Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a $b\bar{b}$ pair in $pp$ collisions at 13 TeV using the ATLAS detector

The ATLAS Collaboration

Abstract

A search for the decay of a Standard Model Higgs boson into a $b\bar{b}$ pair when produced in association with a $W$ or $Z$ boson has been performed with the ATLAS detector. Data were collected in proton-proton collisions from Run 2 of the Large Hadron Collider at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 13.2 fb$^{-1}$. Final states considered contain 0, 1 and 2 charged leptons (electrons or muons), targeting the decays: $Z \rightarrow \nu\nu$, $W \rightarrow \ell\nu$, and $Z \rightarrow \ell\ell$. For $m_H = 125$ GeV the ratio of the observed signal strength to the SM expectation is found to be $\mu = 0.21^{+0.36}_{-0.35}$ (stat.) $\pm 0.36$ (syst.). This corresponds to an observed significance of 0.42 standard deviations compared with an expected sensitivity of 1.94. The analysis procedure has been validated by measuring the yield of $(W/Z)Z$ with $Z \rightarrow b\bar{b}$, where the ratio of the observed yield to that expected in the Standard Model was found to be $0.91 \pm 0.17$ (stat.)$^{+0.32}_{-0.27}$ (syst.), corresponding to a significance of 3.0 standard deviations compared to an expected significance of 3.2.
1. Introduction

For over four decades the prediction of the existence of a neutral spin-0 Higgs boson [1–3] was unconfirmed by experiment. The discovery of a new boson with a mass of about 125 GeV and consistent with the predictions of the Standard Model (SM) [4–7] was announced by the ATLAS [8] and CMS [9] experiments in July 2012. Further measurements [10–12] have since been conducted strengthening the hypothesis that the new particle is a Higgs boson. Multiple decay paths of the Higgs boson are susceptible to measurement at the Large Hadron Collider (LHC) [13] and it is essential to measure as many of these as precisely as possible to verify the SM hypothesis or establish evidence for new phenomena beyond the SM.

The process with the largest predicted branching fraction (58%) [14] for a SM Higgs boson of mass 125 GeV is \( H \rightarrow b \bar{b} \). At the LHC the overwhelming backgrounds arising from multi-jet production make a fully inclusive search extremely challenging. The production modes where the Higgs boson is produced together with a \( W \) or \( Z \) boson [15] (where \( V \) is used to jointly denote \( W \) or \( Z \)) provide a promising alternative despite having a cross section more than an order of magnitude lower than the dominant gluon-gluon fusion production mode. The leptonic decays of the \( W \) and \( Z \) boson lead to relatively clean signatures that can be used to significantly suppress the contributions from background processes and allow for an efficient triggering strategy.

Searches for a Higgs boson in the \( b \bar{b} \) decay mode have been conducted at the Tevatron by the CDF and DZero collaborations, reporting an excess of events with a significance of 2.8 standard deviations for a Higgs boson with mass 125 GeV [16]. ATLAS and CMS have reported results from Run 1 each using approximately 25 fb\(^{-1}\) of integrated luminosity of proton-proton (\( pp \)) collisions at centre-of-mass energies of \( \sqrt{s} = 7 \) and 8 TeV [17, 18]. Excesses of events consistent with a Higgs boson with mass 125 GeV were observed with significances of 1.4 and 2.1 standard deviations respectively. The LHC combination of the Run 1 ATLAS and CMS analyses resulted in observed and expected significances of 2.6 and 3.7 standard deviations respectively for the \( H \rightarrow b \bar{b} \) decay channel [12].

This note reports on the continuing search for a SM Higgs boson produced in association with a \( W \) or \( Z \) boson and decaying to a \( b \bar{b} \) pair using the ATLAS detector in Run 2 of the LHC using \( pp \) collisions collected at an increased centre-of-mass energy of \( \sqrt{s} = 13 \) TeV, representing an integrated luminosity of 13.2 fb\(^{-1}\). Three main categories of events are studied, containing exactly 0, 1 or 2 charged leptons (electrons and muons) in the final state, targeting the \( Z \rightarrow \nu \nu \), \( W \rightarrow \ell \nu \), and \( Z \rightarrow \ell \ell \) decay modes of the vector bosons respectively. Jets consistent with originating from the decay \( H \rightarrow b \bar{b} \) are identified using a \( b \)-tagging algorithm. Events are then further categorised according to the number of jets and event kinematics to improve the sensitivity.

A binned maximum likelihood fit, referred to as the “global likelihood fit”, is used to extract the signal yield. Systematic uncertainties on the signal and background modelling are implemented as deviations from the nominal model scaled by nuisance parameters that are profiled in the fit. The final discriminating variables are a collection of multivariate discriminants, one for each analysis category, based on boosted decision trees (BDTs) incorporating a selection of kinematic variables including the \( b \bar{b} \) invariant mass. The methodology is validated by measuring the \((W/Z)Z\) with \( Z \rightarrow b \bar{b} \) diboson signal.
2. ATLAS detector

ATLAS [19] is a general-purpose detector designed for a broad programme of particle physics measurements. An inner tracking detector, located within a 2 T axial magnetic field generated by a superconducting solenoid, is used to measure the trajectories and momenta of charged particles. The innermost layers consisting of high-granularity silicon pixel detectors instrument a pseudo-rapidity\(^1\) range \(|\eta| < 2.5\). A new innermost silicon pixel layer, the Insertable B-Layer [20] (IBL), was added to the detector between Run 1 and Run 2. The IBL improves the experiment’s ability to identify displaced vertices and thereby significantly improves the \(b\)-tagging performance [21]. Beyond the pixel detectors also with coverage within \(|\eta| < 2.5\) are silicon strip detectors. Outside of this within a range \(|\eta| < 2.0\) additional straw tube tracking detectors are placed. These provide measurements of transition radiation that contribute to the identification of electrons. The calorimeter system consists of a number of different technologies with coverage out to \(|\eta| < 4.9\). The liquid-argon electromagnetic calorimeter is divided into three regions: barrel (\(|\eta| < 1.475\)), endcap (1.375 < \(|\eta| < 3.2\)), and forward (3.1 < \(|\eta| < 4.9\)). The hadronic calorimetry surrounds the electromagnetic calorimeters and extends out to a range of \(|\eta| < 4.9\), using scintillator tiles or liquid argon as the active materials. The outermost layers of the detector are taken up by the muon spectrometer that measures the trajectories of muons bending in the field of three large air-core toroidal magnets. High precision tracking can be achieved out to \(|\eta| < 2.7\) and there are additional chambers for fast triggering within the range \(|\eta| < 2.4\). A two-level trigger system [22] is used to reduce the recorded data rate to a level of around 1 kHz. The first level is a hardware implementation that makes use of only a subset of the total available information to make fast decisions to accept or reject an event, aiming to reduce the rate to around 100 kHz, and the second level is the software based High-Level Trigger that provides the remaining rate reduction.

3. Data and simulated samples

The data used in this analysis were collected at a centre of mass energy of 13 TeV during the 2015 and early 2016 running periods, and correspond to integrated luminosities of 3.2 ± 0.1 and 10 ± 0.4 fb\(^{-1}\) respectively [23]. In the combined dataset the events recorded typically have between 10 and 30 additional inelastic \(pp\) collisions (pile-up) in each event. The data were collected using missing transverse energy \(E_T^{\text{miss}}\) triggers for the 0- and 1-lepton channels and single lepton triggers for the 1- and 2-lepton channels. Events are only selected if they are of good quality and if the detector was known to be operating well.

Monte Carlo (MC) simulated events are used to model the SM background and \(VH \rightarrow Vb\bar{b}\) signal processes. All simulated processes are normalised to the best available theoretical predictions for their cross sections. Difficulties in modelling the QCD multi-jet contribution using simulated samples lead to it being estimated using data driven methods for the 1-lepton channel. This background is negligible in the other channels. All simulated samples make use of a GEANT 4 [24] based detector simulation [25] and are reconstructed with the standard ATLAS reconstruction software. The effects of pile-up from

---

\(^{1}\) ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the \(z\)-axis coinciding with the axis of the beam pipe. The \(x\)-axis points from the IP towards the centre of the LHC ring, and the \(y\)-axis points upward. Cylindrical coordinates \((r,\phi)\) are used in the transverse plane, \(\phi\) being the azimuthal angle around the \(z\)-axis. The pseudo-rapidity is defined in terms of the polar angle \(\theta\) as \(\eta = -\ln \tan(\theta/2)\). The distance in \((\eta,\phi)\) coordinates, 

\[
\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2},
\]

is also used to define cone sizes. Transverse momentum and energy are defined as \(p_T = p \sin \theta\) and \(E_T = E \sin \theta\), respectively.
multiple interactions in the same and neighbouring bunch crossings are modelled by overlaying minimum bias events, simulated using the soft QCD processes of Pythia 8.186 [26] with the A2 [27] tune and MSTW2008LO [28] parton density functions (PDF). For all simulated samples, except for those generated using Sherpa, the EvtGen v1.2.0 program [29] is used to describe the properties of the bottom and charm hadron decays.

