



ATLAS NOTE

ATLAS-CONF-2013-108

November 28, 2013



Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector

The ATLAS Collaboration

Abstract

A search for the Higgs boson with a mass of about 125 GeV decaying into a pair of τ leptons is performed with a data sample of proton-proton collisions, corresponding to an integrated luminosity of $\mathcal{L} = 20.3 \text{ fb}^{-1}$, collected with the ATLAS detector at the LHC at a centre-of-mass energy of $\sqrt{s} = 8 \text{ TeV}$. Final states in all τ decay combinations (both hadronic and leptonic) are examined. The observed (expected) deviation from the background-only hypothesis corresponds to a significance of 4.1 (3.2) standard deviations, and the measured signal strength is $\mu = 1.4^{+0.5}_{-0.4}$. This is evidence for the existence of $H \rightarrow \tau^+\tau^-$ decays, consistent with the Standard Model expectation for a Higgs boson with $m_H = 125 \text{ GeV}$.



1 Introduction

The observation of a new particle with a mass near 125 GeV by the ATLAS and CMS experiments [1, 2] in the search for the Standard Model (SM) Higgs boson [3–8] is a great success of the Large Hadron Collider (LHC) physics programme at the CERN laboratory. To date, however, evidence for Higgs-boson decays into fermionic final states is not conclusive.

With a branching ratio of 6.3% [9], $H \rightarrow \tau^+\tau^-$ is among the leading decay modes for a SM Higgs boson with a mass of 125 GeV. This decay mode can provide a direct measurement of the coupling of the Higgs boson to fermions, thereby testing an important prediction of the theory. The observation of the $H \rightarrow \tau^+\tau^-$ decay mode would be strong evidence that fermions acquire their mass through the Higgs mechanism.

This note presents a search for the SM Higgs boson in the $H \rightarrow \tau_{\text{lep}}^+\tau_{\text{lep}}^-$, $H \rightarrow \tau_{\text{lep}}^+\tau_{\text{had}}^-$, and $H \rightarrow \tau_{\text{had}}^+\tau_{\text{had}}^-$ final states¹ with the ATLAS detector [10], using the full dataset collected in proton-proton (pp) collisions at $\sqrt{s} = 8$ TeV in 2012 and corresponding to an integrated luminosity of $\mathcal{L} = 20.3 \text{ fb}^{-1}$.

The search is designed to be sensitive to the SM Higgs boson produced through gluon fusion (ggF) [11], vector boson fusion (VBF) [12], and associated production (VH) with $V = W$ or Z decaying hadronically. All these mechanisms can give rise to jet signatures, particularly in the VBF case, where two high-energy jets with a large pseudorapidity separation are produced.

The results presented in this note supersede those of Ref. [13] which correspond to an analysis of part of the 2012 dataset with a luminosity of $\mathcal{L} = 13.0 \text{ fb}^{-1}$, and the full 2011 dataset. The most recent results from the CMS collaboration on the search for the SM Higgs boson in the $\tau^+\tau^-$ channel can be found in Ref. [14].

2 Analysis Strategy

The ATLAS detector is described in Section 3, with details on the data and simulated samples given in Section 4. Object identification cuts are applied as described in Section 5 and Section 6 details pre-selection cuts for all three channels: $\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$. The events are further classified into *categories* optimized for the different SM Higgs production mechanisms, as described in Section 7. The category definitions depend to a certain extent on the channel because of the different background compositions.

Following preselection and categorization, the $H \rightarrow \tau^+\tau^-$ signal is still overwhelmed by a variety of background sources. A boosted decision tree (BDT) multivariate analysis technique is used to discriminate signal from background [15–17]. Separate BDTs are trained for each channel in each category as described in Section 8, based on input variables that have differing distributions for signal and background. The background estimates are described in Section 9. The BDT output distribution (BDT score) is then used as the final discriminant.

Since a BDT is sensitive to the correlations among variables, it is necessary to demonstrate that these correlations are well modelled. For that reason, the BDT output is computed in several signal-depleted control regions, and the agreement between the data and the background model is confirmed, as described in Section 10.

To ensure that analysis choices are not biased, discriminating variables which might on their own reveal the presence of a signal were blinded throughout the analysis optimization. This blinding was applied to variables such as $m_{\tau\tau}$ and the final BDT score in signal sensitive bins, but not to basic kinematic distributions such as the transverse momentum of τ leptons.

¹ τ_{lep} and τ_{had} denote leptonically and hadronically decaying τ leptons, respectively. Charge-conjugate decay modes are implied. Throughout the remainder of this note, a simplified notation without the particle charges is used.

Systematic uncertainties relevant to the analysis are mentioned in Section 11, and the signal extraction procedure is detailed in Section 12. The results of this search are given in Section 13.

3 The ATLAS Detector

The ATLAS detector [10] is a cylindrical² multi-purpose detector at the LHC. The detector subsystem closest to the interaction point, the Inner Detector (ID), provides precise position and momentum measurements of charged particles. It covers the pseudorapidity range $|\eta| < 2.5$ and provides full azimuthal coverage. It consists of three subdetectors arranged in a coaxial geometry around the beam axis: the silicon pixel detector, the silicon microstrip detector and the straw-tube transition-radiation tracker. A solenoid magnet generates a 2 T magnetic field in which the ID is immersed.

Electromagnetic calorimetry in the region $|\eta| < 3.2$ is based on a high-granularity, lead/liquid-argon (LAr) sampling technology. Hadronic calorimetry uses a scintillating-tile/steel detector covering the region $|\eta| < 1.7$ and a copper/LAr detector in the region $1.5 < |\eta| < 3.2$. The most forward region of the detector $3.1 < |\eta| < 4.9$ is equipped with a dedicated forward calorimeter, measuring both electromagnetic and hadronic energies using copper/LAr and tungsten/LAr modules.

A large stand-alone Muon Spectrometer (MS) is the outermost part of the detector. It consists of three large air-core superconducting toroidal magnet systems. The deflection of the muon trajectories in the magnetic field is measured in three layers of precision drift tube chambers for $|\eta| < 2$. In higher η regions ($2.0 < |\eta| < 2.7$), two layers of drift tube chambers are used in combination with one layer of cathode strip chambers in the innermost endcap wheels of the MS. Three layers of resistive plate chambers in the barrel ($|\eta| < 1.05$) and three layers of thin gap chambers in the endcaps ($1.05 < |\eta| < 2.4$) provide the muon trigger and also measure the muon trajectory in the non-bending plane of the spectrometer magnets.

A three-level trigger system [18] is used to select events in real time. 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 300 Hz by two software-based trigger levels, Level-2 and the Event Filter.

4 Data and Simulated Samples

This search uses pp collision data at $\sqrt{s} = 8$ TeV collected in 2012. After requiring that all detector systems are operational, the dataset used corresponds to an integrated luminosity of $\mathcal{L} = 20.3 \text{ fb}^{-1}$. The triggers used by each channel are given in Table 1.

The simulated event samples, based on Monte Carlo (MC) techniques and a full description of the the ATLAS detector [19] with GEANT4 [20], are listed below. These samples include the simulation of pile-up activity in the same or nearby bunch crossings. The MC samples are re-weighted to reproduce the observed distribution of the mean number of interactions per bunch crossing in the data.

The simulation of signal events produced via the gluon-fusion and VBF-production mechanisms is performed using the POWHEG [21–23] event generator based on next-to-leading order (NLO) QCD calculations. Soft-gluon resummation up to next-to-next-to-leading logarithm order [24] is adopted. The finite quark-mass effects are taken into account in POWHEG [25]. The parton shower, hadronization and underlying event simulations are provided by PYTHIA [26, 27]. The CT10 [28] parton distribution

²ATLAS 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 $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

Trigger	p_T threshold(s) [GeV]	$\tau_{\text{lep}}\tau_{\text{lep}}$	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$
Electron	24	•	•	
Muon	24		•	
Di-electron	12 ; 12	•		
Di-muon	18 ; 8	•		
Electron + Muon	12 ; 8	•		
Electron + τ_{had}	18 ; 20		•	
Muon + τ_{had}	15 ; 20		•	
Di- τ_{had}	29 ; 20			•

Table 1: Triggers used for each channel. When more than one trigger is used, a logical OR of the triggers is taken and the trigger efficiency is calculated accordingly. The electron+ τ_{had} and muon+ τ_{had} triggers are used for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel at preselection, but not in the VBF and boosted categories as defined in Section 7.

function (PDF) is used. The associated production (VH) samples are generated at leading order (LO) in QCD using PYTHIA with the CTEQ6L1 [29] PDF set. The signal samples are normalized to cross sections computed at next-to-next-to-leading order (NNLO) QCD and taken from Ref. [9].

Background simulation samples use several generators, as described below, each interfaced with HERWIG [30] (with one exception noted below) to provide the parton shower and hadronization. JIMMY [31] provides the modeling of the underlying event. The samples for W/Z +jets events are generated with ALPGEN [32]. This generator employs the MLM matching scheme [33] between the hard process (calculated with LO matrix elements for up to five jets) and the parton shower. The $\tau_{\text{lep}}\tau_{\text{had}}$ channel uses these samples interfaced with PYTHIA rather than HERWIG. The $t\bar{t}$ samples are produced with MC@NLO [34] with NLO accuracy. Single-top events are generated with AcerMC [35]. The diboson (WW , WZ , ZZ) MC samples are generated using HERWIG for the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel. In the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, HERWIG is used for the WZ and ZZ samples, while the WW sample is obtained using ALPGEN interfaced to HERWIG. The loop-induced $gg \rightarrow WW$ processes are generated using gg2WW [36]. The PDF set used with the AcerMC, ALPGEN and HERWIG event generators is CTEQ6L1, while CT10 is used for the generation of events with MC@NLO and gg2WW.

TAUOLA [37] performs the tau decay and PHOTOS [38] provides additional photon radiation from charged leptons for all samples described. The normalization for these backgrounds is either estimated from data control regions, as described in Section 9, or NLO cross sections are used.

The main and largely irreducible $Z/\gamma^* \rightarrow \tau^+\tau^-$ background is modelled with selected $Z/\gamma^* \rightarrow \mu^+\mu^-$ data events, where the muon tracks and associated calorimeter cells are replaced by the corresponding signatures of the decay of τ leptons. The two τ leptons are simulated by TAUOLA, matched to the kinematics of the data muons they replace. Here, the τ polarization and spin correlations are modelled with the TAUOLA program and the $\mu - \tau$ mass difference is taken into account. Thus, only the τ decays (both hadronic and leptonic) and the corresponding detector response are taken from the simulation, whereas the underlying event kinematics and all other properties, including pile-up effects, are obtained from the data. This hybrid sample will be referred to as embedded [13] data in the following. The embedding procedure is extensively validated, e.g. by replacing the muons in selected $Z/\gamma^* \rightarrow \tau^+\tau^-$ data events by simulated muons instead of τ leptons. This test is sensitive to systematic effects intrinsic to the method, for instance due to the subtraction of calorimeter cell energy associated to the data muons. The validation does not reveal any bias in the embedding procedure [13] beyond the associated systematic uncertainties.

5 Object and Event Reconstruction

The reconstruction and identification of leptons, jets, and missing transverse momentum are performed using standard ATLAS algorithms briefly described below.

5.1 Electrons

Electron candidates are reconstructed from a cluster in the EM calorimeter and a matching ID track. They are selected if they pass the *Medium* [39] identification criteria³, have transverse energy greater than 15 GeV and are in the region $|\eta| < 2.47$. Candidates found in the calorimeter transition region ($1.37 < |\eta| < 1.52$) are not considered. Typical electron efficiencies, after these selection cuts, range between 80% and 90% depending on η and p_T . Additional isolation criteria, based on tracking and calorimeter information, are used to suppress the background from misidentified jets or from semileptonic decays of charm and bottom hadrons.

5.2 Muons

Muon candidates are reconstructed from the association of an ID-track and an MS-track [40]. The momentum is evaluated from their combination. Muon candidates are selected if they have a transverse momentum greater than 10 GeV and are in the region $|\eta| < 2.5$. Typical muon efficiencies, after these selection cuts, are approximately 90%. Further isolation criteria are required to suppress the background from misidentified jets or from semileptonic decays of charm and bottom hadrons.

5.3 Jets

Jets are reconstructed by the anti- k_r algorithm [41, 42] with a distance parameter $R = 0.4$, taking topological clusters [43] in the calorimeters as inputs. The local hadronic calibration scheme [44] and the jet energy scale [44] (JES) are used to calibrate energy deposits from hadrons based on calorimeter signals only. Jets are required to be reconstructed in the range $|\eta| < 4.5$ and to have a minimum transverse momentum of 30 GeV except in the $\tau_{\text{had}}\tau_{\text{had}}$ channel where they are required to have a minimum transverse momentum of 35 GeV for $|\eta| > 2.4$ and in counting b -tagged jets in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel where the transverse momentum threshold is lowered to 25 GeV.

A jet-vertex fraction (JVF) requirement is used to reduce the number of selected jets in the event due to pile-up activity. The JVF is defined as the ratio between the scalar sum of the transverse momenta of the tracks in the jet associated to the primary vertex and the scalar sum of the transverse momentum of the tracks in the jet associated to any vertex in the event. Jets with $|\eta| < 2.4$ and $p_T < 50$ GeV are required to have a JVF exceeding 0.5.

In the pseudorapidity range $|\eta| < 2.5$, b -jets are selected using a tagging algorithm [45]. The b -tagging algorithm used has an efficiency of 60–70% for b -tagged jets in simulated $t\bar{t}$ events [46]. The corresponding light-quark jet misidentification probability is 0.1–0.5%, depending on the jet p_T and η [47].

5.4 Hadronically Decaying τ Leptons

Hadronically decaying τ leptons are reconstructed starting from clusters in the electromagnetic and hadronic calorimeters [48]. 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 track charges. This search uses τ_{had} candidates with $p_T > 20$ GeV and $|\eta| < 2.47$. The τ_{had} candidates are required to

³The electron identification criteria have been reoptimized for 2012 data-taking conditions.

have charge ± 1 , and must be 1- or 3-track (prong) candidates. The two-track sample (where the charge requirement is dropped) is retained for background studies as described in Section 9.1. A BDT τ identification method is used, requiring that the τ_{had} candidate passes the *Medium* [49] identification criteria, corresponding to approximately 55–60% efficiency. Dedicated criteria [49] to suppress τ_{had} candidates from misidentified electrons and muons are also applied. The misidentification probabilities for τ_{had} candidates with $p_T > 20$ GeV have a typical value of 1–2% .

5.5 Object Overlap Removal

When different objects selected according to the criteria mentioned above overlap with each other geometrically (within $\Delta R < 0.2$), only one of them is considered. The overlap is resolved by selecting muon, electron, τ_{had} and jet candidates in this order of priority.

5.6 Missing Transverse Momentum

The signal events are characterized by true missing transverse momentum (E_T^{miss}) due to the presence of the neutrinos from τ decays. In this analysis, the E_T^{miss} reconstruction [50] uses calorimeter cells calibrated according to the reconstructed physics objects to which they are associated. Calorimeter cells are associated with a reconstructed and identified high- p_T parent object in the following order: electrons, photons, hadronically decaying τ -leptons, jets and muons. Calorimeter cells not associated with any other objects are scaled by the soft term vertex fraction and are used in the E_T^{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_T^{miss} in high pile-up conditions [51]. The p_T of muons identified in the events are also taken into account in the E_T^{miss} calculation.

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, a second variable named High p_T Objects E_T^{miss} (HPTO E_T^{miss}) is also used to reject Drell-Yan background. It is built from the high p_T objects: the two leptons and the jets with $p_T > 25$ GeV. The two E_T^{miss} variables are strongly correlated for the signal due to neutrinos in the final state, but only loosely correlated for background from $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$.

5.7 Higgs Candidate Kinematic Reconstruction

The invariant $\tau\tau$ mass ($m_{\tau\tau}^{\text{MMC}}$) is reconstructed using the missing mass calculator (MMC) [52]. This requires solving an underconstrained system of equations for 6 to 8 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 τ leptons in the event, and the invariant mass of the two neutrinos from any leptonic τ decays. This is done by using the constraints from the measured x and y components of E_T^{miss} , and the visible masses of both τ candidates. A scan is then performed over the yet undetermined variables, and each scan point is then weighted by its probability according to the τ decay topologies. The estimator for the $\tau\tau$ mass ($m_{\tau\tau}^{\text{MMC}}$) is then defined as the most probable value of the weighted scan points.

Another important variable is the transverse momentum p_T^H of the Higgs-boson candidate. This quantity is reconstructed using the vector sum of the event E_T^{miss} and the transverse momentum of the visible τ decay products.

6 Preselection

In addition to criteria to ensure the detector was functioning properly, requirements are applied in order to increase the purity and quality of the data sample by rejecting non-collision events such as cosmic rays

and beam halo events. To ensure that the event is the result of hard-scattering, at least one vertex with at least four associated tracks and a position consistent with the beam spot position is required. After these basic criteria are applied, the event preselection varies by channel.

For all channels, the leptons that are considered for overlap removal with τ_{had} candidates need only satisfy *Loose* criteria, 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 channels involving real electrons and muons, $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$, use tighter selections of electron and muon candidates, including isolation criteria. For the τ_{had} candidates considered in the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, the criteria used to reject electrons misidentified as τ_{had} candidates are tightened [49].

Higher p_{T} thresholds are applied to electrons, muons, and τ_{had} candidates according to the trigger condition satisfied by the event, as described in Table 1. For events passing the single-electron or single-muon trigger, an offline requirement of $p_{\text{T}} > 26$ GeV is applied for the lepton that triggered the event. The p_{T} thresholds are unchanged from the general object selection for events passing the di-electron and electron+muon combined triggers. For events passing the di-muon trigger, the leading muon is required to have $p_{\text{T}} > 20$ GeV. For the combined electron+ τ_{had} and muon+ τ_{had} triggers, the τ_{had} candidates must satisfy $p_{\text{T}} > 25$ GeV, while the electron and muon candidates must have transverse momenta exceeding 20 and 17 GeV, respectively. In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, where the events are required to pass the di- τ_{had} trigger, the leading (sub-leading) τ_{had} candidate must satisfy $p_{\text{T}} > 35(25)$ GeV.

6.1 The $\tau_{\text{lep}}\tau_{\text{lep}}$ Channel

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, exactly two isolated leptons of opposite-sign (OS) charges are required. Events containing a τ_{had} candidate are vetoed. The two leptons must satisfy $30 < m_{\tau\tau}^{\text{vis}} < 100$ GeV in the $e^+\mu^-$ channel, and $30 < m_{\tau\tau}^{\text{vis}} < 75$ GeV for the e^+e^- and $\mu^+\mu^-$ channels⁴, thus avoiding Z , charmonium and bottomonium resonances. It is also required that the scalar sum of the transverse momenta $p_{\text{T}}(\ell_1) + p_{\text{T}}(\ell_2)$ exceeds 35 GeV and that $\Delta\phi_{\ell\ell} < 2.5$.

For e^+e^- and $\mu^+\mu^-$ final states, the $E_{\text{T}}^{\text{miss}}$ and HPTO $E_{\text{T}}^{\text{miss}}$ must both be greater than 40 GeV, while for $e^+\mu^-$ only a requirement of $E_{\text{T}}^{\text{miss}} > 20$ GeV is made.

Finally, the fraction of the momentum of each tau lepton carried by its visible decay products⁵ (calculated using the collinear approximation to determine the neutrino momenta [53]), as defined by the equation,

$$x_{\tau 1(2)} = \frac{p_{\text{vis}1(2)}}{p_{\text{vis}1(2)} + p_{\text{mis}1(2)}} \quad (1)$$

are required to satisfy $0.1 < x_{\tau 1}, x_{\tau 2} < 1$.

6.2 The $\tau_{\text{lep}}\tau_{\text{had}}$ Channel

In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, exactly one lepton and one τ_{had} candidate with OS charges, passing the given p_{T} thresholds, are required. To substantially reduce the W +jets background at this stage, events are rejected if the transverse mass⁶ constructed from the lepton and the $E_{\text{T}}^{\text{miss}}$ satisfies $m_{\text{T}} > 70$ GeV.

⁴ $m_{\tau\tau}^{\text{vis}}$ is defined as the invariant mass of the visible decay products of the τ leptons.

⁵ p_{vis} is defined as the total momentum of the visible decay products of the τ lepton. p_{mis} is defined as the momentum of the neutrino reconstructed using the collinear approximation.

⁶ $m_{\text{T}} = \sqrt{2p_{\text{T}}(\ell) \times E_{\text{T}}^{\text{miss}} \times (1 - \cos \Delta\phi)}$, and $\Delta\phi$ is the azimuthal separation between the directions of the lepton and the missing transverse momentum vector.

