Observation of $H \to b\bar{b}$ decays and $VH$ production with the ATLAS detector

The ATLAS Collaboration

A search for the decay of the Standard Model Higgs boson into a $b\bar{b}$ pair when produced in association with a $W$ or $Z$ boson is performed with the ATLAS detector. The analysed data, corresponding to an integrated luminosity of 79.8 fb$^{-1}$, were collected in proton–proton collisions in Run 2 of the Large Hadron Collider at a centre-of-mass energy of 13 TeV. For a Higgs boson mass of 125 GeV, an excess of events over the expected background from other Standard Model processes is found with an observed (expected) significance of 4.9 (4.3) standard deviations. A combination is performed with the results from other searches for the Higgs boson in the $b\bar{b}$ decay mode, which yields an observed (expected) significance of 5.4 (5.5) standard deviations, thus providing direct observation of the Higgs boson decay into $b$-quarks. The measured ratio of the branching fraction for a Higgs boson decaying to $b\bar{b}$ to the Standard Model expectation is $1.01 \pm 0.12$ (stat.)$^{+0.16}_{-0.15}$ (syst.). Additionally, a combination of Run 2 results searching for the Higgs boson produced in association with a vector boson, yields an observed (expected) significance of 5.3 (4.8) standard deviations.

© 2018 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.
1 Introduction

The Higgs boson [1–4] was discovered in 2012 by the ATLAS and CMS Collaborations [5, 6] at a mass of approximately 125 GeV from the analysis of proton-proton (pp) collisions produced by the Large Hadron Collider (LHC) [7]. Since then, the analysis of datasets collected at centre-of-mass energies of 7 TeV, 8 TeV and 13 TeV in Runs 1 and 2 of the LHC has led to the observation of many of the production modes and decay channels predicted by the Standard Model (SM). The bosonic decay channels are well established and have entered an era of precision measurements [8–14]. The decay in τ pairs was first observed in the combination of the ATLAS and CMS analyses [15], then by the individual collaborations [16, 17]. The main Higgs boson production modes, gluon fusion and vector boson fusion (VBF), were already measured following the analysis of the Run-1 dataset, and recently a direct observation of the coupling of the Higgs boson to top quarks has been achieved by the ATLAS [18] and CMS [19] Collaborations through the observation of the associated production of a Higgs boson and a top quark pair (t¯tH).

The dominant decay of the SM Higgs boson is to pairs of b-quarks, with an expected branching ratio of approximately 58% for a mass of \( m_H = 125 \) GeV [20], however very large backgrounds from multi-jet production make a search in the dominant gluon fusion production mode very challenging at the LHC. The most sensitive channels to measure \( H \rightarrow b\bar{b} \) decays are instead the associated production of a Higgs boson and a W or Z boson [21] (VH), where the leptonic decay of the vector boson allows for efficient triggering and significant reduction of the multi-jet backgrounds. As well as probing the dominant decay of the Higgs boson, this measurement allows the overall Higgs boson decay width [22, 23] to be constrained and provides the best sensitivity to the \( ZH \) and \( WH \) production modes, which are important elements in the interpretation of Higgs boson measurements in effective field theories, as discussed in Ref. [24].

Searches in this channel at the Tevatron by the CDF and D0 Collaborations showed an excess of events with a significance of 2.8 standard deviations for a Higgs boson at a mass of 125 GeV [25]. Analysing the 2015 and 2016 datasets and combining with the Run-1 results [26, 27], both the ATLAS and CMS Collaborations reported evidence for Higgs boson production and decay in this channel, with observed significances of 3.6 and 3.8 standard deviations, respectively [28, 29]. Searches for \( H \rightarrow b\bar{b} \) decays have also been conducted in the VBF [30–32] and t¯tH [33–37] channels, and at high \( p_T \) of the Higgs boson [38], but with markedly lower sensitivities.

