathrm{MMC}} < 100$ GeV or $m_{\tau\tau}^{\mathrm{MMC}} > 150$ GeV) for the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ 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.

Process	$\tau_{\text{lep}}\tau_{\text{lep}}$	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$
$Z \to \ell\ell$-enriched	$80 < m_{\tau\tau}^{\text{vis}} < 100$ GeV (same-flavour)		
Top control region	invert b-jet veto	invert b-jet veto and $m_{\text{T}} > 40$ GeV	
Rest category			pass preselection, fail *VBF* and *Boosted* selections
$Z \to \tau\tau$-enriched	$m_{\tau\tau}^{\text{HPTO}} < 100$ GeV	$m_{\text{T}} < 40$ GeV and $m_{\tau\tau}^{\text{MMC}} < 110$ GeV	
Fake-enriched	same sign τ decay products	same sign τ decay products	
W-enriched		$m_{\text{T}} > 70$ GeV	
Mass sideband			$m_{\tau\tau}^{\text{MMC}} < 110$ GeV or $m_{\tau\tau}^{\text{MMC}} > 150$ GeV

Table 6: Summary and definition of the control regions used in the analysis. The definitions shown are modifications on top of the signal region requirements. The variable $m_{\tau\tau}^{\text{HPTO}}$ is the invariant mass of the $\tau\tau$-system obtained using the colinear 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.

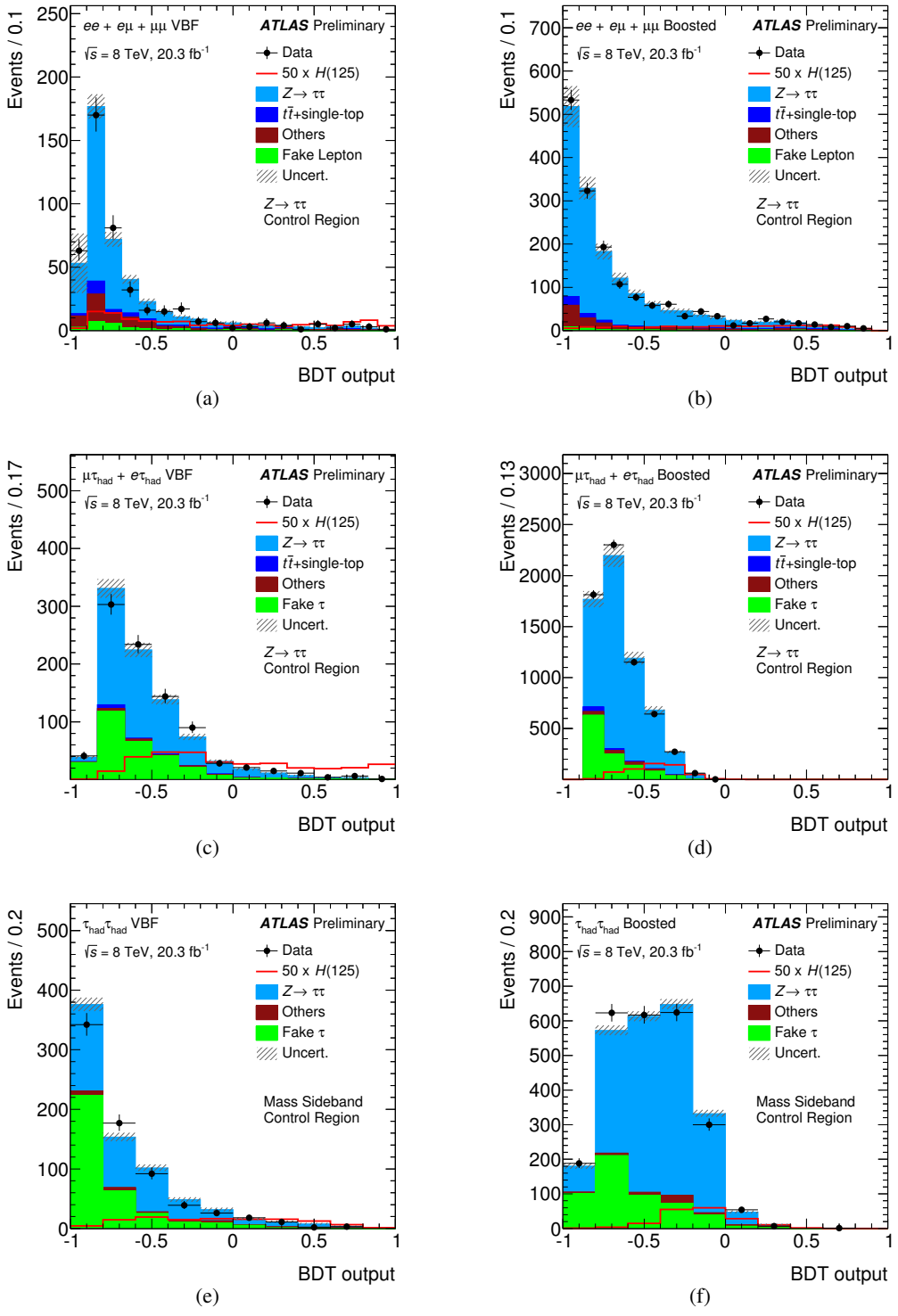


Figure 6: Distributions of the BDT output for data collected at $\sqrt{s} = 8$ TeV, compared to the expected background contributions in important control regions. The distributions are shown for the *VBF* (left) and *Boosted* (right) categories (a,b) for the $Z \to \tau\tau$-enriched control region in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, (c,d) for the $Z \to \tau\tau$-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. The uncertainties on the individual background components are calculated by applying the methods discussed in Section 6. For all uncertainties the effects on both the total signal and background yields and on the shape of the BDT output distribution are evaluated. 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 is given in Table 7. In this table also the dominant sources that affect the shape of the BDT output distribution are marked. All uncertainties are treated either as fully correlated or uncorrelated across channels. The latter are also marked in the table. 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 each systematic uncertainty on the final sensitivity is presented in Table 11.

7.1 Experimental systematic 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 7. In addition, uncertainties on the luminosity affect the number of signal and background events from simulation.

- Luminosity: The uncertainty on the integrated luminosity is ±2.8% for the 8 TeVdataset and ±1.8% for the 7 TeVdataset. 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. [102].

- Efficiencies: The efficiencies for triggering, reconstructing and identifying electrons, muons, and τ_{had} candidates are measured in data using tag-and-probe techniques. The uncertainties on the τ_{had} identification efficiency are ±(2–3)% for 1-prong and ±(3–5)% for 3-prong tau decays. 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 ±2% for jets with transverse momenta up to 100 GeV [103, 104]. 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, η inter-calibration, and detector response and modelling in in-situ jet calibration [34, 35]. 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 $Z \to \tau\tau$ events in data, which can be selected with a satisfactory purity. It is measured with a precision of ±(2–4)% [105]. 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 $E_{\text{T}}^{\text{miss}}$ due to the soft term is also taken into account.

- Energy resolutions: Systematic uncertainties on the energy resolution of taus, electrons, muons, jets, and $E_{\text{T}}^{\text{miss}}$ 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 (insensitive material in the detector, e.g. support structures), in the hadronic shower model and in the underlying event tune have been 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. [106], and affects signal modelling and background components modelled by the simulation. The uncertainty of the resolution on $E_{\mathrm{T}}^{\mathrm{miss}}$ is estimated by evaluating the energy resolution of each of the $E_{\mathrm{T}}^{\mathrm{miss}}$ terms. The largest impact results from the soft term, arising both from the MC modelling and the effects of pile-up. It is evaluated using simulated $Z(\rightarrow \mu\mu)$+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 $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ 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 ±33% (±20%) at 8 TeV and ±10.5% (±13%) at 7 TeV for the *Boosted* (*VBF*) category. The extrapolation uncertainty on 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 ±6%, while the corresponding extrapolation uncertainty for top-quark background sources is $\pm(3-6)$%, obtained from the difference of event yields in the top-quark control regions when using different MC generators. Neither of these extrapolation uncertainties is significant for the final result.

In the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel, an important systematic uncertainty on the background estimation comes from the estimated fake background, for which several sources of systematic uncertainty are considered. The statistical uncertainty on the effective fake factor is ±4.3% (±2.3%) in the 8 TeV *VBF* (*Boosted*) category, and about ±22% (±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 in large parts 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 the fakes-rich region of data where the lepton and τ_{had}candidate have same charge, and very good agreement was observed between data and the method prediction, so that no additional in-situ uncertainty was deemed necessary. In addition, the uncertainties on the normalisation of the $t\bar{t}$ background, obtained from the global fit, are important. The dominant contribution results 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 $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ 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 τ_{had} candidates fail the isolation and opposite-sign charge requirements, is compared with an alternative template derived from a sample where the τ_{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 difference in how it propagates into the various categories gives a difference in yields, which along with the difference in shape between the two templates constitutes the systematic uncertainties 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 $Z \to \mu\mu$ events in data and to the subtraction of the muon energy depositions in the calorimeters. The selection uncertainties are estimated by varying the muon isolation criteria in the selection from the nominal value of $I(p_{\mathrm{T}}, 0.2) < 0.2$ (see Section 4) to tighter ($I(p_{\mathrm{T}}, 0.4) < 0.06$ and $I(E_{\mathrm{T}}, 0.2) < 0.04$) 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 have been taken into account. Due to the combination of single- and dilepton triggers used, they are largest for the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel. All experimental systematics relating to the embedded τ decay products (such as tau energy scale or identification uncertainties) are applied normally. The combined effect of all uncertainties on the signal and background numbers is included in Table 7. Because the $Z \to \tau\tau$ normalisation is determined in the final fit, the impact on the final result is much smaller.