6.3 The $\tau_{\text{had}}\tau_{\text{had}}$ Channel

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, exactly two τ_{had} candidates with OS charges are required. Events with electron or muon candidates are rejected. Trigger requirements in this channel motivate minimum p_{T} thresholds of $p_{\text{T}}(\tau_{\text{had}1}) > 35$ GeV and $p_{\text{T}}(\tau_{\text{had}2}) > 25$ GeV, as mentioned at the beginning of Section 6.

In order to reduce the background from multijet production and to exclude badly modelled events, several additional requirements are applied: the τ_{had} identification criteria are tightened such that at least one of the candidates satisfies the *Tight* criteria, and the separation between the two τ_{had} candidates must satisfy $0.8 < \Delta R(\tau_{\text{had}}, \tau_{\text{had}}) < 2.8$ and $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}}) < 1.5$. The missing transverse momentum requirement is set to $E_{\text{T}}^{\text{miss}} > 20$ GeV and the $E_{\text{T}}^{\text{miss}}$ direction must either be between the two visible τ_{had} candidates in ϕ or the condition $\min[\Delta\phi(\tau, E_{\text{T}}^{\text{miss}})] < \pi/2$ must be fulfilled.

6.4 Kinematic Distributions after Preselection

In Figures 1-3, the distributions for $E_{\text{T}}^{\text{miss}}$, $m_{\tau\tau}^{\text{vis}}$, p_{T}^H , and $\Delta\eta(j_1, j_2)$ (for events with at least two jets) are shown for each channel after the preselection criteria have been applied. Good agreement is shown between the observed data and the predictions from the background modelling, which is discussed in Section 9. The variables $E_{\text{T}}^{\text{miss}}$ and $m_{\tau\tau}^{\text{vis}}$ are important kinematic variables that are correlated with $m_{\tau\tau}^{\text{MMC}}$, while p_{T}^H and $\Delta\eta(j_1, j_2)$ are variables that help define the categorization of events as described in Section 7.

7 Analysis Categories

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

- **VBF:** targeted at the vector boson fusion Higgs production mechanism. This category is characterized by the presence of two jets with a large pseudorapidity separation. Some signal events from the gluon-fusion and VH production mechanisms are also selected in this category.
- **Boosted:** targeted at events with a boosted Higgs boson from the gluon-fusion production mechanism. It includes only events which fail the VBF category definition. Hence, this category selects Higgs boson candidates which have larger p_{T} ($p_{\text{T}}^H > 100$ GeV) and well-measured mass. Some signal events from the VBF and VH production mechanisms are also selected in this category.

While these categories are conceptually identical across the three channels, differences in dominant backgrounds and in dataset size require that the selection criteria differ in each. Table 2 specifies the selection criteria used for each channel. For both categories, the requirement on jets is an inclusive requirement: additional jets aside from those passing the category requirements are not discriminated against. The $\Delta\eta(j_1, j_2)$ requirement is applied using the highest two p_{T} jets in the event. A b -jet veto is used in the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels to suppress top-quark backgrounds. The $\tau_{\text{lep}}\tau_{\text{had}}$ channel imposes an additional cut on events in the VBF category, requiring that $m_{\tau\tau}^{\text{vis}} > 40$ GeV. The events failing this cut are not used in the boosted category. Furthermore, while other triggers were included at preselection, the final $\tau_{\text{lep}}\tau_{\text{had}}$ categories only consider events accepted by the single electron or single muon trigger.

8 Boosted Decision Trees

BDTs are used in each category to extract the Higgs-boson signal from the large number of background events. Decision trees [15] recursively partition the parameter space into multiple regions where signal

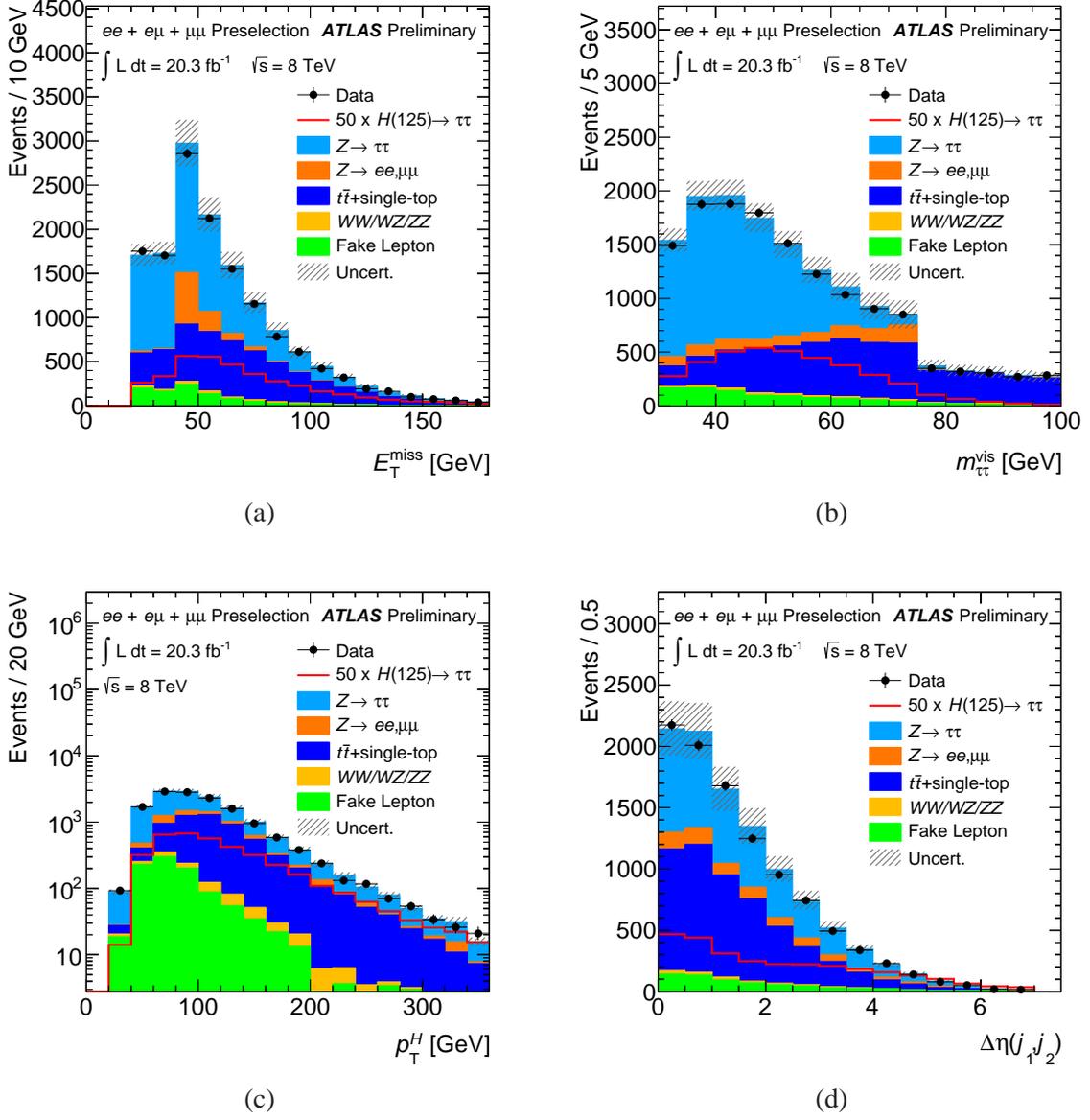
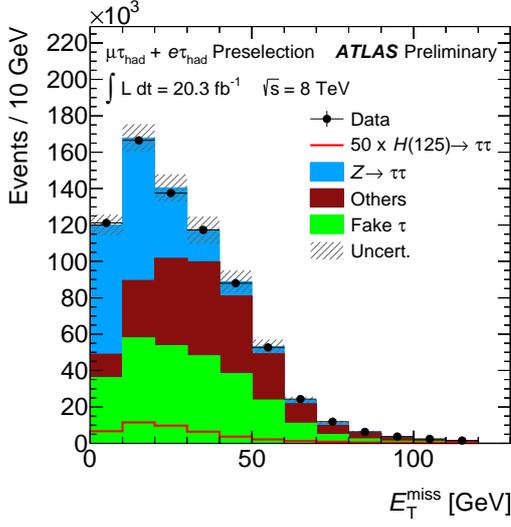
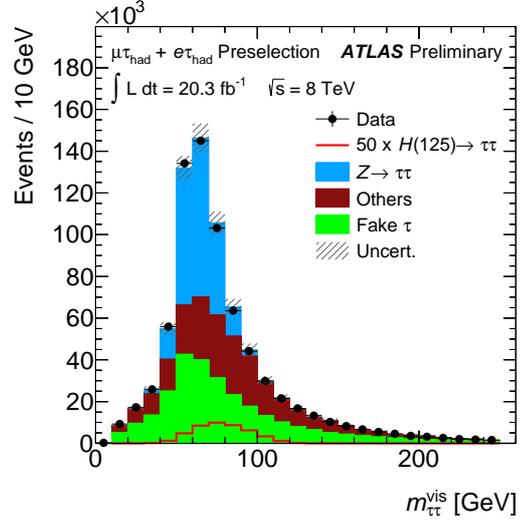


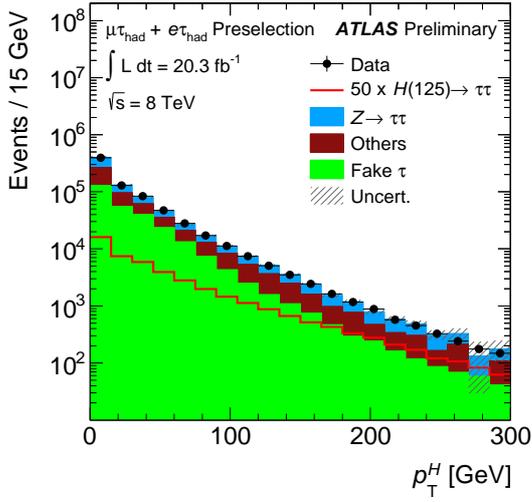
Figure 1: Kinematic distributions for the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel after preselection: (a) $E_{\text{T}}^{\text{miss}}$, (b) $m_{\tau\tau}^{\text{vis}}$, (c) p_{T}^H , and (d) $\Delta\eta(j_1, j_2)$. The background estimate for these distributions is described in Section 9. The $m_{\tau\tau}^{\text{vis}}$ distribution shows a step at 75 GeV due to the difference in cuts between same flavour and different flavour event selection. The $\Delta\eta(j_1, j_2)$ distribution is shown for events with at least two jets. Signal shapes are shown multiplied by a factor of 50. These figures use background predictions made without the global fit defined in Section 12.



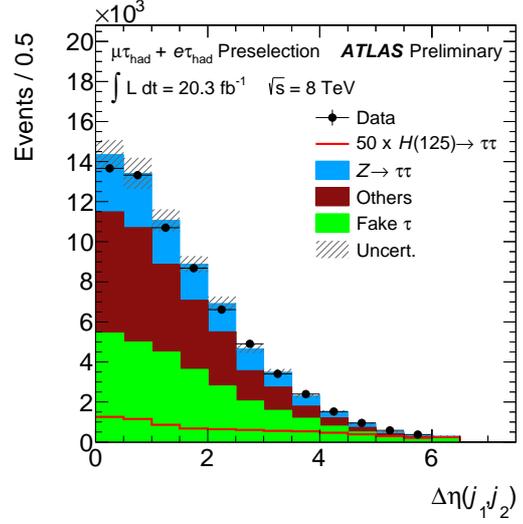
(a)



(b)



(c)



(d)

Figure 2: Kinematic distributions for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel after preselection: (a) E_T^{miss} , (b) $m_{\tau\tau}^{\text{vis}}$, (c) p_T^H , and (d) $\Delta\eta(j_1, j_2)$. The background estimate for these distributions is described in Section 9. The $\Delta\eta(j_1, j_2)$ distribution is shown for events with at least two jets. Signal shapes are shown multiplied by a factor of 50. These figures use background predictions made without the global fit defined in Section 12.

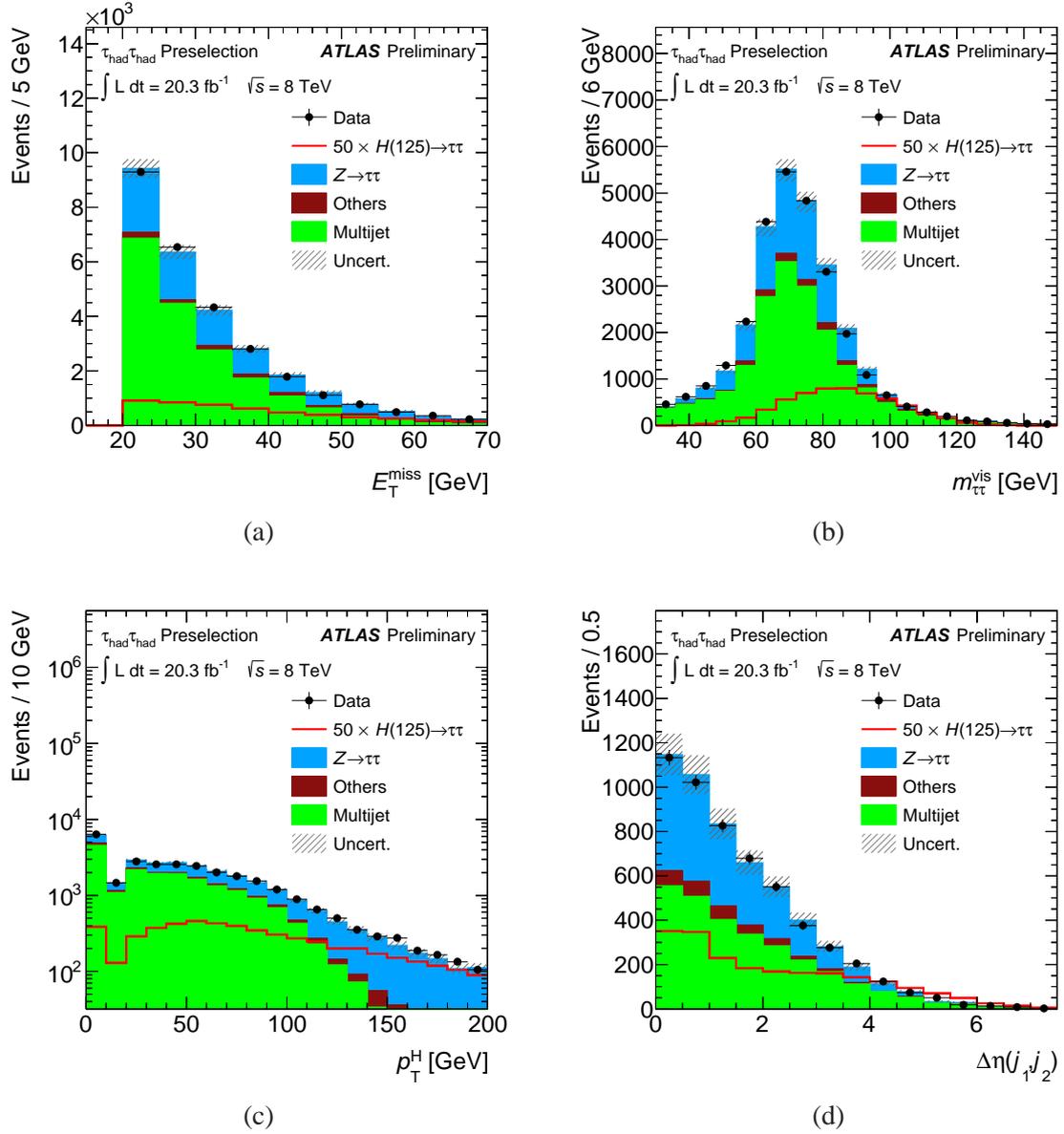


Figure 3: Kinematic distributions for the $\tau_{\text{had}}\tau_{\text{had}}$ channel after preselection: (a) E_T^{miss} , (b) $m_{\tau\tau}^{\text{vis}}$, (c) p_T^H , and (d) $\Delta\eta(j_1, j_2)$. The background estimate for these distributions is described in Section 9. The $\Delta\eta(j_1, j_2)$ distribution is shown for events with at least two jets. Signal shapes are shown multiplied by a factor of 50. These figures use background predictions made without the global fit defined in Section 12.

Category	Selection	$\tau_{\text{lep}}\tau_{\text{lep}}$	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$
VBF	$p_{\text{T}}(j_1)$ (GeV)	40	50	50
	$p_{\text{T}}(j_2)$ (GeV)	30	30	30/35
	$\Delta\eta(j_1, j_2)$	2.2	3.0	2.0
	b -jet veto for jet p_{T} (GeV)	25	30	-
	p_{T}^H (GeV)	-	-	40
Boosted	$p_{\text{T}}(j_1)$ (GeV)	40	-	-
	p_{T}^H (GeV)	100	100	100
	b -jet veto for jet p_{T} (GeV)	25	30	-

Table 2: Selection criteria applied in each analysis category for each channel. The numbers shown are lower thresholds. Only events that fail VBF category selection are considered for the boosted category. The $\Delta\eta(j_1, j_2)$ cut is applied on the two highest p_{T} jets in the event. Events in the $\tau_{\text{lep}}\tau_{\text{had}}$ VBF category must also satisfy $m_{\tau\tau}^{\text{vis}} > 40$ GeV, and those that fail this requirement are not considered for the $\tau_{\text{lep}}\tau_{\text{had}}$ boosted category. The $p_{\text{T}}(j_2)$ threshold in the $\tau_{\text{had}}\tau_{\text{had}}$ channel is 30 (35) GeV for jets within (outside of) $|\eta| = 2.4$.

or background purities have been 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 [16, 17]. After boosting, the final score undergoes a monotonic transformation to spread the scores between -1 and 1. The most signal-like events have scores near 1 while the most background-like have scores near -1.

BDTs trained on a sample of signal and background must be evaluated on an independent sample of events. In this analysis each sample (signal and background) is partitioned into two separate samples A and B, each with separate associated BDTs. The training is then performed with sample A and evaluated on sample B and vice-versa. For events in data, they are evaluated in the same way, such that half use the BDT trained on sample A and half use the BDT trained on sample B. For background model or signal MC events, the final distributions use all sample A events (evaluated using the BDT trained on sample B) and all sample B events (evaluated using BDT trained on sample A). In this way, 100% of each of the original samples appears in the final distributions, but no BDT is ever applied on an event from its own training sample.

Separate BDTs are trained for each analysis category and channel. Separately training by category naturally exploits differences in event kinematics between different Higgs boson production modes. The VBF category is trained with only a VBF signal sample, while the Boosted category is trained with gluon-gluon fusion, VBF, and VH signal samples. The signal samples used for all production processes in the training have $m_H = 125$ GeV. Separately training in each channel allows different discriminating variables to be used to address the differing background compositions in each channel. The details of background models and compositions for each channel are described in Section 9. The BDT input variables for each category are shown in Table 3. Most of these variables have straightforward definitions; those requiring definition are listed below:

- $p_{\text{T}}^{\text{Total}}$: Magnitude of vector sum of the visible components of the τ decay products, the two leading jets and the $E_{\text{T}}^{\text{miss}}$.
- $\text{sum } p_{\text{T}}$: Scalar sum of p_{T} of the visible components of the τ decay products and of the jets.
- $E_{\text{T}}^{\text{miss}}\phi$ centrality: A variable that quantifies the relative angular position of the $E_{\text{T}}^{\text{miss}}$ with respect to the τ decay products in the transverse plane. The transverse plane is transformed such that the

direction of the τ decay products are orthogonal, and that the smaller ϕ angle between the τ decay products defines the positive quadrant of the transformed plane. E_T^{miss} ϕ centrality is defined as the sum of the x and y components of the E_T^{miss} unit vector in this transformed plane.

- sphericity: A variable that describes the isotropy of energy flow. It is based on the quadratic momentum tensor:

$$S^{\alpha\beta} = \frac{\sum_i P_i^\alpha P_i^\beta}{\sum_i |\vec{p}_i|^2}. \quad (2)$$

Both leptons and the selected jets are considered in the computation. In this equation, α and β are the indices of the tensor, and the summation is performed over the momenta of the leptons and the jets in the event. The sphericity of the event is then defined in terms of the two largest eigenvalues of this tensor, λ_2 and λ_3 :

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

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

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

where η_1 and η_2 are the pseudorapidities of the two leading jets. This variable has value 1 when the object is halfway 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. This variable is used for the following BDT inputs: $\ell_1 \times \ell_2$ η centrality (product of the two η centralities), ℓ η centrality, j_3 η centrality and $\tau_{1,2}$ η centrality (η centrality of each τ_{had}). When j_3 η centrality is used, events with only two jets are assigned a dummy value of -0.5 .