For the generation of $t\bar{t}$ and $Wt$ and $s$-channel single top-quark production, the Powheg-Box v2 [30] generator with the CT10 [31] PDF set is used in the matrix-element calculations. Electroweak $t$-channel single top-quark events are generated using the Powheg-Box v1 generator. This generator uses the 4-flavour scheme for the next-to-leading order (NLO) matrix-element calculations together with the fixed 4-flavour PDF set CT10f4. The parton shower, fragmentation, and the underlying event are simulated using Pythia 6.428 [32] with the CTEQ6L1 [33] PDF set and the corresponding Perugia 2012 tune (P2012) [34]. The top-quark mass is set to 172.5 GeV. Events are filtered such that at least one $W$ boson in the event decays leptonically. The $t\bar{t}$ sample is normalised to its next-to-NLO (NNLO) cross-section including the resummation of soft gluon emission at next-to-leading-logarithmic accuracy using Top++2.0 [35]. The $s$ [36], $t$ [37] and $Wt$ [38] channel samples are normalised to NLO cross-sections. Samples produced with alternative generators are used to assign uncertainties on the $t\bar{t}$ and single top-quark backgrounds. Alternative samples are generated with varying schemes: different parton showering generator (Powheg+Herwig++) [39], alternative matrix-element generator (Madgraph5 _aMC@NLO+Herwig++ [40]) and varying the choice of generator parameters to increase or decrease the amount of radiation.

Events containing $W$ or $Z$ bosons with jets ($V$ + Jets) are simulated using the Sherpa 2.2.0 [41] generator. Matrix elements are calculated for up to two partons at NLO and four partons at leading order (LO) using the Comix [42] and OpenLoops [43] matrix-element generators and merged with the Sherpa parton shower [44] using the ME+PS@NLO prescription [45]. The CT10 PDF set is used in conjunction with a dedicated parton shower tuning developed by the Sherpa authors. The $V$ + Jets samples are produced with a simplified scale setting prescription in the multi-parton matrix elements, to improve the event generation speed. A theory-based re-weighting of the jet multiplicity distribution is applied at event level, which is derived from an event generation with the strict scale prescription. The $V$ + Jets events are normalized to the NNLO QCD theoretical cross-sections [46].

Alternative versions of the Sherpa samples are generated to evaluate modelling uncertainties, where the factorisation, renormalisation, CKKW merging and resummation scales have been varied. These samples are generated using Sherpa 2.1.1 but it has been verified that the relative variations due to the scale variations are consistent with those found for Sherpa 2.2.0. Alternative samples are also generated using Madgraph5 _aMC@NLO v2.2.2 at LO interfaced to the Pythia 8.186 parton shower model. The A14 [47] tune is used together with the NNPDF23LO PDF set [48].

Diboson processes ($WW$, $WZ$ and $ZZ$) are generated using Sherpa 2.1.1 which calculates up to 1 ($ZZ$) or 0 ($WW$, $WZ$) additional partons at NLO and up to 3 additional partons at LO. Alternative Sherpa samples are generated with the factorisation, normalisation and resummation scales varied. Additional diboson samples are produced using the Powheg-Box v2 generator, interfaced to either the Pythia 8.2 [49] or the Herwig++ parton shower generators. The CT10NLO PDF set is used for the hard process, while the CTEQ6L1 PDF set is used for the parton shower.

Simulated events for $qq \to VH$ production are generated at leading order for the hard-scatter and leading-log for the parton shower using Pythia 8.186, using the A14 tune and NNPDF23LO PDF set. The samples for $gg \to ZH$ (gluon-induced) production are simulated using the Powheg-Box v2 matrix-element generator interfaced with Pythia 8.2, using the A inicial tune with CT10NLO PDF sets.
The mass of the Higgs boson is fixed to 125 GeV and the $b\bar{b}$ branching fraction is fixed at 58%. The inclusive cross sections [51–57] are calculated at NNLO (QCD) and NLO (EW). For the gluon-induced $ZH$ production the cross section is calculated at NLO+NLL (QCD) [58–61]. This is then subtracted from the inclusive cross section to estimate the quark-induced cross section. An additional NLO (EW) scale factor is applied as a function of the vector boson transverse momentum to take into account effects not already modelled in the simulation. This makes use of the $VH$ differential cross section computed with Hawk [62, 63]. The simulated samples include all final states where the Higgs boson decays to $b\bar{b}$ and the vector boson to a leptonic final state, including those with a $\tau$-lepton.

4. Object and event selection

Broadly, events are selected where two high transverse momentum $b$-tagged jets are reconstructed together with 0, 1 or 2 charged leptons, either electrons or muons. The channels are then categorised using the number of jets in the events. A set of kinematic observables is used as input to a multivariate algorithm for each analysis category to provide a discriminant that is used in the signal extraction fits.

4.1. Object reconstruction

Charged particle tracks are reconstructed by fitting track hypotheses to the collections of clustered hits left by the passage of charged particles through the inner detector. The algorithm is seeded from clusters in the pixel layers of the inner detector and has a sophisticated approach to resolving ambiguities in hit assignments that arise in the high multiplicity environment.

Interaction vertices from each bunch crossing are reconstructed using tracks in the inner detector. Candidate vertices are required to be consistent with the known envelope of the beamspot and have at least two associated tracks with transverse momenta $p_T > 0.4$ GeV. The reconstructed vertex with the largest sum of the squared transverse momenta of the associated tracks is selected as the primary vertex.

Electrons are reconstructed [64, 65] as noise-suppressed clusters of energy in the calorimeter using a sliding window algorithm with a matching charged track from the inner detector. Energy calibration is driven by data from reference processes such as $Z \rightarrow e^+e^-$. Identification requirements are used to reduce the number of jets and non-prompt electrons incorrectly identified as prompt electrons. These include a likelihood discriminant based method, using both shower shape information and track properties, a requirement that the charged track be isolated within the inner detector, and additional requirements on the matched track properties to reject tracks arising from pile-up events. Three basic electron categories are defined in this analysis: loose, medium and tight. Loose electrons pass the basic requirements listed above and have $p_T > 7$ GeV, $|\eta| < 2.47$. Medium electrons pass the loose requirements and additionally have $p_T > 25$ GeV. Tight electrons are required to pass tighter requirements on the likelihood and isolation to further suppress backgrounds from multi-jet processes.

Muons are reconstructed [66] from matching tracks in the inner detector and the muon spectrometer. Full coverage of both detectors extends to $|\eta| < 2.5$, with additional acceptance using the muon spectrometer alone out to $|\eta| < 2.7$. As with electrons three basic categories of selection are defined: loose, medium and tight. Loose muon candidates are required to have $p_T > 7$ GeV with relatively loose requirements applied on muon track parameters to reject pile-up and cosmic muons. Loose quality requirements are applied including isolation requirements in the tracking detectors and calorimeters to reject non-prompt
muons. Medium muons pass the loose requirements and have $p_T > 25$ GeV and $|\eta| < 2.5$. Tight muons pass the medium selection and must pass tighter identification and isolation requirements to further reduce the multi-jet background.

Hadronically decaying $\tau$-leptons ($n_{\text{had}}$) are reconstructed [67, 68] using the anti-$k_t$ algorithm [69] with a radius parameter of 0.4. They are required to have exactly 1 or 3 matching charged tracks within a cone around the jet axis of $\Delta R = 0.2$. To reject fake $\tau$-leptons from jets, a multivariate approach using boosted decision trees is employed, making use of information from the calorimeters and the tracking detectors.

Jets are reconstructed from noise suppressed energy clusters in the calorimeter [70] with the anti-$k_t$ algorithm with a radius parameter of 0.4. The energies of reconstructed jets are calibrated using a jet energy scale correction (JES) derived from both simulation and in-situ calibration using data [71]. Jet cleaning criteria are applied to identify those jet candidates likely to have arisen from non-collision sources and noise and any event containing such a jet is removed [72]. Additional selection is applied to jets with transverse momenta less than 60 GeV and $|\eta| < 2.4$ to remove jets likely to originate from additional collisions within the same bunch crossing based on a likelihood discriminant, using information about the primary vertex and charged tracks associated with the jet [73]. Two jet categories, forward and signal, are defined: forward jets have $p_T > 30$ GeV and $2.5 \leq |\eta| < 4.5$ and signal jets have $p_T > 20$ GeV and $|\eta| < 2.5$.

An updated multivariate $b$-tagging algorithm [74], designed to fully exploit the improved tracking capabilities from the introduction of the IBL [21], has been developed for Run 2. The improved algorithm makes use of information about the jet kinematics, the properties of tracks within jets, and the presence of displaced secondary vertices to provide light-flavour and $c$-jet rejection factors of 380 and 12 respectively for a 70% $b$-jet efficiency, as measured using a simulated $t\bar{t}$ sample.

Jets in the simulated samples are labelled according to which hadrons with $p_T > 5$ GeV are found within a cone around the jet axis of size $\Delta R = 0.3$. If a $b$-hadron is found the jet is labelled as a $b$-jet. If no $b$-hadron is found but a $c$-hadron is present then the jet is labelled as a $c$-jet. Otherwise the jet is labelled as a light jet (i.e. originating from a $u$, $d$, or $s$ quark or from a gluon). Simulated $V +$ Jets samples are labelled according to the flavour label of the two jets chosen for the reconstructed Higgs boson candidate: $V + bb, V + bc, V + cc, V + bl, V + cl, V + l$ (short for $V + ll$). An inclusive category $V +$ HF is defined as containing the first four: $V + bb, V + bc, V + cc, V + bl$.

