ATLAS CONF Note

ATLAS-CONF-2018-021

Cross-section measurements of the Higgs boson decaying to a pair of tau leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

5th June 2018

A measurement of total production cross sections of the Higgs boson in proton–proton collisions is presented in the $H \to \tau\tau$ decay channel. The analysis is performed using 36.1 fb^{-1} of data recorded by the ATLAS experiment at the Large Hadron Collider at a center-of-mass energy of $\sqrt{s} = 13$ TeV. All combinations of leptonic ($\tau \to \ell\nu\overline{\nu}$ with $\ell = e, \mu$) and hadronic ($\tau \to$ hadrons ν) tau decays are considered. The $H \to \tau\tau$ signal over the expected background from other Standard Model processes is established with an observed (expected) significance of 4.4 (4.1) standard deviations. Combined with results using data taken at 7 and 8 TeV center-of-mass energies, the observed (expected) significance amounts to 6.4 (5.4) standard deviations and constitutes an observation of $H \to \tau\tau$ decays. Using the data taken at $\sqrt{s} = 13$ TeV, the total cross section in the $H \to \tau\tau$ decay channel is measured to be $3.71 \pm 0.59\,(\text{stat.})\,^{+0.87}_{-0.74}\,(\text{syst.})$ pb, for a Higgs boson of mass 125 GeV assuming the relative contributions of its production modes as predicted by the Standard Model. Total cross sections in the $H \to \tau\tau$ decay channel are determined separately for vector boson fusion production and gluon–gluon fusion production to be $\sigma^{\text{VBF}}_{H\to\tau\tau} = 0.28 \pm 0.09\,(\text{stat.})\,^{+0.11}_{-0.09}\,(\text{syst.})$ pb and $\sigma^{gg\text{F}}_{H\to\tau\tau} = 3.0 \pm 1.0\,(\text{stat.})\,^{+1.6}_{-1.2}\,(\text{syst.})$ pb, respectively. All measurements are in agreement with Standard Model expectations.

1 Introduction

The ATLAS and CMS collaborations discovered in 2012 [1, 2] a particle consistent with the Standard Model (SM) [3–5] Higgs boson [6–10]. Several properties of this particle, such as its coupling strengths, spin and charge-parity (CP) quantum numbers, were studied with 7 and 8 TeV center-of-mass energy ($\sqrt{s}$) proton–proton collision data delivered by the Large Hadron Collider (LHC) in 2011 and 2012, respectively, denoted as 'Run 1'. These results rely predominantly on studies of the bosonic decay modes [11–14] and have not shown any significant deviations from the SM expectations.

The coupling of the Higgs boson to the fermionic sector has been established with the observation of the $H \to \tau\tau$ decay mode with a signal significance of 5.5σ from a combination of ATLAS and CMS results [15–17] using LHC Run 1 data. A measurement performed by the CMS collaboration with Run 2 data at $\sqrt{s} = 13$ TeV reached a significance of 4.9σ using $35.9\,\text{fb}^{-1}$ of integrated luminosity and 5.9σ combined with data from Run 1 [18]. While the Higgs boson coupling to other fermions such as top quarks has been observed [19, 20] and there is evidence of its coupling to bottom quarks [21, 22], only upper limits exist on its coupling to muons [23, 24] and the $H \to \tau\tau$ decay mode has been the only accessible leptonic decay mode. It was also used to constrain CP violation in the production via vector-boson fusion (VBF) [25] and is unique in that it provides sensitivity to CP violation in the Higgs boson coupling to leptons [26].

This note presents cross-section measurements of Higgs bosons that decay to a pair of tau leptons in proton–proton (pp) collisions at $\sqrt{s} = 13$ TeV using data collected by the ATLAS experiment in 2015 and 2016, corresponding to an integrated luminosity of $36.1\,\text{fb}^{-1}$. All combinations of leptonic ($\tau \to \ell\nu\overline{\nu}$ with $\ell = e, \mu$) and hadronic ($\tau \to$ hadrons ν) tau decays are considered.[1] The corresponding three analysis channels are denoted by $\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ and are composed of different dominant backgrounds. While $Z \to \tau\tau$ is a dominant background in all channels, the relative contributions from other backgrounds from top-quark and vector-boson decays, as well as from misidentified leptonic or hadronic tau decays, vary considerably. Two analysis categories are defined that are predominantly sensitive to Higgs bosons produced via VBF and gluon–gluon fusion (ggF). A maximum likelihood fit is performed on data using distributions of the reconstructed di-tau mass in signal regions (SRs), simultaneously with event yields from control regions (CRs) that are included to constrain normalizations of major backgrounds estimated from simulation. The dominant and irreducible $Z \to \tau\tau$ background is estimated from simulation. This is different from the search for $H \to \tau\tau$ decays in Run 1 [15], which used the embedding technique [27]. A reliable modeling of this background is therefore of critical importance for this analysis. Validation regions (VRs) based on $Z \to \ell\ell$ events are studied, but not included in the fit, to verify as precisely as possible the modeling of the $Z \to \tau\tau$ background.

The article is organized as follows. Section 2 describes the ATLAS detector. This is followed in Section 3 by a description of the dataset and Monte Carlo (MC) simulated samples employed by this measurement. Section 4 details the reconstruction of particles and jets. The event selection for each channel and event category as well as signal, control and validation regions are discussed in Section 5. Background estimation techniques and the systematic uncertainties for the analysis are described in Sections 6 and 7, respectively. The signal extraction procedure and the results of the Higgs cross-section measurement in the $H \to \tau\tau$ decay mode are presented in Section 8.

[1] Throughout this article the inclusion of charge-conjugate decay modes is implied. The symbol ℓ is used to denote electrons and muons, also referred to as 'light leptons'.

2 The ATLAS detector

The ATLAS experiment [28] at the LHC is a multi-purpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4π coverage in solid angle[2]. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadron calorimeters, and a muon spectrometer. The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of a silicon pixel detector, which has an additional innermost layer (positioned at a radial distance of 3.3 cm from the beam line) that has been installed since the end of Run 1 (IBL) [29], and a silicon microstrip detector surrounding the pixel detector, both covering $|\eta| < 2.5$, followed by a transition radiation straw-tube tracker covering $|\eta| < 2$. This is surrounded by a 2 T axial magnetic field provided by the solenoid. Lead/liquid-argon sampling calorimeters provide electromagnetic energy measurements with high granularity. A hadron iron/scintillator-tile calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The end-cap and forward regions are instrumented with liquid-argon calorimeters for both the electromagnetic and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large air-core toroid superconducting magnets with eight coils each. Its bending power is in the range from 2.0 T m to 6.0 T m.

Events are selected using a two-level trigger system. The first-level trigger is implemented in hardware and uses a subset of the detector information to filter events that are then processed by a software-based high-level trigger. This further reduces the average recorded collision rate to approximately 1 kHz.

3 Data and simulated samples

The data used in this analysis are taken from proton–proton collisions at the LHC where proton bunches are collided every 25 ns. A combination of several triggers for single light leptons, two light leptons and two hadronically-decaying tau leptons are used to record the data for the analysis, depending on the analysis channel (see Section 5.1). After data quality requirements, the samples used for this measurement consist of 3.2 fb^{-1} of data recorded in 2015, with an average of 14 interactions per bunch crossing, and 32.9 fb^{-1} recorded in 2016, with an average of 25 interactions per bunch crossing.

Samples of signal and background processes are simulated using various Monte Carlo generators as summarized in Table 1. The signal contributions considered include the following three processes for Higgs boson production at the LHC: ggF, VBF and associated production of a Higgs boson with a vector boson (VH) where all decay modes for the $H \to \tau\tau$ process are included. Other Higgs production processes like associated production with a top–antitop quark pair, bottom–antibottom quark pair and with a single top quark are found to be negligible. Higgs decays to WW are considered background and simulated likewise for these production processes. The mass of the Higgs boson is assumed to be 125 GeV [30].

Higgs production by ggF is simulated with the Powheg v2 [31–34] NNLOPS program [35] at next-to-leading order (NLO) accuracy in quantum chromodynamics (QCD) using the MiNLO approach [36],

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

Table 1: Monte Carlo generators used to describe all signal and background processes together with the corresponding PDF set and the model of parton shower, hadronization and underlying event (UEPS). In addition, the order of the total cross-section calculation is given. The total cross section for VBF production is calculated at approximate-NNLO QCD. More details are given in the text.

Process	Monte Carlo generator	PDF	UEPS	Cross-section order
ggF	Powheg-Box v2	PDF4LHC15 NNLO	Pythia 8.212	N^{3}LO QCD + NLO EW
VBF	Powheg-Box v2	PDF4LHC15 NLO	Pythia 8.212	~NNLO QCD + NLO EW
VH	Powheg-Box v2	PDF4LHC15 NLO	Pythia 8.212	NNLO QCD + NLO EW
W/Z + jets	Sherpa 2.2.1	NNPDF30NNLO	Sherpa 2.2.1	NNLO
$VV/V\gamma^*$	Sherpa 2.2.1	NNPDF30NNLO	Sherpa 2.2.1	NLO
$t\bar{t}$	Powheg-Box v2	CT10	Pythia 6.428	NNLO+NNLL
Wt	Powheg-Box v1	CT10F4	Pythia 6.428	NLO

and reweighted to next-to-next-to-leading order (NNLO) in QCD in the Higgs rapidity. The VBF and VH production processes are simulated at NLO accuracy in QCD using Powheg-Box with the MiNLO approach. For these signal samples, the simulation is interfaced to the Pythia 8.212 [37] model of parton shower, hadronization and underlying event (UEPS). To estimate the impact of UEPS uncertainties, the ggF, VBF and VH samples are also simulated with the Herwig 7.0.3 [38, 39] UEPS model. The PDF4LHC15 [40] parameterization of the parton distribution functions (PDFs) is used for these production processes. The AZNLO [41] set of tuned parameters is used, with the CTEQ6L1 [42] PDF set, for the modeling of non-perturbative effects. Photos++ version 3.52 [43] is used for QED emissions from electroweak (EW) vertices and charged leptons.

The overall normalization of the ggF process is taken from a next-to-next-to-next-to-leading-order (N^{3}LO) QCD calculation with NLO EW corrections applied [44–47]. Production by VBF is normalized to an approximate-NNLO QCD cross section with NLO EW corrections applied [48–50]. The VH samples are normalized to cross sections calculated at NNLO in QCD, with NLO EW radiative corrections applied [51–53].

Background samples of EW production of W/Z bosons from VBF, W/Z boson production with associated jets and di-boson production processes are simulated with the Sherpa 2.2.1 [54] generator. Matrix elements are calculated using the Comix [55] and OpenLoops [56] matrix-element generators and merged with the Sherpa UEPS model [57] using the ME+PS@NLO prescription [58]. For W and Z production with associated jets the matrix elements are calculated for up to two partons at NLO and four partons at LO precision. Their inclusive cross sections are normalized to NNLO calculations from Fewz [59, 60]. For di-boson production, the matrix elements are calculated for up to one additional parton at NLO and up to three additional partons at LO precision. For all samples the NNPDF30NNLO [61] PDF set is used together with the Sherpa UEPS model. In particular, the dominant $Z \to \tau\tau$ background is estimated using these simulations of Z-boson production.

The impact of UEPS uncertainties, and other modeling uncertainties such as LO/NLO precision comparison for leading jets, on the main background from $Z \to \tau\tau$ is studied in an alternative sample which is simulated using Madgraph5_aMC@NLO 2.2.2 [62] at leading order interfaced to the Pythia 8.186 UEPS model. The A14 set of tuned parameters [63] is used together with the NNPDF23LO PDF set [64].

For the generation of $t\bar{t}$ production, the Powheg-Box v2 [31–33, 65] generator with the CT10 PDF sets in the matrix element calculations is used. The predicted $t\bar{t}$ cross section is calculated with the Top++2.0

program to NNLO in perturbative QCD, including soft-gluon resummation to next-to-next-to-leading-log order [66]. Single top-quark production of Wt is simulated using the POWHEG-BOX v1 [67, 68] generator. This generator uses the four-flavor scheme for the NLO matrix -element calculations together with the fixed four-flavor PDF set CT10F4. For all top-quark production processes, top-quark spin correlations are preserved (for the t-channel, top quarks are decayed using MadSpin [69]). The parton shower, hadronization, and the underlying event are simulated using PYTHIA 6.428 [70] with the CTEQ6L1 PDF set and the corresponding Perugia 2012 set of tuned parameters [71]. The top mass is assumed to be 172.5 GeV. The EvtGen v.1.2.0 program [72] is used for the properties of b- and c-hadron decays.

For all samples, a full simulation of the ATLAS detector response [73] using the GEANT4 program [74] was performed. The effect of multiple pp interactions in the same and neighbouring bunch crossings (pile-up) is included by overlaying minimum-bias events simulated with PYTHIA 8.186 using the MSTW2008LO PDF [75] and the A2 [76] set of tuned parameters on each generated signal and background event. The number of overlaid events is chosen such that the distribution of the average number of interactions per pp bunch crossing in the simulation matches that observed in data.

4 Object reconstruction

Electron candidates are reconstructed from energy deposits in the electromagnetic calorimeter associated with a charged-particle track measured in the inner detector. The electron candidates are required to pass the 'loose' likelihood-based identification selection of Ref. [77, 78], to have transverse momentum $p_{\mathrm{T}} > 15$ GeV and to be in the fiducial volume of the inner detector, $|\eta| < 2.47$. The transition region between the barrel and end-cap calorimeters ($1.37 < |\eta| < 1.52$) is excluded. The trigger efficiency for single electrons selected in the analysis ranges between 90–95% [79]. Electron candidates are ignored if they share their reconstructed track with a muon candidate defined below or if their angular distance to a jet is within $0.2 < \Delta R < 0.4$.

Muon candidates are constructed by matching an inner detector track with a track reconstructed in the muon spectrometer [80]. The muon candidates are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$ and to pass the 'loose' muon identification requirements of Ref. [80]. The trigger efficiency for single muons selected in the analysis is close to 80% (70%) in the barrel in the 2016 (2015) dataset and 90% in the end-caps [79]. Muon candidates are ignored if their angular distance to a jet is $\Delta R < 0.4$ with the following exceptions: If $\Delta R < 0.2$ or the muon track is associated to the jet, and if the jet has either less than three tracks or less than twice the transverse momentum of the muon candidate, the jet is removed instead. This recovers efficiency for muons which radiate a hard bremsstrahlung photon in the calorimeter.