9 Background Estimation

The background models are derived from a mixture of simulated samples and data. The normalization of background contributions generally relies on comparing the simulated samples of individual backgrounds to data in regions which have little signal contamination. The differences in background composition of the three channels necessitates different strategies for the background estimation. Common to all three is the dominant $Z \rightarrow \tau^+ \tau^-$ background which is taken from τ -embedded $Z \rightarrow \mu^+ \mu^-$ data, described in Section 4.

In the $\tau_{\text{lep}} \tau_{\text{lep}}$ channel, a non-isolated lepton region is used to model multijet, W +jets, and semi-leptonic $t\bar{t}$ backgrounds, while other contributions are estimated using simulated samples.

The $\tau_{\text{lep}} \tau_{\text{had}}$ channel measures misidentification factors (called “fake-factors”, described below) for evaluating contributions from multijet and W +jets backgrounds where a jet can be misidentified as a τ_{had} candidate. Other remaining backgrounds are accounted for using MC simulation normalized in control regions.

The $\tau_{\text{had}} \tau_{\text{had}}$ channel strategy is to model the multijet background with so-called not-opposite-sign (notOS) data, with τ_{had} candidates being required not to have opposite charges, and then to normalize the multijet and $Z \rightarrow \tau^+ \tau^-$ samples simultaneously as described in Section 9.1. All other backgrounds and their normalizations are taken from MC simulation.

The treatment of each background contribution is now described, highlighting differences among the three channels.

Variable	VBF			Boosted		
	$\tau_{\text{lep}}\tau_{\text{lep}}$	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$	$\tau_{\text{lep}}\tau_{\text{lep}}$	$\tau_{\text{lep}}\tau_{\text{had}}$	$\tau_{\text{had}}\tau_{\text{had}}$
$m_{\tau\tau}^{\text{MMC}}$	•	•	•	•	•	•
$\Delta R(\tau, \tau)$	•	•	•		•	•
$\Delta\eta(j_1, j_2)$	•	•	•			
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		•	•	•	•	•
x_{τ_1} and x_{τ_2}						•
$m_{\tau\tau, j_1}$				•		
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}		•			•	
$\min(\Delta\eta_{\ell_1\ell_2, \text{jets}})$	•					
$j_3 \eta$ centrality	•					
$\ell_1 \times \ell_2 \eta$ centrality	•					
$\ell \eta$ centrality		•				
$\tau_{1,2} \eta$ centrality			•			

Table 3: Discriminating variables used for each channel and category. The filled circles identify which variables are used in each decay mode. Note that variables such as $\Delta R(\tau, \tau)$ are defined either between the two leptons, between the lepton and τ_{had} , or between the two τ_{had} candidates, depending on the decay mode.

9.1 Data-Driven Backgrounds

The $\tau_{\text{lep}}\tau_{\text{lep}}$ channel treats all backgrounds from a misidentified lepton together, which accounts for multijet, W +jets and semileptonic $t\bar{t}$ processes. A control sample is created for this purpose by inverting isolation selections on one of the two leptons and subtracting contributions from other electroweak backgrounds (dileptonic $t\bar{t}$, $Z \rightarrow e^+e^-$, $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow \tau_{\text{lep}}\tau_{\text{lep}}$, diboson) obtained from MC simulations. A template is created from this sample and normalized by fitting the p_T distribution of the sub-leading lepton at an early stage in the preselection requirements defined in Section 6.1.

The $\tau_{\text{lep}}\tau_{\text{had}}$ channel uses the “fake-factor” [13] method to derive estimates for multijet and W +jets backgrounds that pass the $\tau_{\text{lep}}\tau_{\text{had}}$ selection due to a misidentified τ_{had} candidate. The fake-factor method requires a sample enriched in fake τ_{had} candidates defined by applying an event selection equivalent to the VBF and boosted category definitions except that the candidate must fail the *Medium* [49] τ_{had} identification requirement. The fake-factor is defined by the ratio of identified τ_{had} candidates to non-identified τ_{had} candidates. This is determined separately for samples dominated by quark and gluon jets using a separate W +jets control region which is dominated by quark jets defined by high m_T and a loose lepton region dominated by gluon jets. The derived fake-factors, weighted by the expected relative W +jets/multijets composition, also determined in the same control regions, are then applied to the fake τ_{had} events to calculate the expected multijet and W +jets background. The $p_T(\tau_{\text{had}})$ dependence of these fake-factors is shown in Figure 4.

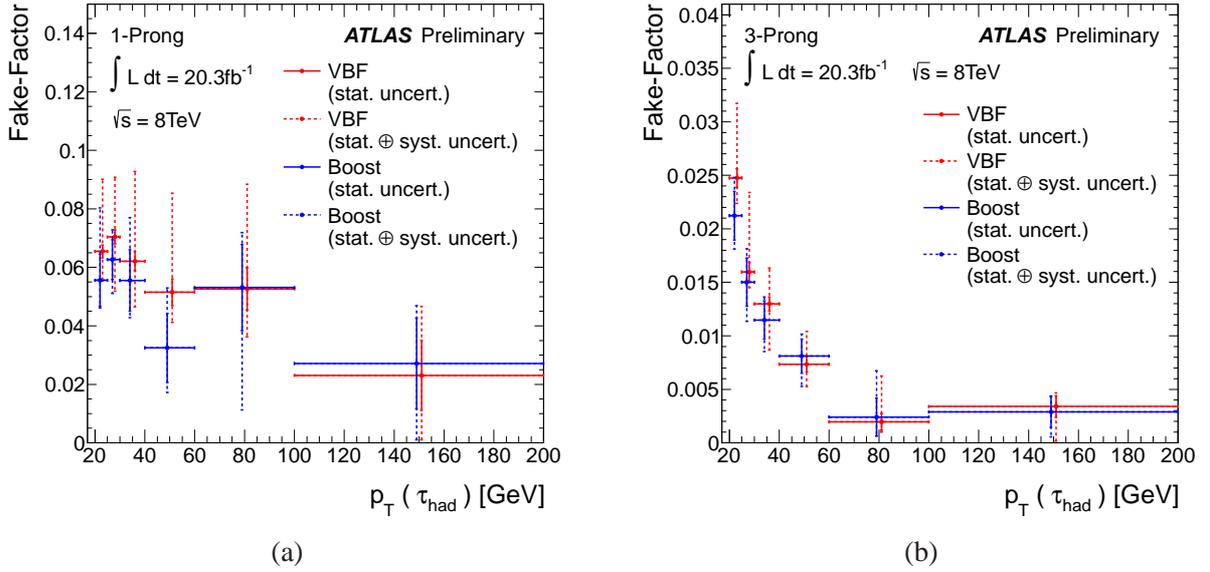


Figure 4: Fake-factors used to derive estimates for multijet and W +jets backgrounds in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. Factors are plotted as a function of the p_T of the τ_{had} candidate for the VBF and boosted categories: (a) for 1-prong τ_{had} candidates (b) for 3-prong τ_{had} candidates.

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the notOS data is used as a multijet template. For validation plots and as a starting point for the global fit, the normalization of the sample is determined by performing a simultaneous fit of multijet (modeled by notOS data) and $Z \rightarrow \tau^+\tau^-$ (modelled by embedding) templates to preselection data. The variable used in the fit is $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}})$, shown in Figure 5. The final normalizations for multijet and $Z \rightarrow \tau^+\tau^-$ are floated in the global fit in this channel, as described in Section 12.

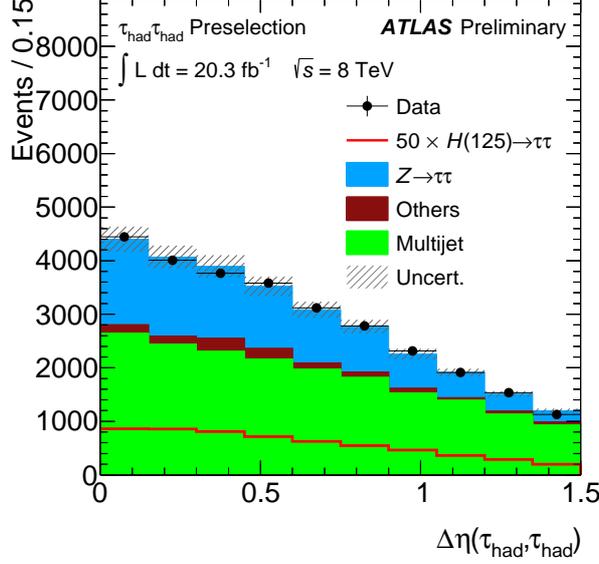


Figure 5: The $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}})$ distribution for the $\tau_{\text{had}}\tau_{\text{had}}$ channel after preselection. A fit to this distribution determines the starting values of the $Z \rightarrow \tau^+\tau^-$ and multijet normalizations used in the global fit defined in Section 12. The signal shape is shown multiplied by a factor of 50.

9.2 $Z \rightarrow \tau^+\tau^-$ Background

$Z \rightarrow \tau^+\tau^-$ decays constitute the dominant irreducible background. The contribution of this background is estimated using the τ -embedded $Z \rightarrow \mu^+\mu^-$ data sample described in Section 4. The normalization for this process is taken from the final fit described in Section 12 where it is a free parameter. These normalizations are independent for the $\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$, and $\tau_{\text{had}}\tau_{\text{had}}$ channels.

9.3 $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ Background

The Drell-Yan $Z/\gamma^* \rightarrow e^+e^-, \mu^+\mu^-$ background is important for the final states with two same-flavour leptons. It also makes contributions to the other channels. As described below, ALPGEN MC simulation samples are used and correction factors to account for any discrepancies with data are applied.

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, the ALPGEN MC simulation is normalized to the data in the Z -mass control region, $80 \text{ GeV} < m_{\tau\tau}^{\text{vis}} < 100 \text{ GeV}$, for each category and for $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ events separately.

The $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ backgrounds are accounted for in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel using MC simulation, with separate treatments for τ_{had} candidates originating 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 electrons [54, 55]. This is particularly important for $Z \rightarrow e^+e^-$ events with a misidentified τ_{had} candidate originating from a true electron. In the second case, a dilepton control region is used, where the MC simulation is normalized to data. Different normalizations are used for the VBF and Boosted categories, due to differing kinematic cuts.

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

9.4 W+jets Background

Events with W bosons and jets are a background to all channels since the W boson can decay to e , μ or τ and feed into all signatures when accompanied by a jet which provides a falsely identified τ_{had} or lepton candidate. This process can also contribute when there is a semi-leptonic heavy quark decay that provides an identified lepton.

As stated in Section 9.1, in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel and $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the W +jets contributions are determined from data-driven methods. The $\tau_{\text{had}}\tau_{\text{had}}$ channel uses MC simulation samples to evaluate, both $W \rightarrow \tau_{\text{had}}\nu$ +jets and $W(\rightarrow \ell\nu)$ +jets backgrounds. The dominant contribution is from $W \rightarrow \tau_{\text{had}}\nu$ decays.

9.5 Top-quark Background

Backgrounds that arise from $t\bar{t}$ and single-top quark production are estimated from MC simulation in the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels. The normalization is obtained from data control regions defined by inverting the b -tag veto (and requiring large m_T for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel) to enhance the top backgrounds and suppress the signal contribution. This background is negligible for $\tau_{\text{had}}\tau_{\text{had}}$ but is included using MC simulation.

For the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, a correction to the $\Delta\phi_{\ell\ell}$ distribution is determined in the top-quark control region, defined by inverting the b -tag requirement, by comparing simulation with data, and applied to the MC events in the signal region. Better agreement is seen in all kinematic variables in the top-quark background control region after this correction. The b -tag jet multiplicity distribution in the $\tau_{\text{lep}}\tau_{\text{lep}}$ top-quark background control region, after this correction is applied, is shown in Figure 6 for the VBF and boosted categories.

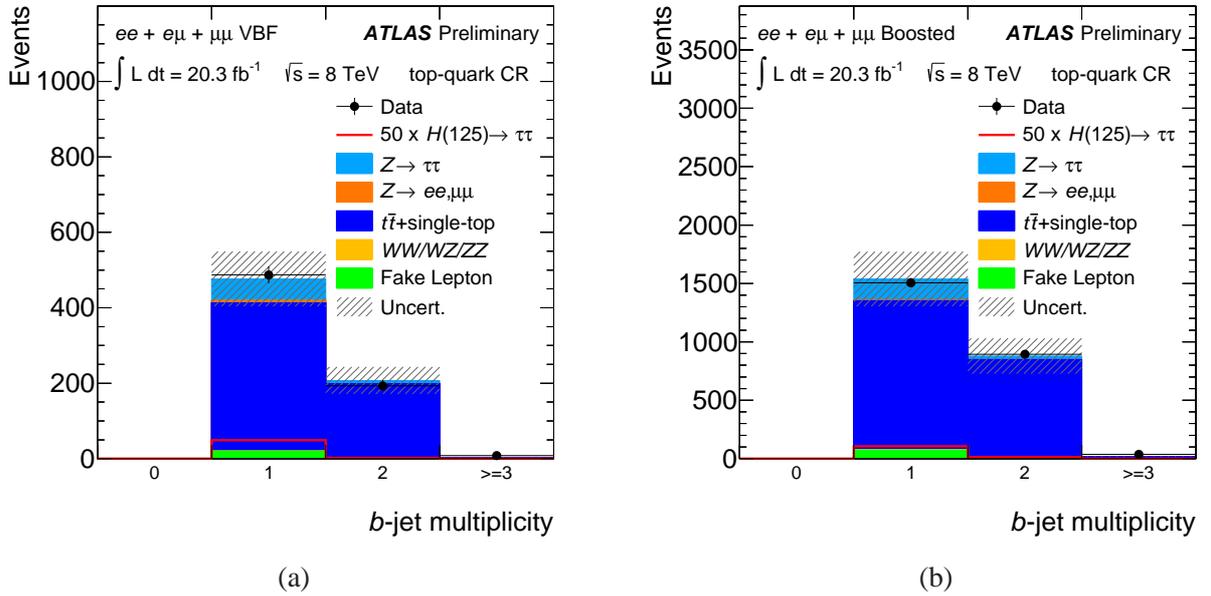


Figure 6: The b -tag jet multiplicity distributions, for jets with $p_T > 25$ GeV, for the top-quark background control regions in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel for the (a) VBF and (b) boosted categories. These figures use background predictions made without the global fit defined in Section 12. The signal shapes are shown multiplied by a factor of 50.

9.6 Diboson Background

The production of W^+W^- , ZZ and $W^\pm Z$ pairs with subsequent decays to leptons or jets contributes to the background, especially in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel. In each channel, the contribution is estimated from MC simulation.

9.7 Overlap with Other Higgs Boson Decays

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, there is a non-negligible contribution from $H \rightarrow W^+W^-$ with both W bosons decaying leptonically. Therefore, the process $H \rightarrow W^+W^-$, with a Higgs boson mass at $m_H = 125$ GeV, is considered as a background in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, and is estimated using MC simulation. The cross section for this process, as for the signal, is taken from the calculation in Ref. [9].

10 Background Model Validation

The background model described in Section 9 was decided upon before looking into bins sensitive to the presence of signal, in order to ensure that the presence (or absence) of $H \rightarrow \tau^+\tau^-$ events plays no role in the final background estimates.

Since each channel has a different background composition, separate control regions are defined for each channel in order to validate the background model.

The following control regions are defined for the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, modifying the signal regions with the following cuts:

- **$Z \rightarrow \ell\ell$ -enriched:** for same-flavour events, the $m_{\tau\tau}^{\text{vis}}$ selection is changed to $80 \text{ GeV} < m_{\tau\tau}^{\text{vis}} < 100 \text{ GeV}$.
- **$Z \rightarrow \tau\tau$ -enriched:** $m_{\tau\tau}^{\text{HPTO}} < 100 \text{ GeV}$, where $m_{\tau\tau}^{\text{HPTO}}$ is the invariant mass, obtained using the collinear approximation, calculated only with high p_T objects.
- **$t\bar{t}$ -enriched:** invert b -jet veto.
- **Fake-enriched:** same sign lepton events.
- **Low BDT score.**

The following control regions have been defined for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, modifying the signal regions with the following cuts:

- **$Z \rightarrow \tau\tau$ -enriched:** $m_T < 40 \text{ GeV}$ and $m_{\tau\tau}^{\text{MMC}} < 110 \text{ GeV}$.
- **W -enriched:** $m_T > 70 \text{ GeV}$.
- **$t\bar{t}$ -enriched:** invert b -jet veto and $m_T > 50 \text{ GeV}$.
- **Low BDT score.**

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the following regions are used, modifying the signal regions with the following cuts:

- **Mass sideband:** $m_{\tau\tau}^{\text{MMC}} < 100 \text{ GeV}$ or $m_{\tau\tau}^{\text{MMC}} > 140 \text{ GeV}$.
- **Multijet-enriched:** inverted signal region $\Delta\eta$ selection: $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}}) > 1.5$.

- **Rest category:** events that pass preselection but fail the VBF and boosted category selections. This region is used in the global likelihood fit to determine the $Z \rightarrow \tau^+\tau^-$ and the multijet background normalizations.
- **Low BDT score.**

The $Z \rightarrow \tau\tau$ -enriched and mass sideband control regions are not orthogonal to the signal region. Nevertheless, the signal is suppressed strongly through the additional criteria imposed for these control regions. The low BDT score region is defined as the background-dominated region of the BDT distribution where 30% or less of the signal yield would be located.

The BDT score, $m_{\tau\tau}^{\text{MMC}}$ and other distributions in these background-dominated control regions were scrutinized closely. Other kinematic variables for the selected objects and the BDT input variables were also examined in the entire signal region. Figure 7 shows the BDT score distributions for the $Z \rightarrow \ell\ell$ -enriched control region in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, the $Z \rightarrow \tau\tau$ -enriched control region in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, and the mass sideband control regions for the $\tau_{\text{had}}\tau_{\text{had}}$ channel.

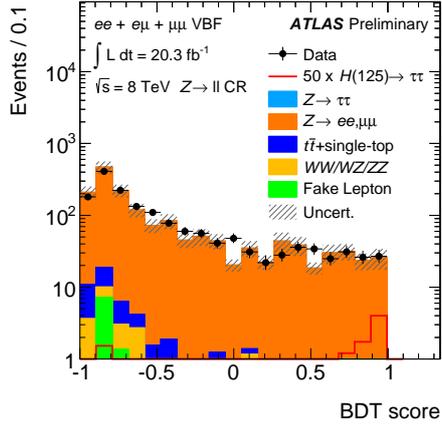
11 Systematic Uncertainties

Systematic uncertainties are evaluated for the following effects in signal and background samples. Their inclusion in the profile likelihood global fit is described in Section 12.

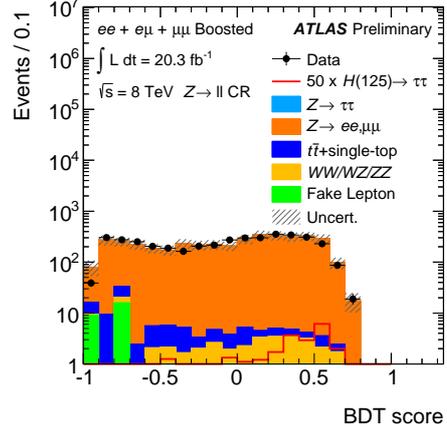
11.1 Experimental Systematic Uncertainties

The following experimental systematic uncertainties on the event yields and BDT distributions are considered, and are treated either as fully correlated or uncorrelated across channels and categories as appropriate.