For the $V + cl, V + l$, and $WW$ samples a truth tagging method is used to simulate the $b$-tagging algorithm, in order to reduce the statistical uncertainties on the simulation. This method uses measurements of the $b$-tagging efficiencies and the truth labelling of jets described above to calculate appropriate event weights. This is in contrast with direct tagging, where the $b$-tagging algorithm is applied to the simulated events and events failing a selection based on the $b$-tagging discriminant are removed. Direct tagging is used for all the other samples.

Two additional energy corrections are applied to $b$-tagged jets. The muon-in-jet correction is used where one or more muons are matched to a jet within $\Delta R < 0.4$. When more than one muon is found, the closest to the jet axis is used. The muon 4-vector is added to that of the jet and the energy deposited in the calorimeter by the muon is removed. For the 0- and 1-lepton channels the second correction, $Pt\text{Reco}$, is a scaling of the jet 4-vector as a function of the jet $p_T$ after the muon-in-jet correction, derived in simulation by comparing the calibrated jet energy to the energy of matching truth jets. Truth jets are constructed in simulated samples by running the same jet algorithms used for standard jet reconstruction, using as input

$\text{The rejection factor is defined as the inverse of the efficiency.}$
all the stable hadrons associated with the hadronization of a generated parton plus any muons or neutrinos produced within the jet from semi-leptonic decays. The correction is derived from $ZH \rightarrow \ell\ell b\bar{b}$ signal events and corrects for biases in the response arising from effects due to resolution, out-of-cone energy and unreconstructed neutrinos and folds in the expected signal kinematics. This results in a correction of around 12% or 24% at low $p_T$ and plateauing at high $p_T$ at around 1% or 6% for hadronic or semi-leptonic $b$-jets respectively. In the 2-lepton channel, where the full $ZH \rightarrow \ell\ell b\bar{b}$ event kinematics can be reconstructed, an improvement in the Higgs boson mass reconstruction is achieved through the use of a likelihood based kinematic fit, instead of the PtReco correction.

In the 0- and 1-lepton channels the presence of undetected neutrinos leads to a momentum imbalance in the event. This is measured as the missing transverse momentum $E_{\text{miss}}^T$, defined as the negative vector sum of the transverse momentum of physics objects in the event (electrons, muons, and jets). Additionally a soft term [75] is added based on well-reconstructed tracks originating from the primary vertex that are not already included in any of the physics objects. An overlap removal algorithm is applied to prevent double counting of reconstructed objects. If a reconstructed muon and electron share the same matched ID track then the electron is removed. The closest jet to an electron within a cone of radius $\Delta R = 0.2$ around the electron is removed. Any electrons reconstructed within $\Delta R < 0.4$ around the jet axis of a surviving jet are then removed. If a jet is reconstructed within $\Delta R < 0.2$ around a muon and the jet has fewer than three associated tracks then the jet is removed. Where a muon is reconstructed within a cone of radius $\Delta R = 0.4$ around the jet axis of any surviving jets then the muon is removed. Jets that are reconstructed within a cone of radius $\Delta R = 0.2$ around the axis of a $\tau_{\text{had}}$ candidate are removed.

Corrections are applied to the simulation to account for small differences with the collision data for the trigger, reconstruction and identification efficiencies of the reconstructed jets, muons and electrons as well as differences in the modelling of their energy and momenta. Dedicated simulation-to-data correction factors are also measured using data-based analyses, to correct the simulated $b$-tagging performance for $b$, $c$ and light-flavour jets separately. A dedicated simulation-to-data correction factor is also derived for the $E_{\text{miss}}^T$ trigger. This is derived using single-muon triggered events, which provide an unbiased data sample with respect to the $E_{\text{miss}}^T$ trigger as muons are not included in the $E_{\text{miss}}^T$ calculation at trigger level. Three control regions predominantly composed of $W + \text{Jets}$, $Z + \text{Jets}$ and $t\bar{t}$ events are used to measure the efficiency of the $E_{\text{miss}}^T$ trigger as a function of the offline $E_{\text{miss}}^T$. This is then compared to the efficiencies measured in simulated $W + \text{Jets}$, $Z + \text{Jets}$ and $t\bar{t}$ events to provide the correction factor.

### 4.2. Event categorisation and selection

Events passing detector quality requirements and the appropriate $E_{\text{T}}^{\text{miss}}$ or single lepton triggers are subdivided into multiple categories according to lepton multiplicity, vector boson transverse momentum, and jet multiplicity. Events in the 0-, 1-, and 2-lepton channels contain exactly 0, 1 or 2 electron or muon candidates meeting the loose selection criteria and at least two signal jets, of which exactly two must be $b$-tagged. The Higgs boson candidate is reconstructed from the two $b$-tagged jets and the highest $p_T$ (leading) $b$-tagged jet is required to have $p_T > 45$ GeV. To improve the sensitivity of the analysis a further categorisation is applied according to a quantity denoted by $p_T^V$. This is defined as $E_{\text{T}}^{\text{miss}}$ in the 0-lepton channel, the vectorial sum of the $E_{\text{T}}^{\text{miss}}$ and the lepton transverse momentum in the 1-lepton channel, and in the 2-lepton channel as the transverse momentum of the 2-lepton system. In the 0- and 1-lepton channels there is only one region $p_T^V \geq 150$ GeV. In the 2 lepton channel two regions are used: $p_T^V < 150$ GeV
and \( p_T^V \geq 150 \) GeV. Events are further split into two subdivisions according to jet multiplicity. In the 0- and 1-lepton channels events are considered with exactly two or exactly three jets. Events with four or more jets are rejected in these channels to reduce the large backgrounds arising from \( t\bar{t} \) production. In the 2-lepton channel a similar pattern is followed but extra sensitivity is gained by accepting events with higher jet multiplicities due to the lower level of the \( t\bar{t} \) background, thus the categories become either exactly two jets or three or more jets. For simplicity these two selections are just referred to as the 2 and 3-jet categories for all three lepton selections in the remainder of the note.

The event selections for the three lepton channels are detailed below and summarized in Table 1.

### 4.2.1. Zero lepton selection

Events are selected using \( E_T^{\text{miss}} \) triggers with thresholds of 70 and 90 GeV. The 70 GeV trigger is used for the 2015 data. The 90 GeV trigger is used for 2016, where the higher pile-up conditions necessitated an increase in the trigger threshold. Events are required to have no loose leptons present and a missing transverse energy of \( E_T^{\text{miss}} > 150 \) GeV. A selection based on the scalar sum of the transverse momentum of the two or three signal jets, \( S_T \), is used to remove poorly modelled regions of the phase-space arising from a non-trivial dependence of the trigger on the jet multiplicity. For events with only 2 signal jets present, the leading forward jet, if present, is used as the third jet in the sum. For 2-jet events the requirement is \( S_T > 120 \) GeV and \( S_T > 150 \) GeV is required for 3-jet events.

In order to suppress the multi-jet background four further angular selections are applied.

- \( \Delta \Phi(E_T^{\text{miss}}, E_T^{\text{miss}, \text{trk}}) < 90^\circ \)
- \( \Delta(\text{jet}_1, \text{jet}_2) < 140^\circ \)
- \( \Delta \Phi(E_T^{\text{miss}}, h) > 120^\circ \)
- \( \min(\Delta \Phi(E_T^{\text{miss}, \text{jets}})) > 20^\circ \)

Here \( \Delta \Phi(a, b) \) indicates the difference in azimuthal angle between quantity \( a \) and \( b \), \( \text{jet}_1 \) and \( \text{jet}_2 \) are the two leading signal jets, \( h \) is the direction of the reconstructed Higgs boson candidate. \( E_T^{\text{miss}} \) is defined as the missing transverse momentum calculated from the negative vector sum of the reconstructed transverse momenta of charged tracks in the inner detector. The final selection is a requirement on the azimuthal angle between the \( E_T^{\text{miss}} \) and the closest of the three leading signal jets (the leading forward jet is used if only two signal jets are present).

### 4.2.2. One lepton selection

For the electron sub-channel, events are selected using single electron triggers with thresholds of 24, 60 and 120 GeV for the 2015 data and with an increased threshold of 140 GeV for the highest threshold trigger in 2016. The lowest threshold trigger includes an isolation requirement. This is removed for the second-highest threshold trigger and the identification requirements are relaxed for the highest threshold trigger. The muon sub-channel uses the same missing transverse energy triggers as the 0-lepton channel. As muons are not included in the \( E_T^{\text{miss}} \) calculation at trigger level, in signal events where a muon is present, this trigger is effectively selecting on the \( p_T^V \). This trigger has a higher efficiency than the single muon trigger, which has limited muon trigger chamber coverage in some regions of the detector. Events are required to contain exactly one tight lepton and no additional loose leptons. In the electron sub-channel
where it is easier for a jet to fake the lepton an additional selection of $E_T^{\text{miss}} > 30$ GeV is applied to reduce multi-jet backgrounds.