In the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ and $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ signal regions events are only selected if the selected electron and muon candidates pass their respective 'medium' identification criteria. The reconstruction and identification efficiency for muons with the 'medium' identification requirement has been measured in $Z \to \mu\mu$ events [80]. It is well above 98% over the full phase space, except for $|\eta| < 0.1$ where the reconstruction efficiency is about 70%. The combined identification and reconstruction efficiency for 'medium' electrons ranges from 80–90% in the p_{T} range from 10 GeV to 80 GeV as measured in $Z \to ee$ events [78]. In addition, the electrons and muons are required to fulfill the 'gradient' isolation criterion, which requires that there are no additional high-p_{T} tracks in a cone around the track and no significant energy deposits in a cone around the calorimeter clusters of the object after correcting for pile-up. The size of the respective cones depends on the p_{T} of the light lepton. This isolation requirement rejects about 10% of light leptons for low-p_{T} and less than 1% for $p_{\mathrm{T}} > 60$ GeV [78, 80].

Jets are reconstructed from topological clusters in the calorimeter using the anti-k_t algorithm [81, 82], with a radius parameter value $R = 0.4$. To reject jets from pile-up a 'Jet Vertex Tagger' (JVT) [83] algorithm is used for jets with $p_{\mathrm{T}} < 50\,\mathrm{GeV}$ and $|\eta| < 2.4$. It employs a multivariate technique that relies on jet-tracking and calorimeter-cluster-shape variables to determine the likelihood that the jet originates from pile-up. Similarly, pile-up jets in the forward region are suppressed with a 'forward JVT' [84] algorithm, relying in this case only on calorimeter-cluster-shape variables, which is applied to all jets with $p_{\mathrm{T}} < 50\,\mathrm{GeV}$ and $|\eta| > 2.5$. In the pseudorapidity range $|\eta| < 2.5$, b-jets are selected using a multivariate algorithm [85, 86]. A working point is chosen that corresponds to an efficiency of approximately 85% for b-jets and rejection factors of 2.8 and 28 for c-jets and light-flavor jets, respectively, in simulated $t\bar{t}$ events. A jet is ignored if it is within $\Delta R < 0.2$ with respect to an electron or hadronically-decaying tau candidate.

Leptonic tau decays are reconstructed as electrons and muons. The reconstruction of the visible decay products of hadronic tau decays ($\tau_{\text{had-vis}}$) [87] starts with a reconstructed jet that has $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.5$. As in the case of electron reconstruction the transition region between the barrel and end-cap calorimeters is excluded. To discriminate $\tau_{\text{had-vis}}$ from jets initiated by light-quarks or gluons, an identification algorithm using multivariate techniques is applied to $\tau_{\text{had-vis}}$ candidates. They have to pass the 'loose' identification requirement of Ref. [87]. In addition, the $\tau_{\text{had-vis}}$ candidates are required to have $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, to have one or three associated tracks and an absolute electric charge of one. Their energy is reconstructed by multivariate regression techniques using information of the associated tracks and calorimeter clusters as well as the average number of collisions recorded. The trigger efficiency per $\tau_{\text{had-vis}}$ selected in the analysis is 95% and 85% for 1-prong and 3-prong taus, respectively [88]. $\tau_{\text{had-vis}}$ candidates are ignored if they are within $\Delta R < 0.2$ of a muon or electron candidate. In addition, if the candidate overlaps within $\Delta R < 0.4$ with an electron candidate, it is required to have a low electron identification likelihood score [78]. The requirement on the likelihood score corresponds to a $\tau_{\text{had-vis}}$ efficiency measured in $Z \to \tau\tau$ decays of 95% [87].

In the $\tau_{\text{lep}}\tau_{\text{had}}$ signal regions, events are only selected if the $\tau_{\text{had-vis}}$ candidate passes the 'medium' identification requirement, corresponding to an efficiency of 55% and 40% for real 1-prong and 3-prong $\tau_{\text{had-vis}}$, respectively [87]. In addition, if a 1-prong $\tau_{\text{had-vis}}$ candidate and an electron candidate are selected, a dedicated multivariate algorithm to reject electrons misidentified as $\tau_{\text{had-vis}}$ is applied to suppress $Z \to ee$ events. In the $\tau_{\text{had}}\tau_{\text{had}}$ signal regions both selected $\tau_{\text{had-vis}}$ candidates have to fulfill the 'tight' identification requirement, which corresponds to a selection efficiency of 45% for real 1-prong $\tau_{\text{had-vis}}$ and 30% for real 3-prong $\tau_{\text{had-vis}}$ [87].

The missing transverse momentum vector is calculated as the negative vectorial sum of the p_{T} of the fully calibrated and reconstructed physics objects [89]. This procedure includes a soft term, which is calculated based on the inner detector tracks originating from the vertex associated to the hard-scattering process that are not associated with any of the reconstructed objects. The missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) is defined as the modulus of this vector.

The Higgs boson candidate is reconstructed from the visible decay products of the tau leptons and from the $E_{\mathrm{T}}^{\mathrm{miss}}$ which is assumed to originate from the final state neutrinos. The di-tau invariant mass ($m_{\tau\tau}^{\mathrm{MMC}}$) is determined using the missing mass calculator (MMC) [90]. The standard deviation of the reconstructed di-tau mass is 17.0 GeV, 15.3 GeV and 14.7 GeV for signal events selected in the $\tau_{\text{lep}}\tau_{\text{lep}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels, respectively. The p_{T} of the Higgs boson candidate ($p_{\mathrm{T}}^{\tau\tau}$) is computed as the vector sum of the transverse momenta of the visible decay products of the tau leptons and the missing transverse momentum vector.

Table 2: Summary of the triggers used to select events for the three analysis channels during 2015 and 2016 data-taking and the corresponding p_{T} requirements applied in the analysis. For the electron+muon trigger the first number corresponds to the electron p_{T} requirement, the second to the muon p_{T} requirement. For the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel, at least one high-p_{T} jet in addition to the two $\tau_{\text{had-vis}}$ candidates is required for the 2016 dataset (see text).

Analysis channel	Trigger	Analysis p_{T} requirement [GeV] 2015	2016
$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ & $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	Single electron	25	27
	Single muon	21	27
$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$	Di-electron	15 / 15	18 / 18
	Di-muon	19 / 10	24 / 10
	Electron+muon	18 / 15	18 / 15
$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$	Di-$\tau_{\text{had-vis}}$	40 / 30	40 / 30

5 Event selection and categorization

In addition to data quality criteria which ensure that the detector was functioning properly, events are rejected if they contain reconstructed jets not associated to real energy deposits that can arise from hardware problems, beam conditions or cosmic showers. To further increase the purity and quality of the data sample by rejecting non-collision events originating from cosmic rays and beam-halo events, at least one reconstructed primary vertex is required with at least two associated tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$. The primary vertex is chosen as the pp vertex candidate with the highest sum of the squared transverse momenta of all associated tracks.

The triggers and event selection for the three analysis channels are described in Section 5.1. Selected events are categorized into exclusive signal regions, with enhanced signal-to-background ratios. In addition, control regions (CRs) are defined where a specific background is dominant and thereby allow the adjustment of simulated predictions for the background contribution to the observed data. Signal and control regions, which are included in the fit described in Section 8, are described in Section 5.2 together with validation regions (VRs) used to validate the simulation of Z + jets.

5.1 Event selection

Depending on the trigger, transverse momentum requirements are applied to selected electron, muon, and $\tau_{\text{had-vis}}$ candidates. They are summarised in Table 2 and their per-object efficiencies are given in Section 4. Due to the increasing luminosity and the different pile-up conditions, the p_{T} thresholds of the triggers have increased during data-taking in 2016, which is taken into account in the p_{T} requirements of the event selection. In the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel, the triggers for multiple light leptons are only used if the highest-p_{T} light lepton does not pass the corresponding single light-lepton trigger p_{T} requirement. This ensures that each trigger selects an exclusive set of events.

All channels require the exact number of identified 'loose' leptons, i.e. electrons, muons and $\tau_{\text{had-vis}}$, as defined in Section 4, corresponding to their respective final state. Events with additional 'loose' leptons are rejected. Both leptons are required to be of opposite charge and they have to fulfill the p_{T} requirements

of the respective trigger shown in Table 2. The selected $\tau_{\text{had-vis}}$ in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel is required to have $p_{\text{T}} > 30\,\text{GeV}$.

Only events with $E_{\text{T}}^{\text{miss}} > 20\,\text{GeV}$ are selected to reject events without neutrinos. In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel with two same-flavor (SF) light leptons this requirement is further tightened to suppress the large $Z \to \ell\ell$ background. For the same reason, requirements are introduced on the invariant mass of two light leptons ($m_{\ell\ell}$) and on the $E_{\text{T}}^{\text{miss}}$ calculated only from the physics objects without the soft track term ($E_{\text{T}}^{\text{miss, hard}}$). Requirements on the angular distance between the visible decay products of the two selected tau lepton decays ($\Delta R_{\tau\tau}$) and the projected angular distance in η ($|\Delta\eta_{\tau\tau}|$) are applied in all channels to reject non-resonant background events. Requirements are applied to the fractions of the tau-lepton momenta carried by each visible decay product $x_i = p_i^{\text{vis}} / \left(p_i^{\text{vis}} + p_i^{\text{miss}}\right)$, where p_i^{vis} and p_i^{miss} are the visible and missing momenta of the ith tau lepton, ordered in descending p_{T}, calculated in the collinear approximation [91], to suppress events with $E_{\text{T}}^{\text{miss}}$ that is incompatible with a di-tau decay. Low transverse mass (m_{T}), calculated from $E_{\text{T}}^{\text{miss}}$ and the momenta of the selected light lepton and $\tau_{\text{had-vis}}$, is required in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel to reject events with leptonic W decays. A requirement on the di-tau mass calculated in the collinear approximation ($m_{\tau\tau}^{\text{coll}}$) of $m_{\tau\tau}^{\text{coll}} > m_Z - 25\,\text{GeV}$ is introduced in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel to suppress events from $Z \to \ell\ell$ and to ensure orthogonality between this measurement and the measurement of $H \to WW^* \to \ell\nu\ell\nu$ [92], which has a similar final state.

All channels require at least one jet (j_1) with $p_{\text{T}}^{j_1} > 40\,\text{GeV}$ as the analysis targets either boosted Higgs bosons or those produced by VBF. Since 2016 the di-$\tau_{\text{had-vis}}$ first-level trigger requires a jet with $p_{\text{T}} > 25\,\text{GeV}$ calibrated at trigger level with $|\eta| < 3.2$ in addition to the two $\tau_{\text{had-vis}}$ candidates. In the $\tau_{\text{had}}\tau_{\text{had}}$ channel the jet p_{T} requirement is thus raised to $p_{\text{T}}^{j_1} > 70\,\text{GeV}$ to achieve uniform trigger selection efficiency as a function of $p_{\text{T}}^{j_1}$. The trigger efficiency for the additional jet ranges from 95–100% for these requirements. In the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, the top-quark background is suppressed by requiring that no jet with $p_{\text{T}} > 25\,\text{GeV}$ is tagged as a b-jet. The event selection for the three analysis channels is summarized in Table 3.

5.2 Signal, control and validation regions

To exploit signal-sensitive event topologies, a 'VBF ' and a 'boosted' analysis category are defined without any overlap in phase space. The VBF category targets events with a Higgs boson produced by VBF and is characterized by the presence of a second high-p_{T} jet ($p_{\text{T}}^{j_2} > 30\,\text{GeV}$). In addition, the two jets are required to be in opposite hemispheres of the detector with a large pseudorapidity separation of $|\Delta\eta_{jj}| > 3$ and their invariant mass (m_{jj}) is required to be larger than 400 GeV. The selected leptons are required to have η-values that lie between those of the two jets ('central leptons'). Although this category is dominated by VBF production, it also includes significant contributions from ggF production, amounting to up to 30% of the total expected Higgs-boson signal.

The boosted category targets events with Higgs bosons produced through ggF with an additional recoiling jet, which is motivated by an on average higher p_{T} of the boson for $H \to \tau\tau$ compared to the largest background from $Z \to \tau\tau$. It contains all events with $p_{\text{T}}^{\tau\tau} > 100\,\text{GeV}$ that do not pass the VBF selection. In addition to events from ggF, the boosted categories contain sizeable contributions from VBF and VH production of 10–20% of the expected signal. Events that pass the event selection, detailed in Table 3, but do not fall into the VBF or boosted categories are not used in the analysis.

Table 3: Summary of the event selection requirements for the three analysis channels that are applied in addition to the respective lepton p_{T} requirements listed in Table 2. $E_{\mathrm{T}}^{\mathrm{miss,\,hard}}$ is an alternative $E_{\mathrm{T}}^{\mathrm{miss}}$ calculated only from the physics objects without the soft-track term.

	$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$		$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$
	$ee/\mu\mu$	$e\mu$		
$N_{e/\mu}^{\mathrm{loose}} = 2$, $N_{\tau_{\mathrm{had\text{-}vis}}}^{\mathrm{loose}} = 0$			$N_{e/\mu}^{\mathrm{loose}} = 1$, $N_{\tau_{\mathrm{had\text{-}vis}}}^{\mathrm{loose}} = 1$	$N_{e/\mu}^{\mathrm{loose}} = 0$, $N_{\tau_{\mathrm{had\text{-}vis}}}^{\mathrm{loose}} = 2$
e/μ : Medium, gradient iso.			e/μ : Medium, gradient iso.	
			$\tau_{\mathrm{had\text{-}vis}}$: Medium	$\tau_{\mathrm{had\text{-}vis}}$: Tight
Opposite charge			Opposite charge	Opposite charge
$m_{\tau\tau}^{\mathrm{coll}} > m_Z - 25\,\mathrm{GeV}$			$m_{\mathrm{T}} < 70\,\mathrm{GeV}$	
$30 < m_{\ell\ell} < 75\,\mathrm{GeV}$	$30 < m_{\ell\ell} < 100\,\mathrm{GeV}$			
$E_{\mathrm{T}}^{\mathrm{miss}} > 55\,\mathrm{GeV}$	$E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$		$E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$	$E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$
$E_{\mathrm{T}}^{\mathrm{miss,\,hard}} > 55\,\mathrm{GeV}$				
$\Delta R_{\tau\tau} < 2.0$			$\Delta R_{\tau\tau} < 2.5$	$0.8 < \Delta R_{\tau\tau} < 2.5$
$\lvert\Delta\eta_{\tau\tau}\rvert < 1.5$			$\lvert\Delta\eta_{\tau\tau}\rvert < 1.5$	$\lvert\Delta\eta_{\tau\tau}\rvert < 1.5$
$0.1 < x_1 < 1.0$			$0.1 < x_1 < 1.4$	$0.1 < x_1 < 1.4$
$0.1 < x_2 < 1.0$			$0.1 < x_2 < 1.2$	$0.1 < x_2 < 1.4$
$p_{\mathrm{T}}^{j_1} > 40\,\mathrm{GeV}$			$p_{\mathrm{T}}^{j_1} > 40\,\mathrm{GeV}$	$p_{\mathrm{T}}^{j_1} > 70\,\mathrm{GeV}$, $\lvert\eta_{j_1}\rvert < 3.2$
$N_{b\text{-jets}} = 0$			$N_{b\text{-jets}} = 0$	

Using $p_{\mathrm{T}}^{\tau\tau}$, $\Delta R_{\tau\tau}$ and m_{jj}, both inclusive categories are split further into 13 exclusive signal regions with different signal-to-background ratios to improve the sensitivity. Table 4 summarizes the analysis categories and signal region definitions. Figure 1 illustrates the signal and background composition in the signal and control regions.