7.3 Theoretical uncertainties

Theoretical systematic uncertainties are estimated for the signal and for all background contributions modelled with the simulation. Since the major background contributions, from $Z \to \tau\tau$ and misidentification of hadronically decaying τ, 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 the *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 [107]. The resulting uncertainties range from ±2% to ±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 the Higgs boson production via gluon fusion, the uncertainties on the cross sections associated with the analysis categories are estimated by varying the renormalisation and factorisation scales around the central values $\mu_{\mathrm{R}}, \mu_{\mathrm{F}} = \sqrt{m_H^2 + p_{\mathrm{T}}^2}$ in the NLO cross-section calculations of the $H + 1$ jet and $H + 2$ jet production. In the calculations appropriate cuts of $p_{\mathrm{T}}^H > 100$ 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 ±24% in the *Boosted* and ±23% in the *VBF* categories. Whereas the *ggF* contribution is dominant in the *Boosted* category, it only contributes of the order of 20% to the signal in the *VBF* category. Since the two categories are exclusive, their anti-correlation is taken into account following the prescription of Ref. [108].

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 of 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 has been evaluated using the MCFM Monte Carlo program [97], which calculates $H + 3\,jets$ at LO. Scale variations induce changes of the *ggF* contribution in the highest BDT bin of about ±30%. They have been 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 POWHEG+PYTHIA to POWHEG+HERWIG for both *VBF* and *ggF* Higgs production modes. Differences in the signal yields range from ±1% to ±8% for the *VBF* and from ±1% to ±9% for *ggF* production, depending on the channel and category. The BDT score distribution of the POWHEG+PYTHIA and POWHEG+HERWIG samples are compatible with each other within statistical uncer-

tainties.

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* POWHEG sample and a MC@NLO [109] *ggF* sample, both generated with the CT10 PDFs, are reweighted to the MSTW2008NLO [110], NNPDF [111] and the CT10 eigen-tunes parametrisation. The largest variation in acceptance for each category is used as a flat 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 *VBF*, VH and *ggF* production modes due to the PDFs is also considered.

Variations in the acceptance for different MC generators are also included, comparing POWHEG+HERWIG samples to MC@NLO+HERWIG for *ggF* and AMC@NLO+HERWIG [112] for *VBF*. The generator modelling uncertainty is around 2% for *ggF* and 4% for *VBF* productions modes.

Finally the uncertainty on the decay branching ratio, BR($H \to \tau\tau$), of ±5.7% [71] 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 choices of the QCD scales, the PDF parametrisation and the underlying-event model are estimated. The results are reported in Table 7.

Table 7: 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. All systematic uncertainties are 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.

Source	Relative signal and background variations [%]											
	$\tau_{\text{lep}}\tau_{\text{lep}}$ *VBF*		$\tau_{\text{lep}}\tau_{\text{lep}}$ *Boosted*		$\tau_{\text{lep}}\tau_{\text{had}}$ *VBF*		$\tau_{\text{lep}}\tau_{\text{had}}$ *Boosted*		$\tau_{\text{had}}\tau_{\text{had}}$ *VBF*		$\tau_{\text{had}}\tau_{\text{had}}$ *Boosted*	
	S	B	S	B	S	B	S	B	S	B	S	B
Experimental												
Luminosity	±2.8	±0.1	±2.8	±0.1	±2.8	±0.1	±2.8	±0.1	±2.8	±0.1	±2.8	±0.1
Tau trigger*	–	–	–	–	–	–	–	–	$^{+7.7}_{-8.8}$	< 0.1	$^{+7.8}_{-8.9}$	< 0.1
Tau identification	–	–	–	–	±3.3	±1.2	±3.3	±1.8	±6.6	±3.8	±6.6	±5.1
Lepton ident. and trigger*	$^{+1.4}_{-2.1}$	$^{+1.3}_{-1.7}$	$^{+1.4}_{-2.1}$	$^{+1.1}_{-1.5}$	±1.8	±0.5	±1.8	± 0.8	–	–	–	–
b-tagging	±1.3	±1.6	±1.6	±1.6	< 0.1	±0.2	±0.4	±0.2	–	–	–	–
τ energy scale†	–	–	–	–	±2.4	±1.3	±2.4	±0.9	±2.9	±2.5	±2.9	±2.5
Jet energy scale and resolution†	$^{+3.4}_{-7.4}$	$^{+2.8}_{-7.3}$	±1.2	±1.4	$^{+9.5}_{-8.7}$	±1.0	±3.9	±0.4	$^{+10.1}_{-8.0}$	±0.3	$^{+5.1}_{-6.2}$	±0.2
$E_{\text{T}}^{\text{miss}}$ soft scale & resolution	$^{+0.0}_{-0.2}$	$^{+0.0}_{-1.2}$	$^{+0.0}_{-0.1}$	$^{+0.0}_{-1.2}$	$^{+0.8}_{-0.3}$	±0.2	±0.4	< 0.1	±0.5	±0.2	±0.1	< 0.1
Background Model												
Modelling of fake backgrounds*†	–	± 1.2	–	±1.2	–	±2.6	–	±2.6	–	±5.2	–	±0.6
Embedding†	–	$^{+3.8}_{-4.3}$	–	$^{+6.0}_{-6.5}$	–	±1.5	–	±1.2	–	±2.2	–	±3.3
$Z \to \ell\ell$ normalisation*	–	± 2.1	–	±0.7	–	–	–	–	–	–	–	
Theoretical												
Higher-order QCD corrections †	$^{+11.3}_{-9.1}$	±0.2	$^{+19.8}_{-15.3}$	±0.2	$^{+9.7}_{-7.6}$	±0.2	$^{+19.3}_{-14.7}$	±0.2	$^{+10.7}_{-8.2}$	< 0.1	$^{+20.3}_{-15.4}$	< 0.1
UE/PS	± 1.8	< 0.1	± 5.9	< 0.1	±3.8	< 0.1	±2.9	< 0.1	±4.6	< 0.1	±3.8	< 0.1
Generator Modelling	±2.3	< 0.1	±1.2	< 0.1	±2.7	< 0.1	±1.3	< 0.1	±2.4	< 0.1	±1.2	< 0.1
EW corrections	±1.1	< 0.1	±0.4	< 0.1	±1.3	< 0.1	±0.4	< 0.1	±1.1	< 0.1	±0.4	< 0.1
PDF †	$^{+4.5}_{-5.8}$	± 0.3	$^{+6.2}_{-8.0}$	± 0.2	$^{+3.9}_{-3.6}$	± 0.2	$^{+6.6}_{-6.1}$	± 0.2	$^{+4.3}_{-4.0}$	± 0.2	$^{+6.3}_{-5.8}$	± 0.1
BR ($H \to \tau\tau$)	± 5.7	–	±5.7	–	±5.7	–	±5.7	–	± 5.7	–	±5.7	–

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, μ, 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 $\mathcal{L}(\mu, \vec{\theta})$ constructed as the product of Poisson probability terms as an estimator for μ.

The impact of systematic uncertainties on the signal and background expectations is described by nuisance parameters, $\vec{\theta}$, which are parametrised by a Gaussian or log-normal constraint. The expected numbers of signal and background events in each bin are functions of $\vec{\theta}$. The test statistic q_μ is then constructed according to the profile likelihood ratio: $q_\mu = -2\ln[\mathcal{L}(\mu, \hat{\hat{\vec{\theta}}})/\mathcal{L}(\hat{\mu}, \hat{\vec{\theta}})]$, where $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the parameters that maximise the likelihood, and $\hat{\hat{\vec{\theta}}}$ are the nuisance parameter values that maximise the likelihood for a given μ. 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. For an assumed Higgs boson mass, there are six BDTs (for the *VBF* and *Boosted* categories of the three channels). The fit also includes the event yields from the $Z \to \ell\ell$ and top control regions in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, from the top control region of the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, and from the $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}})$ distribution in the *Rest* category of the $\tau_{\text{had}}\tau_{\text{had}}$ channel.

The $Z \to \tau\tau$ background is constrained primarily in the signal regions, due to the difference between the BDT distribution for $Z \to \tau\tau$ compared to the signal. For the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the $Z \to \tau\tau$ and multijet background rates are constrained from 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 $Z \to \ell\ell$ 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 by how much the nuisance parameters are pulled away from their nominal values and by how much their uncertainties are constrained. Furthermore, it is important to understand which systematic uncertainties have 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 (nominal) value shifted up or down by its fitted (nominal) 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 7 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 $Z \to \tau\tau$ and top-quark events, and the tau energy scale. The uncertainties on the jet energy scale are decomposed into several uncorrelated components (among others: η intercalibration of different calorimeter regions, jet energy response, and response to jets of different flavour). In addition, theoretical uncertainties on the branching ratio BR ($H \to \tau\tau$) are found to be important. In general, good agreement is found between the prefit and postfit 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 8, with background normalisations, signal normalisation, and nuisance parameters adjusted by the profile likelihood global fit.