This note reports an update of the search for the SM Higgs boson decaying to a b¯b pair in the VH production mode with the ATLAS detector in Run 2 of the LHC presented in Ref. [28], using a much larger dataset of 79.8 fb\(^{-1}\) of pp collision data collected at a centre-of-mass energy of 13 TeV. The analysis selects events in 0-, 1- and 2-lepton channels, based on the number of charged leptons (electrons or muons), to explore the \( ZH \rightarrow \nu\nu b\bar{b} \), WH \( \rightarrow \ell\nu b\bar{b} \) and \( ZH \rightarrow \ell\ell b\bar{b} \) signatures, respectively. Multivariate discriminants, built from variables which describe the kinematics of the selected events, are used to maximise the sensitivity to the Higgs boson signal. Their outputs are combined using a binned maximum likelihood fit, referred to as the global likelihood fit, that allows the signal strength and the background normalisations to be extracted. The result is then combined with that of the previously published analysis of the Run 1 dataset [27], with other searches for b¯b decays of the Higgs boson (in t¯tH [33, 35] and VBF [30, 32] production modes) and with other searches in the VH production mode (in ZZ\(^*\) to four leptons [39] and diphoton [40] decay modes).
2 ATLAS detector

ATLAS [41] is a general-purpose particle detector covering nearly the entire solid angle\(^1\) around the collision point. An inner tracking detector (ID or inner detector in the rest of the note), located within a 2 T axial magnetic field generated by a thin superconducting solenoid, is used to measure the trajectories and momenta of charged particles. The inner layers consisting of high-granularity silicon pixel detectors instrument a pseudo-rapidity range \(|\eta| < 2.5\), including an innermost silicon pixel layer [42] that was added to the detector between Run 1 and Run 2. Silicon strip detectors covering \(|\eta| < 2.5\) are located beyond the pixel detectors. Outside the strip detectors and covering \(|\eta| < 2.0\), there are straw tube tracking detectors, which also provide measurements of transition radiation that are used in electron identification. A calorimeter system surrounds the ID, covering the pseudorapidity range \(|\eta| < 4.9\). Within the region \(|\eta| < 3.2\), electromagnetic calorimetry is provided by barrel (\(|\eta| < 1.475\)) and endcap (1.375 < |\eta| < 3.2) high-granularity lead/liquid-argon (LAr) electromagnetic calorimeters, with an additional thin LAr presampler covering \(|\eta| < 1.8\) to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by a steel/scintillator-tile calorimeter within \(|\eta| < 1.7\), and two copper/LAr hadronic endcap calorimeters extend the coverage to \(|\eta| = 3.2\). The solid angle coverage for \(|\eta|\) between 3.2 and 4.9 is completed with copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements, respectively. The outermost part of the detector is the muon spectrometer, which measures the curved trajectories of muons in the field of three large air-core superconducting toroidal magnets. High precision tracking is performed within the range \(|\eta| < 2.7\) and there are chambers for fast triggering within the range \(|\eta| < 2.4\). A two-level trigger system [43] is used to reduce the recorded data rate. The first level is a hardware implementation aiming to reduce the rate to around 100 kHz, while the software-based High-Level Trigger provides the remaining rate reduction to approximately 1 kHz.

3 Object and event selection

Event topologies characteristic of \(VH, H \rightarrow b\bar{b}\) processes typically contain zero, one or two charged leptons, and two jets containing particles from \(b\)-hadron decays. The object and event selections used follow to a large extent the ones of Ref. [28].

3.1 Object reconstruction

Tracks measured in the inner detector are used to reconstruct interaction vertices [44], of which the one with the highest sum of squared transverse momenta of associated tracks is selected as the primary vertex.

Electrons are reconstructed from topological clusters of energy deposits in the calorimeter [45] and matched to a track in the inner detector. Following Refs. [28, 46], loose electrons are required to have

\(^1\) 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 pseudorapidity 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.
\( p_T > 7 \text{ GeV} \) and \( |\eta| < 2.47 \), to have small impact parameters\(^2\), a loose track isolation requirement, and to pass a “LooseLH” quality criterion computed from shower shape and track quality variables \([47]\). In the 1-lepton channel, tight electrons are selected based on a “TightLH” likelihood requirement and on a more stringent calorimeter-based isolation.

Muons are required to be within the acceptance of the muon spectrometer \( |\eta| < 2.7 \), to have \( p_T > 7 \text{ GeV} \) and to have small impact parameters. Loose muons are selected based on a “loose” quality criteria \([48]\) and on a loose track isolation. In the 1-lepton channel, tight muons follow “medium” quality criteria and a stricter track isolation.

Hadronically decaying \( \tau \)-leptons \([49, 50]\) are required to have \( p_T > 20 \text{ GeV} \) and \( |\eta| < 2.5 \), to be outside of the transition region between the calorimeters \( 1.37 < |\eta| < 1.52 \), and to pass a “medium” quality criterion \([50]\). They are only used in the analysis to distinguish jets from \( \tau \) leptons.