Six control regions are defined to constrain the normalization of the dominant backgrounds in regions of phase space where their purity is high. Their definitions are summarized in Table 5. Two $Z \to \ell\ell$ CRs, which both consist of more than 90% of $Z \to \ell\ell$ events, are defined by applying the same selection as for the SF $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ VBF and boosted inclusive regions, respectively, but with the $m_{\ell\ell}$ requirement modified to $80 < m_{\ell\ell} < 100\,\mathrm{GeV}$. The top quark background is characterized by the presence of b-jets. Four separate top CRs are defined by inverting the b-jet veto in the inclusive VBF and boosted categories for each of the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ and $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channels. The top CRs in the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel consist of about 80% of top-quark events. For the top CRs in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel, the requirement on $m_{\mathrm{T}} < 70\,\mathrm{GeV}$ is replaced by $m_{\mathrm{T}} > 40\,\mathrm{GeV}$ to further enhance the purity to about 70% in the VBF top CR and about 60% in the boosted top CR. No such control regions are defined for the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel since the top and $Z \to \ell\ell$ backgrounds are negligible in this case.

One validation region is defined for each signal region ('$Z \to \tau\tau$ VRs') to validate the event yields and kinematic distributions of simulated $Z \to \tau\tau$ events. The $Z \to \tau\tau$ VRs are composed of $Z \to \ell\ell$ events with similar kinematics as the $Z \to \tau\tau$ background in the respective signal regions. This is achieved by starting with an event selection that is based on the SF $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel preselection with the following differences that account for the selection of light leptons instead of decay products from tau leptons: The $m_{\tau\tau}^{\mathrm{coll}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss,\,hard}}$ requirements are dropped and the $m_{\ell\ell}$ requirement is inverted to $m_{\ell\ell} > 80\,\mathrm{GeV}$. The other requirements on tau-lepton decays are replaced with requirements on the two light leptons. In particular, the requirements on $p_{\mathrm{T}}^{\tau\tau}$ are substituted by the p_{T} of the Z boson computed from the p_{T} of the light leptons ($p_{\mathrm{T}}^{\ell\ell}$). Requirements on jets are unchanged since they define the shape of most kinematic

Table 4: Definition of the VBF and boosted analysis categories and of their respective signal regions (SRs). The selection criteria, which are applied in addition to those described in Table 3, are listed for each channel. The VBF high-p_{T} SR is only defined for the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel, resulting in a total of seven VBF SRs and six boosted SRs. All signal regions are exclusive and their yields add up to those of the corresponding VBF and boosted inclusive regions. In particular, the SRs marked with 'otherwise' contain all events that pass the corresponding inclusive event selection for an analysis channel but do not pass any of the other SR selections.

Signal Region		Inclusive	$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$	$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$
VBF	High-$p_{\mathrm{T}}^{\tau\tau}$	$p_{\mathrm{T}}^{j_2} > 30\,\mathrm{GeV}$ $\lvert\Delta\eta_{jj}\rvert > 3$ $m_{jj} > 400\,\mathrm{GeV}$ $\eta_{j_1} \cdot \eta_{j_2} < 0$ Central leptons	—		$p_{\mathrm{T}}^{\tau\tau} > 140\,\mathrm{GeV}$ $\Delta R_{\tau\tau} < 1.5$
	Tight		$m_{jj} > 800\,\mathrm{GeV}$	$m_{jj} > 500\,\mathrm{GeV}$ $p_{\mathrm{T}}^{\tau\tau} > 100\,\mathrm{GeV}$	Not VBF high-p_{T} $m_{jj} > (1550 - 250 \cdot \Delta\eta_{jj})\,\mathrm{GeV}$
	Loose		Otherwise		
Boosted	high-$p_{\mathrm{T}}^{\tau\tau}$	Not VBF $p_{\mathrm{T}}^{\tau\tau} > 100\,\mathrm{GeV}$	$p_{\mathrm{T}}^{\tau\tau} > 140\,\mathrm{GeV}$ $\Delta R_{\tau\tau} < 1.5$		
	Low-$p_{\mathrm{T}}^{\tau\tau}$		Otherwise		

Table 5: Definitions of the six control regions (CRs) used to constrain the $Z \to \ell\ell$ and top backgrounds to the event yield in data in the $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ and $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channels. 'SF' denotes a selection of same-flavor light leptons.

Region	Selection
$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ VBF $Z \to \ell\ell$ CR	$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ VBF incl. selection, $80 < m_{\ell\ell} < 100\,\mathrm{GeV}$ (SF)
$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ boosted $Z \to \ell\ell$ CR	$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ boosted incl. selection, $80 < m_{\ell\ell} < 100\,\mathrm{GeV}$ (SF)
$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ VBF top CR	$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ VBF incl. selection, inverted b-jet veto
$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ boosted top CR	$\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ boosted incl. selection, inverted b-jet veto
$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ VBF top CR	$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ VBF incl. selection, inverted b-jet veto, $m_{\mathrm{T}} > 40\,\mathrm{GeV}$
$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ boosted top CR	$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ boosted incl. selection, inverted b-jet veto, $m_{\mathrm{T}} > 40\,\mathrm{GeV}$

distributions for Z-boson production similarly in the SRs and the $Z \to \tau\tau$ VRs. More than 99% of the selected events are from $Z \to \ell\ell$ in all $Z \to \tau\tau$ VRs.

6 Background estimation

The final-state topologies of the three analysis channels have different background compositions which necessitates different strategies for the background estimation. In each SR, the number of expected background events and the associated kinematic distributions are derived from a mixture of data-driven methods and simulation.

Background contributions with prompt, light leptons and $\tau_{\text{had-vis}}$ are estimated from simulation. If their contribution is significant, their normalization is constrained by the observed event yields in CRs. For smaller contributions of this type, their normalisation is entirely taken from the theoretical cross sections with the precision in QCD listed in Table 1. This includes di-boson processes and a small contribution

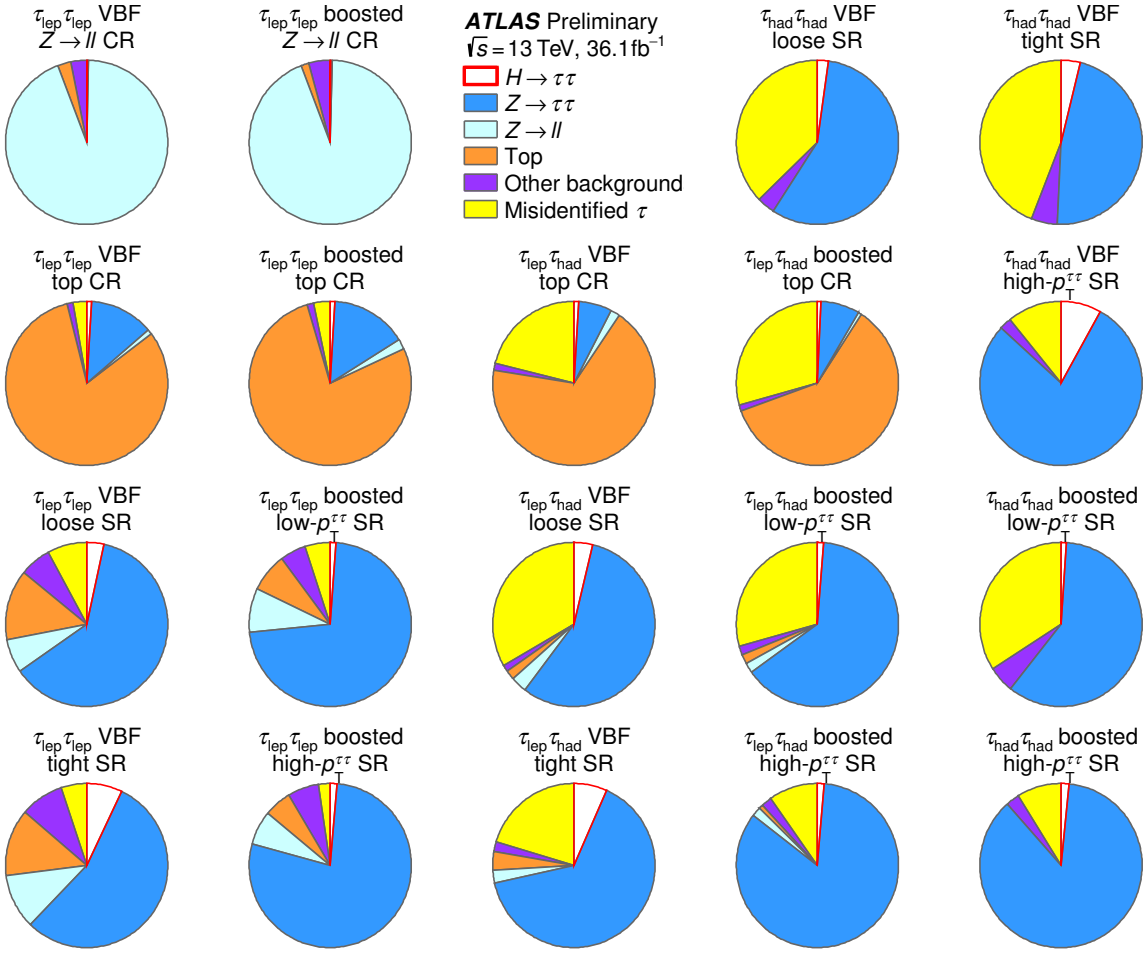


Figure 1: Expected signal and background composition in six control regions (CRs) and the 13 signal regions (SRs) used in the analysis.

from EW production of W/Z bosons from VBF. Contributions from light- and heavy-flavor jets that are misidentified as prompt, light leptons or $\tau_{\text{had-vis}}$ (labeled as 'fake-ℓ ' and 'fake-$\tau_{\text{had-vis}}$' backgrounds, respectively, and collectively as 'misidentified τ ', throughout this paper) are estimated using data-driven methods. The contamination from $H \to WW^*$ decays is treated as a background in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, while it is negligible in other channels.

For the background sources that have their normalization constrained using data, Table 6 shows the normalization factors and their uncertainties obtained from the fit (see Section 8). For simulated backgrounds, the factors compare the background normalizations to values determined from their theoretical cross sections. The normalization factor for the data-driven fake-$\tau_{\text{had-vis}}$ background scales the event yield of the template of events that fail the opposite-charge requirement (see Section 6.4). The $Z \to \tau\tau$ normalization is constrained by data in the $m_{\tau\tau}^{\text{MMC}}$ distributions of the signal regions. Systematic uncertainties are the dominant contribution to the normalization factor uncertainties.

Table 6: Normalization factors for backgrounds that have their normalization constrained using data in the fit including all statistical and systematic uncertainties described in Section 7, but without uncertainties in total simulated cross sections extrapolated to the selected phase space. Systematic uncertainties are the dominant contribution to the normalization factor uncertainties. Also shown are the analysis channels the factors are applied to.

Background	Channel	Normalization factors	
		VBF	Boosted
$Z \to \ell\ell$ (CR)	$\tau_{\text{lep}}\tau_{\text{lep}}$	$0.87^{+0.34}_{-0.30}$	$1.25^{+0.29}_{-0.24}$
Top (CR)	$\tau_{\text{lep}}\tau_{\text{lep}}$	1.19 ± 0.09	1.06 ± 0.05
Top (CR)	$\tau_{\text{lep}}\tau_{\text{had}}$	$1.53^{+0.30}_{-0.27}$	1.12 ± 0.07
Fake-$\tau_{\text{had-vis}}$ (data-driven)	$\tau_{\text{had}}\tau_{\text{had}}$	0.88 ± 0.12	
$Z \to \tau\tau$ (fit in each SR)	$\tau_{\text{lep}}\tau_{\text{lep}}$,$\tau_{\text{lep}}\tau_{\text{had}}$,$\tau_{\text{had}}\tau_{\text{had}}$	$1.04^{+0.10}_{-0.09}$	1.11 ± 0.05

6.1 $Z \to \tau\tau$ background validation

The Drell-Yan process $pp \to Z/\gamma^* \to \tau\tau$ is a dominant irreducible background in all analysis categories and contributes between 50–90% of the total background depending on the signal region. The separation between the Drell-Yan and the $H \to \tau\tau$ signal processes is limited by the $m_{\tau\tau}^{\text{MMC}}$ resolution.

The modeling of this important background is validated using $Z \to \tau\tau$ VRs that consist of $Z \to \ell\ell$ events. In Fig. 2, the observed distributions of several variables are compared to simulation normalized to the event yield in data. The selected observables correspond to either variables correlated with $m_{\tau\tau}^{\text{MMC}}$ ($p_{\text{T}}^{\ell_1}$ and $p_{\text{T}}^{\ell_2}$), or to major variables used for categorization ($p_{\text{T}}^{\ell\ell}$, $\Delta R_{\ell\ell}$, $\Delta\eta_{jj}$ and m_{jj}), or to variables to which different requirements are applied in each decay channel ($p_{\text{T}}^{j_1}$). Generally, the SHERPA simulation describes the shape of data within the experimental and theoretical uncertainties (see Section 7), with the exception of a slight trend in the ratio of data to simulation as a function of $\Delta\eta_{jj}$ and m_{jj} shown in Fig. 2. These trends have no impact on the modelling of $m_{\tau\tau}^{\text{MMC}}$. Reweighting the simulation with the observed m_{jj} distribution, which is an important variable for VBF categorization, has negligible impact on the measurement. In the fit, the normalization of the $Z \to \tau\tau$ background is correlated across the decay channels and constrained by data in the $m_{\tau\tau}^{\text{MMC}}$ distributions of the signal regions associated to the boosted and VBF categories, independently. As shown in Table 6, it is constrained to ±5% in the boosted and to ±9% in the VBF categories. The relative acceptance of events among the signal regions within a category is validated by applying the corresponding event selection criteria to the $Z \to \tau\tau$ VRs. The expected relative acceptance from simulation agrees with data within uncertainties for all regions. Figures 7 and 8 show the good modeling of the $Z \to \tau\tau$ $m_{\tau\tau}^{\text{MMC}}$ distribution in all signal regions. Additional uncertainties in the relative acceptances and on the shape of the $m_{\tau\tau}^{\text{MMC}}$ distributions in the signal regions are evaluated from theoretical and experimental uncertainties described in Section 7.