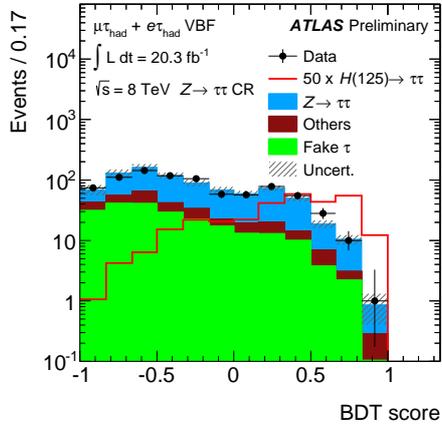
- **Luminosity:** The uncertainty on the integrated luminosity is $\pm 2.8\%$. It is derived following the same methodology as that detailed in Ref. [56], from a preliminary calibration of the luminosity scale derived from beam-separation scans performed in November 2012.
- **Tau identification efficiency:** The MC simulation is corrected, as described in Ref. [49], so that the τ_{had} identification efficiency matches that measured in the data. The corrections and their corresponding uncertainties are obtained using tag and probe studies. To assign a systematic uncertainty to this evaluation, the τ_{had} identification efficiency is varied within these uncertainties.
- **Tau energy scale (TES):** The TES calculation and its uncertainty are described in Ref. [57]. To obtain the systematic uncertainty due to the TES, upwards and downwards variations are applied to the τ_{had} energy scale. These systematic variations are also propagated to the $E_{\text{T}}^{\text{miss}}$ measurement.
- **Jet energy:** The jet energy scale (JES) uncertainties arise from several independent sources [58]. Since several backgrounds are evaluated from data, the JES uncertainty enters mainly through the lesser backgrounds which are determined from MC simulation. The systematic uncertainty due to the jet energy resolution (JER) is obtained by smearing every jet by the uncertainty in the resolution as determined by the in-situ measurement described in Ref. [59].
- **b -tagging efficiency:** The b -tagging efficiencies are corrected to match those measured in data, using a sample of jets containing muons [46]. Systematic uncertainties on these corrections are considered.



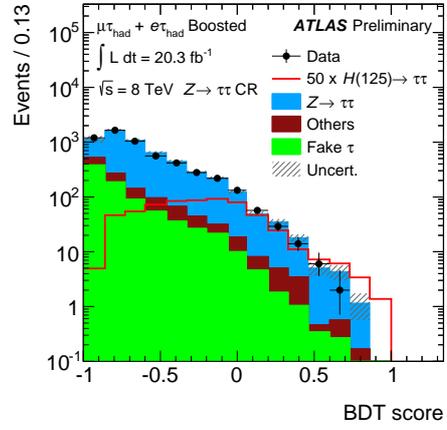
(a)



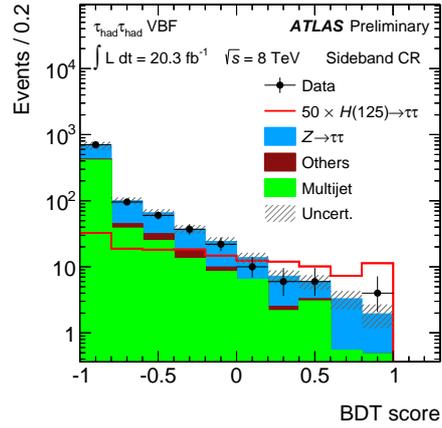
(b)



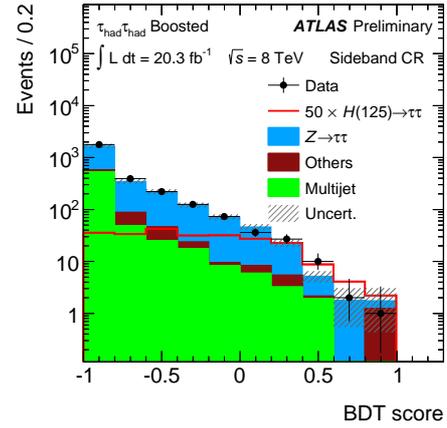
(c)



(d)



(e)



(f)

Figure 7: BDT score distributions for the $Z \rightarrow \ell\ell$ -enriched control region in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel (top), $Z \rightarrow \tau\tau$ -enriched control region in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel (middle), and mass sideband control region in the $\tau_{\text{had}}\tau_{\text{had}}$ channel (bottom), for the VBF (left) and Boosted (right) categories. The signal shapes are shown multiplied by a factor of 50. These figures use background predictions made without the global fit defined in Section 12.

- Lepton energy/momentum resolution: The electron energy and the muon momentum are smeared according to the resolutions measured in data. This systematic uncertainty plays a very minor role on the final result.
- Lepton reconstruction/identification/isolation/trigger efficiency: Corrections are obtained for each of these types of efficiencies using tag-and-probe measurements. These are then applied to MC samples to ensure agreement with data. The uncertainties are propagated to the final result by varying the corrections by one standard deviation.
- E_T^{miss} : Systematic uncertainties on the energy scales of all objects affect the E_T^{miss} and so it is recalculated after each of these variations is applied. In addition, the scale of the soft E_T^{miss} term for energy outside reconstructed objects and the resolution uncertainties are considered independently [51].
- Embedding method: The selection of the $Z \rightarrow \mu^+\mu^-$ events for the embedding method contributes to a systematic uncertainty related to the isolation requirement which is applied to the muons. The uncertainty is calculated by varying the isolation requirement. In addition, the uncertainty from the subtraction of the calorimeter cell energy associated with the muons is considered [13].

11.2 Background Modelling Uncertainties

This section details the most significant systematic uncertainties on the background estimation techniques described in Section 9.

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, systematic uncertainties on the shape and normalization of fake lepton backgrounds are estimated from comparisons of the differences in same-sign lepton events that pass and fail the lepton isolation criteria. The overall $Z \rightarrow \tau^+\tau^-$ background normalization is left free in the likelihood fit described in Section 12, but an additional systematic is applied to the predicted difference in normalization between different lepton flavour and same lepton flavour events. The uncertainty on the $Z \rightarrow \ell\ell$ normalization is obtained by varying the cuts on the $m_{\tau\tau}^{\text{vis}}$ window that define the control region for this background. The uncertainty on the normalization of top-quark backgrounds is derived from the difference of event yields in observed events and predictions from MC simulation in the top-quark control regions.

In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, an important systematic uncertainty on the background estimation comes from the estimated fake-factors, based upon the difference in the determination of these factors in two different regions dominated by gluon- and quark-initiated jets, respectively. In addition, the uncertainties on the normalization of $Z \rightarrow \ell\ell$ and top-quark backgrounds, obtained from the global fit described in Section 12 are important. Dominant uncertainties associated with the normalization of these backgrounds come from systematic uncertainties on b -tagging efficiencies and the jet energy scale, along with statistical uncertainties on the observed data in the respective control regions.

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the multijet template, derived from the notOS sample, is compared with a same-sign sample and the difference in both the shape and normalization obtained from the $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}})$ fit is taken as a systematic uncertainty.

11.3 Theoretical Uncertainties

Theoretical systematic uncertainties are estimated for signal samples and for any backgrounds modelled using MC simulations.

Uncertainties on the signal cross sections are assigned from four sources: QCD scale uncertainty, modelling of differential cross section in p_T^H , modelling of the underlying event, and uncertainties due to the choice of PDFs.

Uncertainties resulting from the choice of QCD scale are calculated following the prescription extensively discussed in Refs. [60] and [61]. This prescription has been shown to provide a realistic estimate of higher order contributions to analyses which classify events by the number of jets. In order to apply the procedure, the scale uncertainties for the inclusive multijet cross sections need to be computed together with the event fractions in each analysis category. While the inclusive cross section uncertainty for the ggF process is only 8%, the uncertainty in the exclusive bins used in this analysis ranges from 22–26% and can be as high as 74% for boosted events with at least two jets.

A theoretical systematic uncertainty is also assigned due to modelling of the ggF differential cross section $d\sigma/dp_T^H$. The default $gg \rightarrow H$ simulation in ATLAS is performed using the POWHEG generator. The MC@NLO generator implements the same matrix element but quark mass effects on $d\sigma/dp_T^H$ of the two generators are known to be quite different. The difference between the output of these two generators is assigned as an additional systematic uncertainty. Care has been taken to compare these two calculations at the same QCD scale, to avoid double-counting the uncertainty from that source. Variations of 17%–29% are observed depending on channel and category.

Underlying event effects have been computed comparing the Perugia 2011C underlying event tuning with the AUET2B tune [62]. Discrepancies of $\approx 6\%$ and $\approx 30\%$ for the signal yields in VBF and ggF production in the VBF category have been found, and are assigned as underlying event uncertainties.

A systematic uncertainty due to the choice of PDF set is also applied. MCFM [63] was used to verify that the variation of the differential cross sections associated with each analysis category due to different PDF sets (comparing MSTW [64], NNPDF [65] and CT10) are smaller than or equal to the inclusive variation. A PDF uncertainty of 7.5% is assigned in all categories for gluon fusion production, and 2.8% is assigned for VBF and VH production.

While most major backgrounds are normalized in data control regions, some smaller background sources are normalized by their theoretical cross sections. QCD scale and PDF uncertainties are assigned on these minor backgrounds and typically range from 3–4%.

12 Signal Extraction Procedure

The parameter of interest in this analysis is the signal strength parameter μ which is defined as the ratio of the measured cross section times branching ratio normalized to the Standard Model cross section times the branching ratio for $H \rightarrow \tau^+\tau^-$ [9]. 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 inputs to the determination of the maximum likelihood (global fit) are as follows: BDT score distributions in the six signal regions (VBF and boosted categories for each channel), event yields in the top and $Z \rightarrow \ell\ell$ control regions for the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels in both VBF and boosted categories, and the Rest category control region for the $\tau_{\text{had}}\tau_{\text{had}}$ channel, binned in $\Delta\eta(\tau_{\text{had}}, \tau_{\text{had}})$.

The impact of systematic uncertainties on the signal and background expectations is described by nuisance parameters, $\vec{\theta}$, which are parametrized 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: $q_\mu = -2 \ln(\mathcal{L}(\mu, \hat{\vec{\theta}})/\mathcal{L}(\hat{\mu}, \hat{\vec{\theta}}))$, where $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the parameters that maximize the likelihood, and $\hat{\vec{\theta}}$ are the nuisance parameter values that maximize 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 $Z \rightarrow \tau^+\tau^-$ background rate is allowed to float freely in the global fit, and is constrained primarily in the signal regions, due to the difference between the BDT distribution for $Z \rightarrow \tau^+\tau^-$ compared to the

signal. The top and $Z \rightarrow \ell\ell$ background rates 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 included in the fit. For the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the $Z \rightarrow \tau^+\tau^-$ and multijet background rates are constrained from the simultaneous fit of the two signal regions and the Rest category control region. The normalizations of other backgrounds are constrained within their uncertainties as described in Section 11.

Figure 8 shows the BDT score distributions for all signal regions in each channel, with background normalizations, signal normalization, and nuisance parameters adjusted by the profile likelihood global fit. Agreement between data and the model predictions for background together with a Standard Model Higgs boson of $m_H = 125$ GeV is observed within the uncertainties. The fitted numbers of signal and background events in the three rightmost bins in each category and channel are shown in Tables 4–6, together with the observed number of events in the data.

13 Results

The fitted value of the signal strength parameter for each individual channel and for the combined result, all obtained from the global fit, are shown in Figure 9. The fitted value of the signal strength parameter from the likelihood fit is $\mu = 1.43_{-0.29}^{+0.31}(\text{stat.})_{-0.30}^{+0.41}(\text{syst.})$ for $m_H = 125$ GeV. Theory systematic uncertainties on μ , which are included in the quoted systematic uncertainty, amount to $_{-0.15}^{+0.26}$. The dependence of $-2 \ln \mathcal{L}(\mu, \hat{\theta})$ on the signal strength parameter is shown in Figure 10. The likelihood as a function of the signal strength for the individual channels is also shown.

The probability p_0 of obtaining a result at least as signal-like as observed in the data if no signal is present is calculated using the test statistic $q_{\mu=0} = -2 \ln(\mathcal{L}(0, \hat{\theta})/\mathcal{L}(\hat{\mu}, \hat{\theta}))$ in the asymptotic approximation [66]. For $m_H=125$ GeV, the observed p_0 value is 2.0×10^{-5} , which corresponds to a deviation from the background-only hypothesis of 4.1 standard deviations. This can be compared to an expected p_0 value of 6.6×10^{-4} (3.2 standard deviations). This is direct evidence for $H \rightarrow \tau^+\tau^-$ decays.

Figure 11 shows the expected and observed data, in bins of $\log(S/B)^7$, for all signal region bins. Here, S/B is the signal-to-background ratio calculated assuming $\mu = 1$ 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, are shown 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 visualize the compatibility of this excess of events above background predictions, with the SM Higgs boson at $m_H = 125$ GeV, a weighted distribution of events as a function of $m_{\tau\tau}^{\text{MMC}}$ is shown in Figure 12. The events are weighted by a factor of $\ln(1 + S/B)$. 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 with $m_H = 110$ GeV and $m_H = 150$ GeV. This illustrates that the signal is compatible with a Higgs boson of 125 GeV mass.

The dominant uncertainties on the measurement of the signal strength parameters include: statistical uncertainties in the data from the signal regions, the theoretical systematic uncertainty on the differential cross section $d\sigma/dp_T^H$ for ggF production, the uncertainty on the normalization of the $Z \rightarrow \ell\ell$ and top backgrounds in the $\tau_{\text{lep}}\tau_{\text{had}}$ VBF and boosted categories, and the uncertainty on the JES calibration in different detector η regions. Table 7 shows the contribution of each of the significant sources to the uncertainty of the measured signal strength, by assessing their impact $\Delta\mu$ on the measured value of μ , when varying the effect by one standard deviation.

⁷In this document, log always refers to log base 10.

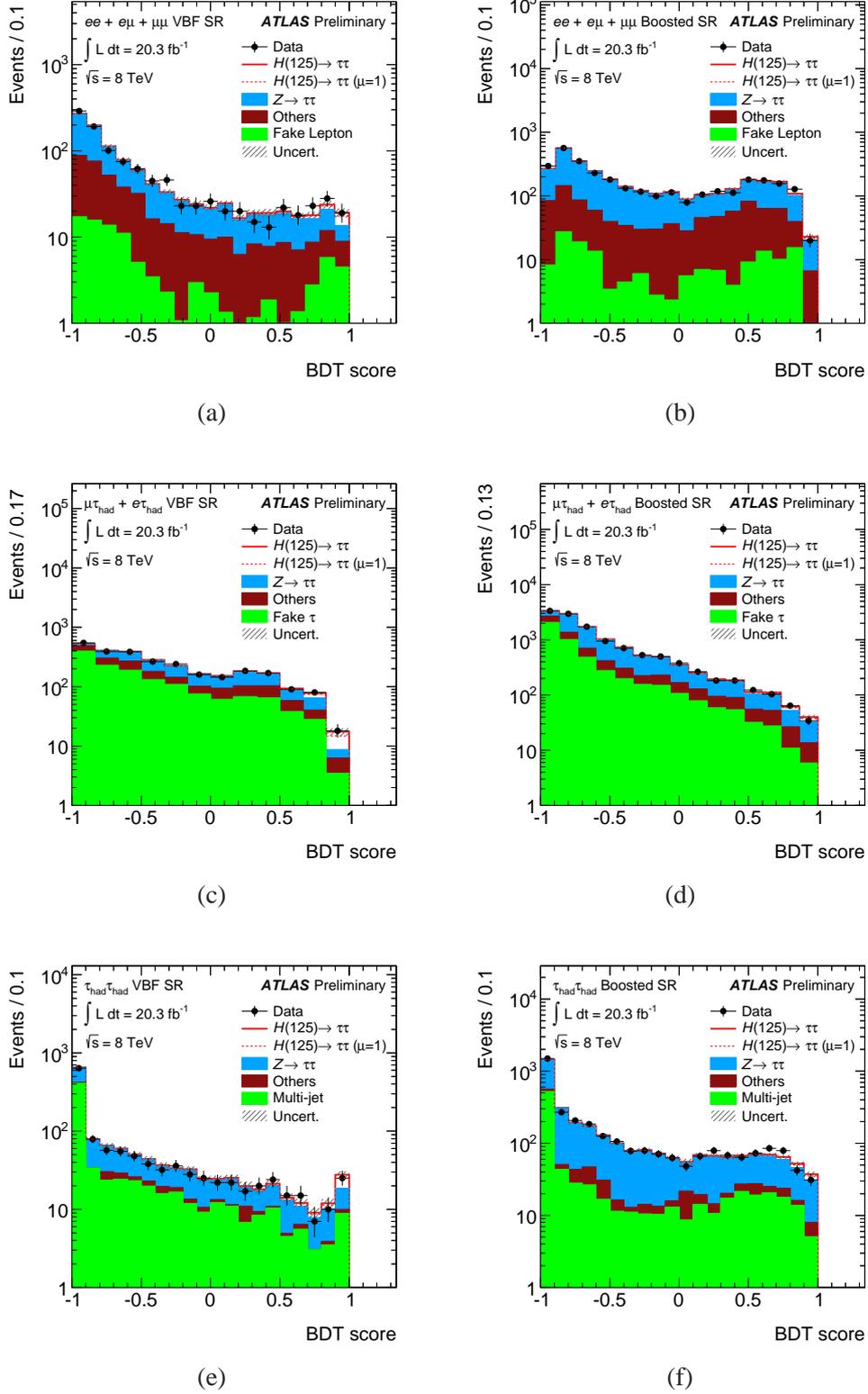


Figure 8: BDT distributions for the VBF (left) and boosted (right) category signal regions for the $\tau_{lep}\tau_{lep}$ (top), $\tau_{lep}\tau_{had}$ (middle), and $\tau_{had}\tau_{had}$ (bottom) channels. The Higgs boson signal ($m_H = 125 \text{ GeV}$) is shown stacked with a signal strength of $\mu = 1$ (dashed line) and $\mu = 1.4$ (solid line). The background predictions come from the global fit (that gives $\mu = 1.4$). The size of the statistical and normalization systematics is indicated by the hashed band. The binning in (e) and (f) is coarser than that used in the global fit and shown in Table 6.

Table 4: Predicted event yields in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel for $m_H = 125$ GeV in the three highest bins of the BDT distributions. The background normalizations, signal normalization, and uncertainties reflect the preferred values from the global fit. The uncertainties on the total background and total signal reflect the full statistical+systematic uncertainty, while the uncertainties on the individual background components reflect the full systematic uncertainty only.

Process/Category	VBF			Boosted		
BDT score bin edges	0.684-0.789	0.789-0.895	0.895-1.0	0.667-0.778	0.778-0.889	0.889-1.0
ggF	0.53 ± 0.26	0.8 ± 0.4	0.7 ± 0.4	5.3 ± 2.1	5.2 ± 2.0	1.7 ± 0.7
VBF	1.15 ± 0.35	2.0 ± 0.6	5.0 ± 1.5	1.01 ± 0.33	1.5 ± 0.5	0.67 ± 0.22
WH	< 0.05	< 0.05	< 0.05	0.71 ± 0.22	0.64 ± 0.20	0.16 ± 0.05
ZH	< 0.05	< 0.05	< 0.05	0.36 ± 0.11	0.32 ± 0.10	0.06 ± 0.02
$Z \rightarrow \tau^+\tau^-$	7.6 ± 0.8	9.0 ± 0.9	4.6 ± 0.6	97 ± 7	61.5 ± 3.2	13.6 ± 1.3
Fake	2.8 ± 0.7	5.8 ± 2.0	4.5 ± 1.7	10.1 ± 3.1	15 ± 5	0.79 ± 0.29
Top	4.0 ± 0.9	2.9 ± 0.7	1.8 ± 0.4	28 ± 7	15 ± 4	3.5 ± 0.9
Others	1.97 ± 0.26	3.3 ± 0.4	2.7 ± 0.4	24.7 ± 1.9	8.8 ± 0.6	2.34 ± 0.24
Total Background	16.3 ± 1.5	20.9 ± 2.4	13.5 ± 2.4	160 ± 7	101 ± 4	20.2 ± 1.8
Total Signal	1.7 ± 0.5	2.9 ± 0.9	5.7 ± 1.7	7.4 ± 2.4	7.7 ± 2.5	2.6 ± 0.8
S/B	0.10	0.14	0.42	0.05	0.08	0.13
Data	23	28	19	156	128	20

Table 5: Predicted event yields in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel for $m_H = 125$ GeV in the three highest bins of the BDT distributions. The background normalizations, signal normalization, and uncertainties reflect the preferred values from the global fit. The uncertainties on the total background and total signal reflect the full statistical+systematic uncertainty, while the uncertainties on the individual background components reflect the full systematic uncertainty only.