The results on the number of fitted signal and background events, split into the various contributions, are summarised in Tables 8, 9 and 10 for the three channels separately for the datasets collected at 8 TeV

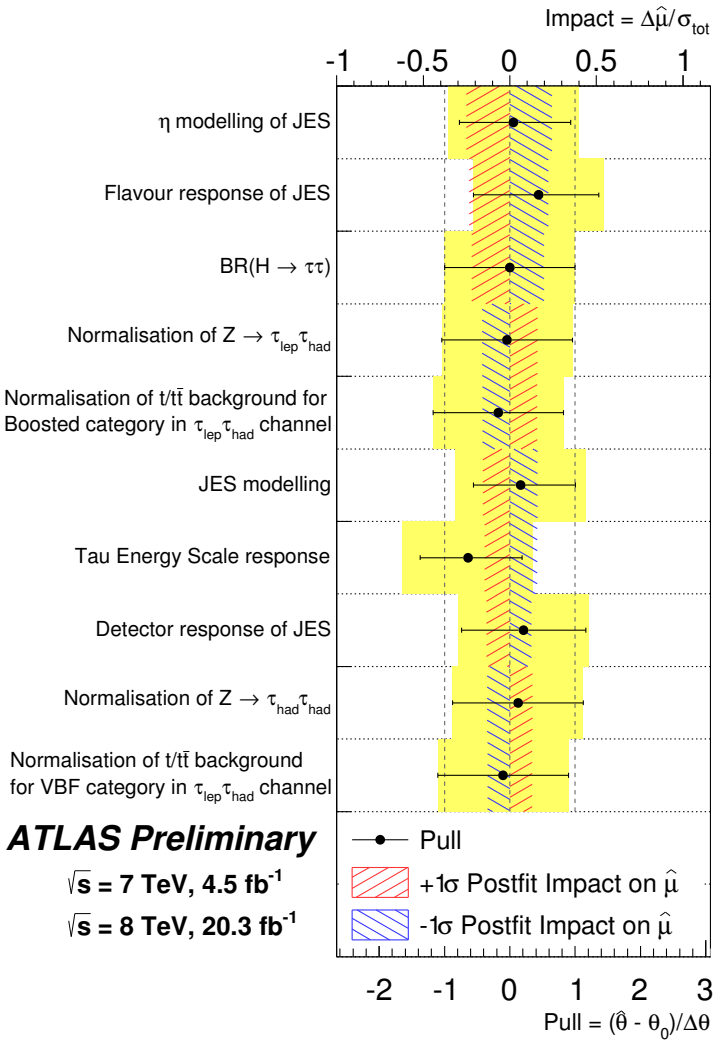


Figure 7: Impact of systematic uncertainties on the fitted signal-strength parameter $\hat{\mu}$ for the combined fit for all channels and both centre-of-mass energies. The systematic uncertainties are listed in decreasing order of their impact on $\hat{\mu}$ on the y-axis. The hatched blue and red boxes show the variations of $\hat{\mu}$ with respect to the total error on μ, σ_{tot}, referring to the top x-axis, when fixing the corresponding individual nuisance parameter θ 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 θ_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.

centre-of-mass energy. In addition to the total number of events, the expected numbers of events in the two highest BDT output bins are given. The numbers of events observed in the data are 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.

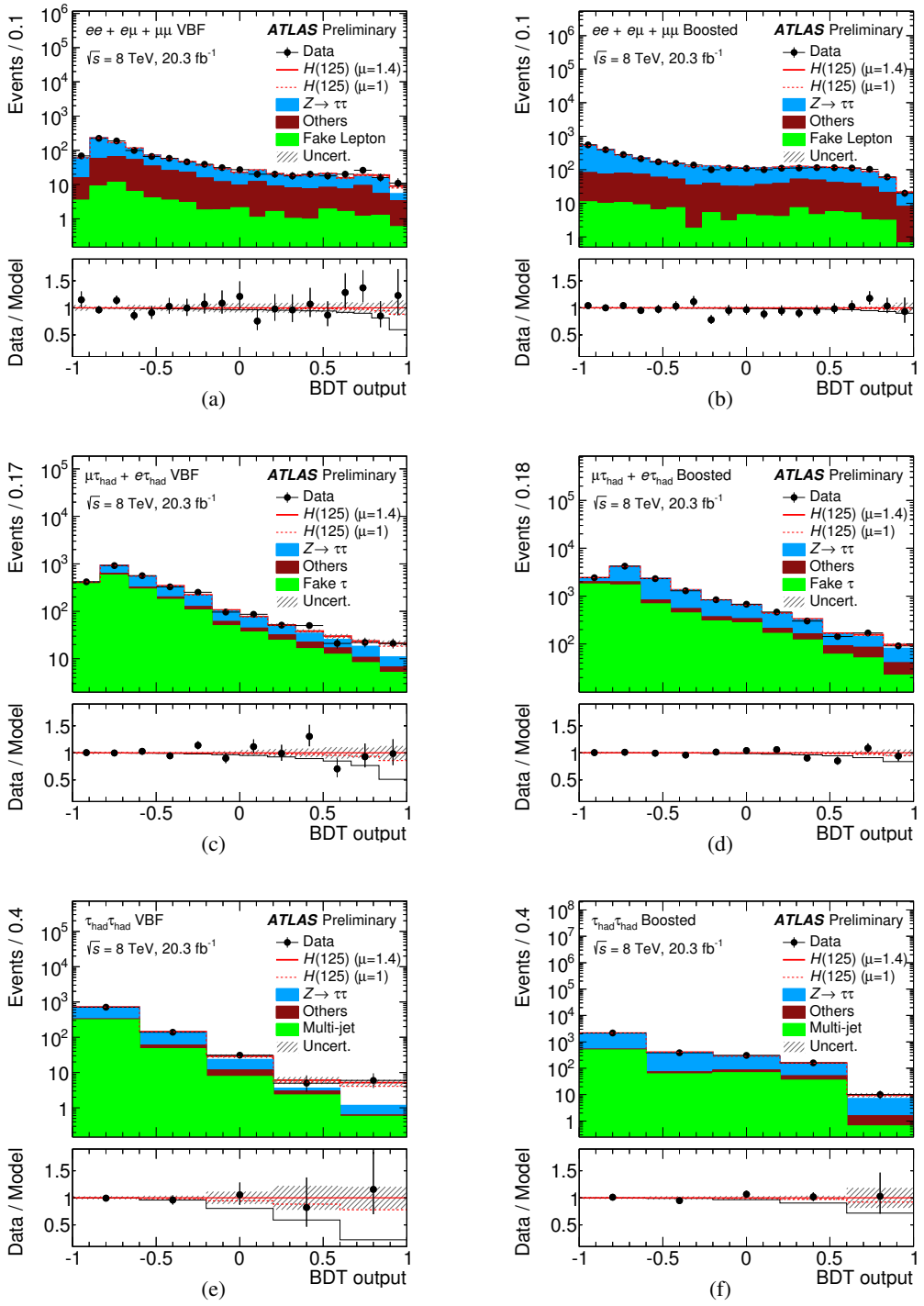


Figure 8: Distributions of the BDT discriminants for the data taken at $\sqrt{s} = 8$ TeV in the signal regions of the *VBF* (left) and *Boosted* (right) categories for the $\tau_{\text{lep}}\tau_{\text{lep}}$ (top), $\tau_{\text{lep}}\tau_{\text{had}}$ (middle), and $\tau_{\text{had}}\tau_{\text{had}}$ (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 between the data and 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 represents the changes in the model when $\mu = 1.0$ or background only are assumed respectively.

Table 8: 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 uncertainties reflect the preferred values from the global fit. The uncertainties on the total background and total signal reflect the full statistical and systematic uncertainty, while the uncertainties on the individual background components reflect the full systematic uncertainty only.

Process/Category	*VBF*			*Boosted*		
BDT output bin	all bins	second last bin	last bin	all bins	second last bin	last bin
$Z \to \tau\tau$	594 ± 26	9.8 ± 1.0	2.0 ± 0.3	2210 ± 80	34.0 ± 2.3	11.3 ± 1.3
Fake backgrounds	58 ± 12	1.3 ± 0.6	0.6 ± 0.4	110 ± 40	3.1 ± 1.4	0.7 ± 0.4
Top	132 ± 18	0.8 ± 0.1	0.9 ± 0.2	380 ± 50	9.6 ± 1.3	3.8 ± 0.6
Others	187 ± 28	2.9 ± 0.5	1.3 ± 0.3	379 ± 32	7.7 ± 0.7	3.2 ± 0.4
$ggF: H \to WW$ ($m_H = 125$ GeV)	2.9 ± 0.7	0.1 ± 0.1	0.1 ± 0.1	7.7 ± 2.2	0.5 ± 0.1	0.2 ± 0.1
$VBF: H \to WW$	3.3 ± 0.3	0.4 ± 0.1	0.4 ± 0.1	1.6 ± 0.2	< 0.1	< 0.1
$WH: H \to WW$	0.10 ± 0.01	< 0.1	< 0.1	0.9 ± 0.1	< 0.1	< 0.1
$ZH: H \to WW$	< 0.1	< 0.1	< 0.1	0.6 ± 0.1	< 0.1	< 0.1
$ggF: H \to \tau\tau$ ($m_H = 125$ GeV)	9.9 ± 3.5	0.74 ± 0.26	0.35 ± 0.14	21 ± 8	2.4 ± 0.9	1.3 ± 0.5
$VBF: H \to \tau\tau$	13 ± 4	2.7 ± 0.7	3.3 ± 0.9	5.4 ± 1.5	0.9 ± 0.3	0.5 ± 0.2
$WH: H \to \tau\tau$	0.24 ± 0.07	< 0.1	< 0.1	3.8 ± 1.1	0.4 ± 0.1	0.2 ± 0.1
$ZH: H \to \tau\tau$	0.13 ± 0.04	< 0.1	< 0.1	1.9 ± 0.5	0.20 ± 0.06	0.11 ± 0.03
Total signal	23 ± 6	3.5 ± 0.9	3.6 ± 1.0	32 ± 10	4.0 ± 1.2	2.1 ± 0.7
Total background	977 ± 26	15.2 ± 1.7	5.4 ± 1.2	3080 ± 50	54.9 ± 3.3	19.4 ± 2.1
Data	1014	16	11	3095	61	20

Table 9: The predicted post-fit event yields in the $\tau_{\text{lep}}\tau_{\text{had}}$ 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 uncertainties reflect the preferred values from the global fit. The uncertainties on the total background and total signal reflect the full statistical and systematic uncertainty, while the uncertainties on the individual background components reflect the full systematic uncertainty only.