6.2 $Z \to \ell\ell$ background

Z-boson decays into light leptons are a significant background for the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels where mismeasured $E_{\text{T}}^{\text{miss}}$ can bias the $m_{\tau\tau}^{\text{MMC}}$ of light-lepton pairs to be of similar values as the expected signal. The observed event yields in the $Z \to \ell\ell$ CRs constrain the normalization of simulated $Z \to \ell\ell$ events in the

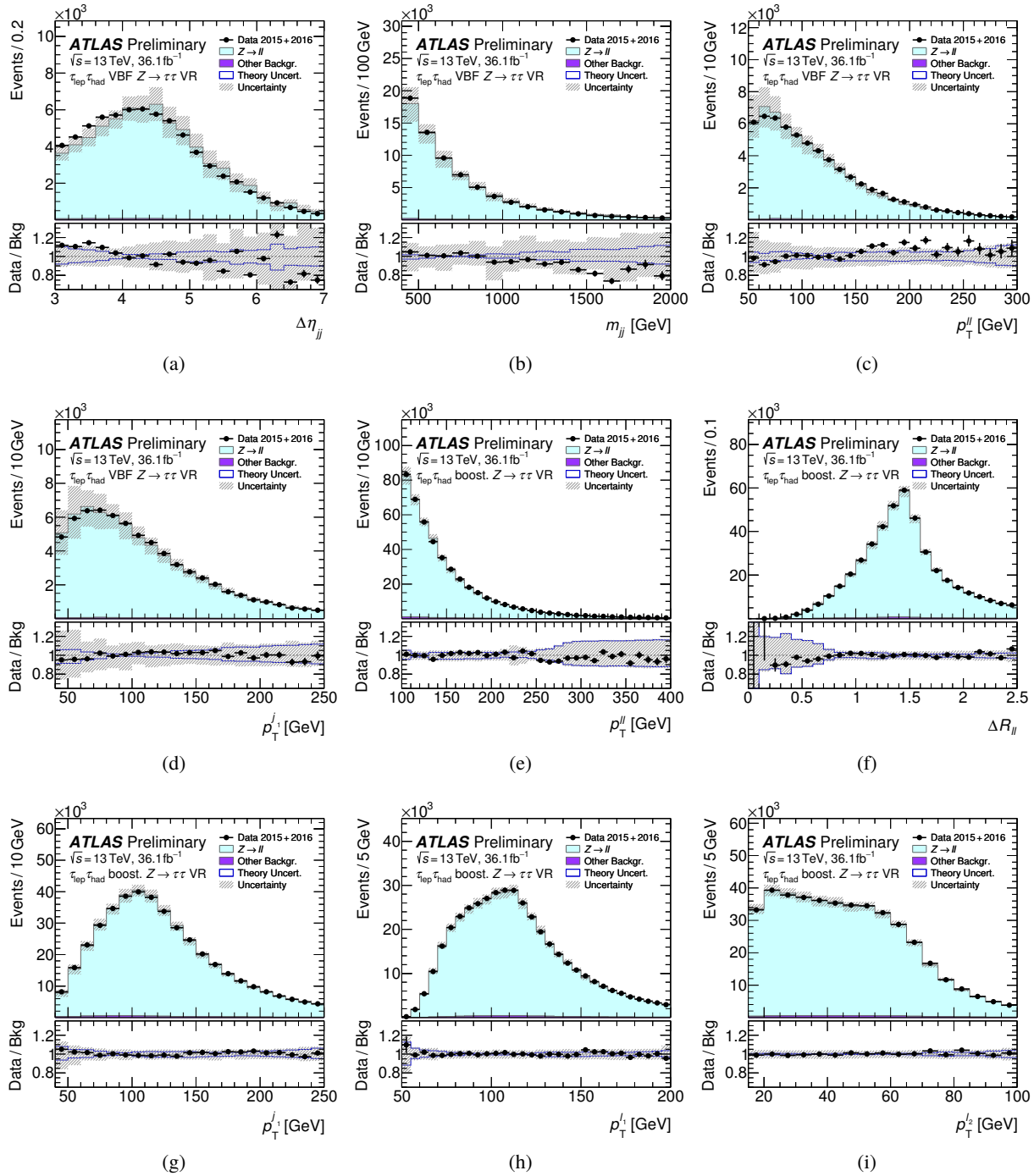
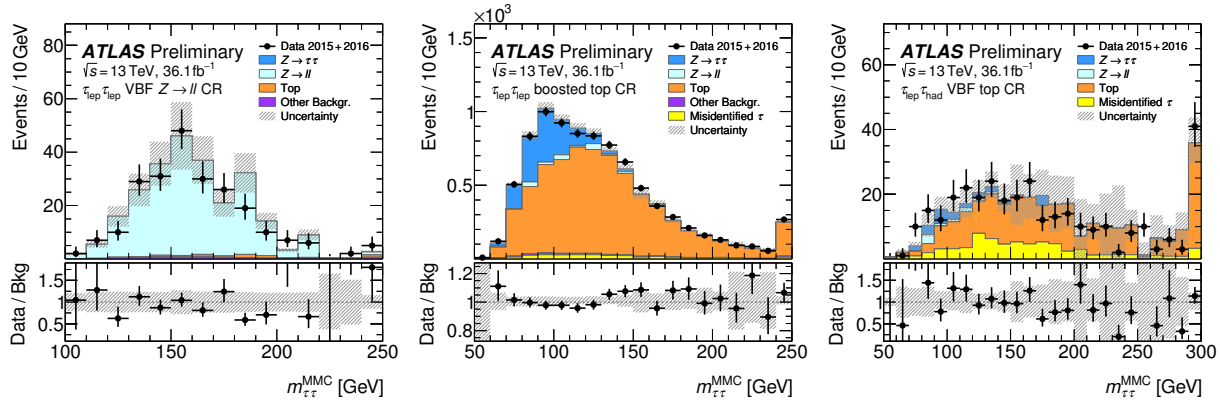


Figure 2: Observed and expected distributions of (a) the angular distance between of the two highest p_{T} jets ($\Delta\eta_{jj}$), (b) the invariant mass of the two highest-p_{T} jets (m_{jj}), (c) the p_{T} of the di-lepton system ($p_{\mathrm{T}}^{\ell\ell}$) and (d) the p_{T} of the highest-p_{T} jet ($p_{\mathrm{T}}^{j_1}$) in the $Z \to \tau\tau$ validation region (VR) corresponding to the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ VBF inclusive category as well as (e) the angular distance of the light leptons ($\Delta R_{\ell\ell}$), (f) the p_{T} of the di-lepton system ($p_{\mathrm{T}}^{\ell\ell}$), (g) the p_{T} of the highest-p_{T} jet ($p_{\mathrm{T}}^{j_1}$), (h) the p_{T} of the highest-p_{T} light lepton ($p_{\mathrm{T}}^{\ell_1}$) and (i) the p_{T} of the second highest-p_{T} light lepton ($p_{\mathrm{T}}^{\ell_2}$) in the $Z \to \tau\tau$ VR corresponding to the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ boosted inclusive category. The predictions in these validation regions are not computed by the fit and are normalized to the event yield in data. The size of the combined statistical, experimental and theoretical uncertainties are indicated by the hatched bands. The ratios of the data to the background model are shown in the lower panels together with the theoretical uncertainties in the SHERPA simulation of $Z \to \ell\ell$, which are indicated by the blue lines.

$\tau_{\text{lep}}\tau_{\text{lep}}$ channel to ±40% in the VBF and to ±25% in the boosted categories, as shown in Table 6. The good modeling of the $m_{\tau\tau}^{\text{MMC}}$ distribution in the $\tau_{\text{lep}}\tau_{\text{lep}}$ VBF $Z \to \ell\ell$ CR is shown in Fig. 3a. In other channels, the contribution from $Z \to \ell\ell$ events is normalized to its theoretical cross section. In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, $Z \to \ell\ell$ background contributes primarily through $Z \to ee$ decays where an electron is misidentified as a $\tau_{\text{had-vis}}$ candidate. Due to the dedicated electron veto algorithm applied to selected 1-prong $\tau_{\text{had-vis}}$ candidates (see Section 5.1) this background is comparably small. This and other backgrounds from light leptons misidentified as $\tau_{\text{had-vis}}$ in this channel are estimated from simulation with the probability for electrons misidentified as $\tau_{\text{had-vis}}$ candidates in simulation scaled to match the misidentification probability observed in data [87].



(a) (b) (c)

Figure 3: For the control regions (CRs) defined in Section 5, comparisons between data and predictions as computed by the fit for the reconstructed di-tau invariant mass ($m_{\tau\tau}^{\text{MMC}}$). Shown are (a) the $\tau_{\text{lep}}\tau_{\text{lep}}$ VBF $Z \to \ell\ell$ control region (CR), (b) the $\tau_{\text{lep}}\tau_{\text{lep}}$ boosted Top CR and (c) the $\tau_{\text{lep}}\tau_{\text{had}}$ VBF Top CR. The size of the combined statistical, experimental and theoretical uncertainties in the background are indicated by the hatched band. The ratios of the data to the background model are shown in the lower panels.

6.3 Top quark background

The production of $t\bar{t}$ pairs or single top quarks is a significant background ('top background') for the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, due to the production of prompt light leptons with associated $E_{\text{T}}^{\text{miss}}$ in the top-quark decay chain $t \to Wb$, $W \to \ell\nu, \tau\nu$. Events where a selected tau-lepton decay product is misidentified, are estimated using data-driven methods that are discussed in Section 6.4. The remaining top background is estimated from simulation. In the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels the normalization of simulated top background is additionally constrained by the absolute event yields in their respective top CRs to ±30% in the $\tau_{\text{lep}}\tau_{\text{had}}$ VBF top CR and less than ±10% in the other top CRs, as shown in Table 6. Figures 3b and 3c show $m_{\tau\tau}^{\text{MMC}}$ distributions in the $\tau_{\text{lep}}\tau_{\text{lep}}$ boosted top CR and the $\tau_{\text{lep}}\tau_{\text{had}}$ VBF top CR, respectively.

6.4 Backgrounds from misidentified τ

Apart from the small contribution from light leptons misidentified as $\tau_{\text{had-vis}}$ described in Section 6.2, hadronic jets can be misidentified as $\tau_{\text{had-vis}}$, electrons and muons. These sources of background contribute

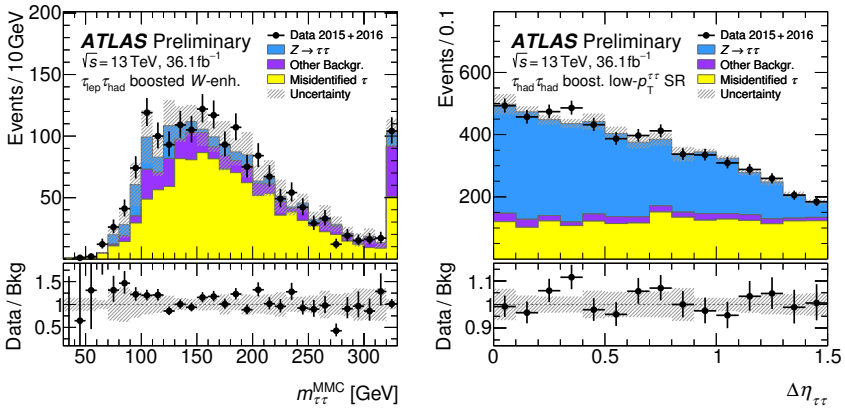
up to half of the total background depending on the signal region and are estimated with data-driven techniques. Since the background sources are dependent on the event topology, specific methods are applied to each individual channel.

In the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel, the main sources of the fake-ℓ background are multijet, W bosons in association with jets and semi-leptonically decaying $t\bar{t}$ events. All these background sources are treated together. Fake-ℓ regions are defined in data by inverting the isolation requirement of the light lepton with the second-highest p_{T} in each signal region and, if it is an electron, by relaxing its identification criteria to 'loose'. Fake-ℓ templates are created from these samples by subtracting top and $Z \to \ell\ell$ backgrounds that produce real light leptons, estimated from simulation. The normalization of each template is then scaled by a factor that corrects for the inverted-isolation requirement. These correction factors are computed for each combination of lepton flavor from events that pass the $\tau_{\text{lep}}\tau_{\text{lep}}$ selection but have same-charge light leptons, subtracting simulated top and $Z \to \ell\ell$ backgrounds. Fake-ℓ background in the top CRs is estimated following the same procedure.

Systematic uncertainties in the shape and normalization of this background depend on the p_{T} of the second-highest-p_{T} lepton and are estimated as follows. A closure test on the background estimate is performed using events where the leptons are required to have the same charge and amounts to an uncertainty ranging between 20–65%. An uncertainty of the heavy-flavor contents is estimated by using isolation correction factors that are computed from samples selected with inverted b-jet requirement which amounts to up to 50%. Minor contributions come from the uncertainty in the fractional composition of the fake-ℓ background in top quark decays, multijets events and W-boson production.