### 4.2.3. Two lepton selection

Events are selected in the electron sub-channel using the same electron triggers as for the 1-lepton channel. For the muon sub-channel triggers with thresholds of 20 and 40 GeV are used for 2015 data and 24 and 50 GeV for 2016 data. As with the electron triggers the lowest threshold triggers include an isolation requirement that is removed for the higher threshold trigger. Exactly two loose leptons of the same flavour are required to be reconstructed, one of which must also pass the medium requirements. In the muon sub-channel an opposite sign requirement is also applied. This is not used in the electron sub-channel where the charge mis-identification rate is higher. The invariant mass of the dilepton system must be consistent with the Z boson mass: $71 < m_{ll} < 121$ GeV. This window is relatively large as this observable is an input to the final multivariate discriminant and is asymmetric to take into account the corresponding asymmetry in the multi-jet backgrounds.
<table>
<thead>
<tr>
<th>Selection</th>
<th>0-lepton</th>
<th>1-lepton</th>
<th>2-lepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger</td>
<td>$E_T^{\text{miss}}$</td>
<td>$E_T^{\text{miss}}$ (µ sub-channel)</td>
<td>Lowest unprescaled single lepton</td>
</tr>
<tr>
<td>Leptons</td>
<td>0 loose lepton</td>
<td>1 tight lepton</td>
<td>2 loose leptons (≥ 1 medium lepton)</td>
</tr>
<tr>
<td>Lepton pair</td>
<td>-</td>
<td>-</td>
<td>Same flavour opposite-charge for µµ</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$</td>
<td>&gt; 150 GeV</td>
<td>&gt; 30 GeV (e sub-channel)</td>
<td>-</td>
</tr>
<tr>
<td>$m_{ll}$</td>
<td>-</td>
<td>-</td>
<td>71 &lt; $m_{ll}$ &lt; 121 GeV</td>
</tr>
<tr>
<td>$S_T$</td>
<td>&gt; 120 (2 jets), &gt; 150 GeV (3 jets)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Jets</td>
<td>Exactly 2 or 3 signal jets</td>
<td>Exactly 2 or ≥ 3 signal jets</td>
<td>-</td>
</tr>
<tr>
<td>$b$-jets</td>
<td>2 $b$-tagged signal jets</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Leading jet $p_T$</td>
<td>-</td>
<td>-</td>
<td>&gt; 45 GeV</td>
</tr>
<tr>
<td>min$\Delta\phi(E_T^{\text{miss}}, \text{jet})$</td>
<td>&gt; 20°</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\Delta\phi(E_T^{\text{miss}} \cdot h)$</td>
<td>&gt; 120°</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\Delta\phi(\text{jet1, jet2})$</td>
<td>&lt; 140°</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\Delta\phi(E_T^{\text{miss}} \cdot E_T^{\text{miss}}_{\text{trk}})$</td>
<td>&lt; 90°</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$p_T^\ell$ regions</td>
<td>[0, 150] GeV (2-lepton), [150, ∞] GeV</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1: Summary of the event selection in the 0-, 1- and 2-lepton channels.
4.3. Multivariate discriminant

Multivariate discriminants making use of BDTs are constructed, trained and evaluated in each of the 0-, 1- and 2-lepton channels separately for events with two and three jets. In the 0 and 1-lepton channels, only events with $p^V_T > 150$ GeV are used. For the 2-lepton channel two BDTs are used in each of the jet multiplicity categories, one for $p^V_T < 150$ GeV and one for $p^V_T > 150$ GeV. The full list of categories for each channel is outlined in Table 2.

Two versions of the BDTs are trained. One to separate the $(VH, H \rightarrow b \bar{b})$ signal from the sum of the expected background processes, referred to as BDT\textsubscript{$VH$}, and another to separate the $(VZ, Z \rightarrow b \bar{b})$ diboson from the sum of the expected background processes, referred to as BDT\textsubscript{$VZ$}. The input variables used to construct the BDTs are chosen in order to maximise the separation, while avoiding the use of variables not improving the performance significantly. Starting from the dijet mass ($m_{bb}$), additional variables are tried one at a time and the one yielding the best separation gain is kept. This procedure is repeated until adding more variables does not result in a significant performance gain. The final sets of variables for the different channels are listed in Table 3. The $b$-tagged jets belonging to the dijet system are labelled in decreasing $p_T$ as $b_1$ and $b_2$, and their separation in pseudorapidity is $|\Delta \eta(b_1, b_2)|$. In 3-jet events, the third jet is labelled as jet$_3$ and the mass of the 3-jet system is denoted $m_{bbj}$. The angular separation, in the transverse plane, of the vector boson and the dijet system of $b$-tagged jets and their pseudorapidity separation are denoted $\Delta \phi(V, bb)$ and $|\Delta \eta(V, bb)|$, respectively. In the 0-lepton channel, $H_T$ is defined as the scalar sum of the transverse momenta of all jets and $E_T^{\text{miss}}$. In the 1-lepton channel, the angle between the lepton and the closest $b$-tagged jet in the transverse plane is denoted $\min[\Delta \phi(\ell, b)]$. The $W$ boson transverse mass is defined as $m_W^T = \sqrt{2p^\ell_T E_T^{\text{miss}} (1 - \cos(\Delta \phi(\ell, E_T^{\text{miss}})))}$ where $p^\ell_T$ is the lepton transverse momentum.

In the 1-lepton channel two variables are used to improve the rejection of the $t\bar{t}$ background: the rapidity difference between the $W$ and Higgs bosons, $|\Delta Y(V, H)|$ and, under the hypothesis that the event is $t\bar{t}$ the reconstructed top-quark mass, $m_{\text{Top}}$. To construct each of these variables an estimate of the 4-vector of the neutrino in the $W$ boson decay is required. The vector $E_T^{\text{miss}}$ is assumed to give the transverse components and then $p^\nu_z$ can be determined up to a possible two-fold ambiguity by constraining the mass of the lepton + neutrino to be consistent with the known $W$ boson mass. The top quark is then reconstructed by considering the reconstructed $W$ boson and one of the two $b$ tagged jets. The choice of the two possible $p^\nu_z$ and $b$-tagged jet is made such that the value of $m_{\text{Top}}$ is minimised.

The other variables are defined in the previous sections.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 $b$-tagged jets</td>
</tr>
<tr>
<td></td>
<td>$p^V_T &lt; 150$ GeV</td>
</tr>
<tr>
<td></td>
<td>2 jets</td>
</tr>
<tr>
<td>0-lepton</td>
<td>-</td>
</tr>
<tr>
<td>1-lepton</td>
<td>-</td>
</tr>
<tr>
<td>2-lepton</td>
<td>BDT</td>
</tr>
</tbody>
</table>

Table 2: The distributions used in the global likelihood fit for all the categories in each channel.
<table>
<thead>
<tr>
<th>Variable</th>
<th>0-lepton</th>
<th>1-lepton</th>
<th>2-lepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_T^V )</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>( E_{\text{miss}} )</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>( p_T^1 )</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>( p_T^2 )</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>( m_{bb} )</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>( \Delta R(b_1, b_2) )</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>(</td>
<td>\Delta\eta(b_1, b_2)</td>
<td>)</td>
<td>×</td>
</tr>
<tr>
<td>( \Delta\phi(V, bb) )</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>(</td>
<td>\Delta\eta(V, bb)</td>
<td>)</td>
<td>×</td>
</tr>
<tr>
<td>( H_T )</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \min[\Delta\phi(f, b)] )</td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>( m_T^W )</td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>( m_{ll} )</td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>( m_{\text{Top}} )</td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>(</td>
<td>\Delta Y(V, H)</td>
<td>)</td>
<td></td>
</tr>
</tbody>
</table>

 Only in 3-jet events

| \( p_T^{\text{jet}} \) | × | × | × |
| \( m_{bbj} \)          | × | × | × |

Table 3: Variables used in the multivariate analysis for the 0-, 1- and 2-lepton channels.

5. Multi-jet background modelling

The different processes contributing to the signal and background have been summarised in Section 3. Multi-jet backgrounds arising from strong interactions are produced with large cross sections and so have the potential to contribute a large proportion of the overall background for this analysis. Additional information regarding these backgrounds broken down by channel is discussed below.

5.1. 0-lepton channel

As described in Section 4, specific criteria are applied in the event selection to suppress the contributions from multi-jet backgrounds. The potential multi-jet contamination was studied in detail using samples of simulated and data events prior to the application of the anti-multi-jet criteria in a variety of observables for events with 0, 1 and 2 \( b \)-tagged jets. The multi-jet contribution is modelled for these studies using Pythia8 MC samples with the A14 tune and NNPDF2.3LO PDFs and used to estimate both the overall multi-jet contribution and the contribution in the region of \( m_{bb} \) close to the Higgs boson mass. Simulation and data generally agree well but a conservative approach of inflating the background prediction where there are signs of possible mis-modelling is taken. The overall contribution is <1% of the total background and <10% of the estimated Higgs boson signal yield in the region \( 60 \leq m_{bb} \leq 160 \) GeV. This is a small enough contribution that this background is neglected in the global likelihood fit.
5.2. 1-lepton channel

Both the electron and muon channels have contributions from multi-jet events faking the isolated lepton signature. The multi-jet background in the muon channel is predominately due to the semi-leptonic decay of heavy-flavour hadrons to a muon. The size of the multi-jet background in the electron channel, which is mostly due to jets or photon conversions faking leptons or semi-leptonic heavy-flavour hadron decays, results in a larger background. Given the different sources of the multi-jet background events in the two channels the backgrounds are estimated separately, but using a similar procedure.