Process/Category	VBF			Boosted		
	BDT score bin edges	0.5-0.667	0.667-0.833	0.833-1.0	0.6-0.733	0.733-0.867
ggF	2.2 ± 0.9	3.5 ± 1.5	1.2 ± 0.6	7.7 ± 2.9	6.3 ± 2.3	5.5 ± 2.1
VBF	4.1 ± 1.2	9.2 ± 2.7	7.5 ± 2.2	1.7 ± 0.5	1.5 ± 0.5	1.3 ± 0.4
WH	< 0.05	< 0.05	< 0.05	0.95 ± 0.29	0.85 ± 0.26	0.81 ± 0.25
ZH	< 0.05	< 0.05	< 0.05	0.42 ± 0.13	0.47 ± 0.14	0.41 ± 0.12
$Z \rightarrow \tau^+\tau^-$	28.6 ± 1.4	25.0 ± 1.6	2.41 ± 0.35	48.3 ± 3.4	26.1 ± 2.7	18.4 ± 2.0
Fake	37.7 ± 1.8	27.9 ± 2.1	3.5 ± 0.5	27 ± 4	10.8 ± 1.8	5.8 ± 1.4
Top	6.5 ± 0.7	4.1 ± 0.8	1.5 ± 0.4	7.0 ± 0.9	5.7 ± 0.8	2.23 ± 0.33
Diboson	2.9 ± 0.4	3.0 ± 0.5	0.23 ± 0.04	4.8 ± 0.5	4.0 ± 0.5	1.69 ± 0.23
$Z \rightarrow \ell\ell(j \rightarrow \tau_{\text{had}})$	8.7 ± 1.7	3.3 ± 0.5	0.40 ± 0.10	3.8 ± 0.5	0.71 ± 0.07	< 0.05
$Z \rightarrow \ell\ell(\ell \rightarrow \tau_{\text{had}})$	2.8 ± 1.2	1.9 ± 1.2	0.7 ± 0.6	9.4 ± 1.9	4.9 ± 1.1	3.8 ± 1.2
Total Background	87.2 ± 2.7	65 ± 5	8.7 ± 2.5	101 ± 6	52 ± 4	32 ± 4
Total Signal	6.3 ± 1.8	12.7 ± 3.5	8.7 ± 2.4	10.7 ± 3.3	9.2 ± 2.8	8.0 ± 2.5
S/B	0.07	0.20	1.0	0.11	0.18	0.25
Data	90	80	18	103	64	34

Table 6: Predicted event yields in the $\tau_{\text{had}}\tau_{\text{had}}$ channel for $m_H = 125$ GeV in the three highest bins of the BDT distributions. The background normalizations, signal normalization, and uncertainties reflect the preferred values from the global fit. The uncertainties on the total background and total signal reflect the full statistical+systematic uncertainty, while the uncertainties on the individual background components reflect the full systematic uncertainty only.

Process/Category	VBF			Boosted		
	0.85-0.9	0.9-0.95	0.95-1.0	0.85-0.9	0.9-0.95	0.95-1.0
ggF	0.39 ± 0.17	0.35 ± 0.16	2.0 ± 0.9	2.2 ± 0.8	2.5 ± 1.0	2.3 ± 0.9
VBF	0.57 ± 0.18	0.72 ± 0.22	5.9 ± 1.8	0.55 ± 0.17	0.61 ± 0.19	0.57 ± 0.17
WH	< 0.05	< 0.05	< 0.05	0.34 ± 0.11	0.40 ± 0.12	0.44 ± 0.14
ZH	< 0.05	< 0.05	< 0.05	0.22 ± 0.07	0.22 ± 0.07	0.22 ± 0.07
$Z \rightarrow \tau^+\tau^-$	3.2 ± 0.6	3.4 ± 0.7	5.3 ± 1.0	15.7 ± 1.7	12.3 ± 1.8	9.7 ± 1.6
Multijet	3.3 ± 0.6	2.9 ± 0.6	5.9 ± 0.9	5.2 ± 0.6	3.7 ± 0.5	1.40 ± 0.22
Others	0.38 ± 0.09	0.49 ± 0.12	0.64 ± 0.13	1.49 ± 0.27	2.8 ± 0.5	0.07 ± 0.02
Total Background	6.9 ± 1.3	6.8 ± 1.3	11.8 ± 2.6	22.4 ± 2.5	18.8 ± 2.8	11.2 ± 1.9
Total Signal	0.97 ± 0.29	1.09 ± 0.31	8.0 ± 2.2	3.3 ± 1.0	3.8 ± 1.2	3.6 ± 1.1
S/B	0.14	0.16	0.67	0.15	0.2	0.32
Data	6	6	19	20	16	15

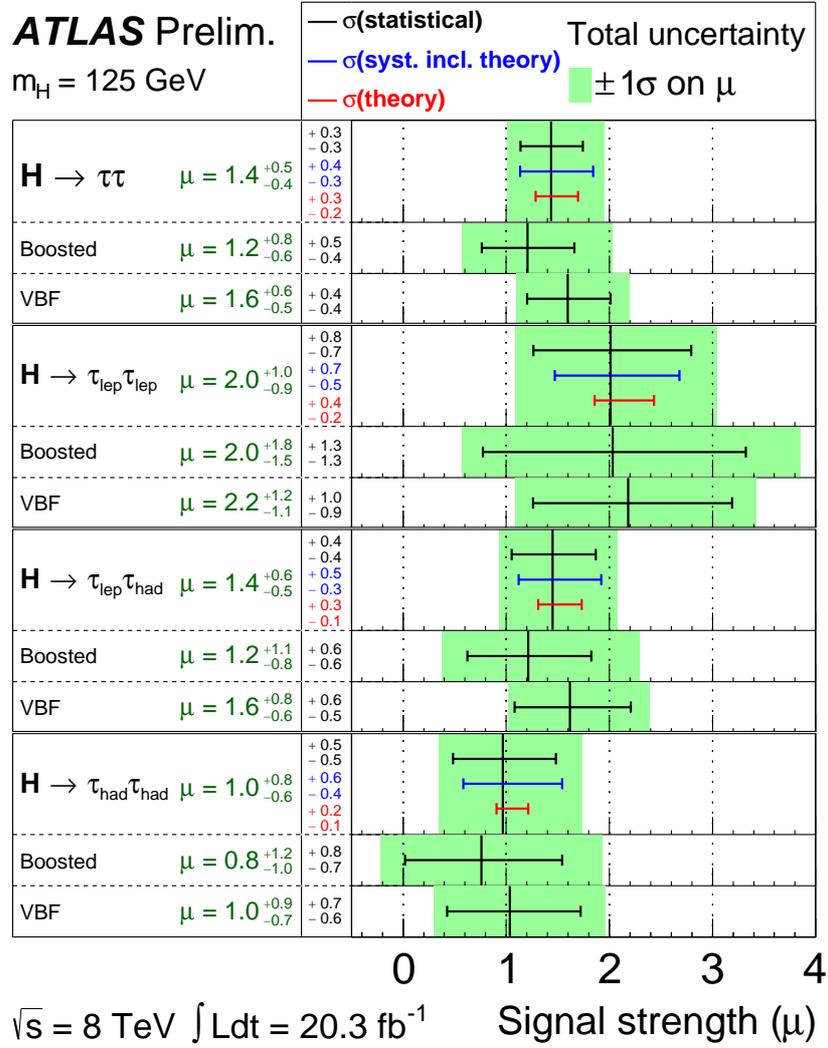


Figure 9: The best-fit value for the signal strength μ in the individual channels and the combination. The total $\pm 1\sigma$ uncertainty is indicated by the shaded green band, with the individual contributions from the statistical uncertainty (top, black), the total (experimental and theoretical) systematic 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.

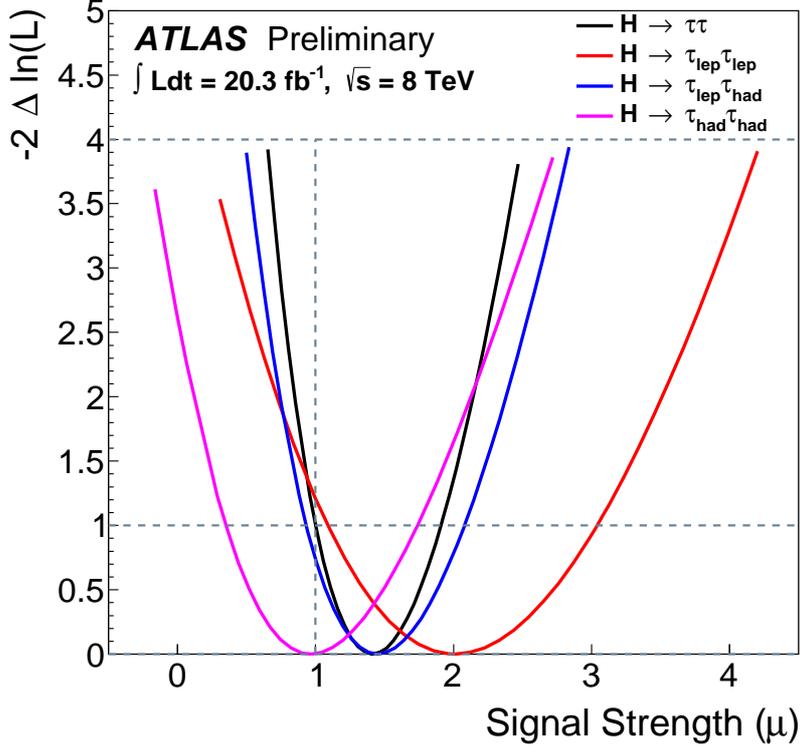


Figure 10: The likelihood dependence on the the signal strength parameter μ , where the y -axis is the deviation from the maximum likelihood: $-2\Delta \ln \mathcal{L}$.

The normalization uncertainties for $Z \rightarrow \ell\ell$ and top-quark backgrounds in the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels are decorrelated across categories, while the $Z \rightarrow \tau^+\tau^-$ normalization uncertainties are correlated for the categories in each respective channel. The global fit also constrains the normalization for $Z \rightarrow \tau^+\tau^-$ more strongly than for the $Z \rightarrow \ell\ell$ and top-quark backgrounds, as the low BDT score region is dominated by $Z \rightarrow \tau^+\tau^-$ events.

Two-dimensional contours in the plane of $\mu_{\text{ggF}} \times B/B_{\text{SM}}$ and $\mu_{\text{VBF+VH}} \times B/B_{\text{SM}}$ [67] are shown in Figure 13 for $m_H = 125$ GeV, where B and B_{SM} are the hypothesised and the SM branching ratios for $H \rightarrow \tau^+\tau^-$. The best-fit values are $\mu_{\text{ggF}} \times B/B_{\text{SM}} = 1.1^{+1.3}_{-1.0}$ and $\mu_{\text{VBF+VH}} \times B/B_{\text{SM}} = 1.6^{+0.8}_{-0.7}$.

14 Conclusions

A search for a Higgs boson of mass 125 GeV decaying into the $\tau^+\tau^-$ final state has been performed using the full ATLAS 2012 dataset, corresponding to an integrated luminosity of $\mathcal{L} = 20.3 \text{ fb}^{-1}$ of pp collisions at a centre-of-mass energy of 8 TeV. A signal has been measured with a significance of 4.1 standard deviations, compared with an expected significance of 3.2 standard deviations. This constitutes direct evidence of the decay of the Higgs boson to fermions. The observed signal strength $\mu = 1.4^{+0.5}_{-0.4}$ is compatible with the Standard Model expectation.

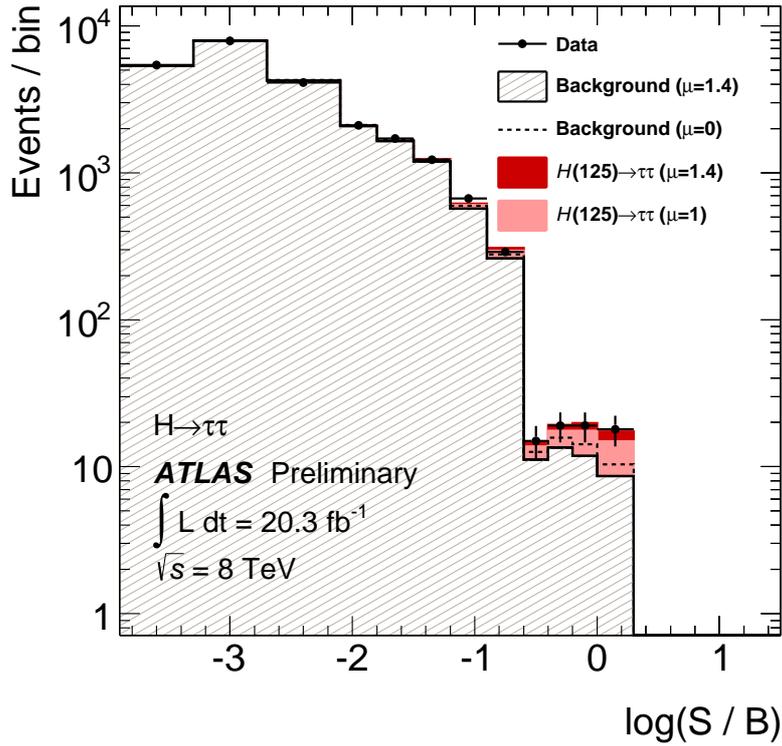


Figure 11: Event yields as a function of $\log(S/B)$, where S (signal yield) and B (background yield) are taken from each event's bin in the BDT. All events in the BDTs are included. The predicted background is obtained from the global fit (with $\mu = 1.4$), and signal yields are shown for $m_H = 125$ 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)	0.30
$Z \rightarrow \ell\ell$ normalization ($\tau_{\text{lep}}\tau_{\text{had}}$ boosted)	0.13
$ggF d\sigma/dp_T^H$	0.12
JES η calibration	0.12
Top normalization ($\tau_{\text{lep}}\tau_{\text{had}}$ VBF)	0.12
Top normalization ($\tau_{\text{lep}}\tau_{\text{had}}$ boosted)	0.12
$Z \rightarrow \ell\ell$ normalization ($\tau_{\text{lep}}\tau_{\text{had}}$ VBF)	0.12
QCD scale	0.07
di- τ_{had} trigger efficiency	0.07
Fake backgrounds ($\tau_{\text{lep}}\tau_{\text{lep}}$)	0.07
τ_{had} identification efficiency	0.06
$Z \rightarrow \tau^+\tau^-$ normalization ($\tau_{\text{lep}}\tau_{\text{had}}$)	0.06
τ_{had} energy scale	0.06

Table 7: The important sources of uncertainty on the measured signal strength parameter μ , given as absolute uncertainties on μ .

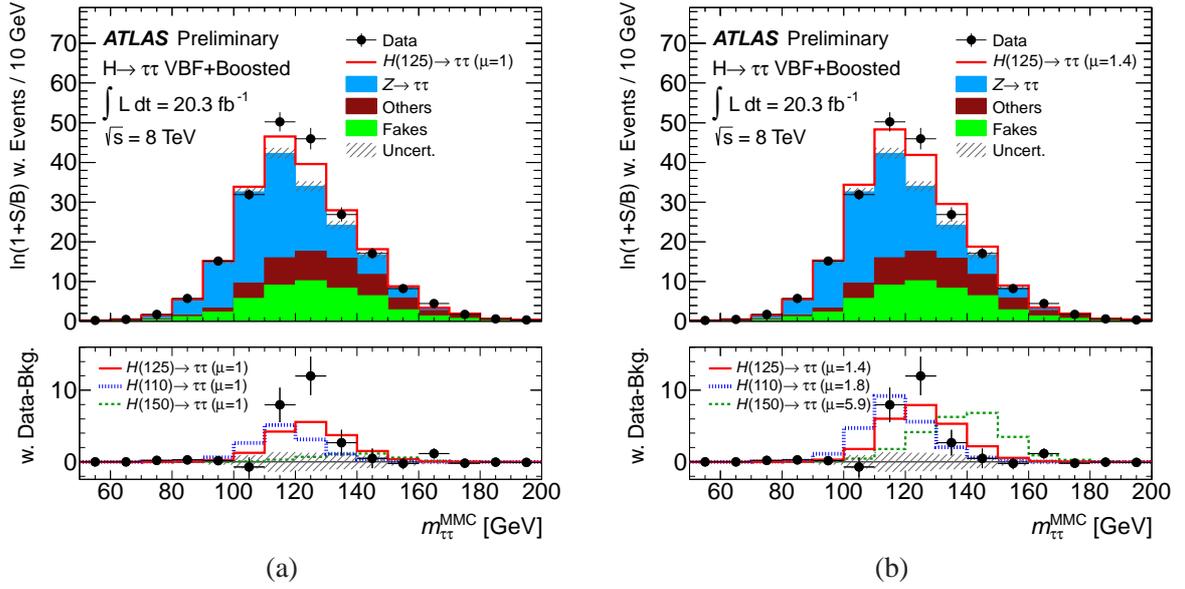


Figure 12: Distributions for $m_{\tau\tau}^{\text{MMC}}$ where events are weighted by $\ln(1 + S/B)$ for all channels. These weights are determined by the signal and background predictions for each BDT bin. The bottom panel in each plot shows the difference between weighted data events and weighted background 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, the $m_H = 110$ GeV (blue) and $m_H = 150$ GeV (green) signals are also shown. (a) Shows all signal strengths set to the Standard Model expectation while (b) shows the signal strengths set to their best fit values.

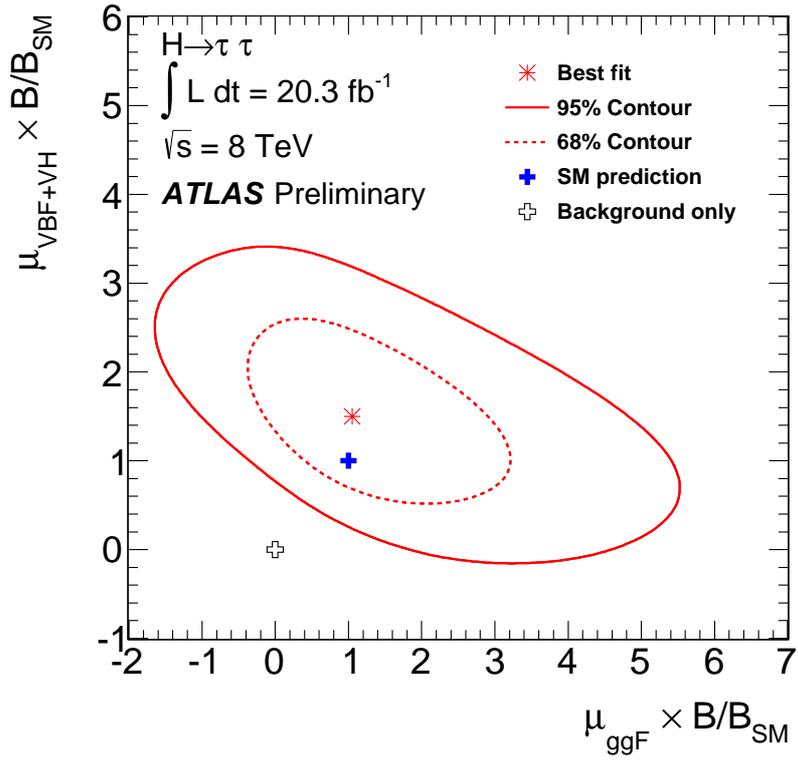


Figure 13: Likelihood contours for the $H \rightarrow \tau\tau$ channel in the $(\mu_{\text{ggF}} \times B/B_{\text{SM}}, \mu_{\text{VBF+VH}} \times B/B_{\text{SM}})$ plane are shown for the 68% and 95% CL by dashed and solid lines, respectively, for $m_H = 125 \text{ GeV}$. The SM expectation and the one corresponding to background-only hypothesis are shown by a filled plus and an open plus symbol, respectively. The best-fit to the data is shown for the case when both the μ_{ggF} and $\mu_{\text{VBF+VH}}$ are unconstrained.

References

- [1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv:1207.7214 [hep-ex].
- [2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv:1207.7235 [hep-ex].
- [3] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [4] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.
- [5] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.
- [7] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. D **145** (1966) 1156.
- [8] T. W. B. Kibble, *Symmetry breaking in non-Abelian gauge theories*, Phys. Rev. **155** (1967) 1554.
- [9] S. Dittmaier et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, arXiv:1101.0593 [hep-ph].
- [10] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [11] S. Alioli et al., *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **0904** (2009) 002, arXiv:0812.0578 [hep-ph].
- [12] 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 [hep-ph].
- [13] ATLAS Collaboration, *Search for the Standard Model Higgs boson in $H \rightarrow \tau^+\tau^-$ decays in proton-proton collisions with the ATLAS detector*, ATLAS-CONF-2012-160. <https://cds.cern.ch/record/1493624>.
- [14] CMS Collaboration, *Search for the Standard-Model Higgs boson decaying to tau pairs in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV*, CMS-PAS-HIG-13-004. <https://cds.cern.ch/record/1528271>.
- [15] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Chapman & Hall, New York, 1984.
- [16] J. Friedman, *Stochastic gradient boosting*, Comput. Stat. Data Anal. **38** (2002) 367.
- [17] 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.
- [18] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2010*, Eur. Phys. J. C **72** (2010) 1849.