Jets are reconstructed with the anti-\( k_T \) algorithm \([51]\) with radius parameter \( R = 0.4 \). A Jet Vertex Tagger (JVT) \([52]\) is used to remove jets associated to pile-up vertices\(^3\) for \( p_T < 60 \text{ GeV} \) and \( |\eta| < 2.4 \). Jet cleaning criteria identify jets arising from non-collision backgrounds or noise in the calorimeters \([53]\), with events containing such jets removed. Jets are required to have \( p_T > 20 \text{ GeV} \) in the central region \( (|\eta| < 2.5) \), and \( p_T > 30 \text{ GeV} \) outside of the tracker acceptance \( (2.5 < |\eta| < 4.5) \). In the central region, they are tagged as containing \( b \)-hadrons by requiring the jet to have a large MV2c10 multivariate discriminant \([54]\) output value, with the selection tuned to produce an average efficiency of 70% for \( b \)-jets in simulated \( t \bar{t} \) events.

Simulated jets are labelled as \( b \)-, \( c \)- or light-jets according to which hadrons with \( p_T > 5 \text{ GeV} \) are found within a cone of size \( \Delta R = 0.3 \) around their axis \([55]\). Simulated \( V + \text{jets} \) events are categorised depending on the labels of the jets that form the Higgs boson candidate: \( V + ll \) when they are both light-jets, \( V + c l \) when there is one \( c \)-jet and one light-jet, and \( V + \text{HF} \) in all other cases, mainly two \( b \)-jets. Owing to the large rejection of light jets achieved by the MV2c10 discriminant, simulated \( V + ll \), \( V + c l \) and \( WW \) events are not subjected to the \( b \)-tagging requirement, but rather they are weighted by the probability that their jets pass the \( b \)-tagging selection \([28]\).

In addition to the standard jet energy scale (JES) calibration \([56, 57]\), \( b \)-tagged jets receive additional flavour-specific corrections to improve their energy scale and resolution: the four-momentum of muons found within \( \Delta R = 0.4 \) is added to that of the jet, and a residual correction is applied to equalise the response of jets with leptonic or hadronic decays of the heavy flavour hadrons. In the 2-lepton channel, a per-event kinematic likelihood uses the full reconstruction of the event kinematics to improve the estimate of the energy of the \( b \)-jets. The corrections improve the resolution of the \( m_{bb} \) mass distribution by up to 40%.

The missing transverse momentum \( E_T^{\text{miss}} \) is reconstructed as the negative vector sum of the momenta of leptons, including hadronically decaying \( \tau \) leptons, and jets, and of a soft term built from the additional tracks matched to the primary vertex \([58, 59]\). The magnitude of \( E_T^{\text{miss}} \) is referred to as \( E_T^{\text{miss}} \).

An overlap removal procedure is applied to avoid any double-counting between the reconstructed leptons, including the hadronically decaying \( \tau \) leptons, and jets.

---

\(^2\) Impact parameters are defined with respect to the primary vertex position, where the beam line is used to approximate the primary vertex position in the transverse plane.

\(^3\) Collisions other than those from the hard scatter are referred to as pile-up vertices.
3.2 Event selection and categorisation

Events are categorised in the 0-, 1- and 2-lepton channels depending on the number of selected electrons and muons, to target the $ZH \rightarrow vvbb, WH \rightarrow ℓνbb$ and $ZH \rightarrow ℓℓbb$ signatures, respectively. In all channels, events are required to have exactly two $b$-tagged jets, which form the Higgs boson candidate. The highest $b$-tagged jet $p_T$ should be greater than 45 GeV. Events are categorised as 2-jet or 3-jet depending on whether additional, untagged jets are present. In the 0- and 1-lepton channels, only one such jet is allowed, while in the 2-lepton channel any number of jets is accepted in the 3-jet category.

The reconstructed vector boson transverse momentum $p_T^V$ corresponds to $E_{T}^{miss}$ in the 0-lepton channel, to the vectorial sum of $E_{T}^{miss}$ and of the charged lepton transverse momentum in the 1-lepton channel, and to the transverse momentum of the 2-lepton system in the 2-lepton channel. As the signal-to-background ratio increases for large Higgs boson transverse momenta, the analysis focuses on a high-$p_T^V$ category defined as $p_T^V > 150$ GeV. In the 2-lepton channel, the sensitivity is increased by the addition of a medium-$p_T^V$ category with $75$ GeV $< p_T^V < 150$ GeV.