Control regions are constructed, separately for each of the 2- and 3-jet categories, enriched in fake leptons that are kinematically close to but not overlapping with the corresponding signal regions. Correction factors are then applied to map each control region into the appropriate signal region to estimate the yield and shape of the multi-jet background. The control regions are selected using the nominal signal selection but with the requirement that the lepton should fail the tight requirements but pass a similar selection with looser identification and isolation requirements (loose-not-tight). For determining the background normalisation control regions with two $b$-tagged signal jets are used, whereas to estimate the shape of the background a larger sample with a relaxed requirement of one $b$-tagged jet is used.

To correctly extrapolate the background measured in control regions to the signal region, the non-multi-jet (referred to as EW) backgrounds with real leptons are subtracted based on MC estimates and two correction factors are applied. The first correction factor takes into account the fact that only one $b$-tag is required for the shape control region, but two are required for the signal region and is derived from measurements of the $b$-tagging efficiencies. The second factor corrects for the differing lepton selection requirements. It is measured from an additional sample of data, highly enriched in di-jet events, selected by requiring exactly one signal jet and exactly one lepton meeting the loose requirements. The scale factor is derived from the ratio of the number of tight leptons to the number of loose-not-tight leptons after subtracting the contribution from the simulated EW backgrounds with real leptons and is measured in bins of $p_T$, $|\eta|$, and $E_T^{\text{miss}}$ (the latter for the electron sub-channel only). The determination is quite sensitive to the modelling of the EW backgrounds and so a data driven constraint on the normalisation of these backgrounds is obtained by performing a fit to the data in the region $150 < E_T^{\text{miss}} < 250$ GeV.

The final background template is constructed in the relevant final discriminant variable by taking the shape and normalisation from the control regions and applying the correction factors discussed above.

5.3. 2-lepton channel

Requiring two isolated leptons with a dilepton invariant mass compatible with that of a $Z$ boson strongly suppresses the contributions from multi-jet events. The residual contribution is estimated using a fit to a sample of events where the two lepton candidates have the same sign charge (SS). The fit model includes expected contributions from EW backgrounds from simulation and an exponential model for the multi-jet background. An estimate is then made of the fraction of the background in a mass window around the $Z$ boson peak in the signal region that could be attributed to multi-jet events based on the assumption that the opposite sign and same sign events are symmetric for the multi-jet background. Inside a mass window $71 < m_{ll} < 121$ GeV the fraction of the background coming from multi-jet events is estimated to be 0.3% and 1.4% for the muon and electron sub-channels respectively. This is small enough to have a negligible impact on the signal extraction and so is not included in the statistical analysis.
6. Systematic uncertainties

There are several sources of systematic uncertainty that will impact the measurement. These can broadly be grouped into four categories: those of an experimental nature, those relating to the multi-jet background estimation, those relating to the modelling of the simulated background and Higgs boson signal samples.

6.1. Experimental uncertainties

The dominant sources of experimental uncertainty arise from the flavour tagging simulation-to-data efficiency correction factors and the jet energy scale and resolution corrections. A list of the leading systematic uncertainties ranked by their impact on the measurement of the signal strength $\mu$ can be found in Figure 5. The flavour tagging simulation-to-data efficiency correction factors are measured separately for $b$, $c$ and light-flavour jets [76]. The correction factors for $c$ and light-flavour jets are based upon measurements carried out with the Run 1 data, with Run 1 to Run 2 extrapolation uncertainties. A dedicated Run 2 correction factor has been derived for $b$-jets using $t\bar{t}$ events. All three correction factors have many sources of uncertainty and are decomposed into uncorrelated components, which are then treated independently resulting in 4 uncertainties for both $b$ and $c$ jets and 6 for light-flavour jets.

Uncertainties on the jet energy scale (JES) and resolution (JER) [71] are estimated from 13 TeV data. The many sources of uncertainties on the JES correction are decomposed into nineteen uncorrelated components which are treated as independent sources. Uncertainties on the reconstruction, identification, isolation and trigger efficiencies of muons and electrons, along with the uncertainty on their energy scale and resolution, are also evaluated based upon 13 TeV data. These are found to have only a small impact on the result. The uncertainty on the energy scale and resolution of the jets and leptons are propagated to the calculation of the $E_T^{\text{miss}}$, which also has additional dedicated uncertainties on the scale, resolution and efficiency of the tracks not associated to any of the reconstructed objects, along with the modelling of the underlying event. An uncertainty is applied on the simulation-to-data $E_T^{\text{miss}}$ trigger scale factor, relating to the statistical uncertainty on the measured scale factor and differences between the scale factor determined on $W +$ Jets, $Z +$ Jets and $t\bar{t}$ events. The uncertainty on the luminosity is 2.1% for the 2015 and 3.7% for the 2016 data, giving an uncertainty of 2.9% on the combined dataset. It is derived, following a methodology similar to that detailed in [23], from a preliminary calibration of the luminosity scale using $x$-$y$ beam-separation scans performed in August 2015 and May 2016.

6.2. Simulated background uncertainties

Modelling uncertainties are derived on the simulated backgrounds and broadly cover three areas: normalisation, acceptance differences that affect the relative normalisation between analysis regions with a common background normalisation and the shapes of the most important kinematic variables. These uncertainties are derived either from truth-particle comparisons based upon the nominal and alternative samples (outlined in Section 3) using the RIVET framework [77] or from comparisons to data in control regions. When normalisation and acceptance uncertainties are evaluated all the nominal and alternative samples are normalised to the same production cross-section. Such uncertainties are generally evaluated by adding the difference between the nominal and appropriate alternative samples in quadrature. Shape uncertainties are considered in each of the analysis regions separately, with the samples scaled to have the same normalisation in each region. The uncertainty is taken from the alternative generator which has the largest shape difference compared to the nominal sample. Shape uncertainties are only derived.
on the $m_{bb}$ and $p_T^V$ variables, as it was found that it is sufficient to only consider systematic variations of these variables to cover the full variation when evaluating the BDT$_{VH}$ discriminant for an alternative generator. A summary of the systematic uncertainties affecting the modelling of the background samples is given in Table 4 and the specific details of how the uncertainties are calculated is provided below for each simulated background sample.

**V + Jets** The V + Jets backgrounds are subdivided into three different components based upon the jet flavour labels of the leading two jets in the event. The main background contributions: $bb$, $bc$, $bl$ and $cc$, are jointly considered as the V + HF background, with the overall normalisation freely floating in the global likelihood fit, as detailed in Section 7. The remaining flavour components, $V + c l$ and $V + l$ make up less than ~1% of the total background in each analysis region; therefore only uncertainties on the total normalisation of these backgrounds are included. Acceptance uncertainties are evaluated on the relative normalisations of the different regions that share a common floating normalisation parameter, in the case of $W + HF$ this is a 2-to-3 jet and a 0-to-1 lepton ratio uncertainty. For $Z + HF$ there is a 2-to-3 jet ratio uncertainty for both the 0- and 2-lepton regions, along with a 0-to-2 lepton ratio uncertainty. Uncertainties are also evaluated on the relative normalisation of the four heavy-flavour components that make up the V + HF background. These are evaluated as uncertainties on the $bc$, $cc$ and $bl$ yield compared to the dominant $bb$ yield. These are evaluated separately for the 0- and 1-lepton channels in the case of $W + HF$ and separately for the 0-lepton, 2-lepton 2-jet and 2-lepton 3-jet regions in the case of $Z + HF$. The normalisation and acceptance uncertainties are both calculated by adding the difference between the central Sherpa 2.1.0 sample and its associated scale variations in quadrature. Shape uncertainties are estimated on $m_{bb}$ and $p_T^V$ for $Z + HF$ by comparing the $Z + Jets$ background to data in signal-depleted regions with a very high $Z + Jets$ purity, specifically the 0, 1 and 2 tag regions, with the $m_{bb}$ region around the Higgs boson mass removed in the 2 tag case. For $W + HF$ due to the lack of a dedicated control region, comparisons are made between the Sherpa 2.2.1 nominal sample and the alternative Madgraph5_aMC@NLO sample.