- [19] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv:1005.4568 [hep-ex].
- [20] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.
- [21] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP (2004) 040, arXiv:hep-ph/0409146.
- [22] 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 [hep-ph].
- [23] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **1006** (2010) 043, arXiv:1002.2581 [hep-ph].
- [24] S. Catani et al., *Soft gluon resummation for Higgs boson production at hadron colliders*, JHEP **0307** (2003) 028, arXiv:hep-ph/0306211.
- [25] E. Bagnaschi et al., *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*, JHEP **1202** (2012) 088, arXiv:1111.2854 [hep-ph].
- [26] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.
- [27] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820 [hep-ph].
- [28] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002) 012, arXiv:hep-ph/0201195.
- [29] H-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ph].
- [30] G. Corcella et al., *HERWIG 6.5 release note*, arXiv:hep-ph/0210213.
- [31] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. C **72** (1996) 637, arXiv:hep-ph/9601371.
- [32] M. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:hep-ph/0206293.
- [33] 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, arXiv:hep-ph/0108069.
- [34] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, arXiv:hep-ph/0204244.
- [35] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*, arXiv:hep-ph/0405247.
- [36] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer, *Gluon-induced W-boson pair production at the LHC*, JHEP **0612** (2006) 046, arXiv:hep-ph/0611170.
- [37] S. Jadach et al., *The tau decay library TAUOLA: Version 2.4*, Comput. Phys. Commun. **76** (1993) 361.

- [38] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS Interface in C++: Technical and Physics Documentation*, arXiv:1011.0937 [hep-ph].
- [39] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. C **72** (2012) 1909, arXiv:1110.3174 [hep-ex].
- [40] ATLAS Collaboration, *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data*, ATLAS-CONF-2013-088. <https://cds.cern.ch/record/1580207>.
- [41] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].
- [42] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. B **641** (2006) 57, arXiv:hep-ph/0512210.
- [43] W. Lampl et al., *Calorimeter Clustering Algorithms : Description and Performance*, ATL-LARG-PUB-2008-002, ATL-LARG-PUB-2008-003. <https://cds.cern.ch/record/1099735>.
- [44] 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 [hep-ex].
- [45] ATLAS Collaboration, *Commissioning of high performance b-tagging algorithms with the ATLAS detector*, ATLAS-CONF-2011-102. <http://cds.cern.ch/record/1369219>.
- [46] 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. <http://cds.cern.ch/record/1435197/>.
- [47] ATLAS Collaboration, *Search for supersymmetry in final states with jets, missing transverse momentum and a Z boson at $\sqrt{s} = 8\text{ TeV}$ with the ATLAS detector*, ATLAS-CONF-2012-152. <http://cds.cern.ch/record/1493491>.
- [48] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment – Detector, Trigger and Physics*, CERN-OPEN-2008-020, arXiv:0901.0512 [hep-ex].
- [49] ATLAS Collaboration, *Identification of Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, ATLAS-CONF-2013-064. <https://cds.cern.ch/record/1562839>.
- [50] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 7\text{ TeV}$ with ATLAS*, Eur. Phys. J. C **72** (2012) 1, arXiv:1108.5602 [hep-ex].
- [51] 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. <https://cds.cern.ch/record/1570993>.
- [52] A. Elagin, P. Murat, A. Pranko, and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, arXiv:1012.4686 [hep-ex].

- [53] R. Ellis et al., *Higgs decay to $\tau^+\tau^-$: A possible signature of intermediate mass Higgs bosons at high energy hadron colliders*, Nucl. Phys. B **297** (1988) 221.
- [54] 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 [hep-ex].
- [55] ATLAS Collaboration, *Measurement of the $Z \rightarrow \tau\tau$ Cross Section with the ATLAS Detector*, Phys. Rev. D **84** (2011) 112006, arXiv:1108.2016 [hep-ex].
- [56] 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 [hep-ex].
- [57] 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. <https://cds.cern.ch/record/1544036>.
- [58] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS 2011 data*, ATLAS-CONF-2013-004. <https://cds.cern.ch/record/1509552>.
- [59] 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 [hep-ex].
- [60] S. Dittmaier et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, arXiv:1201.3084 [hep-ph].
- [61] I. Stewart and F. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, Phys. Rev. D. **85** (2012) , arXiv:1107.2117 [hep-ph].
- [62] ATLAS Collaboration, *ATLAS tunes of PYTHIA6 and Pythia8 for MC11*, ATL-PHYS-PUB-2011-009. <https://cdsweb.cern.ch/record/1363300>.
- [63] J. M. Campbell, R. K. Ellis, and C. Williams, *Vector boson pair production at the LHC*, JHEP **1107** (2011) 018, arXiv:1105.0020 [hep-ph].
- [64] A. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, arXiv:0901.0002 [hep-ph].
- [65] NNPDF Collaboration, R. Ball et al., *Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology*, Nucl. Phys. B **849** (2011) 296, arXiv:1101.1300 [hep-ph].
- [66] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv:1007.1727 [physics.data-an].
- [67] A. David et al. (LHC Higgs Cross Section Working Group), *LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle*, arXiv:1209.0040 [hep-ph].
- [68] 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 [hep-ph].
- [69] ATLAS Collaboration, *Search for the bb decay of the Standard Model Higgs boson in associated W/ZH production with the ATLAS detector*, ATLAS-CONF-2013-079. <https://cds.cern.ch/record/1563235>.

A Additional Figures

A.1 BDT Input Variables

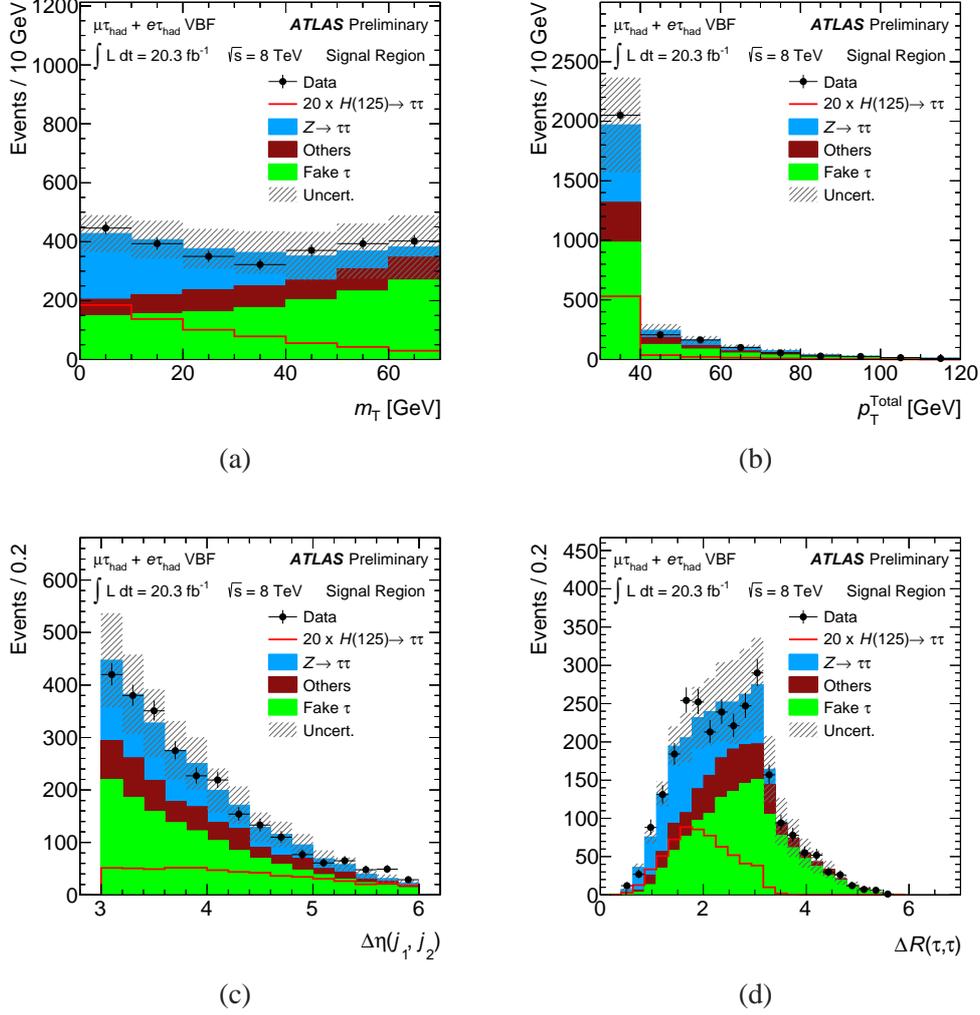


Figure 14: BDT input variables for the $\tau_{\text{lep}}\tau_{\text{had}}$ VBF category obtained in the signal region. The following distributions are shown: (a) m_T , (b) p_T^{Total} , (c) $\Delta\eta(j_1, j_2)$ and (d) $\Delta R(\tau, \tau)$. The signal shapes are shown multiplied by a factor of 20. These figures use background predictions made without the global fit defined in Section 12. The lowest bin in the p_T^{Total} distribution acts as an underflow bin, and includes events with $p_T^{\text{Total}} < 30 \text{ GeV}$.

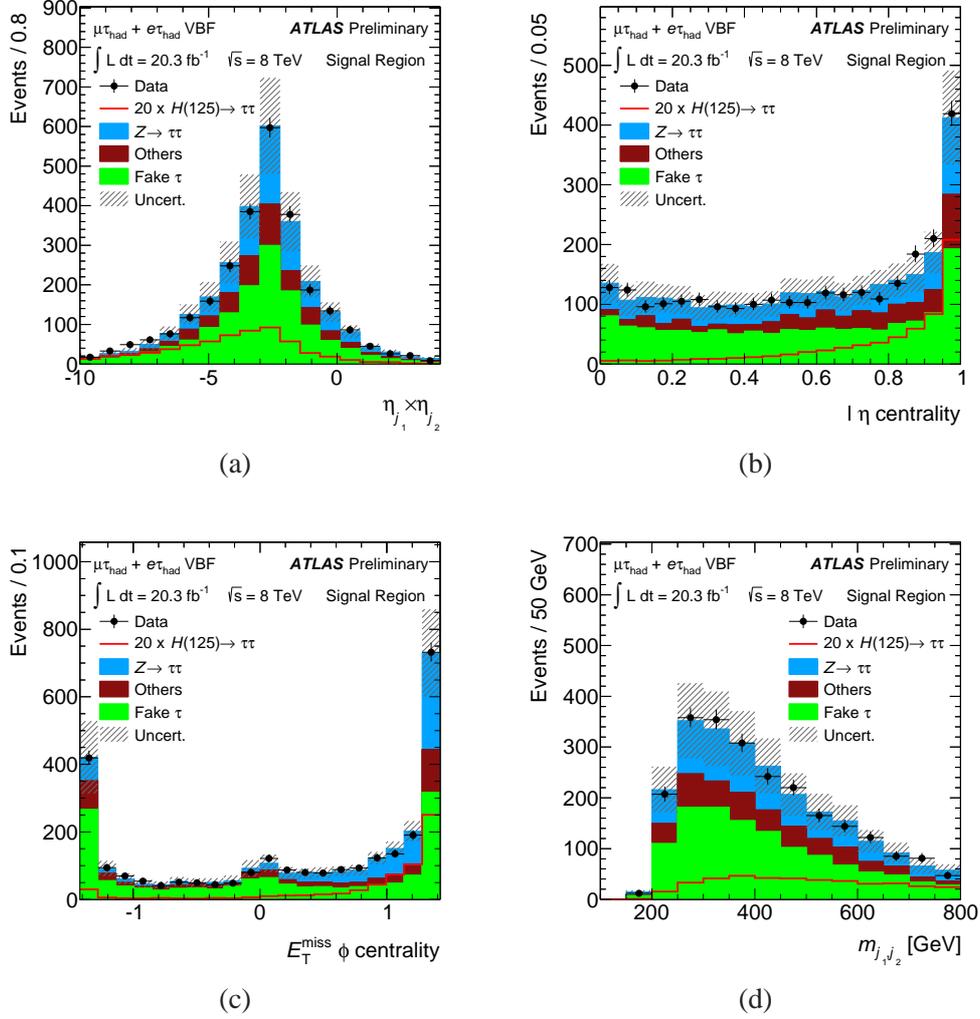


Figure 15: BDT input variables for the $\tau_{\text{lep}}\tau_{\text{had}}$ VBF category obtained in the signal region. The following distributions are shown: (a) $\eta_{j_1} \times \eta_{j_2}$, (b) $\ell \eta$ centrality, (c) $E_{\text{T}}^{\text{miss}} \phi$ centrality and (d) m_{j_1, j_2} . The signal shapes are shown multiplied by a factor of 20. These figures use background predictions made without the global fit defined in Section 12.

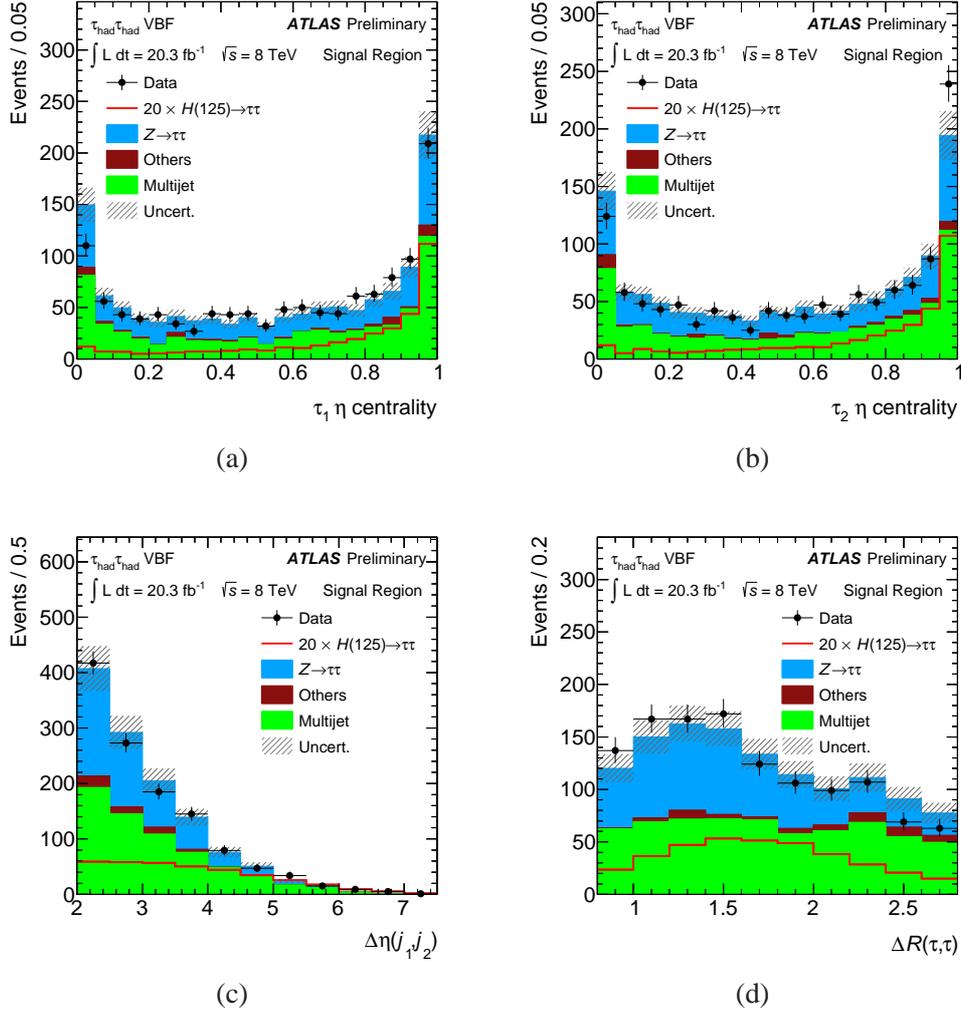


Figure 16: BDT input variables for the $\tau_{had}\tau_{had}$ VBF category obtained in the signal region. The following distributions are shown: (a) $\tau_1 \eta$ centrality, (b) $\tau_2 \eta$ centrality, (c) $\Delta\eta(j_1, j_2)$ and (d) $\Delta R(\tau, \tau)$. The signal shapes are shown multiplied by a factor of 20. These figures use background predictions made without the global fit defined in Section 12.

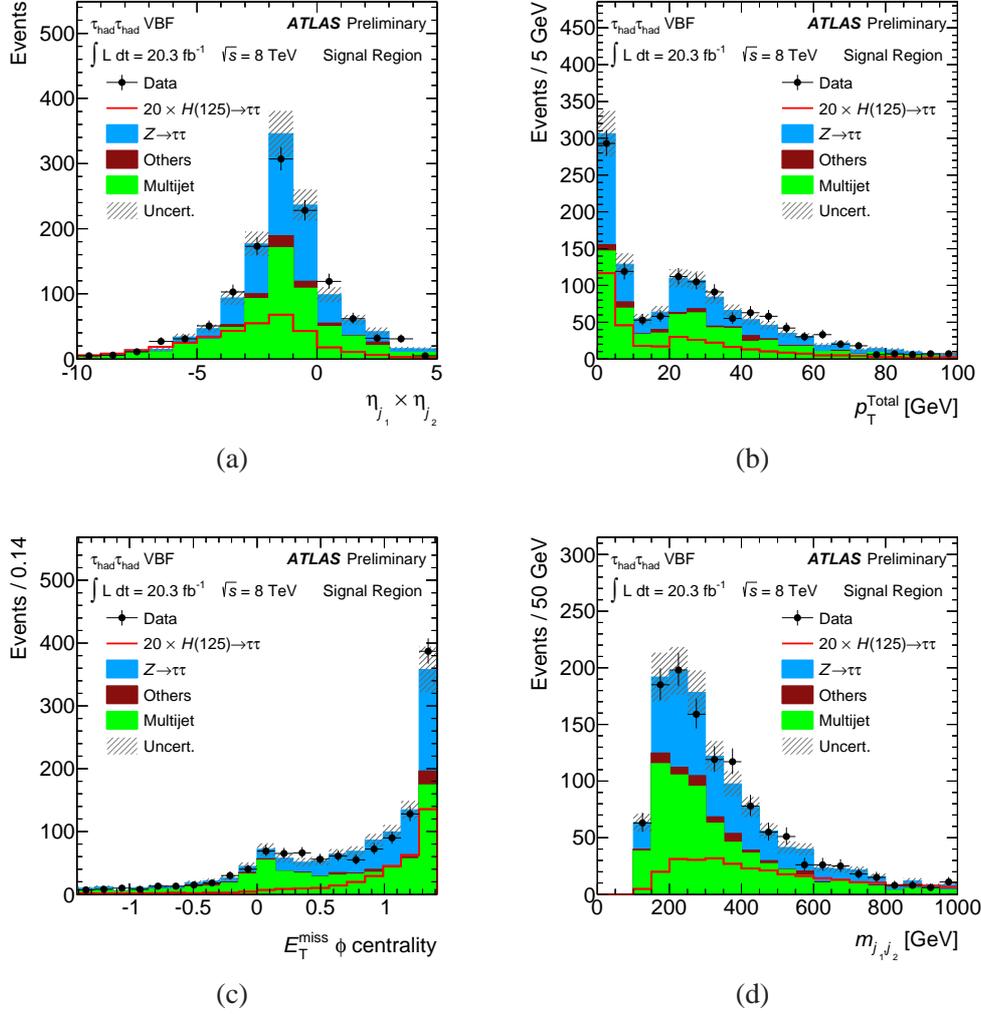


Figure 17: BDT input variables for the $\tau_{\text{had}}\tau_{\text{had}}$ VBF category obtained in the signal region. The following distributions are shown: (a) $\eta_{j_1} \times \eta_{j_2}$, (b) p_T^{Total} , (c) $E_T^{\text{miss}} \phi$ centrality and (d) m_{j_1, j_2} . The signal shapes are shown multiplied by a factor of 20. These figures use background predictions made without the global fit defined in Section 12.