Process/Category	*VBF*			*Boosted*		
BDT output bin	all bins	second last bin	last bin	all bins	second last bin	last bin
Fake backgrounds	1680 ± 50	8.2 ± 0.9	5.2 ± 0.7	5630 ± 160	50.8 ± 2.5	22.2 ± 1.8
$Z \to \tau\tau$	878 ± 29	7.6 ± 0.9	4.3 ± 0.7	6220 ± 170	57.6 ± 2.7	41.1 ± 3.2
Top	85 ± 16	0.1 ± 0.4	0.5 ± 0.4	380 ± 50	11 ± 4	4.6 ± 1.6
$Z \to \ell\ell(\ell \to \tau_{\text{had}})$	46 ± 21	1.2 ± 0.8	0.4 ± 0.3	200 ± 50	13 ± 5	8 ± 4
diboson	66 ± 12	1.1 ± 0.4	0.5 ± 0.2	430 ± 50	9.8 ± 2.5	5.2 ± 1.8
$ggF: H \to \tau\tau$ ($m_H = 125$ GeV)	17 ± 6	1.0 ± 0.4	1.2 ± 0.6	60 ± 20	9.1 ± 3.1	10.1 ± 3.4
$VBF: H \to \tau\tau$	31 ± 8	4.5 ± 1.1	9.2 ± 2.3	15 ± 4	2.5 ± 0.7	2.8 ± 0.7
$WH: H \to \tau\tau$	0.5 ± 0.4	< 0.1	< 0.1	9.1 ± 2.4	1.4 ± 0.4	2.0 ± 0.5
$ZH: H \to \tau\tau$	0.2 ± 0.1	< 0.1	< 0.1	4.5 ± 1.2	0.8 ± 0.2	0.9 ± 0.2
Total signal	48 ± 12	5.5 ± 1.3	10.4 ± 2.5	89 ± 26	14 ± 4	16 ± 4
Total background	2760 ± 40	18.2 ± 2.3	10.9 ± 2.8	12870 ± 110	143 ± 7	82 ± 6
Data	2830	22	21	12952	170	92

Table 10: 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 uncertainties reflect the preferred values from the global fit. The uncertainties on the total background and total signal reflect the full statistical and systematic uncertainty, while the uncertainties on the individual background components reflect the full systematic uncertainty only.

Process/Category	*VBF*			*Boosted*		
BDT output bin	all bins	second last bin	last bin	all bins	second last bin	last bin
Fake backgrounds	370 ± 18	2.3 ± 0.9	0.6 ± 0.3	644 ± 26	35 ± 4	0.6 ± 0.3
Others	37 ± 5	0.7 ± 0.2	< 0.1	90 ± 12	16.1 ± 2.1	0.9 ± 0.2
$Z \to \tau\tau$	475 ± 16	0.6 ± 0.7	0.6 ± 0.4	2230 ± 70	92 ± 4	5.4 ± 1.6
$ggF: H \to \tau\tau$ ($m_H = 125$ GeV)	8.1 ± 2.7	0.7 ± 0.2	0.6 ± 0.2	21 ± 8	9.0 ± 3.3	1.6 ± 0.6
$VBF: H \to \tau\tau$	12.0 ± 3.2	1.8 ± 0.5	3.4 ± 0.9	6.3 ± 1.7	2.9 ± 0.7	0.5 ± 0.1
$WH: H \to \tau\tau$	0.3 ± 0.1	< 0.1	< 0.1	4.0 ± 1.1	1.9 ± 0.5	0.4 ± 0.1
$ZH: H \to \tau\tau$	0.2 ± 0.1	< 0.1	< 0.1	2.4 ± 0.6	1.1 ± 0.3	0.2 ± 0.1
Total signal	21 ± 5	2.5 ± 0.6	4.0 ± 1.0	34 ± 10	15 ± 4	2.7 ± 0.8
Total backgrounds	882 ± 18	3.6 ± 1.3	1.2 ± 1.1	2960 ± 50	143 ± 6	7.0 ± 1.8
Data	892	5	6	3020	161	10

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.

The overall combination of all measurements results is $\mu = 1.40\ ^{+0.43}_{-0.37}$ for $m_H = 125$ GeV. At the value of the Higgs boson mass obtained from the combination of the ATLAS $H \to \gamma\gamma$ and $H \to ZZ^*$ measurements [113], $m_H = 125.36$ GeV, the obtained signal strength for the $H \to \tau\tau$ analysis is:

$$\mu = 1.42\ ^{+0.27}_{-0.26}(\text{stat.})\ ^{+0.32}_{-0.24}(\text{syst.})\ \pm 0.10(\text{theory syst.}).$$

The systematic uncertainties are split into two groups, theoretical uncertainties on the inclusive Higgs boson production cross section and $H \to \tau\tau$ branching ratio, and all other systematic uncertainties including all experimental effects as well as theoretical uncertainties on the signal region acceptance from the QCD scale and PDF choice. The results for each individual channel and for each category as well as for their combination are shown in Figure 9. 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 $q_{\mu=0} = -2\ln(\mathcal{L}(0,\hat{\hat{\vec{\theta}}})/\mathcal{L}(\hat{\mu},\hat{\vec{\theta}}))$ in the asymptotic approximation [114]. For $m_H = 125.36$ GeV, the observed p_0 value is 3.0×10^{-6}, which corresponds to a deviation from the background-only hypothesis of 4.5σ. This can be compared to an expected significance of 3.5σ. This provides evidence at the level of 4.5σ for the decay of the 125 GeV Higgs boson into tau leptons.

In Figure 10 the expected and observed event numbers are shown, 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 expectation is shown for signal yields for both $\mu = 1$ and the best-fit value $\mu = 1.4$ for $m_H = 125$ GeV on top of the background prediction taken also from the best-fit values. The background expectation where the signal strength parameter has been fixed to $\mu = 0$ is also shown for comparison.

In order 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 11. The events are weighted by a factor of $\ln(1 + S/B)$. The excess of events in these mass distributions 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.

As discussed in Section 8, the dominant uncertainties on the measurement of the signal strength parameters include statistical uncertainties in 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. In Table 11 the contributions of each of these significant sources to the uncertainty of the measured signal strength are summarised.

It should be noted that the $Z \to \tau\tau$ embedded sample normalisation uncertainties are correlated for 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

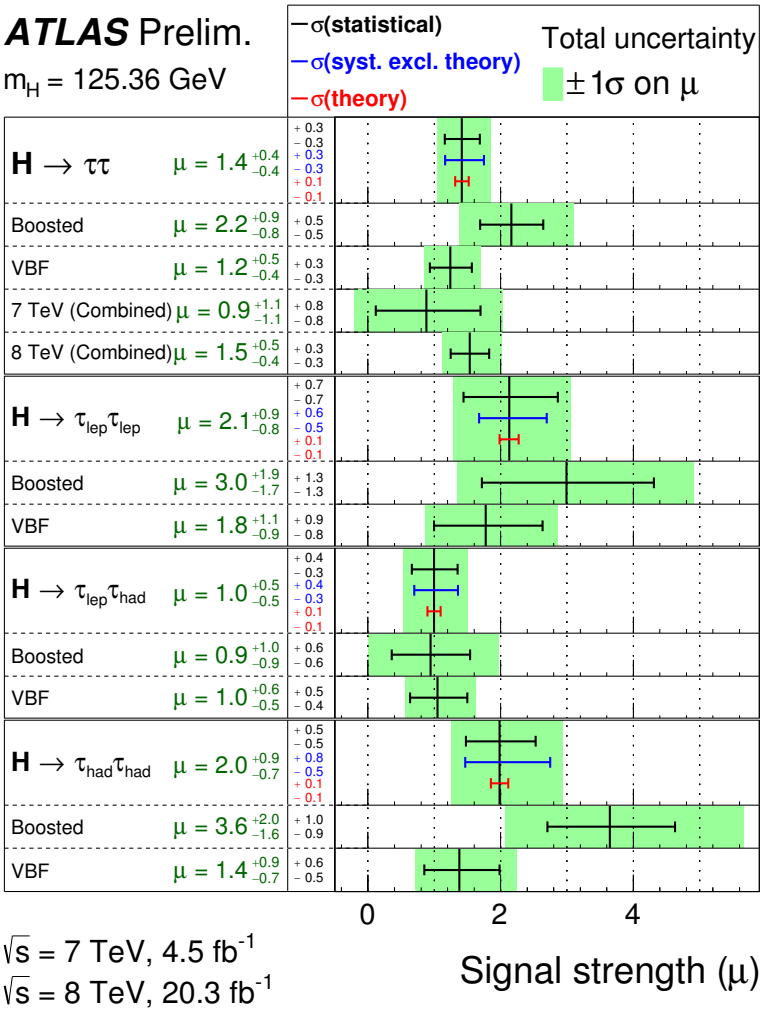


Figure 9: The best-fit value for the signal strength μ in the individual channels and their combination for the full ATLAS datasets at $\sqrt{s}$ = 7 TeVand $\sqrt{s}$ = 8 TeV. The total $\pm 1\sigma$ uncertainty is indicated by the shaded green band, with the individual contributions from the statistical uncertainty (top, black), the total (experimental and theoretical) 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.