In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, a 'fake-factor' method is used to derive estimates for fake-$\tau_{\text{had-vis}}$ events, composed mainly of multijet events and W boson production in association with jets. A fake-factor is defined as the ratio of the number of events where the highest-p_{T} jet is identified as 'medium' $\tau_{\text{had-vis}}$ candidate to the number of events with a highest-p_{T} jet that passes a very loose but fails the 'medium' $\tau_{\text{had-vis}}$ identification. Fake-factors depend on the p_{T} and track multiplicity of the $\tau_{\text{had-vis}}$ candidate and on the type of parton initiating the jet. Therefore, they are computed depending on these quantities, independently in quark-jet dominated 'W-enhanced' regions and gluon-jet dominated 'multijet-enhanced regions'. The W-enhanced regions are defined by inverting the $m_{\text{T}} < 70\,\text{GeV}$ requirement and the multijet-enhanced regions are defined by inverting the light-lepton isolation, relative to the inclusive boosted and VBF selections. Backgrounds from Z boson production with associated jets and semi-leptonically decaying $t\bar{t}$ have fake-factors similar to those found in backgrounds from W bosons and their contributions are negligible. The fake-factors range between 0.15–0.25 for 1-prong and between 0.01–0.04 for 3-prong $\tau_{\text{had-vis}}$. To obtain the fake-$\tau_{\text{had-vis}}$ background estimate for the signal regions, these fake-factors are first weighted by the multijets-to-W fraction. The weighted fake-factors are then applied to events in regions defined by the selections of the corresponding signal regions, except that the highest-p_{T} $\tau_{\text{had-vis}}$ candidate passes a very loose and fails the 'medium' $\tau_{\text{had-vis}}$ identification ('anti-ID' regions). The relative multijets contribution in each anti-ID region is estimated from the yield in events that fail the light-lepton isolation requirement, multiplied by a factor that corrects for this requirement. The multijet contribution varies by more than 50% and depends on the lepton p_{T} and on the $\Delta\phi$ between $\tau_{\text{had-vis}}$ and $E_{\text{T}}^{\text{miss}}$. The good agreement between data and background estimates is shown in Fig. 4a for the main discriminant of the analysis, $m_{\tau\tau}^{\text{MMC}}$, in the boosted W-enhanced region.

The dominant contribution to the uncertainties in this background originates from the statistical uncertainty in the individual fake-factors of up to 10% in the boosted signal regions and up to 35% in the VBF signal regions. Minor contributions originate from the statistical uncertainty in the anti-ID regions and uncertainties in the fractional size of the multijet contribution to the fake-$\tau_{\text{had-vis}}$ background.



(a) (b)

Figure 4: Observed distributions and predictions computed by the fit for (a) $m_{\tau\tau}^{\text{MMC}}$ in the W-enhanced region of the $\tau_{\text{lep}}\tau_{\text{had}}$ boosted inclusive category, and (b) $\Delta\eta$ between the two $\tau_{\text{had-vis}}$, for events in the boosted low-$p_{\text{T}}^{\tau\tau}$ signal region (SR) of the $\tau_{\text{had}}\tau_{\text{had}}$ channel. The size of the combined statistical, experimental and theoretical uncertainties in the background are indicated by the hatched bands. The ratios of the data to the background model are shown in the lower panels.

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the multijets background is modeled using a template extracted from data that pass the signal region selections, but where the $\tau_{\text{had-vis}}$ candidates are allowed to have two tracks and required to fail the opposite-charge requirement (nOC region). The contribution of events with true tau leptons from other SM processes is subtracted from this template using simulation. The template is then reweighted using scale factors dependent on the difference in ϕ between the $\tau_{\text{had-vis}}$ candidates ($\Delta\phi_{\tau\tau}$). These scale factors are derived by comparing the template from an nOC selection to a region obtained by requiring the $\tau_{\text{had-vis}}$ pair to have opposite charge and the second-highest-p_{T} $\tau_{\text{had-vis}}$ to fail the 'tight' but pass 'medium' identification requirements. As the yield of events that pass these identification requirements is small, the scale factors are derived from events that pass the $\tau_{\text{had}}\tau_{\text{had}}$ selection with looser $\Delta\eta_{\tau\tau}$ and $\Delta R_{\tau\tau}$ requirements to gain statistical power. The normalization of the multijet background is constrained in the fit by data in the $m_{\tau\tau}^{\text{MMC}}$ distribution in the signal regions. For this, a normalization factor is defined and correlated across all $\tau_{\text{had}}\tau_{\text{had}}$ signal regions. Figure 4b shows good agreement between data and background predictions in the distribution of the $\Delta\eta$ between the two $\tau_{\text{had-vis}}$, which has a quite different shape for the multijets than for the $Z \to \tau\tau$ process. In this figure, events are selected that pass the $\tau_{\text{had}}\tau_{\text{had}}$ boosted low-$p_{\text{T}}^{\tau\tau}$ selection. Contributions from other backgrounds such as W with associated jets range from 2–5% in the $\tau_{\text{had}}\tau_{\text{had}}$ SRs.

The event yield of this background is constrained by data to ±15% in the signal regions as shown in Table 6. The dominant contribution to the uncertainties that affect the $m_{\tau\tau}^{\text{MMC}}$ shape originates from the statistical uncertainties in the $\Delta\phi_{\tau\tau}$ scale factors and amount to 8%. The systematic uncertainty in these scale factors is estimated by comparing them to scale factors computed from the nOC region and a CR defined by requiring opposite-charge $\tau_{\text{had-vis}}$ to pass 'loose' but not 'medium' identification. Minor contributions arise from the uncertainty in the extrapolation from the nOC requirement and the uncertainty from the subtraction of simulated backgrounds. The combination of these uncertainties lead to a total 10% uncertainty in the $m_{\tau\tau}^{\text{MMC}}$ template shape.

7 Systematic uncertainties

The expected signal and background yields in the various signal and control regions as well as the shape of the $m_{\tau\tau}^{\text{MMC}}$ distributions in the signal regions are affected by systematic uncertainties. These are discussed below, grouped into three categories: theoretical uncertainties in signal, theoretical uncertainties in background and experimental uncertainties. The uncertainties in backgrounds from misidentified tau leptons, which are estimated using data-driven techniques, are discussed in Section 6.4. The effects of all uncertainties are included in the fit model described in Section 8.

7.1 Theoretical uncertainties in signal

The procedures to estimate the uncertainty in the Higgs production cross sections follow the recommendations by the LHC Higgs Cross Section Working Group [93]. They are briefly summarized below. Uncertainties are evaluated separately for their impact on the total cross section, their impact on the acceptance in different SRs, and on the shape of the $m_{\tau\tau}^{\text{MMC}}$ distribution in each SR.

The cross section of ggF production in association with an exclusive number of additional jets has large uncertainties from higher-order QCD corrections [94]. In this analysis, the boosted and VBF categories almost exclusively select ggF events with one and two additional jets, respectively. To take this effect into account, nine uncertainty sources are included. Four sources account for uncertainties in the jet multiplicities due to missing higher-order corrections: Two sources account for yield uncertainties due to factorization and renormalization scale variations and two sources account for migration uncertainties of zero to one and one to at least two jets in the event, respectively, using the STWZ [95] and BLPTW [95–97] predictions as an input. Three uncertainty sources parameterize modeling uncertainties in the Higgs-boson p_{T}, two of which encapsulate the migration uncertainty between the intermediate and high-p_{T} regions of events with at least one jet, and one which encapsulates the uncertainty in the loop corrections due to the top-quark mass uncertainty, where the difference between the LO and NLO predictions is taken as an uncertainty due to missing higher-order corrections. Two sources account for the acceptance uncertainties of ggF production in the VBF phase space from selecting exactly two and at least three jets, respectively. Their estimation uses an extension of the Stewart-Tackmann method [98, 99]. The resulting acceptance uncertainties from these nine sources range from 1–10%, with the dominant uncertainties due to the modeling of the Higgs p_{T} distribution in all SRs, to the scale variations in the boosted SRs, and to the acceptance uncertainties in the VBF signal regions.

For VBF and VH production cross sections, the uncertainties due to missing higher-order QCD corrections are estimated by varying the factorization and renormalization scales by factors of two around the nominal scale. The resulting uncertainties in the total cross section are below 1% for VBF and WH production and below 5% for ZH production. The uncertainties in the acceptance in the different SRs are of about 1% for VBF production in all categories. For VH production the relative acceptance uncertainty ranges between −10% and +20% in VBF SRs. It is below 10% in boosted SRs.

Uncertainties related to the simulation of the underlying event, hadronization and parton shower for all signal samples are estimated by comparing the acceptance when using the default UEPS model from Pythia 8.212 with an alternative UEPS model from Herwig 7.0.3. The resulting acceptance uncertainties range from 2–26% for ggF production and from 2–18% for VBF production, depending on the signal region. The PDF uncertainties are estimated using 30 eigenvector variations and two α_s variations that are evaluated independently with respect to the default PDF set PDF4LHC15 [40]. The total uncertainty

due to these variations is 5% or less depending on the SR and the Higgs production mode. Finally, an uncertainty in the $H \to \tau\tau$ decay branching ratio of 1% [93] affects the signal rates. All sources of theoretical uncertainties in the signal expectation are correlated across SRs.

7.2 Theoretical uncertainties in backgrounds

Uncertainties from missing higher order corrections, the PDF parameterization, underlying-event modeling and from parton-shower modeling are also considered for the dominant $Z \to \tau\tau$ background. Since its overall normalization is constrained separately in the boosted and VBF SRs, only the relative event migration and $m_{\tau\tau}^{\text{MMC}}$ shape uncertainties are considered. Unlike the overall normalization, they are treated as uncorrelated between the three analysis channels and the boosted and VBF SRs to describe the differences in the corresponding event selections. An additional parameter encapsulates the relative change in acceptance between the three analysis channels for each source of uncertainty, independently for boosted and VBF SRs to describe the correlated impact of the respective uncertainty. The largest sources of uncertainties are due to the CKKW matching [100] which are evaluated depending on the number of true jets and the Z-boson p_{T}. They are below 5% depending on the SR. The uncertainty in the measured cross section for electroweak Z production with two associated jets [101] is found to be small compared to the other uncertainties in Z-boson production.

The top-quark background normalization in the $\tau_{\text{lep}}\tau_{\text{lep}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels as well as the $Z \to \ell\ell$ background normalization in the $\tau_{\text{lep}}\tau_{\text{lep}}$ channel are constrained by data in dedicated CRs. All other simulated background contributions are normalized to their Monte Carlo prediction. For all simulated background contribution other than $Z \to \tau\tau$, no theory uncertainties are considered as their impact is small with respect to the uncertainties in the dominating backgrounds from $Z \to \tau\tau$ and misidentified leptons.

7.3 Experimental uncertainties

Experimental systematic uncertainties result from uncertainties in efficiencies for triggering, object reconstruction and identification, as well as from uncertainties in the energy scale and resolution of jets, $\tau_{\text{had-vis}}$, light leptons and $E_{\text{T}}^{\text{miss}}$. These uncertainties affect both event yields and the shape of the $m_{\tau\tau}^{\text{MMC}}$. The dominant experimental uncertainties in the final result are related to jet and $\tau_{\text{had-vis}}$ reconstruction. The impact of the electron- and muon-related uncertainties [79, 80, 102] on the measurement are generally small. Uncertainties of the integrated luminosity affect the number of predicted signal and background events, with the exception of processes that are normalized to data, see Table 6. This uncertainty is 2.1% for the combined 2015+2016 dataset, derived using a methodology similar to that detailed in Ref. [103], performed in August 2015 and May 2016.

The uncertainties of the $\tau_{\text{had-vis}}$ identification efficiency are in the range of 2–4.5% for the reconstruction efficiency [104], 3–14% for the trigger efficiency (depending on the $\tau_{\text{had-vis}}$ p_{T}), 5–6% for the identification efficiency and 3–14% for the rate at which an electron is misidentified as $\tau_{\text{had-vis}}$ (depending on the $\tau_{\text{had-vis}}$ η) [87]. The uncertainties of the b-tagging efficiencies are measured in dedicated calibration analyses [85] and are decomposed into uncorrelated components. Uncertainties in the efficiency to pass the JVT and forward JVT requirements are also considered [84, 105]. Simulated events are corrected for differences in these efficiencies between data and simulation and the associated uncertainties are propagated through the analysis.

The uncertainties of the $\tau_{\text{had-vis}}$ energy scale [87] are determined by fitting the Z-boson mass in $Z \to \tau\tau$ events, reconstructed using the visible tau decay products. The precision amounts to 2–3%, which is dominated by the uncertainty of background modeling. Additional uncertainties based on the modeling of the calorimeter response to single particles are added for $\tau_{\text{had-vis}}$ with $p_{\text{T}} > 50\,\text{GeV}$ [106]. The jet energy scale and its uncertainty are derived by combining information from test-beam data, LHC collision data and simulation [107]. The uncertainties from these measurements are factorized into eight principal components. Additional uncertainties that are considered are related to jet flavor, pile-up corrections, η-dependence, and high-p_{T} jets, yielding a total of 20 independent sources. The uncertainties amount to 1–6% per jet, depending on the jet p_{T}. The jet energy resolution uncertainties [108] are divided into 11 independent components and amount to 1–6%.

Since systematic uncertainties of the energy scales of all objects affect the reconstructed $E_{\text{T}}^{\text{miss}}$, this is recalculated after each variation is applied. The scale uncertainty of $E_{\text{T}}^{\text{miss}}$ due to the energy in the calorimeter cells not associated with physics objects is also taken into account [109]. The uncertainty of the resolution of $E_{\text{T}}^{\text{miss}}$ arises from the energy resolution uncertainties of each of the $E_{\text{T}}^{\text{miss}}$ terms and from the modeling and the effects of pile-up on the soft term (see Section 4).

8 Results

A maximum likelihood fit is performed on data to extract the parameter of interest defined as $\sigma_{H\to\tau\tau} \equiv \sigma_H \cdot \mathcal{B}(H \to \tau\tau)$ where σ_H is the total cross section of the considered Higgs boson production processes ggF, VBF and VH and where $\mathcal{B}(H \to \tau\tau)$ is the $H \to \tau\tau$ branching fraction. In this fit, the relative contributions from the various Higgs production processes are assumed as predicted by the SM.

A probability model is constructed that describes the $m_{\tau\tau}^{\text{MMC}}$ distributions in the 13 signal regions and the event yields in six control regions. The latter are included to constrain the normalizations of the dominant backgrounds. Each signal region is modeled by a product of Poisson distributions, where each such distribution describes the expected event count in intervals of $m_{\tau\tau}^{\text{MMC}}$. Each control region is modeled by a single Poisson distribution that describes the total expected event count in that region. Signal and background predictions depend on systematic uncertainties, which are parameterized as nuisance parameters and are constrained using Gaussian or log-normal probability distributions. The latter are used for normalization factors (see Table 6) to ensure that they are always positive. The dependence of the predictions on systematics-related nuisance parameters is modeled with an interpolation approach between yields obtained at different fixed systematic settings. A smoothing procedure is applied to remove occasional large local fluctuations in the $m_{\tau\tau}^{\text{MMC}}$ distribution templates which encode systematic uncertainties of some background processes in certain regions. For the measurement of $\sigma_{H\to\tau\tau}$, all theoretical uncertainties are included, except those related to the inclusive signal cross section, and are correlated as described in Section 7.1. The experimental uncertainties are fully correlated across categories and the background modeling uncertainties are generally uncorrelated, with the exception of the normalization factors as described in Section 6. Estimates of the parameter of interest and the confidence interval are calculated with the profile likelihood ratio [110] test statistic, whereas the test statistic $\tilde{q}_0$ is used to compute the significance of the deviation from the background-only hypothesis.