**tt** Uncertainties are derived from comparing the nominal sample (Powheg+Pythia6) to alternative samples with different: parton-shower generation (Powheg+Herwig++), matrix element generation (Madgraph5_aMC@NLO+Herwig++) and settings on the nominal generator designed to increase or decrease the amount of radiation. Due to the different regions of phase space being probed, uncertainties are derived separately in, and are considered uncorrelated between, the 0+1 and 2-lepton channels. The normalisation of the $t\bar{t}$ background is a freely floating parameter in the global likelihood fit, with separate parameters implemented for the 0+1 and 2-lepton channels; therefore no uncertainty is assigned on the overall $t\bar{t}$ normalisation. Acceptance uncertainties are evaluated on the relative normalisations of the different regions that share a common normalisation parameter, in this case separate 2-to-3 jet and 0-to-1 lepton ratio uncertainties in the 0+1 and 2-lepton regions. Shape uncertainties are derived on the $p_T^V$ and $m_{bb}$ variables in the 0+1 and 2-lepton regions separately.

**Single top** Uncertainties for the $Wt$ and $t$-channels are derived on the normalisation, acceptance and the shapes of the $m_{bb}$ and $p_T^V$ distributions. The $s$-channel contribution only has the normalisation uncertainty derived as its contribution is negligible overall. The nominal samples (Powheg+Pythia6) are compared to alternative samples, which are similar to those used in the $t\bar{t}$ case using different: parton-shower generation (Powheg+Herwig++, $t$-channel), matrix element generation (Madgraph5_aMC@NLO+Herwig++, $t$-channel), estimates of the single-top interference from using a diagram subtraction scheme ($Wt$) instead of the nominal diagram removal scheme and settings on the nominal generator designed to maximise or minimise the amount of radiation. The normalisation uncertainties take into account variations of the renormalisation and factorisation scale, $\alpha_s$ and PDFs. Uncertainties on the acceptance in the 2 and 3-jet
regions are derived by comparing the alternative generators and summing the difference with the nominal sample in quadrature. Shape uncertainties are derived on the $m_{bb}$ and $p_T^{V}$ distributions.

**Diboson** The diboson backgrounds are composed of three distinct processes, $WZ$, $WW$ and $ZZ$. Given the small contribution from $WW$ only a normalisation uncertainty is assigned. The more important contributions from the $WZ$ and $ZZ$ backgrounds have uncertainties derived on the overall normalisation, the relative normalisation acceptance between regions with a common normalisation parameter and on the $m_{bb}$ and $p_T^{V}$ shapes. Uncertainties are derived by comparing the nominal sample ($\text{Sherpa}$) to the alternative samples with varied factorisation, normalisation and resummation scales. An additional uncertainty is evaluated on the parton-shower generation by comparing the difference between the two Powheg samples using either Pythia8 and Herwig++ as the parton shower generator. Finally a generator comparison is made by comparing the Sherpa and Powheg+Pythia8 samples, which use different matrix element and parton shower generation amongst other differences. To estimate normalisation and acceptance uncertainties two systematic uncertainties are evaluated. Firstly, one is estimated by taking the quadrature sum of the uncertainties arising from the Sherpa scale and parton-showering variations. A second systematic uncertainty is estimated from comparing the Sherpa and Powheg+Pythia8 samples. In all cases the approach leading to the largest uncertainty is used. Acceptance uncertainties are derived on the 0-to-1 lepton channels ratio and the 2-to-3 jet regions ratio for $WZ$. In the $ZZ$ case the acceptance uncertainties are derived on the 0-to-2 lepton channels ratio and 2-to-3 jet regions ratio. Shape uncertainties are derived by comparing the impact of the individual scale, parton-showering and generator variations and taking the largest shape difference as the systematic uncertainty.

6.3. Multi-jet background uncertainties

For the 0 and 2 lepton channels it was verified, using simulated samples and data-driven studies respectively, that the multi-jet background made up less than 1% of the total background in all analysis regions. Due to the small contribution of this background it was neglected in both channels. Given the larger contribution of multi-jet events in the 1-lepton channel, this background component is estimated from data as outlined in Section 5. Three sources of uncertainty are considered on the estimate of the multi-jet background in the electron channel: the relative amount of fake leptons from jets versus those from semi-leptonic heavy flavour hadron decays, the amount of non-multi-jet background in the reversed isolation region and the impact of the energy scale applied to the ‘fake-electrons’ on the $E_{T}^{\text{miss}}$ calculation. In the muon channel the sources of uncertainty are due to the size of the non-multi-jet background in the reversed isolation region and the bias caused by the different $E_{T}^{\text{miss}}$ that isolated and non-isolated muon events may have.

6.4. Signal uncertainties

The signal samples are normalised to the inclusive cross sections as described in Section 3. Uncertainties on the calculations are assigned following the recommendations of the LHC Higgs Cross Section working group [61, 78–80]. An uncertainty is estimated by varying the factorisation and renormalisation scale factors independently between a third and three times their nominal values. For $WH$ production this gives an uncertainty of 0.7%. For $ZH$ production it is somewhat larger at around 3% due to the contributions from the gluon fusion process. To assess uncertainties independently for the quark-quark and gluon-gluon initiated processes it is assumed that the uncertainty on $qq \rightarrow ZH$ is identical to $WH$ at 0.7%. This assumption then leads to an estimate of the uncertainty on the $gg \rightarrow ZH$ cross section of 27%. The PDF and $\alpha_S$ contributions to the uncertainty are 1.9% for $WH$, 1.6% for $qq \rightarrow ZH$ [61], and 5% for
A further uncertainty is assessed to account for missing higher order EW effects as the maximum of either: 1%, the relative NLO EW corrections squared or the uncertainty on the photon induced cross section relative to the total VH cross section. The uncertainty on the $H \rightarrow b \bar{b}$ branching ratio is 1.7%, which takes into account missing higher order QCD and EW effects as well as uncertainties on the Higgs boson decay width.

Acceptance uncertainties are estimated using event samples generated using Madgraph5_aMC@NLO v2.2.2 at LO interfaced to the Pythia 8.186 parton shower model. The A14 tune is used together with the PDF4LHC2015 NLO PDF set [81]. The uncertainty is estimated by varying the renormalisation and factorisation scales between a half and twice their nominal values independently and taking the envelope of the variations. Following the scheme advocated in [82], the uncertainties are decomposed into three components that correctly take into account the use of exclusive jet bins. The uncertainties for $q\bar{q} \rightarrow VH$ in the 2-jet category are 4%, 5%, and 1.4% for the 0-, 1- and 2-lepton channels respectively and 2.7%, 4.7%, and 1.4% for the 3-jet category. The uncertainties for $gg \rightarrow ZH$ are assumed to be the same as for $q\bar{q} \rightarrow VH$.

The PDF and $\alpha_S$ contributions to the acceptance uncertainty are estimated following the PDF4LHC prescription [81, 83] and for $q\bar{q} \rightarrow VH$ are: 0.6%, 0.7% and 0.3% for the 0-, 1- and 2-lepton selections respectively. An uncertainty of 3% is used for the $gg \rightarrow ZH$ contribution. Contributions from the modelling of the underlying event, parton showering and hadronisation (UE/PS/HAD) are estimated using the eigen-tunes of the A14 tune[47]. This procedure leads to uncertainties correlated between the 2- and 3-jet categories of 7.5%, 7.5% and 6.5% for the 0-, 1- and 2-lepton selections, with additional uncertainties for the 3-jet category of 5%, 6% and 4% respectively.

Where significant variations are observed in the modelling of the $p_{T}^{V}$ and $m_{bb}$ distributions under: the scale variations, PDF and $\alpha_S$ variations; variations accounting for missing higher order EW effects; or variations associated with UE/PS/HAD modelling, shape uncertainties are assessed by taking the envelope of the observed variations and propagating it through to the relevant final discriminant distributions. Variations from the same source affecting the same distribution are correlated in the signal extraction fits. For the scale, EW, and UE/PS/HAD variations a shape systematic is assessed for both $p_{T}^{V}$ and $m_{bb}$ for all analysis regions. For PDF and $\alpha_S$ contributions to the shape variation no significant impact is seen for $m_{bb}$ and so shape variations are assessed only for their impact on $p_{T}^{V}$.