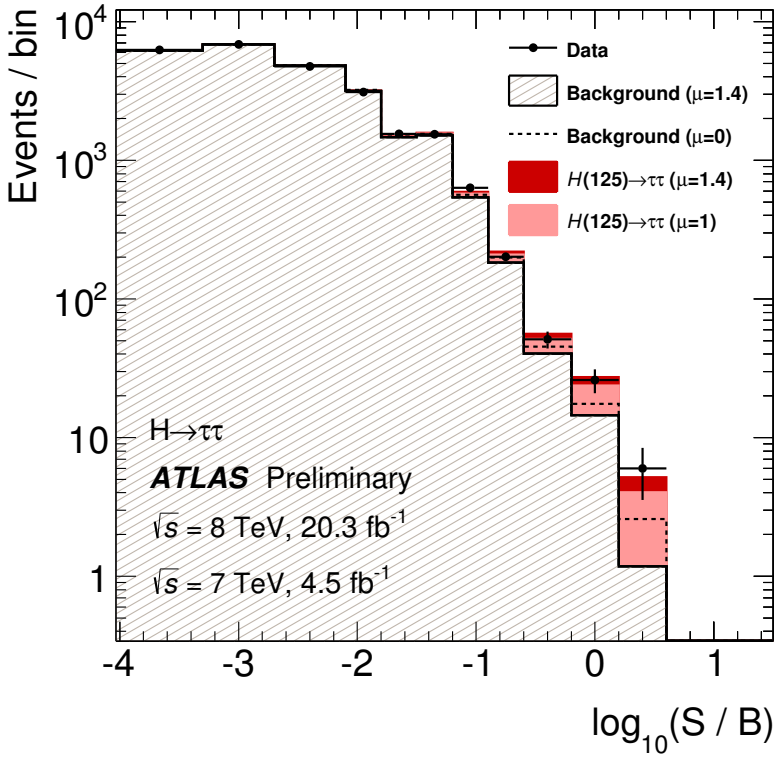


Figure 10: Event yields as a function of $\log_{10}(S/B)$, where S (signal yield) and B (background yield) are taken from the BDT output bin of each event, assuming $\mu = 1.4$. Events in all categories are included. The predicted background is obtained from the global fit (with $\mu = 1.4$) and signal 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, but fixing $\mu = 0$.

Source of Uncertainty	Uncertainty on μ
Signal region statistics (data)	$^{+0.27}_{-0.26}$
Jet energy scale	± 0.16
Tau energy scale	± 0.07
Tau identification	± 0.06
Background normalisation	± 0.12
Background estimate stat.	± 0.10
BR ($H \to \tau\tau$)	± 0.08
Parton shower/Underlying event	± 0.04
PDF	± 0.03

Table 11: Important sources of uncertainty on the measured signal strength parameter μ. The contributions are given as absolute uncertainties on the best-fit value of $\mu = 1.42$. Various sub-components have been combined.

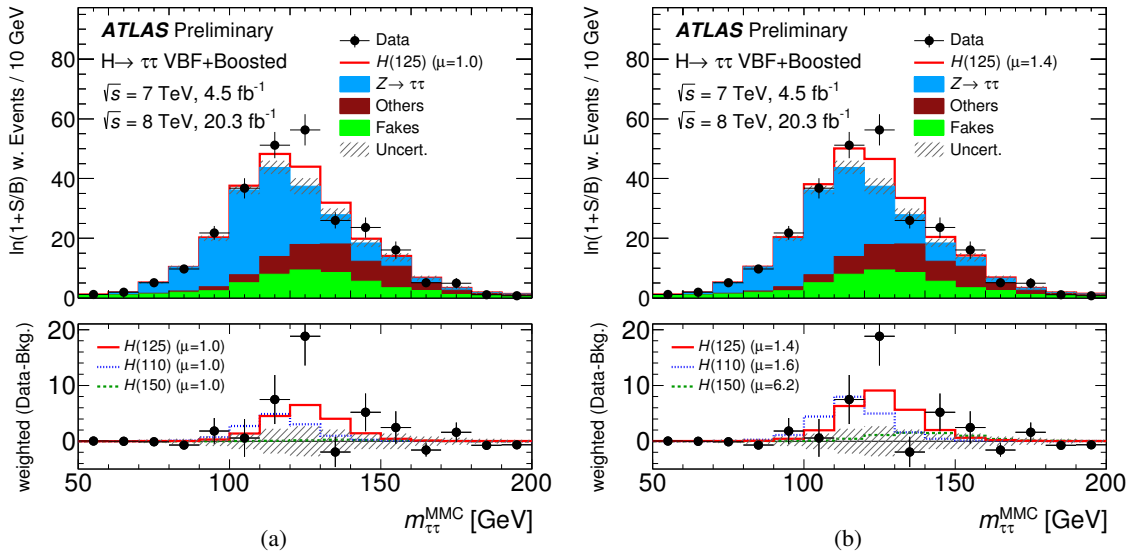


Figure 11: Distributions of $m_{\tau\tau}^{\mathrm{MMC}}$ where events are weighted by $\ln(1 + S/B)$ for all channels. These weights are determined by the signal (S) and background (B) predictions for each BDT bin. The bottom panel in each plot shows the difference between weighted data events and weighted 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 ($\mu = 1.4$). 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 μ 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.

fitted separating vector-boson-mediated processes, VBF and VH, from gluon-mediated processes, ggF. Two signal strength parameters, $\mu^{\tau\tau}_{ggF}$ and $\mu^{\tau\tau}_{VBF+VH}$, which scale the SM-predicted rates to those observed, are introduced. The two-dimensional likelihood contours in the plane of $\mu^{\tau\tau}_{ggF}$ and $\mu^{\tau\tau}_{VBF+VH}$ [115] are shown in Figure 12 for $m_H = 125.36$ GeV. The best-fit values are

$$\mu^{\tau\tau}_{ggF} = 1.93\ ^{+0.78}_{-0.77}(\text{stat.})\ ^{+1.19}_{-0.80}(\text{syst.})\ \pm 0.29(\text{theory syst.})$$

and

$$\mu^{\tau\tau}_{VBF+VH} = 1.24^{+0.48}_{-0.45}(\text{stat.})\ ^{+0.31}_{-0.28}(\text{syst.})\ \pm 0.08(\text{theory syst.}),$$

in agreement with the predictions from the Standard Model.

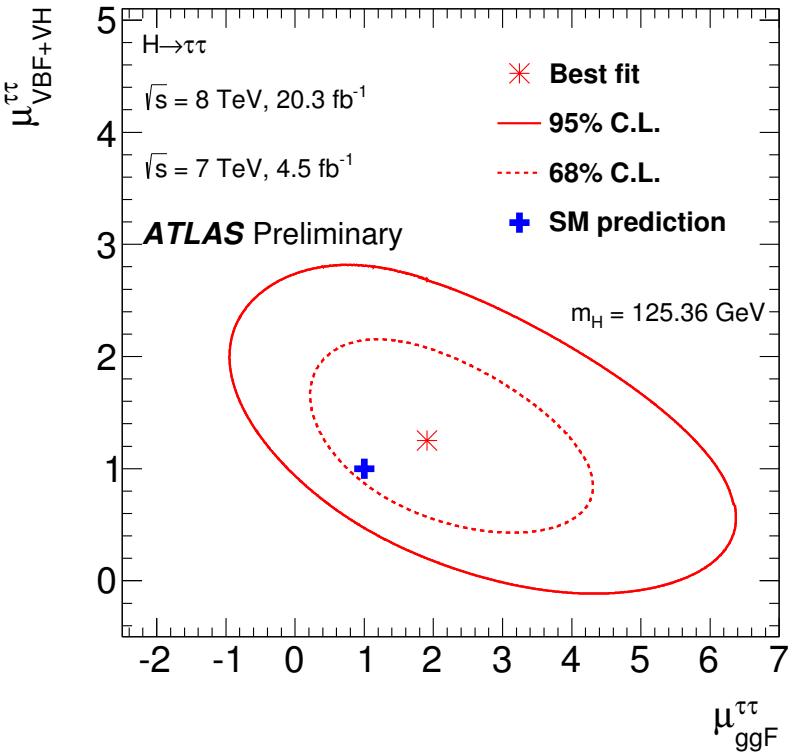


Figure 12: Likelihood contours for the combination of all channels in the ($\mu^{\tau\tau}_{ggF}$, $\mu^{\tau\tau}_{VBF+VH}$) plane. The 68% and 95% CL contours are shown as dashed and solid lines respectively, for $m_H = 125.36$ GeV. The SM expectation is shown by a filled plus symbol. The best fit to the data is shown for the case when both the $\mu^{\tau\tau}_{ggF}$ and $\mu^{\tau\tau}_{VBF+VH}$ are unconstrained.

10 Cut-based analysis

The search for the SM Higgs boson presented above has been cross-checked for the dataset collected at $\sqrt{s} = 8$ TeV in an analysis where cuts on kinematic variables were 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 [19]. 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 $Z \to \tau\tau$ 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 also 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 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 similar categorisation is performed as in the multivariate analysis, i.e. *VBF* and *Boosted* categories are defined. In order to increase the separation power, sub-categories are introduced for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels. These sub-categories 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 sub-categories have a small number of events, but a high signal-to-background ratio. Although the combined sensitivity is dominated by the few highly sensitive sub-categories, the others are important not just to increase the sensitivity but also to constrain the various background components.

An overview on the defined categories in the three channels is given in Table 12. 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 the *VBF* production. In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel tight and loose *VBF* sub-categories are defined, via cuts on the mass of the di-jet system, m_{jj}, and p_{T}^{H}, the transverse momentum of the Higgs boson candidate (see Table 12). 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 τ_{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 sub-category with the highest purity is the *VBF High-*p_{T}^{H} sub-category, where tight cuts on p_{T}^{H} and $\Delta R(\tau, \tau)$ reject almost all non-resonant background sources. The other two *VBF*-like sub-categories are distinguished by a different signal-to-background ratio due to a tighter selection applied on the forward jets. For the $\tau_{\text{had}}\tau_{\text{had}}$ channel, *Boosted* sub-categories are also defined. The division is based on the same cuts on p_{T}^{H} and $\Delta R(\tau, \tau)$ 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 sub-categories 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 results on the number of fitted signal and background events are summarised in Table 13 for the three channels.