Two versions of the analysis are carried out, one using a multivariate approach and the other using the dijet-mass as the final discriminant. The initial event selection shown in Table 1 is applied to both versions, with further selections applied for the dijet-mass analysis. The two versions of the analyses also have different event categorisations, with further details outlined below.

0-lepton channel The online selection relies on $E_{T}^{miss}$ triggers with thresholds that varied from 70 GeV to 110 GeV between the 2015 and 2017 data-taking periods. Their efficiency has been measured in $W$+jets and $t\bar{t}$ events in data, resulting in correction factors that are applied to the simulated events, ranging from 1.05 at the offline $E_{T}^{miss}$ threshold of 150 GeV to a negligible correction at an $E_{T}^{miss}$ above 200 GeV. A selection on the scalar sum of the transverse momenta $H_T$ of the jets removes a small part of the phase space where the trigger efficiency depends mildly on the number of jets in the event. Events with any loose lepton are vetoed. High $E_{T}^{miss}$ in multijet events typically arises from mismeasured jets in the calorimeters. Such events are efficiently removed by requiring events to pass a set of criteria on the angular separation of the $E_{T}^{miss}$, jets, and $E_{T}^{miss,\text{trk}}$ (the missing transverse momentum calculated using only tracks reconstructed in the ID and matched to the primary vertex).

1-lepton channel In the electron sub-channel, events are required to pass an OR of single electron triggers with identification and isolation criteria looser than those used in the offline analysis, and $p_T$ thresholds which start at 24 GeV in 2015 and increase to 26 GeV in 2016 and 2017. The muon sub-channel uses the same $E_{T}^{miss}$ triggers and correction factors as the 0-lepton channel, as these triggers effectively select on $p_T^V$ given that muons are not included in the online $E_{T}^{miss}$ calculation and they subsequently perform more efficiently than the single muon triggers in the analysis phase space. Events are required to have exactly one high-$p_T$ tight electron or muon. In the electron sub-channel, where multi-jet production is a significant background, an additional selection of $E_{T}^{miss} > 30$ GeV is applied. Events are categorised in the signal region (SR) or in a control region enriched in $W$+HF events ($W$+HF CR) based on selections on the invariant mass of the two $b$-tagged jets ($m_{bb}$), and on a reconstructed mass of a candidate leptonically decaying top quark ($m_{top}$). The latter is calculated as the invariant mass of the lepton, the reconstructed
Table 1: Summary of the event selection and categorisation in the 0-, 1- and 2-lepton channels.

<table>
<thead>
<tr>
<th>Selection</th>
<th>0-lepton</th>
<th>1-lepton</th>
<th>1-lepton</th>
<th>2-lepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger</td>
<td>$E_T^{miss}$</td>
<td>Single lepton</td>
<td>$E_T^{miss}$</td>
<td>Single lepton</td>
</tr>
<tr>
<td>Leptons</td>
<td>0 loose leptons with $p_T &gt; 7$ GeV</td>
<td>1 tight electron $p_T &gt; 27$ GeV</td>
<td>1 tight muon $p_T &gt; 25$ GeV</td>
<td>2 loose leptons with $p_T &gt; 7$ GeV $\geq 1$ lepton with $p_T &gt; 27$ GeV</td>
</tr>
<tr>
<td>$E_T^{miss}$</td>
<td>$&gt; 150$ GeV</td>
<td>$&gt; 30$ GeV</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$m_{\ell\ell}$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>$81$ GeV $&lt; m_{\ell\ell} &lt; 101$ GeV</td>
</tr>
<tr>
<td>Jets</td>
<td>Exactly 2 / Exactly 3 jets</td>
<td>Exactly 2 / ≥ 3 jets</td>
<td>Exactly 2 / ≥ 3 jets</td>
<td></td>
</tr>
<tr>
<td>Jet $p_T$</td>
<td>$&gt; 20$ GeV for $</td>
<td>\eta</td>
<td>&lt; 2.5$ and $&gt; 30$ GeV for $2.5 &lt;</td>
<td>\eta</td>
</tr>
<tr>
<td>$b$-jets</td>
<td>Exactly 2 $b$-tagged jets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leading $b$-tagged jet $p_T$</td>
<td>$&gt; 120$ (2 jets), $&gt; 150$ GeV (3 jets)</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>$H_T$</td>
<td>$&gt; 100$ (2 jets), $&gt; 150$ GeV (3 jets)</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>$\Delta \phi(E_T^{miss},bb)$</td>
<td>$&gt; 120^\circ$ (2 jets), $&gt; 30^\circ$ (3 jets)</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>$\Delta \phi(b_1,b_2)$</td>
<td>$&lt; 140^\circ$</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>$\Delta \phi(E_T^{miss},E_{T,\text{trk}})$</td>
<td>$&lt; 90^\circ$</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>$p_T^{V}$ regions</td>
<td>$&gt; 150$ GeV</td>
<td>–</td>
<td>75 GeV $&lt; p_T^{V} &lt; 150$ GeV, $&gt; 150$ GeV</td>
<td></td>
</tr>
<tr>
<td>Signal regions</td>
<td>–</td>
<td>$m_{bb} \geq 75$ GeV or $m_{top} \leq 225$ GeV</td>
<td>Same-flavour leptons Opposite-sign charges ($\mu\mu$ sub-channel)</td>
<td></td>
</tr>
<tr>
<td>Control regions</td>
<td>–</td>
<td>$m_{bb} &lt; 75$ GeV and $m_{top} &gt; 225$ GeV</td>
<td>Different-flavour leptons Opposite-sign charges</td>
<td></td>
</tr>
</tbody>
</table>
neutrino$^4$ and the $b$-tagged jet that yields the lowest value. The resulting purity of the $W$+HF region is around 75%.