A summary of the systematic uncertainties affecting the modelling of the signal samples is given in Table 5.
<table>
<thead>
<tr>
<th>Background</th>
<th>Normalisation</th>
<th>0-to-2 lepton ratio</th>
<th>2-to-3 jet ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(\ell)+jets</td>
<td>Z(\ell) normalisation</td>
<td>18%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zc(\ell) normalisation</td>
<td>23%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zbb normalisation</td>
<td>Floating</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zbc-to-Zbb ratio</td>
<td>14-27%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zcc-to-Zbb ratio</td>
<td>7-31%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zbl-to-Zbb ratio</td>
<td>15-38%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-to-2 lepton ratio</td>
<td>26%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-to-3 jet ratio</td>
<td>28% (0-lepton) and 25% (2-lepton)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(p_T^{\ell}, m_{bb})</td>
<td>S</td>
<td></td>
</tr>
</tbody>
</table>

| W\(\ell\)+jets | W\(\ell\) normalisation | 32% |  |
| | Wc\(\ell\) normalisation | 37% |  |
| | Wbb normalisation | Floating |  |
| | Wbl-to-Wbb ratio | 17% (0-lepton) and 31% (1-lepton) |  |
| | Wbc-to-Wbb ratio | 42% (0-lepton) and 21% (1-lepton) |  |
| | Wcc-to-Wbb ratio | 17% (0-lepton) and 31% (1-lepton) |  |
| | 2-to-3 jet ratio | 23% |  |
| | 0-to-1 lepton ratio | 17% |  |
| | \(p_T^{\ell}, m_{bb}\) | S |  |
| | \(t\bar{t}\) (all are decorrelated between the 0+1 and 2-lepton channels) | Floating |  |
| | 2-to-3-jet ratio | 9% (0+1-lepton) and 24% (2-lepton) |  |
| | \(p_T^{\ell}, m_{bb}\) | S |  |

| Single top | Cross section | 4.4% (s-channel), 4.6% (t-channel), 6% (Wt) |  |
| | Acceptance 2-jet | 16% (t-channel), 25% (Wt) |  |
| | Acceptance 3-jet | 19% (t-channel), 32% (Wt) |  |
| | \(m_{bb}, p_T^{\ell}\) | S (\(p_T^{\ell}\) uncorrelated between 2 and 3-jet channels Wt) |  |

| ZZ | Normalisation | 20% |  |
| | 0-to-2 lepton ratio | 30% |  |
| | 2-to-3 jet ratio | 19% |  |
| | \(m_{bb}, p_T^{\ell}\) | S (correlated with WZ uncertainties) |  |

| WZ | Normalisation | 26% |  |
| | 2-to-3 jet ratio | 14% (0-lepton) and 11% (1-lepton) |  |
| | 0-to-1 lepton ratio | 12% |  |
| | \(m_{bb}, p_T^{\ell}\) | S (correlated with ZZ uncertainties) |  |

| WW | Normalisation | 25% |  |

| Multi-jet (1-lepton) | Normalisation | 14-81% (electron), 5-50% (muon) |  |
| | Template variations | S |  |

Table 4: Summary of the systematic uncertainties on the background modelling. An “S” symbol is used when only a shape uncertainty is assessed.
Table 5: Summary of the systematic uncertainties on the signal modelling. An “S” symbol is used when only a shape uncertainty is assessed.
7. Statistical analysis

A statistical fitting procedure based on the Roostats framework [84, 85] is used to extract the signal strength from the data. The signal strength is a parameter, $\mu$, that scales the SM Higgs boson production cross section times branching ratio into $b\bar{b}$. A binned likelihood function is constructed as the product of Poisson-probability terms over the bins of the input distributions involving the numbers of data events and the expected signal and background yields, taking into account the effects of the floating background normalisations and the systematic uncertainties.

The impact of systematic uncertainties on the signal and background expectations is described by nuisance parameters (NPs), $\theta$, which are constrained by Gaussian or log-normal probability density functions, the latter being used for normalisation uncertainties to prevent normalisation factors from becoming negative in the fit. The expected numbers of signal and background events in each bin are functions of $\theta$. For each NP, the prior is added as a penalty term to the likelihood, $\mathcal{L}(\mu, \theta)$, which decreases it as soon as $\theta$ is shifted away from its nominal value. The statistical uncertainties associated with the background predictions from simulation are included through bin-by-bin nuisance parameters.

The test statistic $q_{\mu}$ is then constructed from the profile likelihood ratio

$$ q_{\mu} = -2 \ln \Lambda_{\mu} \text{ with } \Lambda_{\mu} = \mathcal{L}(\mu, \hat{\theta}_{\mu}) / \mathcal{L}(\hat{\mu}, \hat{\theta}), $$

where $\hat{\mu}$ and $\hat{\theta}$ are the parameters that maximise the likelihood with the constraint $0 \leq \hat{\mu} \leq \mu$, and $\hat{\theta}_{\mu}$ are the nuisance parameter values that maximise the likelihood for a given $\mu$. This test statistic is used for exclusion intervals derived with the CL$_{s}$ method [86, 87]. To measure the compatibility of the background-only hypothesis with the observed data, the test statistic used is $q_{0} = -2 \ln \Lambda_{0}$. The results are presented in terms of: the 95% confidence level (CL) upper limit on the signal strength; the probability $p_{0}$ of the background-only hypothesis; and the best-fit signal-strength value $\hat{\mu}$ with its associated uncertainty $\sigma_{\mu}$. The fitted $\hat{\mu}$ value is obtained by maximising the likelihood function with respect to all parameters. The uncertainty $\sigma_{\mu}$ is obtained from the variation of $2 \ln \Lambda_{\mu}$ by one unit, where $\Lambda_{\mu}$ is now defined without the constraint $0 \leq \hat{\mu} \leq \mu$. Expected results are obtained in the same way as the observed results by replacing the data in each input bin by the expectation from simulation with all NPs set to their best-fit values, as obtained from the fit to the data.3

There are 8 categories used in the global likelihood fit. These are the 2-tag regions for the three lepton channels, which are subdivided into 2 and 3-jet high $p_{T}$ categories with equivalent low $p_{T}$ regions for the 2-lepton analysis. The data have sufficient statistical power to constrain the largest background-normalisation NPs, which are left free to float in the fit. This applies to the $t\bar{t}$, $W + \text{HF}$, and $Z + \text{HF}$ processes. The $t\bar{t}$ background is normalised separately in the 0+1 and 2 lepton channels due to the different regions of phase space probed in the two cases. Additional fits are performed: separately for each lepton channel, combined for all channels but with three independent signal strength parameters one for each lepton channel, and combined for all channels with with two independent signal strength parameters for the $ZH$ and $WH$ processes separately.

The full list of systematic uncertainties, along with the correlation scheme, is outlined in Section 6. A ‘ranking’ for the NPs is established by performing the fit again for each NP, with the corresponding NP

3 This type of pseudo-data sample is referred to as an Asimov dataset in ref. [87].
fixed to its fitted value, $\hat{\theta}$, shifted up or down by its fitted uncertainty with all the other parameters allowed to vary to take properly into account the correlations between systematic uncertainties. The magnitude of the shift in the fitted signal strength $\hat{\mu}$ is a measure of the observed impact of the considered NP. The same procedure is repeated, using the nominal values of the NP and of its associated uncertainty to provide its expected impact.

The fit uses templates constructed from the predicted yields for the signal and the various backgrounds in the bins of the input distribution in each region. The systematic uncertainties are encoded in templates of variations relative to the nominal template for each up-and-down ($\pm 1\sigma$) variation. To reduce the impact of statistical fluctuations, which can be introduced when a systematic variation alters the event selection, a smoothing procedure is applied where bins are merged based on the constraints that the statistical uncertainty in each bin should be less than 5%.

8. Results

The results of the Higgs boson search and the diboson analysis are reported below. In the following the the fitted signal-strength parameters are denoted $\mu$ and $\mu_{VZ}$ rather than $\hat{\mu}$ and $\hat{\mu}_{VZ}$.

8.1. Higgs boson search

The factors applied to the nominal background normalisations resulting from the global likelihood fit are shown in Table 6. The post-global likelihood fit signal and background yields are shown in Table 7 for all the analysis regions. Figure 1 shows a selection of characteristic post-fit plots for each of the lepton channels and Fig. 2 shows the BDT output distributions in the most sensitive 2-jet, 2-tag high $p_T^V$ categories.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Scale factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$ 0+1-lepton</td>
<td>0.86 ± 0.13</td>
</tr>
<tr>
<td>$t\bar{t}$ 2-lepton</td>
<td>0.94 ± 0.09</td>
</tr>
<tr>
<td>W + HF</td>
<td>1.59 ± 0.39</td>
</tr>
<tr>
<td>Z + HF</td>
<td>1.04 ± 0.11</td>
</tr>
</tbody>
</table>

Table 6: Factors applied to the nominal normalisations of the $t\bar{t}$, W + HF and Z + HF backgrounds, as obtained from the global likelihood fit. The $t\bar{t}$ background is normalised independently in the 0+1 and 2 lepton channels. The errors include the statistical and systematic uncertainties.
Table 7: The data, background and signal yields along with the total uncertainty. All the background and signal values are evaluated according to the results of the global fit. The $V + HF$ yields includes events from the $V + bb$, $V + bc$, $V + bl$ and $V + cc$ categories.
For all lepton channels combined the observed limit on the ratio of the cross section times branching ratio with respect to the SM expectation for $m_H = 125$ GeV is 1.2, to be compared to an expected limit, in the absence of signal, of $1.0^{+0.4}_{-0.3}$. The probability $p_0$ of obtaining from background alone a result at least as signal-like as the observation is 34% for a tested Higgs boson mass of 125 GeV. In the presence of a Higgs boson with that mass and the SM signal strength, the expected $p_0$ value is 3%. This corresponds to an observed excess with a significance of 0.42 standard deviations, to be compared to an expectation of 1.94 standard deviations. Table 8 shows the expected and observed 95% CL limits, $p_0$ and significance values for the separate lepton channel fits and for the lepton channels combined in the global fit. For all

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Limit</th>
<th>$p_0$</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-lepton</td>
<td>1.4$^{+0.6}_{-0.4}$</td>
<td>2.0</td>
<td>0.07</td>
</tr>
<tr>
<td>1-lepton</td>
<td>2.0$^{+0.8}_{-0.6}$</td>
<td>2.1</td>
<td>0.15</td>
</tr>
<tr>
<td>2-lepton</td>
<td>1.8$^{+0.7}_{-0.5}$</td>
<td>1.7</td>
<td>0.13</td>
</tr>
<tr>
<td>Combined</td>
<td>1.0$^{+0.4}_{-0.3}$</td>
<td>1.2</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 8: The expected and observed 95% CL limits on the ratio of the cross-section times branching ratio with respect to the SM expectation and $p_0$ and significance values for the individual lepton channels and their combination. The expected limits are evaluated assuming the absence of signal and the expected $p_0$ and significance assuming a Higgs boson of 125 GeV mass with the SM signal strength.

channels combined the fitted value of the signal-strength parameter is $\mu = 0.21^{+0.36}_{-0.35}$ (stat.) $\pm 0.36$ (syst). Fits are also performed in the case of the three lepton channels combined, where the signal strengths are floated independently for (i) the $WH$ and $ZH$ production processes, or (ii) the three lepton channels, but leaving all other NPs with the same correlations as the nominal result. The results of these fits are shown in Fig. 3. Figure 4 shows the data, background and signal yields, where final-discriminant bins in all signal regions are combined into bins of $\log(S/B)$. Here, $S$ is the expected signal yield and $B$ is the fitted background yield.