The final results are derived from the combined fit of the $m_{\tau\tau}$ distributions observed in the various

sub-categories. The combined mass distribution for the three channels is shown in Figure 13. A excess in the mass region around 125 GeV is visible.

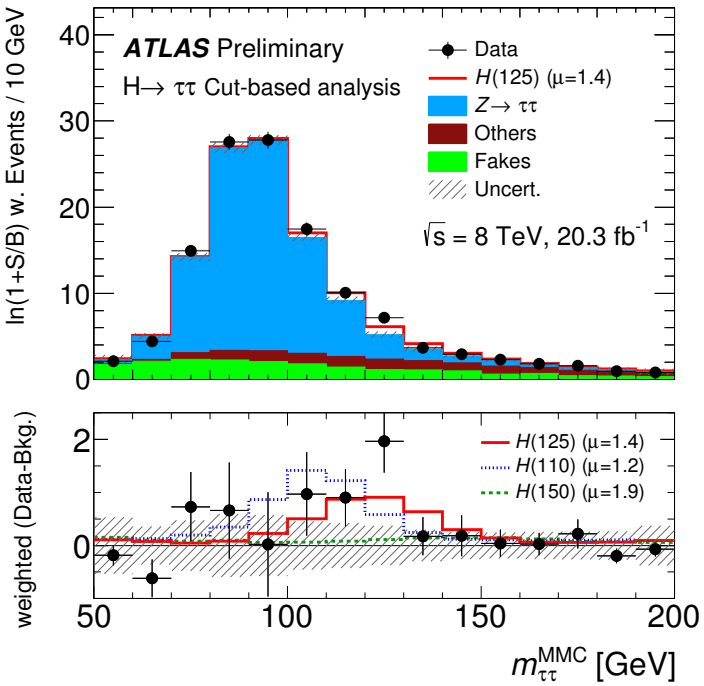


Figure 13: Distribution of the reconstructed $m_{\tau\tau}^{\mathrm{MMC}}$ for the sum of all channels in the cut-based analysis for the data taken at $\sqrt{s} = 8$ TeV. The events are weighted by a factor of $\ln(1 + S/B)$ based on the signal (S) and background (B) yields in each category. The bottom 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 ($\mu = 1.4$). 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 μ values are given in the figure.

The results obtained for the signal strengths extracted for the three analysis channels as well as for their combination for the data taken at $\sqrt{s} = 8$ TeV are given in table 14. Also for this cut-based analysis evidence for $H \to \tau\tau$ decays is obtained and, at $m_H = 125.36$ GeV, a combined signal strength of

$$\mu_{\mathrm{Cut-Based}} = 1.37\ ^{+0.57}_{-0.48}$$

is measured. For comparison, the results obtained in the multivariate analysis for the dataset at $\sqrt{s} = 8$ TeV are also included in Table 14. Good agreement between the results of the two analyses is found for the individual channels as well as for their combination. To quantify further the level of agreement, the correlation ρ and the uncertainties on the difference between the μ values obtained, i.e. $\Delta\mu \pm \delta(\Delta\mu)$, was evaluated using the so-called jackknife technique [116, 117]. Using this method, the correlation between the μ values obtained in the two analyses is found to be in the range between 0.55 and 0.65 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 one 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 in Figure 14 for the cut-based analysis as a function of the mass 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 3.2σ. The expected significance for the cut-based analysis is superimposed on the figure and reaches a significance of 2.5σ at m_H = 125.36 GeV. The corresponding significance values for the multivariate analysis for the dataset at $\sqrt{s}$ = 8 TeV are found to be 4.5σ (observed) and 3.3σ (expected). They are also indicated in the figure.

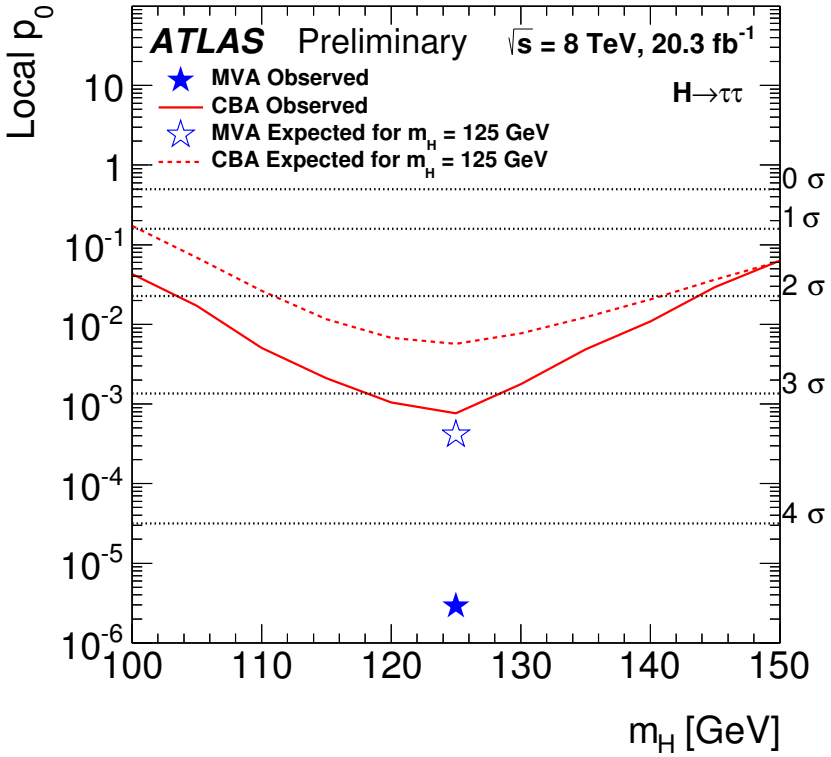


Figure 14: Observed (solid red) and expected (dashed red) p_0 values as a function of m_H for the combination of all channels in the cut-based analysis (CBA) for the data taken at $\sqrt{s}$ = 8 TeV. The expected p_0 values are given for the background-only hypothesis. The corresponding observed and 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.

Channel	*VBF* category selection cuts		
$\tau_{\text{lep}}\tau_{\text{lep}}$	At least two jets with $p_{\text{T}}(j_1) > 40$ GeV and $p_{\text{T}}(j_2) > 30$ GeV $\lvert\Delta\eta_{j_1,j_2}\rvert > 3.0$ $m_{j_1,j_2} > 400$ GeV b jet veto for jets with $p_{\text{T}} > 25$ GeV Jet veto: no additional jet with $p_{\text{T}} > 25$ GeV within $\lvert\eta\rvert < 2.4$		
$\tau_{\text{lep}}\tau_{\text{had}}$	At least two jets with $p_{\text{T}}(j_1) > 40$ GeV and $p_{\text{T}}(j_2) > 30$ GeV $\text{E}_{\text{T}}^{\text{miss}} > 20$ GeV $\lvert\Delta\eta_{j_1,j_2}\rvert > 3.0$ and $\eta(j_1)\cdot\eta(j_2) < 0$, $m_{j_1,j_2} > 300$ GeV $p_{\text{T}}^{\text{total}} = \lvert\vec{p}_{\text{T}}^{\,\ell} + \vec{p}_{\text{T}}^{\,\tau_{\text{had}}} + \vec{p}_{\text{T}}^{\,j_1} + \vec{p}_{\text{T}}^{\,j_2} + \vec{\text{E}}_{T}^{\text{miss}}\rvert < 30$ GeV b jet veto for jets with $p_{\text{T}} > 30$ GeV $\min(\eta_{(j_1)},\eta_{(j_2)}) < \eta_{(\ell)},\eta_{(\tau_{\text{had}})} < \max(\eta_{(j_1)},\eta_{(j_2)})$		
	***VBF* tight** $m_{j_1,j_2} > 500$ GeV $p_{\text{T}}^{H} > 100$ GeV $p_{\text{T}}(\tau_{\text{had}}) > 30$ GeV $m_{\text{vis}} > 40$ GeV	***VBF* loose** Non tight *VBF*	
$\tau_{\text{had}}\tau_{\text{had}}$	At least two jets with $p_{\text{T}}(j_1) > 50$ GeV and $p_{\text{T}}(j_2) > 30$ GeV $\lvert\Delta\eta(\tau_1,\tau_2)\rvert < 1.5$ $\lvert\Delta\eta_{j_1,j_2}\rvert > 2.6$ and $m_{j_1,j_2} > 250$ GeV $\min(\eta_{(j_1)},\eta_{(j_2)}) < \eta_{(\tau_1)},\eta_{(\tau_2)} < \max(\eta_{(j_1)},\eta_{(j_2)})$		
	***VBF* high p_{T}^{H}** $\Delta R(\tau_1,\tau_2) < 1.5$ and $p_{\text{T}}^{\text{H}} > 140$ GeV	***VBF* low p_{T}^{H}, tight** $\Delta R(\tau_1,\tau_2) > 1.5$ or $p_{\text{T}}^{H} < 140$ GeV $m_{j_1,j_2}[\text{GeV}] > -250\lvert\Delta\eta_{\text{j}_1,\text{j}_2}\rvert + 1550$	***VBF* low p_{T}^{H}, loose** $\Delta R(\tau_1,\tau_2) > 1.5$ or $p_{\text{T}}^{H} < 140$ GeV $m_{j_1,j_2}[\text{GeV}] < -250\lvert\Delta\eta_{\text{j}_1,\text{j}_2}\rvert + 1550$

Channel	*Boosted* category selection cuts	
$\tau_{\text{lep}}\tau_{\text{lep}}$	Exclude events passing the *VBF* cuts $p_{\text{T}}^{H} > 100$ GeV b jet veto for jets with $p_{\text{T}} > 25$ GeV	
$\tau_{\text{lep}}\tau_{\text{had}}$	Failing the *VBF* selection $\text{E}_{\text{T}}^{\text{miss}} > 20$ GeV $p_{\text{T}}^{H} > 100$ GeV $p_{\text{T}}(\tau_{\text{had}}) > 30$ GeV b jet veto for jets with $p_{\text{T}} > 30$ GeV	
$\tau_{\text{had}}\tau_{\text{had}}$	Failing the *VBF* selection $\Delta\eta(\tau_1,\tau_2) < 1.5$ $p_{\text{T}}^{H} > 100$ GeV	
	***Boosted* high p_{T}^{H}** $\Delta R(\tau_1,\tau_2) < 1.5$ and $p_{\text{T}}^{H} > 140$ GeV	***Boosted* low p_{T}^{H}** $\Delta R(\tau_1,\tau_2) > 1.5$ or $p_{\text{T}}^{H} < 140$ GeV

Table 12: Summary of the selection cuts used to define the *VBF* and *Boosted* (sub)-categories 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, τ_{had}, jets). The variables are defined in the text.