The observed (expected) significance of the signal excess with respect to the SM-background-only hypothesis computed from the likelihood fit is 4.4 (4.1) standard deviations, compatible with a SM Higgs boson with a mass $m_H = 125\,\text{GeV}$. Combined with the results of the search for $H \to \tau\tau$ using data at 7 and 8 TeV center-of-mass energies [15], the observed (expected) significance amounts to 6.4 (5.4) standard

Table 7: Observed event yields and predictions as computed by the fit in the $\tau_{\text{lep}}\tau_{\text{lep}}$ signal regions. Uncertainties include statistical and systematic uncertainties.

	$\tau_{\text{lep}}\tau_{\text{lep}}$ VBF		$\tau_{\text{lep}}\tau_{\text{lep}}$ boosted	
	Loose	Tight	Low-$p_{\text{T}}^{\tau\tau}$	High-$p_{\text{T}}^{\tau\tau}$
$Z \to \tau\tau$	148 ± 14	105 ± 13	2 992 ± 93	2 701 ± 66
$Z \to \ell\ell$	15.8 ± 5.2	20.6 ± 6.7	357 ± 54	235 ± 31
Top	33.3 ± 6.5	25.2 ± 4.6	319 ± 50	188 ± 29
VV	12.0 ± 2.1	10.8 ± 1.5	194.7 ± 8.5	196.2 ± 8.9
Misidentified τ	18.7 ± 9.7	9.8 ± 4.8	212 ± 93	81 ± 35
$ggF, H \to WW^*$	1.2 ± 0.2	1.4 ± 0.3	11.9 ± 2.6	16.4 ± 1.7
$VBF, H \to WW^*$	1.7 ± 0.2	4.1 ± 0.5	2.9 ± 0.3	2.9 ± 0.3
$ggF, H \to \tau\tau$	2.7 ± 1.0	2.0 ± 1.0	33.5 ± 8.9	33.0 ± 9.3
$VBF, H \to \tau\tau$	5.2 ± 1.5	11.4 ± 3.1	7.6 ± 2.1	8.2 ± 2.3
$WH, H \to \tau\tau$	< 0.1	< 0.1	2.4 ± 0.7	3.1 ± 0.9
$ZH, H \to \tau\tau$	< 0.1	< 0.1	1.3 ± 0.4	1.6 ± 0.5
Total background	231 ± 14	177 ± 12	4 089 ± 65	3 420 ± 57
Total signal	8.1 ± 2.3	13.5 ± 3.7	46 ± 12	47 ± 12
Data	237	188	4124	3444

deviations. In this combination, all nuisance parameters are treated as uncorrelated between the two LHC runs (in particular, the dominant Z to tautau background is estimated very differently in Run 1 and Run 2, as mentioned in Section 1).

For the measurement of $\sigma_{H\to\tau\tau}$ the relative contributions from the various Higgs production processes are assumed as predicted by the SM and the uncertainties related to the total signal cross section are excluded. The measured value of $\sigma_{H\to\tau\tau}$ is $3.71 \pm 0.59\,(\text{stat.})\,^{+0.87}_{-0.74}\,(\text{syst.})\,\text{pb}$, consistent with the SM prediction $\sigma^{\text{SM}}_{H\to\tau\tau} = 3.43 \pm 0.13\,\text{pb}$ [93]. The signal strength $\mu_{H\to\tau\tau}$ is defined as the ratio of the measured signal yield to the Standard Model expectation. It is computed by the fit described above, including uncertainties in the signal cross section and is evaluated to $1.09\,^{+0.18}_{-0.17}\,(\text{stat.})\,^{+0.27}_{-0.22}\,(\text{syst.})\,^{+0.16}_{-0.11}\,(\text{theory syst.})$.

Tables 7 to 9 summarize the fitted signal and background yields expected in each signal region for the $\sigma_{H\to\tau\tau}$ measurement. The signal events are given separately for each production mechanism. Within the uncertainties, good agreement is observed between the data and the model predictions for the sum of background components and a SM Higgs boson of mass $m_H = 125\,\text{GeV}$ with the measured $\sigma_{H\to\tau\tau}$. The $m_{\tau\tau}^{\text{MMC}}$ distributions in all signal regions with background predictions adjusted by the likelihood fit are shown in Figs. 7 and 8.

Table 10 shows a summary of the dominant uncertainties in $\sigma_{H\to\tau\tau}$, grouped by their respective sources and ranked by their fractional impact on the measurement. The dominant systematic uncertainties with the largest impact are shown in Fig. 5, together with a comparison to their nominal values used as input to the fit. Both are ranked by their fractional impact on the measurement of $\sigma_{H\to\tau\tau}$. To compute the impact for each nuisance parameter, a separate fit is performed again with the parameter fixed to its fitted value, and the resulting uncertainty in $\sigma_{H\to\tau\tau}$ is subtracted in quadrature from the uncertainty obtained in the original fit. The dominant uncertainties are related to the limited statistics in the simulated samples,

Table 8: Observed event yields and predictions as computed by the fit in the $\tau_{\text{lep}}\tau_{\text{had}}$ signal regions. Uncertainties include statistical and systematic uncertainties.

	$\tau_{\text{lep}}\tau_{\text{had}}$ VBF		$\tau_{\text{lep}}\tau_{\text{had}}$ boosted	
	Loose	Tight	Low-$p_{\text{T}}^{\tau\tau}$	High-$p_{\text{T}}^{\tau\tau}$
$Z \to \tau\tau$	175 ± 18	319 ± 22	4 159 ± 96	5 313 ± 92
$Z \to \ell\ell$	10.1 ± 3.0	12.6 ± 3.0	130 ± 37	115 ± 16
Top	5.8 ± 1.6	17.9 ± 4.7	119 ± 20	56 ± 10
Misidentified τ	103 ± 16	100 ± 15	1 907 ± 77	617 ± 27
Other backgrounds	4.0 ± 1.6	9.5 ± 1.9	115.2 ± 7.9	129.6 ± 8.9
ggF, $H \to \tau\tau$	4.1 ± 1.3	7.0 ± 2.0	62 ± 16	64 ± 21
VBF, $H \to \tau\tau$	7.5 ± 2.2	25.3 ± 7.1	12.0 ± 3.5	14.2 ± 4.1
WH, $H \to \tau\tau$	< 0.1	0.1 ± 0.0	4.0 ± 1.1	5.3 ± 1.4
ZH, $H \to \tau\tau$	< 0.1	< 0.1	1.8 ± 0.5	2.8 ± 0.8
Total background	299 ± 18	459 ± 23	6 430 ± 88	6 230 ± 92
Total signal	11.7 ± 3.3	32.5 ± 8.4	80 ± 20	86 ± 24
Data	318	496	6556	6347

Table 9: Observed event yields and predictions as computed by the fit in the $\tau_{\text{had}}\tau_{\text{had}}$ signal regions. Uncertainties include statistical and systematic uncertainties.

	$\tau_{\text{had}}\tau_{\text{had}}$ VBF			$\tau_{\text{had}}\tau_{\text{had}}$ boosted	
	Loose	Tight	High-$p_{\text{T}}^{\tau\tau}$	Low-$p_{\text{T}}^{\tau\tau}$	High-$p_{\text{T}}^{\tau\tau}$
$Z \to \tau\tau$	69.0 ± 9.5	103 ± 12	144 ± 12	3 260 ± 130	3 592 ± 85
Misidentified τ	45.1 ± 5.4	96.4 ± 9.2	19.8 ± 2.9	1 870 ± 140	366 ± 54
Other backgrounds	4.4 ± 1.4	11.5 ± 1.7	4.4 ± 0.7	281 ± 21	109.9 ± 9.2
ggF, $H \to \tau\tau$	1.1 ± 0.4	2.0 ± 0.7	3.5 ± 1.1	41 ± 11	49 ± 15
VBF, $H \to \tau\tau$	1.5 ± 0.5	6.4 ± 1.8	11.0 ± 3.0	8.9 ± 3.4	10.6 ± 2.9
WH, $H \to \tau\tau$	< 0.1	< 0.1	< 0.1	3.3 ± 1.0	4.4 ± 1.2
ZH, $H \to \tau\tau$	< 0.1	< 0.1	< 0.1	2.4 ± 0.7	2.8 ± 0.8
Total background	119 ± 10	210 ± 13	168 ± 13	5 411 ± 80	4 068 ± 66
Total signal	2.6 ± 0.8	8.4 ± 2.4	14.6 ± 3.8	56 ± 15	67 ± 18
Data	121	220	179	5455	4103

Table 10: Summary of different sources of uncertainty in decreasing order of their impact on $\sigma_{H\to\tau\tau}$. Their observed and expected fractional impacts, both computed by the fit, are given, relative to the $\sigma_{H\to\tau\tau}$ value. Experimental uncertainties in reconstructed objects combine efficiency and energy/momentum scale and resolution uncertainties. Background statistics includes the bin-by-bin statistical uncertainties of the simulated backgrounds as well as misidentified τ backgrounds which are estimated using data. Background normalization describes the combined impact of all background normalization uncertainties.

Source of uncertainty	Impact $\Delta\sigma/\sigma_{H\to\tau\tau}$ (%)	
	Observed	Expected
Theoretical uncert. on signal	+13.5 / −8.7	+11.9 / −7.7
Background statistics	+11 / −10	+10.2 / −9.8
Jets and $E_{\mathrm{T}}^{\mathrm{miss}}$	+11.5 / −9.3	+10.5 / −8.6
Background normalization	+6.8 / −4.8	+6.6 / −4.6
Misidentified τ	+4.5 / −4.2	+3.7 / −3.4
Theoretical uncert. on background	+4.6 / −3.6	+5.1 / −4.2
Hadronic taus	+4.7 / −3.0	+5.8 / −4.2
Flavour tagging	+3.3 / −2.4	+2.9 / −2.2
Luminosity	+3.3 / −2.3	+3.1 / −2.2
Electrons and muons	+1.2 / −1.0	+1.1 / −0.9
Total systematic uncert.	+24 / −20	+22 / −19
Data statistics	±16	±15
Total	+28 / −26	+27 / −25

the missing higher-order QCD corrections on the signal process cross sections, the jet energy resolution, the $\tau_{\text{had-vis}}$ identification and the normalization of the $Z \to \tau\tau$ background. Figure 5 also shows that in most cases the fitted parameters are in agreement with the nominal values, except the uncertainties related to jet energy resolution and scale to which the $m_{\tau\tau}^{\mathrm{MMC}}$ reconstructed from real di-tau events is sensitive: Selected di-tau events in VBF and boosted categories are characterized by one or more high-p_{T} jets which recoil against the two tau leptons. The main contributions to $E_{\mathrm{T}}^{\mathrm{miss}}$ are thus the neutrinos in the tau lepton decays and the impact of the jet energy resolution when projected onto the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction. Applying both jet energy resolution and scale uncertainties causes a shift in the mean jet p_{T}, which therefore translates directly into a shift of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$. This, in turn, translates into a shift of the reconstructed $m_{\tau\tau}^{\mathrm{MMC}}$ that is constrained by data in the $Z \to \tau\tau$ mass peak.

Results of the fit when only the data of an individual channel or of an individual category are used, are shown in Fig. 6. Also shown is the result from the combined fit and the uncertainty in $\sigma_{H\to\tau\tau}^{\mathrm{SM}}$. All results are consistent with the SM expectation. The simple combination of the individual fit results does not agree exactly with the combined fit result because the values of the nuisance parameters are different. The comparison of the number of event above background predictions with a SM Higgs boson at $m_H = 125\,\mathrm{GeV}$ for the measured $\sigma_{H\to\tau\tau}$, is visualized as a function of $m_{\tau\tau}^{\mathrm{MMC}}$ in Figs. 7 and 8 for all signal regions.