The ranking of the systematic uncertainties in the global likelihood fit are shown in Fig. 5. The NPs are ordered by decreasing post-fit impact on $\mu$. The five systematic uncertainties with the largest impact are the two leading $b$-jet efficiency uncertainties, the leading $c$-jet efficiency uncertainty, the $W + HF$ and $Z + HF$ normalisations.

8.2. Diboson validation

As outlined in Section 4, BDTs have also been trained to select $VZ$ diboson as the signal to validate the techniques and modelling used in the Higgs boson analysis. The validation results are obtained using the maximum likelihood fits described in Section 7 in an identical manner as for the signal fit, but using the BDT$_{VZ}$ output as the final discriminant and with the $VZ$ signal-strength parameter, $\mu_{VZ}$, freely floating. The diboson and Higgs boson BDTs provide sufficient separation between the $VZ$ and $VH$ processes such that they only have a weak correlation in their results. The Higgs boson signal-strength parameter, $\mu$, is set to the SM prediction with a 50% uncertainty assigned on the normalisation. A value
of $\mu_{VZ} = 0.91 \pm 0.17(\text{stat.})^{+0.32}_{-0.27}(\text{syst.})$ is obtained. The $VZ$ signal is observed with a significance of 3.0 standard deviations, to be compared to an expected significance of 3.2 standard deviations.
Figure 1: The $E_{T}^{miss}$ (top left), $m_{T}^{W}$ (middle left), $m_{ll}$ (bottom left) and $m_{bb}$ (right) post-fit distributions in the 0-lepton (top) 1-lepton (middle) and 2-lepton (bottom) channels for 2-jet, 2 $b$-tag events in the high $p_{T}$ region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_{H} = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.
Figure 2: The BDT\textsubscript{VH} post-fit distributions in the 0-lepton (top) 1-lepton (middle) and 2-lepton (bottom) channel for 2-tag events, in the 2-jet (left) and exactly 3 or ≥3-jets for the 2-lepton case (right) categories in the high $p_T$ region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.
Figure 3: The fitted values of the Higgs boson signal-strength parameter, \( \mu \), for \( m_H = 125 \) GeV for the \( WH \) and \( ZH \) processes and their combination (left) and for the 0-, 1- and 2-lepton channels and their combination (right). The individual \( \mu \) values in either the case of the \( (W/Z)H \) processes or individual lepton channels are obtained from a simultaneous fit with the signal strength for each of the processes or lepton channels floated independently.

Figure 4: Event yields as a function of \( \log(S/B) \) for data, background and Higgs boson signal with \( m_H = 125 \) GeV. Final-discriminant bins in all signal regions are combined into bins of \( \log(S/B) \). The signal \( S \) and background \( B \) yields are the expected and fitted values, respectively. The Higgs boson signal contribution is shown as expected for the SM cross section (indicated as \( \mu = 1.0 \)). The pull of the data with respect to the background-only prediction is shown without systematic uncertainties. The solid red line indicates the pull of the prediction for signal (\( \mu = 1.0 \)) and background with respect to the background-only prediction.
Figure 5: Impact of systematic uncertainties on the fitted signal-strength parameter $\mu$. The systematic uncertainties are listed in decreasing order of their impact on $\mu$ on the $y$-axis. The boxes show the variations of $\mu$, referring to the top $x$-axis, when fixing the corresponding individual nuisance parameter $\theta$ to its post-fit value $\hat{\theta}$ modified upwards or downwards by its post-fit uncertainty, and repeating the fit as explained in the text. The hatched and open areas correspond to the upwards and downwards variations, respectively. The filled circles, referring to the bottom $x$-axis, show the deviations of the fitted nuisance parameters $\hat{\theta}$ from their nominal values $\theta_0$, expressed in terms of standard deviations with respect to their nominal uncertainties $\Delta \theta$. The associated error bars show the post-fit uncertainties of the nuisance parameters, relative to their nominal uncertainties. The open circles with their error bars, also referring to the bottom $x$-axis, show the fitted values and uncertainties of the normalisation parameters that are freely floating in the fit. The normalisation parameters have a pre-fit value of one. As explained in Section 6, the jet energy scale and $b$-tagging uncertainties are decomposed into uncorrelated components; the numerical labels refer to such components.
9. Conclusions

A search for the decay of a Standard Model Higgs boson into a $b\bar{b}$ pair when produced in association with a $W$ or $Z$ boson has been performed with the ATLAS experiment using data collected in proton-proton collisions from Run 2 of the Large Hadron Collider. This dataset corresponds to an integrated luminosity of 13.2 fb$^{-1}$ collected at a centre of mass energy of $\sqrt{s} = 13$ TeV. For $m_H = 125$ GeV a 95% CL upper limit is set on the ratio of the cross section times branching ratio with respect to the SM expectation for $pp \rightarrow (W/Z)(H \rightarrow b\bar{b})$ of 1.2, to be compared to the expected limit in the absence of signal of $1.0^{+0.4}_{-0.3}$. The measured signal strength with respect to the SM expectation is found to be $\mu = 0.21^{+0.36}_{-0.35}$ (stat.) $\pm 0.36$ (syst.). This corresponds to a significance of 0.42 standard deviations compared with an expectation of 1.94 and is consistent with previous results in this channel. The analysis procedure has been validated by measuring the yield of $(W/Z)Z$ with $Z \rightarrow b\bar{b}$, where the ratio of the observed yield to that expected in the Standard Model is found to be $0.91 \pm 0.17$ (stat.) $^{+0.32}_{-0.27}$ (syst.), corresponding to a significance of 3.0 standard deviations compared with an expectation of 3.2.

References


29


32

[65] ATLAS Collaboration, Electron efficiency measurements with the ATLAS detector using the 2015 477 LHC proton-proton collision data,


[80] LHC Higgs Cross Section Working Group et al.,
Handbook of LHC Higgs Cross Sections: 3. Higgs Properties,

arXiv: 1510.03865 [hep-ph].

[82] 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].


[84] L. Moneta et al., The RooStats Project, PoS ACAT2010 (2010) 057,

[85] W. Verkerke and D. P. Kirkby, The RooFit toolkit for data modeling,


[87] G. Cowan et al., Asymptotic formulae for likelihood-based tests of new physics,
Appendix

A. Additional post-global fit distributions

Figures 6, 7 and 8–10 show characteristic distributions for the 0, 1 and 2-lepton channel respectively after the global likelihood fit.

Figure 6: The $E_{T}^{miss}$ (left) and $m_{bb}$ (right) post-fit distributions in the 0-lepton channel for 2-tag events, in the 3-jet high $p_{T}^{V}$ region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.
Figure 7: The $m_W$ (left) and $m_{bb}$ (right) post-fit distributions in the 1-lepton channel for 2-tag events, in the 3-jet high $p_T$ region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.
Figure 8: The $m_{bb}$ post-fit distributions in the 2-lepton channel for 2-tag events, in the 2-jet low $p_T^{V}$ (top left), $\geq 3$-jets low $p_T^{V}$ (top right) and $\geq 3$-jets high $p_T^{V}$ regions (bottom). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_{H} = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.
Figure 9: The $m_{ll}$ post-fit distributions in the 2-lepton channel for 2-tag events, in the 2-jet low $p_T$ (top left), ≥ 3-jets low $p_T$ (top right) and ≥ 3-jets high $p_T$ regions (bottom). The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.
Figure 10: The BDT_{VH} post-fit distributions in the 2-lepton channel for 2-tag events, in the 2-jet (left) and ≥3-jets (right) for the low $p_T^{V}$ region. The background contributions after the global likelihood fit are shown as filled histograms. The Higgs boson signal ($m_H = 125$ GeV) is shown as a filled histogram on top of the fitted backgrounds as expected from the SM (indicated as $\mu = 1.0$). The dashed histogram shows the total background as expected from the pre-fit MC simulation. The size of the combined statistical and systematic uncertainty on the sum of the signal and fitted background is indicated by the hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.