$\tau_{\text{lep}}\tau_{\text{lep}}$	*VBF*		*Boosted*	
Total signal	11 ± 4		36 ± 13	
Total background	130 ± 7		3402 ± 58	
Data	152		3428	
$\tau_{\text{lep}}\tau_{\text{had}}$	*Tight VBF*	*Loose VBF*	*Boosted*	
Signal	8.5 ± 3	17 ± 6	50 ± 17	
Background	52 ± 4	398 ± 17	4400 ± 74	
Data	62	407	4435	
$\tau_{\text{had}}\tau_{\text{had}}$	*VBF* high p_{T}^{H}	*VBF* low p_{T}^{H}		*Boosted*	
		tight	loose	high p_{T}^{H}	low p_{T}^{H}
Signal	5.5 ± 2.0	5.1 ± 1.9	3.7 ± 1.3	16 ± 6	19 ± 7
Background	60 ± 5	86 ± 5	156 ± 7	1155 ± 28	2131 ± 41
Data	65	94	157	1204	2121

Table 13: The measured signal and background yields of the cut-based analysis at $\sqrt{s}$ = 8 TeV in the $\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels for m_H = 125 GeV. The normalisations and uncertainties are taken from the global fit. The uncertainties on the predicted yields reflect the full statistical and systematic uncertainties.

	Fitted μ values		
	$\sqrt{s}$	Multivariate analysis	Cut-based analysis
$\tau_{\text{lep}}\tau_{\text{lep}}$	8 TeV	$2.09^{+0.98}_{-0.88}$	$3.03^{+1.34}_{-1.25}$
$\tau_{\text{lep}}\tau_{\text{had}}$	8 TeV	$1.14^{+0.56}_{-0.48}$	$0.69^{+0.67}_{-0.57}$
$\tau_{\text{had}}\tau_{\text{had}}$	8 TeV	$1.77^{+0.93}_{-0.71}$	$1.64^{+0.90}_{-0.74}$
All channels	8 TeV	$1.53^{+0.49}_{-0.41}$	$1.37^{+0.57}_{-0.48}$

Table 14: 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.

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 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 (3.5) standard deviations. This excess is consistent with resulting from $H \to \tau\tau$ decays with $m_H = 125$ GeV. The measured signal strength, normalised to the Standard Model expectation for a Higgs boson of $m_H = 125.36$ GeV is

$$\mu = 1.42\ ^{+0.27}_{-0.26}(\text{stat.})\ ^{+0.32}_{-0.24}(\text{syst.})\ \pm 0.10(\text{theory syst.}).$$

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

The results of the multivariate analysis have been 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.

References

[1] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, *Phys. Rev. Lett.* **13** (1964) 321.

[2] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132.

[3] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508.

[4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, *Phys. Rev. Lett.* **13** (1964) 585.

[5] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, *Phys. Rev.* **145** (1966) 1156.

[6] T. W. B. Kibble, *Symmetry breaking in non-Abelian gauge theories*, *Phys. Rev.* **155** (1967) 1554.

[7] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.

[8] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B 716** (2012) 1, [arXiv:1207.7214].

[9] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B 716** (2012) 30, [arXiv:1207.7235].

[10] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, *Phys. Lett.* **B 726** (2013) 88, [arXiv:1307.1427].

[11] CMS Collaboration, *Observation of a new boson with mass near 125 GeV in pp collisions at* $\sqrt{s}$ *= 7 and 8 TeV*, *JHEP* **1306** (2013) 081, [arXiv:1303.4571].

[12] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, *Phys. Lett.* **B 726** (2013) 120, [arXiv:1307.1432].

[13] CMS Collaboration, *Study of the mass and spin-parity of the Higgs boson candidate via its decays to Z boson pairs*, *Phys. Rev. Lett.* **110** (2013) 081803, [arXiv:1212.6639].

[14] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264.

[15] CMS Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of* τ *leptons*, *JHEP* **1405** (2014) 104, [arXiv:1401.5041].

[16] CMS Collaboration, *Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks*, *Phys. Rev.* **D 89** (2014) 012003, [arXiv:1310.3687].

[17] CMS Collaboration, *Evidence for the direct decay of the 125 GeV Higgs boson to fermions*, *Nature Phys.* **10** (2014) 557, [arXiv:1401.6527].

[18] ATLAS Collaboration, *Search for the* $b\bar{b}$ *decay of the Standard Model Higgs boson in associated* $(W/Z)H$ *production with the ATLAS detector*, (2014), [arXiv:1409.6212].

[19] 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* **1209** (2012) 070, [arXiv:1206.5971].

[20] S. Alioli et al., *NLO Higgs boson production via gluon fusion matched with shower in* POWHEG, *JHEP* **0904** (2009) 002, [arXiv:0812.0578].

[21] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in* POWHEG, *JHEP* **1002** (2010) 037, [arXiv:0911.5299].

[22] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Chapman & Hall, New York, 1984.

[23] J. Friedman, *Stochastic gradient boosting*, *Comput. Stat. Data Anal.* **38** (2002) 367.

[24] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, *J. Comput. Syst. Sci.* **55** (1997) 119.

[25] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[26] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, *Eur. Phys. J.* **C 72** (2012) 1849, [arXiv:1110.1530].

[27] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, *Eur. Phys. J.* **C 74** (2014) 2941, [arXiv:1404.2240].

[28] ATLAS Collaboration, *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data*, ATLAS-CONF-2013-088 (2013). http://cds.cern.ch/record/1580207.

[29] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, (2014), [arXiv:1407.3935].

[30] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [arXiv:0802.1189].

[31] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett.* **B 641** (2006) 57, [hep-ph/0512210].

[32] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002 (2008). http://cds.cern.ch/record/1099735.

[33] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, *Phys. Lett.* **B 659** (2008) 119, [arXiv:0707.1378].

[34] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J.* **C 73** (2013) 2304, [arXiv:1112.6426].

[35] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, (2014), [arXiv:1406.0076].

[36] ATLAS Collaboration, *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC*, *Eur. Phys. J.* **C 73** (2013) 2305, [arXiv:1203.1302].

[37] ATLAS Collaboration, *Commissioning of high performance b-tagging algorithms with the ATLAS detector*, ATLAS-CONF-2011-102 (2011). http://cds.cern.ch/record/1369219.

[38] ATLAS Collaboration, *Measurement of the b-tag efficiency in a sample of jets containing muons with 5 fb^{-1} of data from the ATLAS detector*, ATLAS-CONF-2012-043 (2012). http://cds.cern.ch/record/1435197/.

[39] ATLAS Collaboration, *Search for supersymmetry in final states with jets, missing transverse momentum and a Z boson at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, ATLAS-CONF-2012-152 (2012). http://cds.cern.ch/record/1493491.

[40] ATLAS Collaboration, *Identification of Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, ATLAS-CONF-2013-064 (2013). http://cds.cern.ch/record/1562839.

[41] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s}$ = 7 TeV with ATLAS*, *Eur. Phys. J.* **C 72** (2012) 1, [arXiv:1108.5602].

[42] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at* 8 *TeV*, ATLAS-CONF-2013-082 (2013). http://cds.cern.ch/record/1570993.

[43] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146].

[44] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the* Powheg *method*, *JHEP* **0711** (2007) 070, [arXiv:0709.2092].

[45] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the* Powheg *BOX*, *JHEP* **1006** (2010) 043, [arXiv:1002.2581].

[46] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*, *JHEP* **1202** (2012) 088, [arXiv:1111.2854].

[47] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A brief introduction to* Pythia *8.1*, *Comput. Phys. Commun.* **178** (2008) 852, [arXiv:0710.3820].

[48] H.-L. Lai et al., *New generation of parton distributions with uncertainties from global QCD analysis*, *Phys. Rev.* **D 82** (2010) 074024, [arXiv:1007.2241].

[49] A. Djouadi, M. Spira, and P. Zerwas, *Production of Higgs bosons in proton colliders: QCD corrections*, *Phys. Lett.* **B 264** (1991) 440.

[50] S. Dawson, *Radiative corrections to Higgs boson production*, *Nucl. Phys.* **B 359** (1991) 283.

[51] M. Spira et al., *Higgs boson production at the LHC*, *Nucl. Phys.* **B 453** (1995) 17, [hep-ph/9504378].