Figure 6 illustrates that both the VBF and boosted categories provide good sensitivity, respectively, to VBF and ggF Higgs boson production. A two-parameter fit is therefore performed to determine the cross sections of these production processes by exploiting the sensitivity offered by the use of the event categories in the analyses of the three channels. Two cross-section parameters $\sigma_{H\to\tau\tau}^{gg\mathrm{F}}$ and $\sigma_{H\to\tau\tau}^{\mathrm{VBF}}$ are

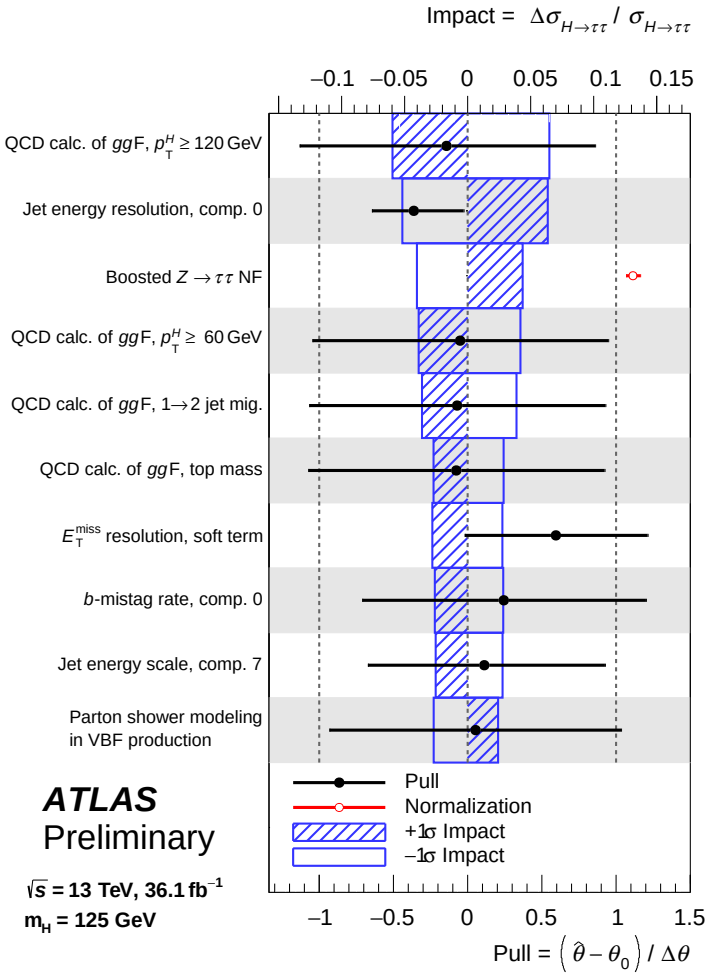
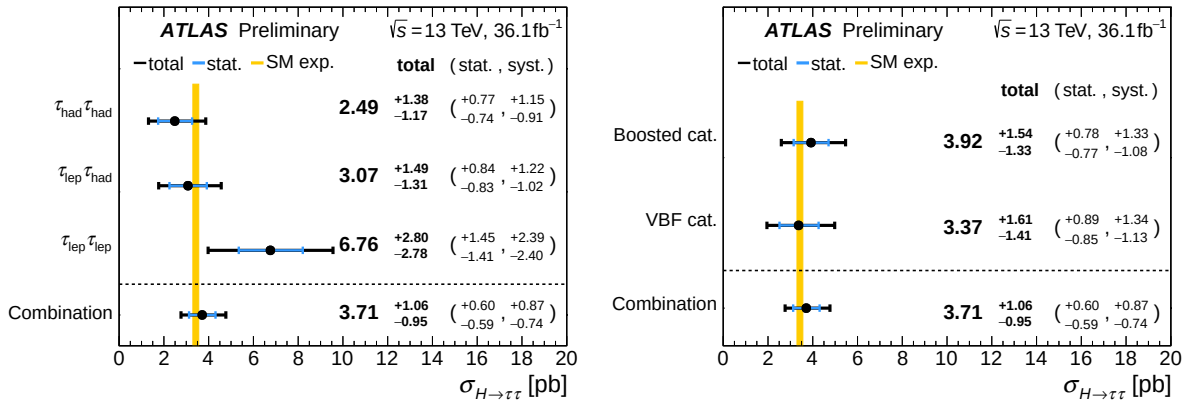


Figure 5: Fractional impact of systematic uncertainties in $\sigma_{H\to\tau\tau}$ as computed by the fit. The systematic uncertainties are listed in decreasing order of their impact on $\sigma_{H\to\tau\tau}$ on the y-axis. The hatched blue and open blue boxes show the variations of $\sigma_{H\to\tau\tau}$ referring to the top x-axis, as described in the text. The filled circles, referring to the bottom x-axis, show the pulls of the fitted nuisance parameters, i.e. the deviations of the fitted parameters $\hat{\theta}$ from their nominal values θ_0, normalized to their nominal uncertainties $\Delta\theta$. The black lines show the uncertainties of the nuisance parameters resulting from the fit. Several sources of uncertainties like the jet energy scale and resolution as well as the b-mistag rate are described by their principal components in the fit. For the boosted $Z \to \tau\tau$ normalization factor the fitted value and the uncertainty resulting from the fit are shown.

introduced and the data are fitted separating the fermion-mediated ggF process from the vector-boson-mediated VBF process while the contributions from other Higgs production processes are set to their predicted SM values. The two-dimensional 68% and 95% confidence level (CL) contours in the plane of $\sigma^{ggF}_{H\to\tau\tau}$ and $\sigma^{VBF}_{H\to\tau\tau}$ are shown in Fig. 9. The best-fit values are $\sigma^{ggF}_{H\to\tau\tau} = 3.0 \pm 1.0\,(\text{stat.})\,^{+1.6}_{-1.2}\,(\text{syst.})$ pb and $\sigma^{VBF}_{H\to\tau\tau} = 0.28 \pm 0.09\,(\text{stat.})\,^{+0.11}_{-0.09}\,(\text{syst.})$ pb, in agreement with the predictions from the Standard Model of $\sigma^{SM}_{ggF,\,H\to\tau\tau} = 3.05 \pm 0.13$ pb and $\sigma^{SM}_{VBF,\,H\to\tau\tau} = 0.237 \pm 0.006$ pb [93]. The two results are strongly anti-correlated (correlation coefficient of −52%), as can be seen in Fig. 9.



(a) (b)

Figure 6: The measured values for $\sigma_{H\to\tau\tau}$ when only the data of (a) individual channels or (b) individual categories are used. Also shown is the result from the combined fit. The total $\pm 1\sigma$ uncertainty in the measurement is indicated by the black error bars, with the individual contribution from the statistical uncertainty in blue. The theory uncertainty in the predicted signal cross section is shown by the yellow band.

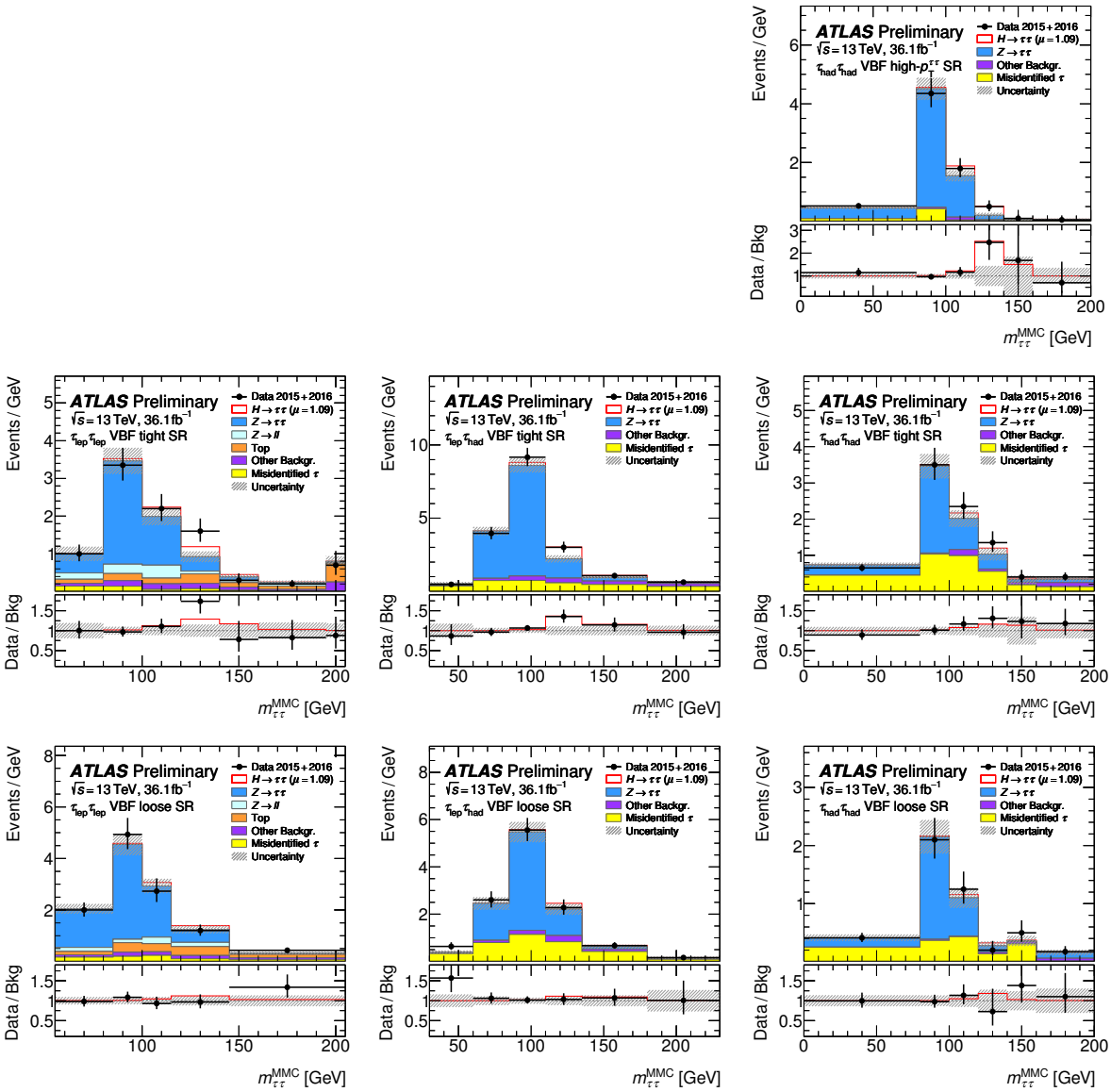


Figure 7: Observed and expected $m_{\tau\tau}^{\text{MMC}}$ distributions as used in the fit in all signal regions (SRs) in the VBF category for the $\tau_{\text{lep}}\tau_{\text{lep}}$ (left), $\tau_{\text{lep}}\tau_{\text{had}}$ (middle) and $\tau_{\text{had}}\tau_{\text{had}}$ (right) analysis channels. The Higgs boson signal ($m_H = 125\,\text{GeV}$) is shown with the solid red line. The signal and background predictions are determined in the likelihood fit. The size of the combined statistical, experimental and theoretical uncertainties in the background are indicated by the hatched bands. The ratios of the observed data to the expectations are shown in the lower panels.

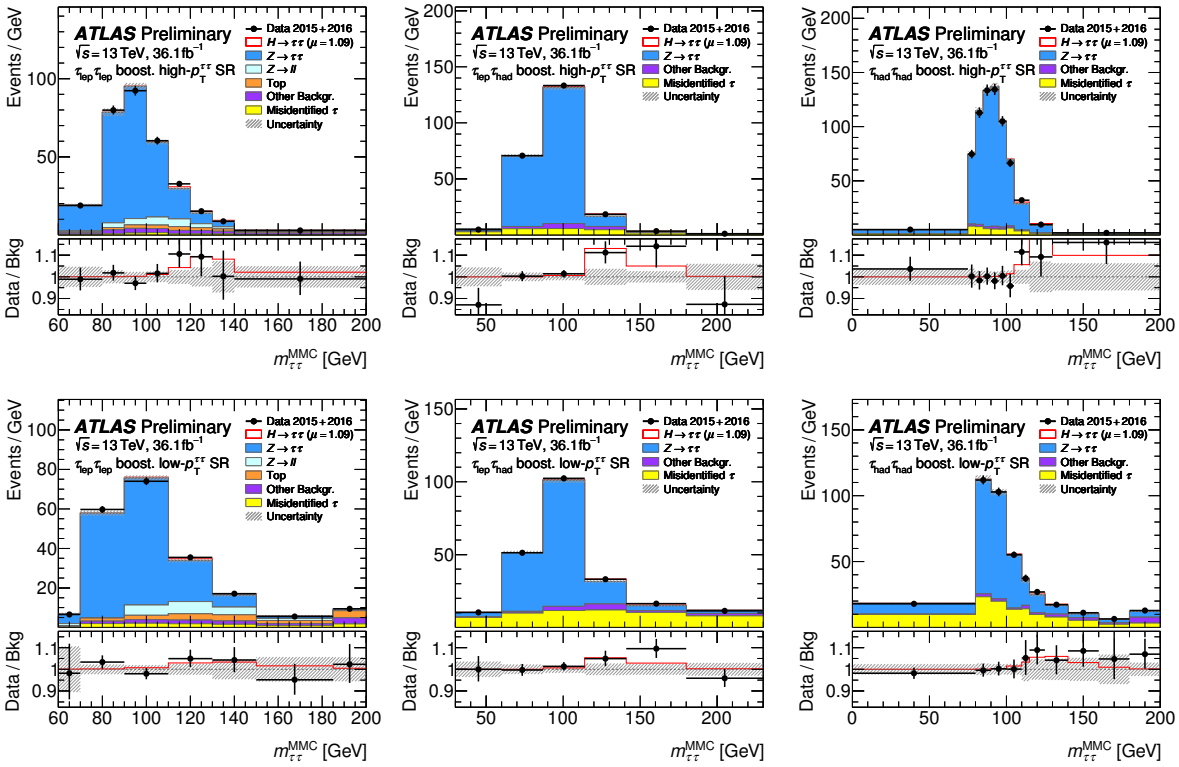


Figure 8: Observed and expected $m_{\tau\tau}^{\text{MMC}}$ distributions as used in the fit in all signal regions (SRs) in the boosted category for the $\tau_{\text{lep}}\tau_{\text{lep}}$ (left), $\tau_{\text{lep}}\tau_{\text{had}}$ (middle) and $\tau_{\text{had}}\tau_{\text{had}}$ (right) analysis channels. The Higgs boson signal ($m_H = 125\,\text{GeV}$) is shown with the solid red line. The signal and background predictions are determined in the likelihood fit. The size of the combined statistical, experimental and theoretical uncertainties in the background are indicated by the hatched bands. The ratios of the observed data to the expectations are shown in the lower panels.

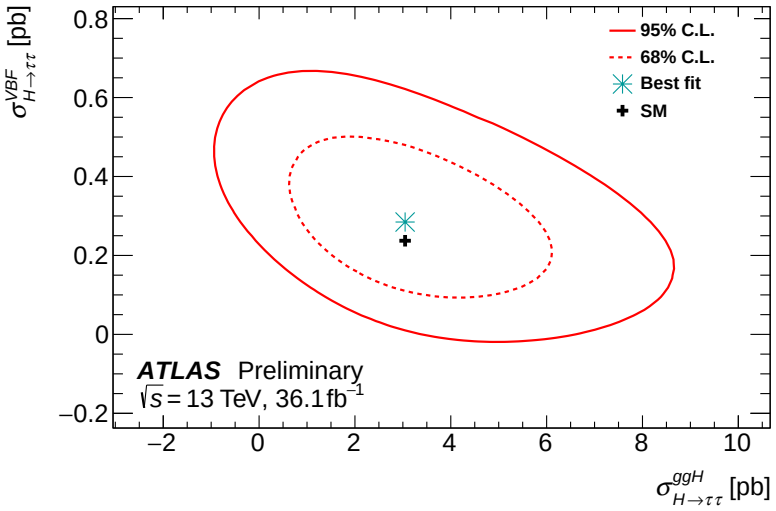


Figure 9: Likelihood contours for the combination of all channels in the $(\sigma_{H\to\tau\tau}^{ggF}, \sigma_{H\to\tau\tau}^{\text{VBF}})$ plane. The 68% and 95% CL contours are shown as dashed and solid lines, respectively, for $m_H = 125\,\text{GeV}$. The SM expectation is indicated by a plus symbol and the best fit to the data is shown as a star.