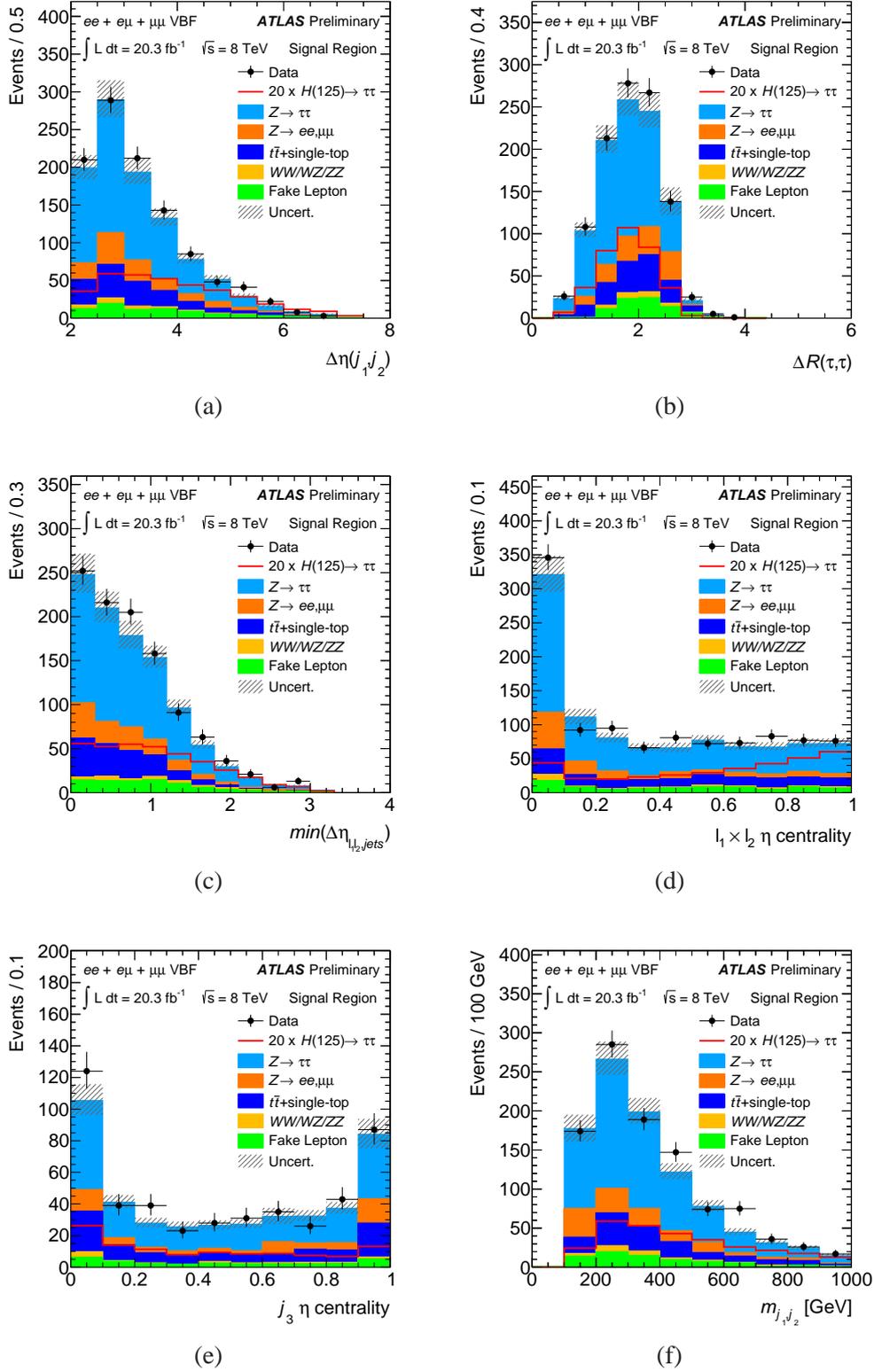


Figure 18: BDT input variables for the $\tau_{\text{lep}}\tau_{\text{lep}}$ VBF category obtained in the signal region. The following distributions are shown: (a) $\Delta\eta(j_1, j_2)$, (b) $\Delta R(\tau, \tau)$, (c) $\min(\Delta\eta_{\ell_1, \ell_2, \text{jets}})$, (d) ℓ η centrality, (e) j_3 η centrality and (f) m_{j_1, j_2} . The signal shapes are shown multiplied by a factor of 20. These figures use background predictions made without the global fit defined in Section 12.

A.2 Additional Mass Figures

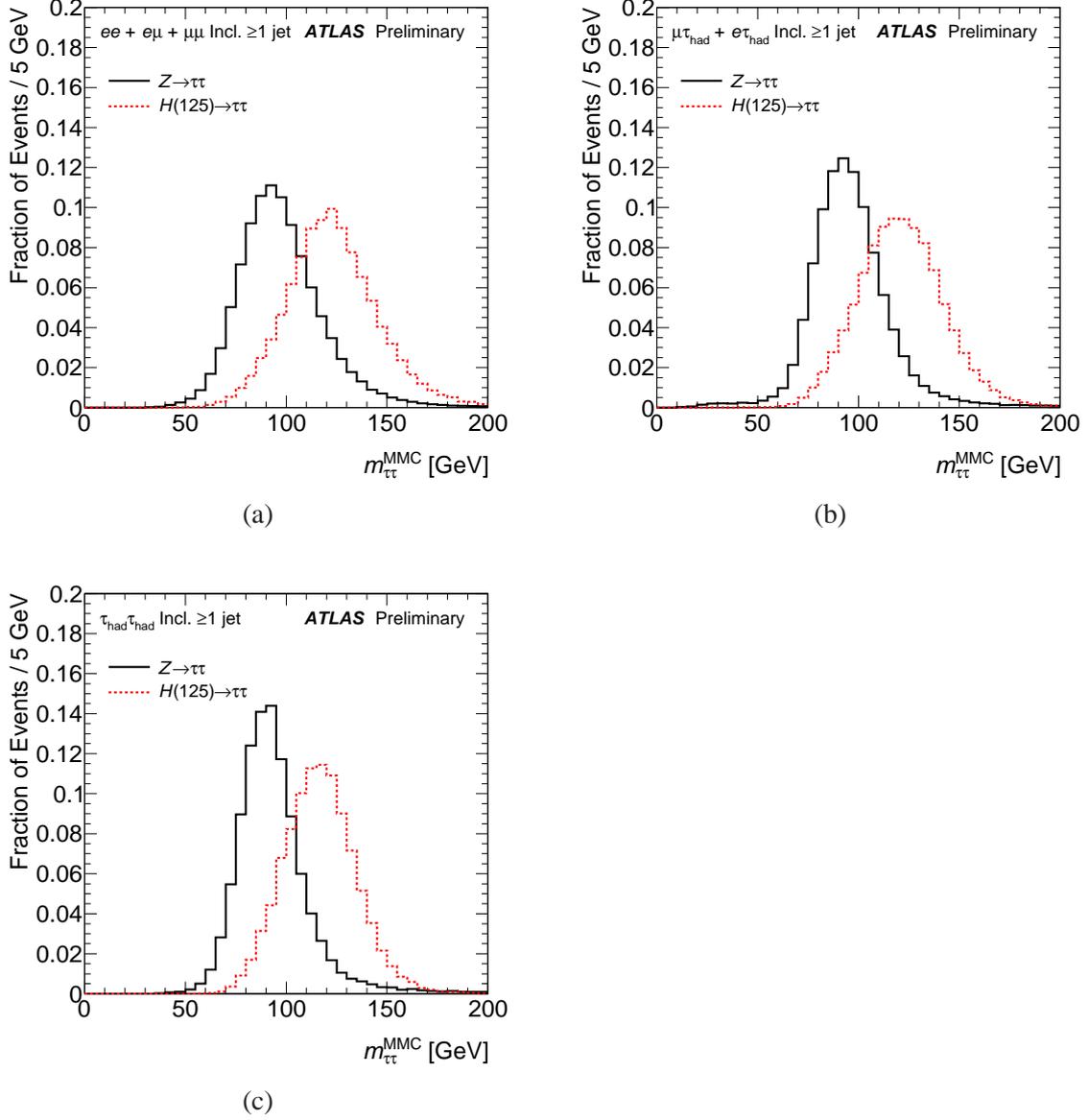


Figure 19: Distributions for $m_{\tau\tau}^{\text{MMC}}$ normalized to unit area for $Z \rightarrow \tau^+\tau^-$, from τ -embedded $Z \rightarrow \mu^+\mu^-$ data, and $H \rightarrow \tau^+\tau^-$ events in the (a) $\tau_{\text{lep}}\tau_{\text{lep}}$, (b) $\tau_{\text{lep}}\tau_{\text{had}}$ and (c) $\tau_{\text{had}}\tau_{\text{had}}$ channels for events that pass preselection requirements and have at least 1 jet. Full width at half maximum (FWHM) values for the $\tau_{\text{lep}}\tau_{\text{lep}}$ distributions are: 42 GeV for $H \rightarrow \tau^+\tau^-$ and 39 GeV for $Z \rightarrow \tau^+\tau^-$. FWHM values for the $\tau_{\text{lep}}\tau_{\text{had}}$ distributions are: 47 GeV for $H \rightarrow \tau^+\tau^-$ and 33 GeV for $Z \rightarrow \tau^+\tau^-$. FWHM values for the $\tau_{\text{had}}\tau_{\text{had}}$ distributions are: 38 GeV for $H \rightarrow \tau^+\tau^-$ and 26 GeV for $Z \rightarrow \tau^+\tau^-$.

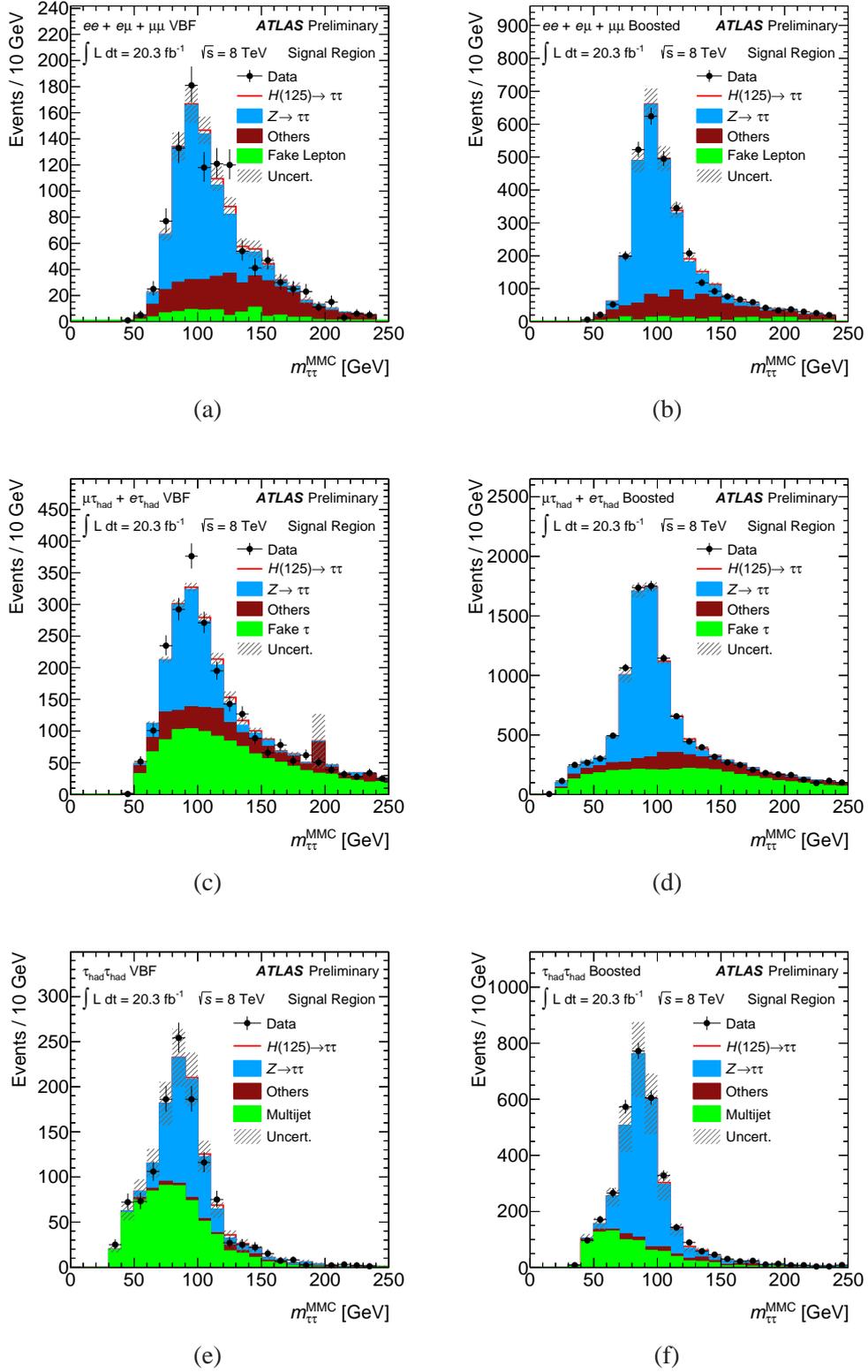


Figure 20: Distributions of $m_{\tau\tau}^{MMC}$ for the following categories: (a) $\tau_{lep}\tau_{lep}$ VBF, (b) $\tau_{lep}\tau_{lep}$ boosted, (c) $\tau_{lep}\tau_{had}$ VBF, (d) $\tau_{lep}\tau_{had}$ boosted, (e) $\tau_{had}\tau_{had}$ VBF and (f) $\tau_{had}\tau_{had}$ boosted. The background and signal predictions are shown from the global fit.

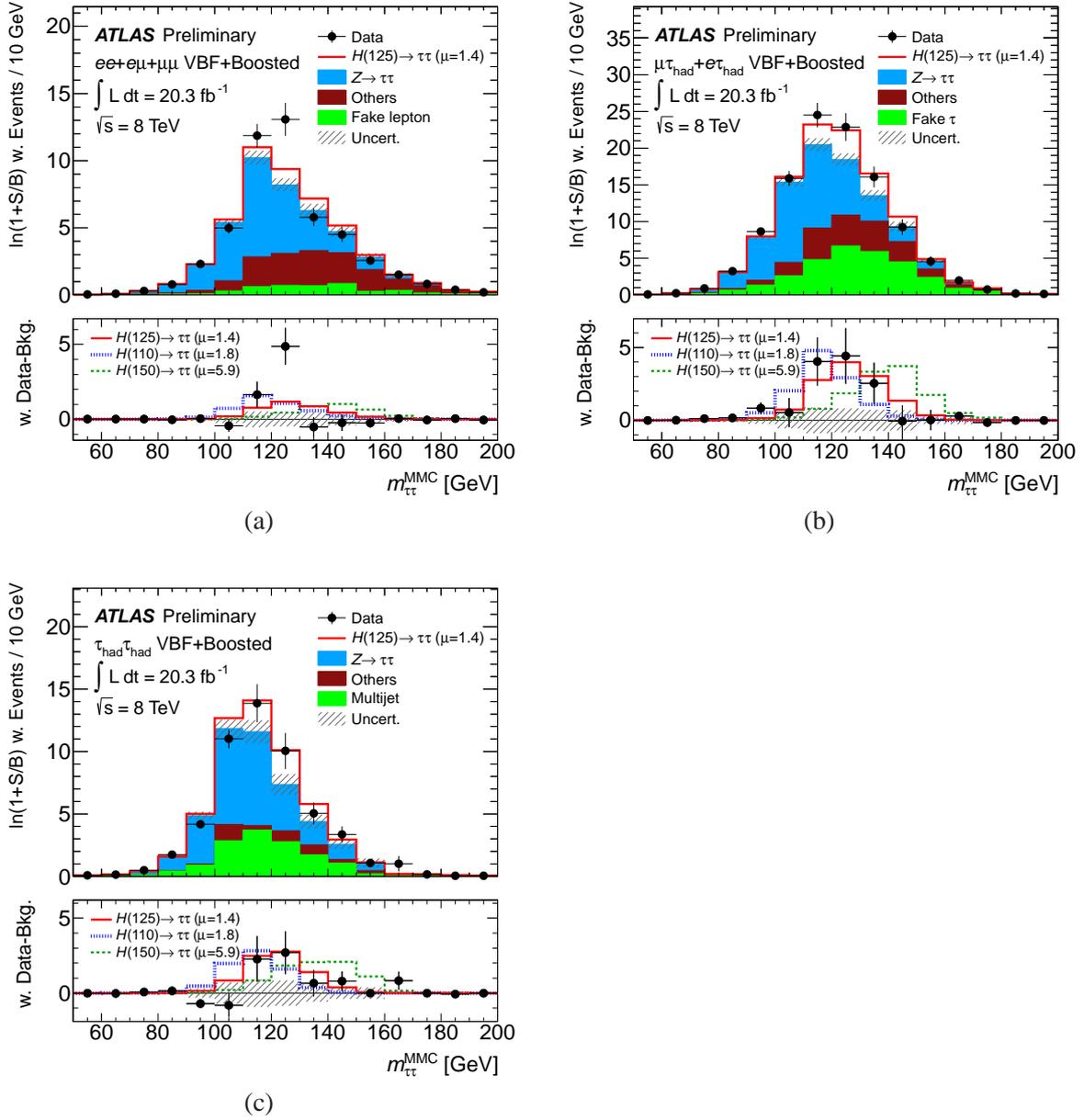


Figure 21: Distributions for $m_{\tau\tau}^{MMC}$ where events are weighted by $\ln(1 + S/B)$ for the $\tau_{lep}\tau_{lep}$ (a), $\tau_{lep}\tau_{had}$ (b), and $\tau_{had}\tau_{had}$ (c) channels. These weights are determined by the signal and background predictions for each BDT bin. Signal is shown stacked on background, with both predictions coming from the global fit yielding a signal strength of $\mu = 1.4$. The bottom panel shows the difference between weighted data events and weighted background events (black points), compared to the weighted signal yields. The $m_H = 125$ GeV signal is plotted with a solid red line, and, for comparison, the $m_H = 110$ GeV (blue) and $m_H = 150$ GeV (green) signals are plotted with a signal strength set to the observed values from the global fit, for those respective signal hypotheses.

A.3 Additional Sensitivity Figures

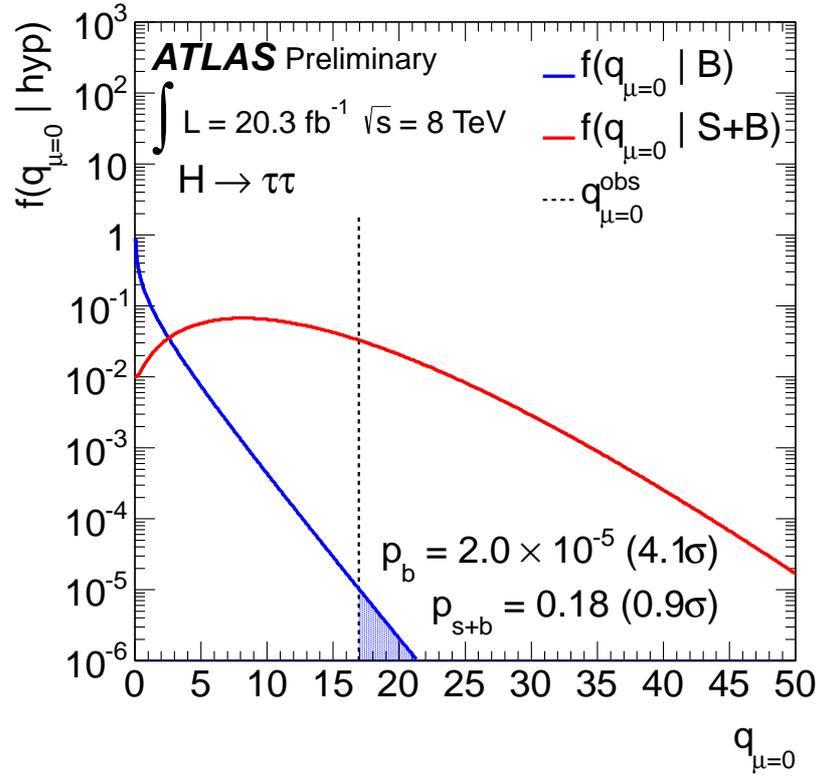


Figure 22: Distribution of the $q_{\mu=0}$ for the background only (blue) and signal+background hypothesis (red), in the asymptotic approximation [66]. The observed value (dotted line) obtained on data is also shown. The probability of the background only hypothesis (p_b) and the probability of the signal plus background hypothesis (p_{s+b}), are computed as the integral of $f(q_{\mu=0} | \text{hyp})$ over $q_{\mu=0} > q_{\mu=0}^{\text{obs}}$.