[52] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, *Phys. Rev. Lett.* **88** (2002) 201801, [hep-ph/0201206].

[53] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, *Nucl. Phys.* **B 646** (2002) 220, [hep-ph/0207004].

[54] V. Ravindran, J. Smith, and W. L. van Neerven, *NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions*, *Nucl. Phys.* **B 665** (2003) 325, [hep-ph/0302135].

[55] S. Catani et al., *Soft gluon resummation for Higgs boson production at hadron colliders*, *JHEP* **0307** (2003) 028, [hep-ph/0306211].

[56] U. Aglietti et al., *Two loop light fermion contribution to Higgs production and decays*, *Phys. Lett.* **B 595** (2004) 432, [hep-ph/0404071].

[57] S. Actis et al., *NLO electroweak corrections to Higgs boson production at hadron colliders*, *Phys. Lett.* **B 670** (2008) 12, [arXiv:0809.1301].

[58] M. Ciccolini, A. Denner, and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC*, *Phys. Rev. Lett.* **99** (2007) 161803, [arXiv:0707.0381].

[59] M. Ciccolini, A. Denner, and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*, *Phys. Rev.* **D 77** (2008) 013002, [arXiv:0710.4749].

[60] K. Arnold et al., *VBFNLO: A parton level Monte Carlo for processes with electroweak bosons*, *Comput. Phys. Commun.* **180** (2009) 1661, [arXiv:0811.4559].

[61] P. Bolzoni et al., *Higgs production via vector-boson fusion at NNLO in QCD*, *Phys. Rev. Lett.* **105** (2010) 011801, [arXiv:1003.4451].

[62] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **0207** (2002) 012, [hep-ph/0201195].

[63] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, *Phys. Lett.* **B 579** (2004) 149, [hep-ph/0307206].

[64] M. Ciccolini, S. Dittmaier, and M. Krämer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, *Phys. Rev.* **D 68** (2003) 073003, [hep-ph/0306234].

[65] M. Grazzini and H. Sargsyan, *Heavy-quark mass effects in Higgs boson production at the LHC*, *JHEP* **1309** (2013) 129, [arXiv:1306.4581].

[66] D. de Florian et al., *Higgs boson production at the LHC: transverse momentum resummation effects in the $H \to 2\gamma$, $H \to WW \to \ell\nu\ell\nu$ and $H \to ZZ \to 4\ell$ decay modes*, *JHEP* **1206** (2012) 132, [arXiv:1203.6321].

[67] J. M. Campbell, R. K. Ellis, and G. Zanderighi, *Next-to-Leading order Higgs + 2 jet production via gluon fusion*, *JHEP* **0610** (2006) 028, [hep-ph/0608194].

[68] A. Banfi et al., *Higgs and Z-boson production with a jet veto*, *Phys. Rev. Lett.* **109** (2012) 202001, [arXiv:1206.4998].

[69] A. Banfi, G. P. Salam, and G. Zanderighi, *NLL+NNLO predictions for jet-veto efficiencies in Higgs-boson and Drell-Yan production*, *JHEP* **1206** (2012) 159, [arXiv:1203.5773].

[70] A. Banfi, P. F. Monni, and G. Zanderighi, *Quark masses in Higgs production with a jet veto*, *JHEP* **1401** (2014) 097, [arXiv:1308.4634].

[71] S. Dittmaier et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, (2012), [arXiv:1201.3084].

[72] S. Jadach et al., *The tau decay library Tauola: Version 2.4*, *Comput. Phys. Commun.* **76** (1993) 361.

[73] Z. Czyczula, T. Przedzinski, and Z. Was, *TauSpinner Program for Studies on Spin Effect in tau Production at the LHC*, *Eur. Phys. J.* **C 72** (2012) 1988, [arXiv:1201.0117].

[74] T. Sjostrand, S. Mrenna, and P. Skands, PYTHIA *6.4 physics and manual*, *JHEP* **05** (2006) 026, [hep-ph/0603175].

[75] G. Corcella et al., HERWIG *6.5 release note*, (2002), [hep-ph/0210213].

[76] N. Davidson, T. Przedzinski, and Z. Was, *Photos Interface in C++: Technical and Physics Documentation*, (2010), [arXiv:1011.0937].

[77] M. Mangano et al., ALPGEN*, a generator for hard multiparton processes in hadronic collisions*, *JHEP* **07** (2003) 001, [hep-ph/0206293].

[78] M. Mangano et al., *Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b}$ + n jets as a case study*, *Nucl. Phys.* **B 632** (2002) 343, [hep-ph/0108069].

[79] T. Binoth et al., *Gluon-induced W-boson pair production at the LHC*, *JHEP* **0612** (2006) 046, [hep-ph/0611170].

[80] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, *Comput. Phys. Commun.* **184** (2013) 919, [hep-ph/0405247].

[81] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J.* **C 70** (2010) 823, [arXiv:1005.4568].

[82] S. Agostinelli et al., *Geant4: A Simulation toolkit*, *Nucl. Instrum. Meth.* **A 506** (2003) 250.

[83] ATLAS Collaboration, *Summary of ATLAS* PYTHIA *8 tunes*, ATL-PHYS-PUB-2012-003 (2012). http://cds.cern.ch/record/1474107/.

[84] S. Heinemeyer, and C. Mariotti, and G. Passarino, and R. Tanaka (Eds.) (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, *CERN-2013-004* (CERN, Geneva, 2013) , [arXiv:1307.1347].

[85] S. Catani et al., *Vector boson production at hadron colliders: A fully exclusive QCD calculation at next-to-next-to-leading order*, *Phys. Rev. Lett.* **103** (2009) 082001.

[86] S. Catani and M. Grazzini, *Next-to-next-to-leading-order subtraction formalism in hadron collisions and its application to Higgs-boson production at the Large Hadron Collider*, *Phys. Rev. Lett.* **98** (2007) 222002.

[87] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **0902** (2009) 007, [arXiv:0811.4622].

[88] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, *Phys. Lett.* **B 710** (2012) 612.

[89] P. Bärnreuther, M. Czakon and A. Mitov, *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, *Phys. Rev. Lett.* **109** (2012) 132001, [arXiv:1204.5201].

[90] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, *JHEP* **1212** (2012) 054, [arXiv:1207.0236].

[91] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, *JHEP* **1301** (2013) 080, [arXiv:1210.6832].

[92] M. Czakon, P. Fiedler, A. Mitov, *The total top quark pair production cross-section at hadron colliders through $O(\alpha_s^4)$*, *Phys. Rev. Lett.* **110** (2013) 252004, [arXiv:1303.6254].

[93] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930, [arXiv:1112.5675].

[94] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, *Phys. Rev.* **D 82** (2010) 054018, [arXiv:1005.4451].

[95] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, *Phys. Rev.* **D 83** (2011) 091503, [arXiv:1103.2792].

[96] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, *Phys. Rev.* **D 81** (2010) 054028, [arXiv:1001.5034].

[97] J. M. Campbell, K. R. Ellis, and C. Williams, *Vector boson pair production at the LHC*, *JHEP* **1107** (2011) 018, [arXiv:1105.0020].

[98] R. Ellis et al., *Higgs decay to $\tau^+\tau^-$: A possible signature of intermediate mass Higgs bosons at high energy hadron colliders*, *Nucl. Phys.* **297** (1988) 221.

[99] A. Elagin et al., *A new mass reconstruction technique for resonances decaying to di-tau*, *Nucl. Instrum. Meth.* **A 654** (2011) 481, [arXiv:1012.4686].

[100] G. Hanson et al., *Evidence for Jet Structure in Hadron Production by e+ e- Annihilation*, *Phys. Rev. Lett.* **35** (1975) 1609.

[101] ATLAS Collaboration, *Measurement of the $Z \to \tau\tau$ cross section with the ATLAS detector*, *Phys. Rev.* **D 84** (2011) 112006, [arXiv:1108.2016].

[102] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, *Eur. Phys. J.* **C 73** (2013) 2518, [arXiv:1302.4393].

[103] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046 (2014). http://cds.cern.ch/record/1741020.

[104] ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, ATLAS-CONF-2014-004 (2014). http://cds.cern.ch/record/1664335.

[105] 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). http://cds.cern.ch/record/1544036.

[106] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s}$ = 7 TeV recorded in 2010 with the ATLAS detector*, *Eur. Phys. J.* **C 73** (2013) 2306, [arXiv:1210.6210].

[107] S. Dittmaier et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, (2011), [arXiv:1101.0593].

[108] I. Stewart and F. Tackmann, *Theory uncertainties for Higgs and other searches using jet bins*, *Phys. Rev.* **D 85** (2012) 034011, [arXiv:1107.2117].

[109] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029, [hep-ph/0204244].

[110] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J.* **C 63** (2009) 189, [arXiv:0901.0002].

[111] NNPDF Collaboration, Richard D. Ball et al., *Impact of heavy quark masses on parton distributions and LHC phenomenology*, *Nucl. Phys.* **B 849** (2011) 296, [arXiv:1101.1300].

[112] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **1407** (2014) 079, [arXiv:1405.0301].

[113] ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, *Phys.Rev.* **D90** (2014) 052004, [arXiv:1406.3827].

[114] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C 1** (2011) 1554, [arXiv:1007.1727].

[115] A. David et al. (LHC Higgs Cross Section Working Group), *LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle*, (2012), [arXiv:1209.0040].

[116] B. Efron and C. Stein, *The Jackknife Estimate of Variance*, *The Annals of Statistics* **9** (1981) no. 3 586–596.

[117] G. Bohm and G. Zech, *Introduction to statistics and data analysis for physicists*. Verlag Deutsches Elektronen-Synchrotron, Hamburg, ISBN 978-3-935702-41-6, 2010.