9 Conclusions

A measurement of total production cross sections of the Higgs boson in proton–proton collisions has been presented in the $H \to \tau\tau$ decay channel. The analysis has been performed using 36.1 fb^{-1} of data recorded by the ATLAS experiment at the LHC at a center-of-mass energy of $\sqrt{s} = 13$ TeV. All combinations of leptonic and hadronic tau decays were considered. An excess of events over the expected background from other Standard Model processes is found with an observed (expected) significance of 4.4 (4.1) standard deviations. Combined with results using data taken at $\sqrt{s}$ of 7 and 8 TeV, the observed (expected) significance amounts to 6.4 (5.4) standard deviations standard deviations and constitutes an observation of $H \to \tau\tau$ decays by the ATLAS experiment. Using the data taken at $\sqrt{s} = 13$ TeV, the $pp \to H \to \tau\tau$ total cross section has been measured to be $3.71 \pm 0.59\,(\text{stat.})\,^{+0.87}_{-0.74}\,(\text{syst.})$ pb, for a Higgs boson of mass 125 GeV. A two-dimensional fit has been performed to separate the vector-boson-mediated VBF process from the fermion-mediated ggF process. The cross sections of the Higgs boson decaying into two tau leptons have been measured to be $\sigma^{\text{VBF}}_{H\to\tau\tau} = 0.28 \pm 0.09\,(\text{stat.})\,^{+0.11}_{-0.09}\,(\text{syst.})$ pb and $\sigma^{gg\text{F}}_{H\to\tau\tau} = 3.0 \pm 1.0\,(\text{stat.})\,^{+1.6}_{-1.2}\,(\text{syst.})$ pb, respectively, for the two production processes. All measurements are consistent with SM predictions.

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] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

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

[5] A. Salam, *Weak and Electromagnetic Interactions*, Proceedings of the 8th Nobel symposium, Ed. N. Svartholm, Almqvist & Wiskell, 1968, Conf. Proc. **C680519** (1968) 367, URL: {http://inspirehep.net/record/53083}.

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

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

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

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

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

[11] ATLAS Collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s}$ = 7 and 8 TeV in the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 6, arXiv: 1507.04548 [hep-ex].

[12] CMS Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, Eur. Phys. J. C **75** (2015) 212, arXiv: 1412.8662 [hep-ex].

[13] ATLAS Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 476, arXiv: 1506.05669 [hep-ex].

[14] CMS Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, Phys. Rev. D **92** (2015) 012004, arXiv: 1411.3441 [hep-ex].

[15] ATLAS Collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, JHEP **04** (2015) 117, arXiv: 1501.04943 [hep-ex].

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

[17] ATLAS and CMS Collaborations,
Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV,
JHEP **08** (2016) 045, arXiv: 1606.02266 [hep-ex].

[18] CMS Collaboration, *Observation of the Higgs boson decay to a pair of τ leptons*,
Phys. Lett. **B779** (2018) 283, arXiv: 1708.00373 [hep-ex].

[19] ATLAS Collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. **D97** (2018) 072003,
arXiv: 1712.08891 [hep-ex].

[20] CMS Collaboration, *Observation of $t\bar{t}H$ production*, (2018), arXiv: 1804.02610 [hep-ex].

[21] ATLAS Collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*,
JHEP **12** (2017) 024, arXiv: 1708.03299 [hep-ex].

[22] CMS Collaboration, *Evidence for the Higgs boson decay to a bottom quark-antiquark pair*,
Phys. Lett. **B780** (2018) 501, arXiv: 1709.07497 [hep-ex].

[23] ATLAS Collaboration, *Search for the Dimuon Decay of the Higgs Boson in pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector*, Phys. Rev. Lett. **119** (2017) 051802,
arXiv: 1705.04582 [hep-ex].

[24] CMS Collaboration,
Search for neutral MSSM Higgs bosons decaying to $\mu^+\mu^-$ in pp collisions at $\sqrt{s}$ = 7 and 8 TeV,
Phys. Lett. B **752** (2016) 221, arXiv: 1508.01437 [hep-ex].

[25] ATLAS Collaboration, *Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector*,
Eur. Phys. J. C **76** (2016) 658, arXiv: 1602.04516 [hep-ex].

[26] S. Berge et al.,
Prospects of constraining the Higgs boson's CP nature in the tau decay channel at the LHC,
Phys. Rev. **D 92** (2015) 096012, arXiv: 1510.03850 [hep-ph].

[27] ATLAS Collaboration, *Modelling $Z \to \tau\tau$ processes in ATLAS with τ-embedded $Z \to \mu\mu$ data*,
JINST **10** (2015) P09018, arXiv: 1506.05623 [hep-ex].

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

[29] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ().

[30] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s}$ = 7 and 8 TeV with the ATLAS and CMS Experiments*,
Phys. Rev. Lett. **114** (2015) 191803, arXiv: 1503.07589 [hep-ex].

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

[32] S. Frixione, P. Nason and C. Oleari,
Matching NLO QCD computations with parton shower simulations: the POWHEG method,
JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[33] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph].

[34] E. Bagnaschi et al.,
Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM,
JHEP **02** (2012) 088, arXiv: 1111.2854 [hep-ph].

[35] K. Hamilton et al., *NNLOPS simulation of Higgs boson production*, JHEP **10** (2013) 222,
arXiv: 1309.0017 [hep-ph].

[36] K. Hamilton, P. Nason and G. Zanderighi,
Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator,
JHEP **05** (2015) 140, arXiv: 1501.04637 [hep-ph].

[37] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A brief introduction to PYTHIA 8.1*,
Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

[38] M. Bahr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639,
arXiv: 0803.0883 [hep-ph].

[39] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196,
arXiv: 1512.01178 [hep-ph].

[40] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. **G43** (2016) 023001,
arXiv: 1510.03865 [hep-ph].

[41] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014) 145,
arXiv: 1406.3660 [hep-ex].

[42] J. Pumplin et al.,
New generation of parton distributions with uncertainties from global QCD analysis,
JHEP **07** (2002) 012, arXiv: hep-ph/0201195.

[43] N. Davidson, T. Przedzinski and Z. Was,
PHOTOS Interface in C++: Technical and Physics Documentation, (2010),
arXiv: 1011.0937 [hep-ph].

[44] C. Anastasiou et al., *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*,
Phys. Rev. Lett. **114** (2015) 212001, arXiv: 1503.06056 [hep-ph].

[45] C. Anastasiou et al.,
High precision determination of the gluon fusion Higgs boson cross-section at the LHC,
JHEP **05** (2016) 058, arXiv: 1602.00695 [hep-ph].

[46] S. Actis et al., *NLO electroweak corrections to Higgs boson production at hadron colliders*,
Phys. Lett. **B670** (2008) 12, arXiv: 0809.1301 [hep-ph].

[47] C. Anastasiou, R. Boughezal and F. Petriello,
Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion,
JHEP **04** (2009) 003, arXiv: 0811.3458 [hep-ph].

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

[49] M. Ciccolini, A. Denner and S. Dittmaier,
Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC,
Phys. Rev. **D77** (2008) 013002, arXiv: 0710.4749 [hep-ph].

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

[51] O. Brein, A. Djouadi and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. **B579** (2004) 149, arXiv: hep-ph/0307206 [hep-ph].

[52] L. Altenkamp et al., *Gluon-induced Higgs-strahlung at next-to-leading order QCD*, JHEP **02** (2013) 078, arXiv: 1211.5015 [hep-ph].

[53] A. Denner et al., *Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK*, JHEP **03** (2012) 075, arXiv: 1112.5142 [hep-ph].

[54] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph].

[55] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph].

[56] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph].

[57] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph].

[58] S. Höche et al., *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph].

[59] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through $O(\alpha_S^2)$*, Phys. Rev. **D74** (2006) 114017, arXiv: hep-ph/0609070 [hep-ph].

[60] C. Anastasiou et al., *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. **D69** (2004) 094008, arXiv: hep-ph/0312266 [hep-ph].

[61] R. Ball et al., *Parton distributions for the LHC Run II*, JHEP **1504** (2014) 040, arXiv: 1410.8849 [hep-ph].

[62] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **1407** (2014) 079, arXiv: 1405.0301 [hep-ph].

[63] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

[64] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244, arXiv: 1207.1303 [hep-ph].

[65] S. Alioli, S.-O. Moch and P. Uwer, *Hadronic top-quark pair-production with one jet and parton showering*, JHEP **01** (2012) 137, arXiv: 1110.5251 [hep-ph].

[66] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph].

[67] S. Alioli et al.,
NLO single-top production matched with shower in POWHEG: s- and t-channel contributions,
JHEP **09** (2009) 111, [Erratum: JHEP02,011(2010)], arXiv: 0907.4076 [hep-ph].

[68] E. Re,
Single-top Wt-channel production matched with parton showers using the POWHEG method,
Eur. Phys. J. **C71** (2011) 1547, arXiv: 1009.2450 [hep-ph].

[69] P. Artoisenet et al.,
Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations,
JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph].

[70] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026,
arXiv: hep-ph/0603175 [hep-ph].

[71] P. Z. Skands, *Tuning Monte Carlo generators: The Perugia tunes*, Phys. Rev. **D82** (2010) 074018,
arXiv: 1005.3457 [hep-ph].

[72] D. J. Lange, *The EvtGen particle decay simulation package*,
Nucl. Instrum. Meth. A **462** (2001) 152.

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

[74] GEANT4 Collaboration, S. Agostinelli et al., *Geant4 – a simulation toolkit*,
Nucl. Instrum. Meth. **A 506** (2003) 250.

[75] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189,
arXiv: 0901.0002 [hep-ph].

[76] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012,
URL: https://cds.cern.ch/record/1474107.

[77] ATLAS Collaboration,
Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data,
Eur. Phys. J. C **74** (2014) 3071, arXiv: 1407.5063 [hep-ex].

[78] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016,
URL: https://cds.cern.ch/record/2157687.

[79] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*,
Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[80] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **76** (2016) 292,
arXiv: 1603.05598 [hep-ex].

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

[82] ATLAS Collaboration,
Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1,
Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex].

[83] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581,
arXiv: 1510.03823 [hep-ex].

[84] ATLAS Collaboration,
Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector,
Eur. Phys. J. C **77** (2017) 580, arXiv: 1705.02211 [hep-ex].

[85] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*,
JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex].

[86] ATLAS Collaboration,
Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run,
ATL-PHYS-PUB-2016-012, 2016, URL: https://cds.cern.ch/record/2160731.

[87] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s}$ = 13 *TeV*,
ATLAS-CONF-2017-029, 2017, URL: https://cds.cern.ch/record/2261772.

[88] ATLAS Collaboration, *The ATLAS Tau Trigger in Run 2*, ATLAS-CONF-2017-061, 2017,
URL: https://cds.cern.ch/record/2274201.

[89] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at* $\sqrt{s}$ = *13 TeV*, (2018),
arXiv: 1802.08168 [hep-ex].

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

[91] R. Ellis et al., *Higgs decay to* $\tau^+\tau^-$*:A possible signature of intermediate mass Higgs bosons at high energy hadron colliders*, Nucl. Phys. **297** (1998) 221.

[92] ATLAS Collaboration,
Observation and measurement of Higgs boson decays to WW^* *with the ATLAS detector*,
Phys. Rev. D **92** (2015) 012006, arXiv: 1412.2641 [hep-ex].

[93] LHC Higgs Cross Section Working Group et al.,
Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables,
CERN-2011-002 (CERN, Geneva, 2011), arXiv: 1101.0593 [hep-ph].

[94] LHC Higgs Cross Section Working Group et al.,
Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector,
CERN-2017-002-M (CERN, Geneva, 2016), arXiv: 1610.07922 [hep-ph].

[95] I. W. Stewart et al., *Jet* p_T *resummation in Higgs production at* $NNLL' + NNLO$,
Phys. Rev. **D89** (2014) 054001, arXiv: 1307.1808 [hep-ph].

[96] X. Liu and F. Petriello,
Reducing theoretical uncertainties for exclusive Higgs-boson plus one-jet production at the LHC,
Phys. Rev. **D87** (2013) 094027, arXiv: 1303.4405 [hep-ph].

[97] R. Boughezal et al., *Combining Resummed Higgs Predictions Across Jet Bins*,
Phys. Rev. **D89** (2014) 074044, arXiv: 1312.4535 [hep-ph].

[98] I. W. Stewart and F. J. Tackmann,
Theory Uncertainties for Higgs and Other Searches Using Jet Bins,
Phys. Rev. **D85** (2012) 034011, arXiv: 1107.2117 [hep-ph].

[99] S. Gangal and F. J. Tackmann,
Next-to-leading-order uncertainties in Higgs+2 jets from gluon fusion,
Phys. Rev. **D87** (2013) 093008, arXiv: 1302.5437 [hep-ph].

[100] L. Lönnblad and S. Prestel, *Matching tree-level matrix elements with interleaved showers*, JHEP **03** (2012) 019, arXiv: 1109.4829 [hep-ph].

[101] ATLAS Collaboration, *Measurement of the cross-section for electroweak production of dijets in association with a Z boson in pp collisions at* $\sqrt{s}$ = 13 *TeV with the ATLAS detector*, Phys. Lett. B **775** (2017) 206, arXiv: 1709.10264 [hep-ex].

[102] ATLAS Collaboration, *Electron identification measurements in ATLAS using* $\sqrt{s}$ = 13 *TeV data with* 50 *ns bunch spacing*, ATL-PHYS-PUB-2015-041, 2015, URL: https://cds.cern.ch/record/2048202.

[103] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s}$ = 8 *TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653, arXiv: 1608.03953 [hep-ex].

[104] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, ATL-PHYS-PUB-2015-045, 2015, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-045.

[105] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: https://cds.cern.ch/record/1700870.

[106] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at* $\sqrt{s}$ = 8 *TeV*, Eur. Phys. J. C **75** (2015) 303, arXiv: 1412.7086 [hep-ex].

[107] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at* $\sqrt{s}$ = 13 *TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex].

[108] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s}$ = 13 *TeV*, ATL-PHYS-PUB-2015-015, 2015, URL: https://cds.cern.ch/record/2037613.

[109] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector in the first proton–proton collisions at* $\sqrt{s}$ = 13 *TeV*, ATL-PHYS-PUB-2015-027, 2015, URL: https://cds.cern.ch/record/2037904.

[110] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, [Erratum: Eur. Phys. J.C73,2501(2013)], arXiv: 1007.1727 [physics.data-an].