A.4 Additional Control Region Figures

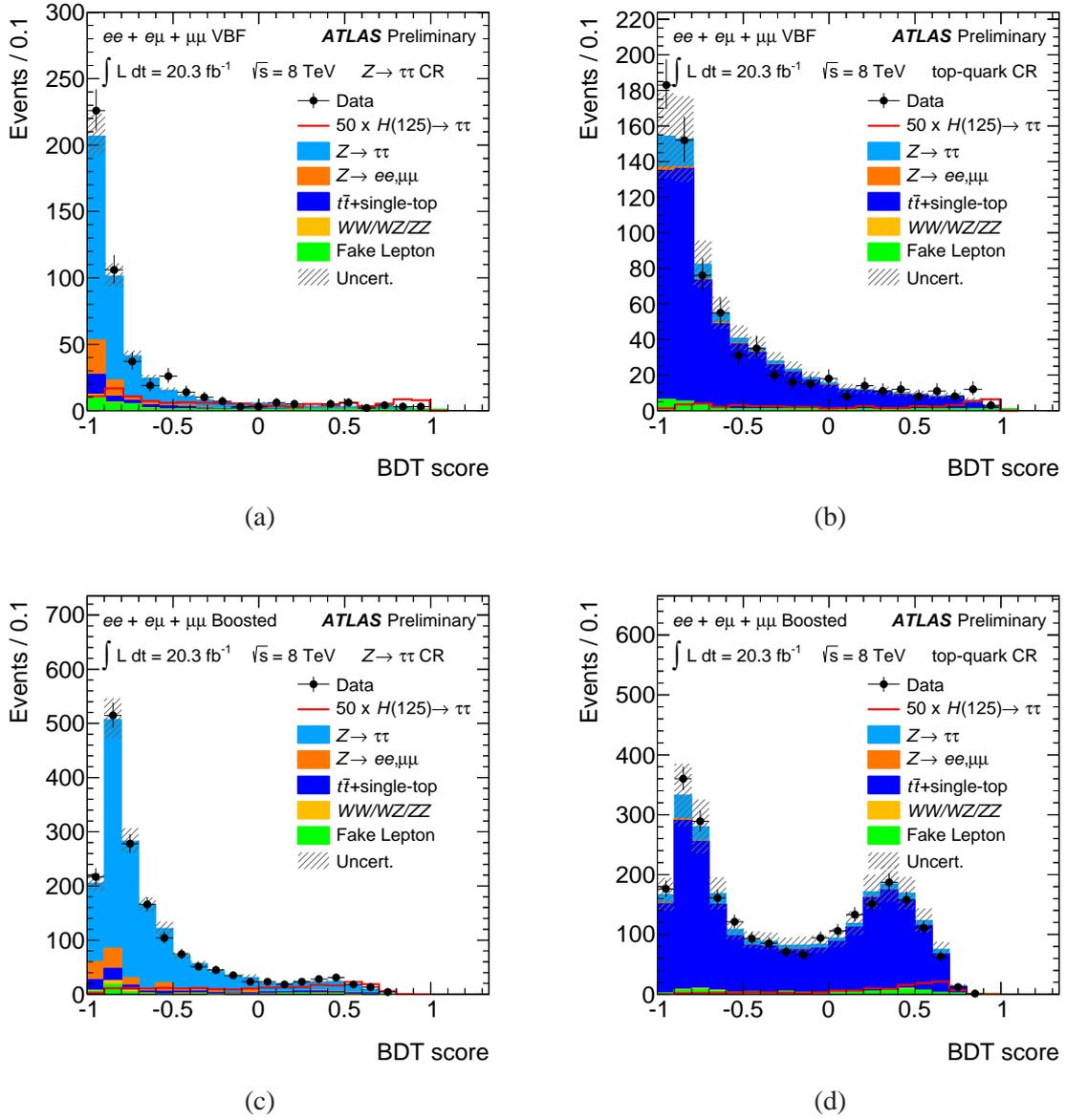
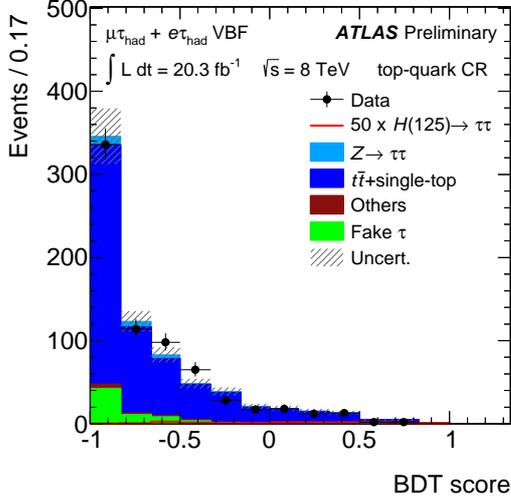
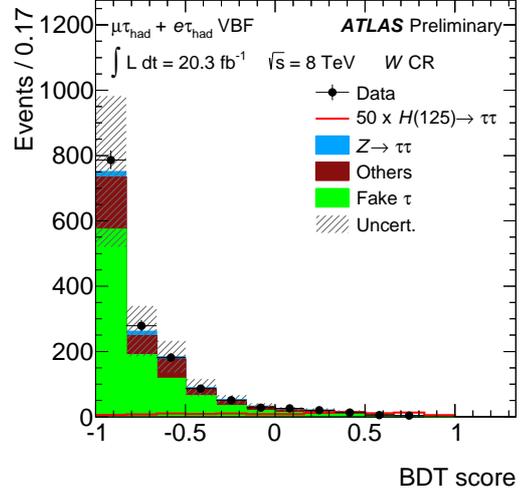


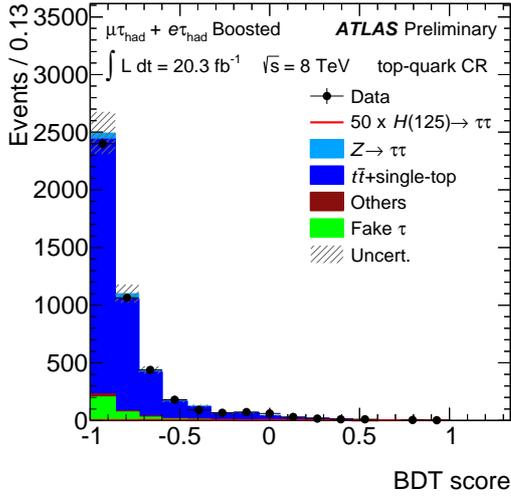
Figure 23: BDT score distributions for the $Z \rightarrow \tau\tau$ -enriched (a, c) and top-quark enriched (b, d) control regions for the VBF (a, b) and Boosted (c, d) categories in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel. The signal shapes are shown multiplied by a factor of 50. These figures use background predictions made without the global fit defined in Section 12.



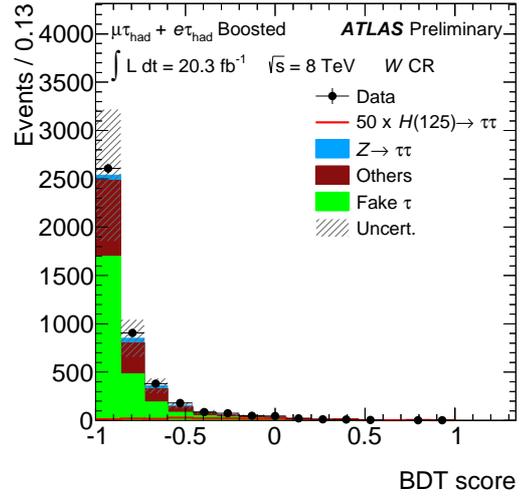
(a)



(b)



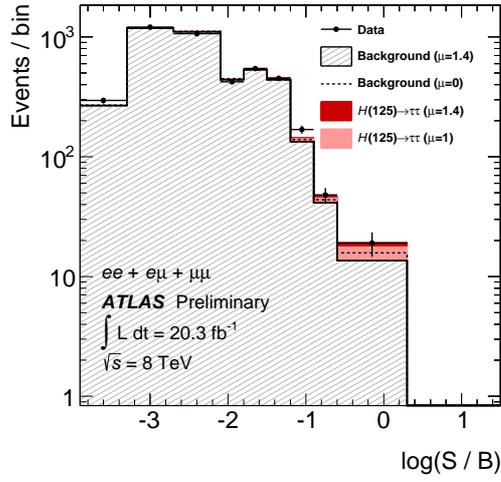
(c)



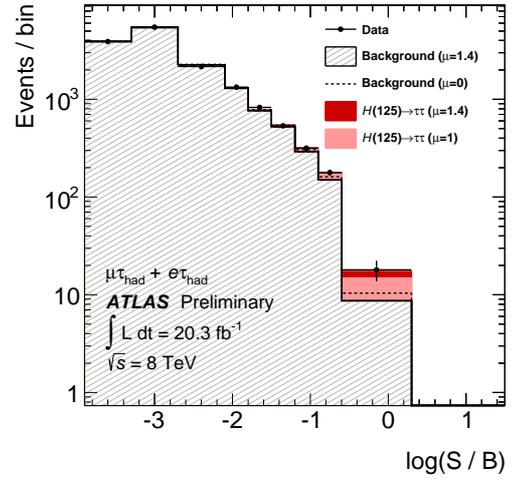
(d)

Figure 24: BDT score distributions for the top-quark enriched (a, c) and W -enriched control regions (b, d) for the VBF (a, b) and Boosted (c, d) categories in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. The signal shapes are shown multiplied by a factor of 50. These figures use background predictions made without the global fit defined in Section 12.

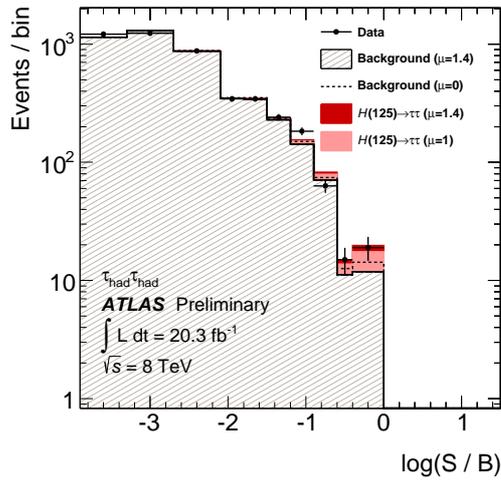
A.5 S/B ordered Event Yields



(a)



(b)



(c)

Figure 25: Event yields as a function of $\log(S/B)$ in all signal regions bins in the $\tau_{\text{lep}}\tau_{\text{lep}}$ (a), $\tau_{\text{lep}}\tau_{\text{had}}$ (b), $\tau_{\text{had}}\tau_{\text{had}}$ (c) channels. The predicted background, obtained from the global fit ($\mu = 1.4$), and signal yields, for $m_H = 125$ GeV, at $\mu = 1$ and $\mu = 1.4$ are shown. The background only distribution (dashed line), is obtained from the global fit, but fixing $\mu = 0$.

A.6 Comparison with other channels

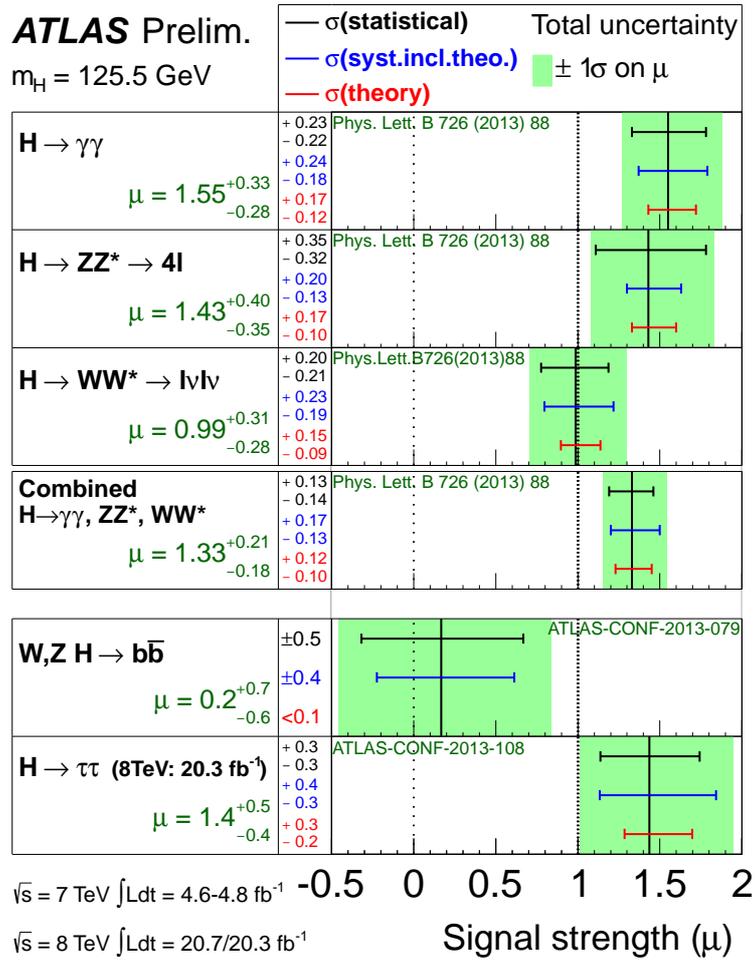


Figure 26: The measured production strengths normalized to the SM expectations, for the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ(*) \rightarrow 4l$, $H \rightarrow WW(*) \rightarrow \ell\nu\ell\nu$ final states and their combination [68] together with the preliminary signal strength measurements for the $H \rightarrow b\bar{b}$ [69] and $H \rightarrow \tau\tau$ final states. The measured production strengths are at $m_H = 125.5$ GeV, except for $H \rightarrow \tau\tau$ which is at $m_H = 125$ GeV. The best-fit values are shown by the solid vertical lines. The total $\pm 1\sigma$ uncertainty is indicated by the shaded band, with the individual contributions from the statistical uncertainty (top), the total (experimental and theoretical) systematic uncertainty (middle), and the theory uncertainty (bottom) on the signal cross section (from QCD scale, PDF, and branching ratios) shown as superimposed error bars.

A.7 Event Displays

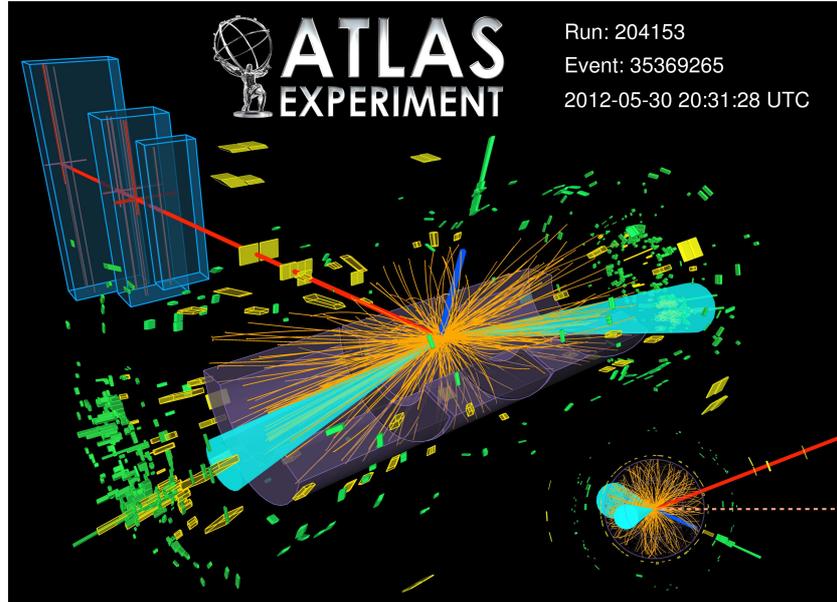


Figure 27: Display of an event selected by the $H \rightarrow \tau_{\text{lep}} \tau_{\text{lep}}$ channel in the VBF category, where one τ decays to a muon and the other to an electron. The electron is indicated by a blue track and the muon indicated by a red track. The approximately horizontal dashed line in the $R - \phi$ view represents the direction of the $E_{\text{T}}^{\text{miss}}$ vector, and there are two VBF jets marked with turquoise cones. The muon p_{T} is 53 GeV, the electron p_{T} is 34 GeV, $E_{\text{T}}^{\text{miss}} = 102$ GeV, $m_{j_1, j_2} = 1.04$ TeV and $m_{\tau\tau}^{\text{MMC}} = 127$ GeV and $\text{BDT} = 0.97$. The S/B ratio in the BDT score bin of this event is 0.42.

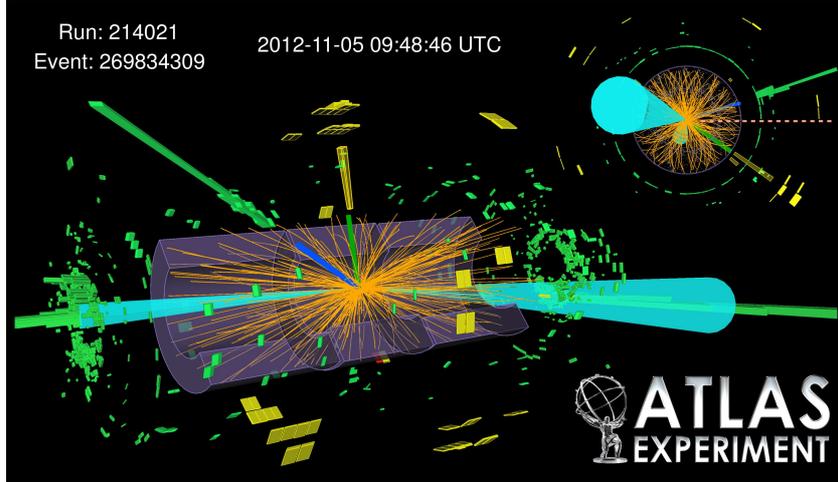


Figure 28: Display of an event selected by the $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ channel in the VBF category, where one τ decays to an electron. The hadronically decaying τ lepton (1 prong decay) is indicated by a green track and the yellow cluster. The electron is indicated by the blue track match to the green cluster. The approximately horizontal dashed line in the $R - \phi$ view represents the direction of the $E_{\text{T}}^{\text{miss}}$ vector, and there are two VBF jets marked with turquoise cones. The electron p_{T} is 56 GeV, the τ_{had} p_{T} is 27 GeV, $E_{\text{T}}^{\text{miss}} = 113$ GeV, $m_{j_1, j_2} = 1.53$ TeV, $m_{\tau\tau}^{\text{MMC}} = 129$ GeV, and the BDT score is 0.99. The S/B ratio in the BDT score bin of this event is 1.0.

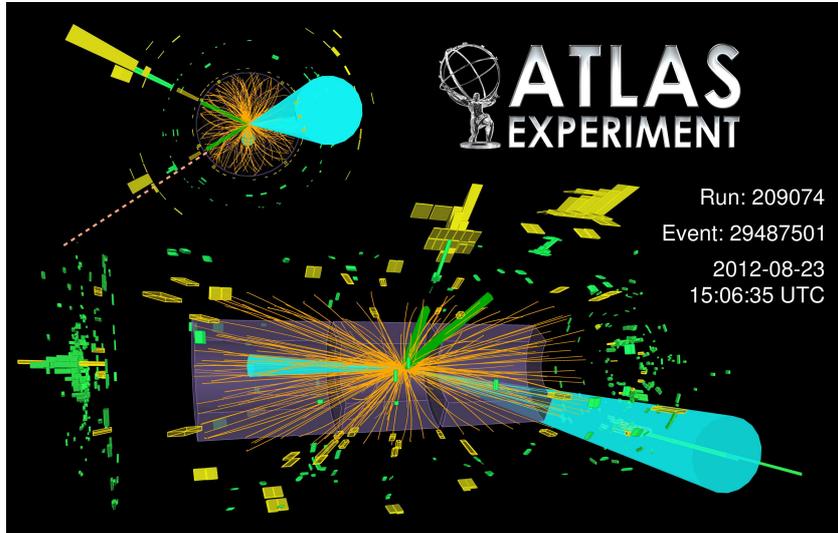


Figure 29: Display of an event selected by the $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ channel in the VBF category, where the two τ decays into hadrons. The hadronically decaying τ leptons are indicated by green tracks. The dashed line in the lower left quadrant of the $R - \phi$ view represents the direction of the $E_{\text{T}}^{\text{miss}}$ vector, and there are two VBF jets marked with turquoise cones. The leading τ_{had} p_{T} is 122 GeV, the sub-leading τ_{had} p_{T} is 67 GeV, $E_{\text{T}}^{\text{miss}} = 72$ GeV, $m_{j_1, j_2} = 1.02$ TeV and $m_{\tau\tau}^{\text{MMC}} = 126$ GeV and BDT= 1.0. The S/B ratio in the BDT score bin of this event is 0.67.