**2-lepton channel** The online selection in the electron sub-channel is the same as in the 1-lepton channel. In the muon sub-channel, a similar OR of single-muon triggers is used, with $p_T$ thresholds increasing with luminosity and ranging from 20 GeV to 26 GeV. Events must have exactly two loose leptons, one of which must have $p_T > 27$ GeV, and the invariant mass of the lepton pair must be compatible with that of the $Z$ boson. Events with leptons of same flavour enter the signal region, while events with one muon and one electron define an $e\mu$ control region which is over 99% pure in $t\bar{t}$ and single top-quark events.

### 3.3 Multivariate analysis

Boosted decision trees (BDT) are trained in eight signal regions, corresponding to two jet categories for the three lepton channels in the high $p_T^{V}$ region, in addition to the two jet categories for the 2-lepton low $p_T^{V}$ region, to yield the final discriminant variables used in the analysis. Two sets of BDTs are constructed with the same input variables and parameters. The nominal one (BDTVH) is designed to separate Higgs boson events from the sum of expected backgrounds, while the second one (BDTVZ) is used to validate the analysis by the extraction of the diboson $VZ$, $Z \rightarrow b\bar{b}$ process against the sum of all other background processes.

The same input variables and BDT settings as those detailed in Ref. [28] are used, with the exception of the 2-lepton channel where the $E_T^{\text{miss}}$ has been replaced with $E_T^{\text{miss}}/\sqrt{S_T}$ (where $S_T$ is the scalar sum of transverse momenta of the charged leptons and jets in the event). Eight to thirteen input variables describing the kinematics of the events are used depending on the regions, of which $m_{bb}$, $p_T^V$ and $\Delta R(b_1, b_2)$ are the most discriminant.

### 3.4 Dijet-mass analysis

A validation of the main multivariate analysis is performed by using the invariant mass of the $b$-tagged jets as the discriminant variable. Additional selections displayed in Table 2 increase the purity of the signal regions and are necessary to improve the sensitivity of this result.

The high-$p_T^{V}$ category is split in two regions $150$ GeV $< p_T^{V} < 200$ GeV and $p_T^{V} > 200$ GeV, with further requirements placed upon the separation of the two $b$-tagged jets in the $(\eta, \phi)$ plane. Selections on the transverse mass of the $W$ boson and on $E_T^{\text{miss}}/\sqrt{S_T}$ reduce the $t\bar{t}$ background in the 1- and 2-lepton channels respectively.

In the 1-lepton channel, the $m_{bb}$ distribution constrains the $W$+HF background sufficiently for the $W$+HF CR and SR to be merged.

---

$^4$ Reconstructed from the $E_T^{\text{miss}}$, with the $W$ mass constraint used to calculate the missing longitudinal component of the neutrino four-vector.
Table 2: Summary of the event selection criteria in the 0-, 1- and 2-lepton channels for the dijet-mass analysis, applied in addition to those described in Table 1 for the multivariate analysis.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Selection</th>
<th>0-lepton</th>
<th>1-lepton</th>
<th>2-lepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m^W_T$</td>
<td>-</td>
<td>&lt; 120 GeV</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$E^\text{miss}_T / \sqrt{S_T}$</td>
<td>-</td>
<td>-</td>
<td>&lt; 3.5 $\sqrt{GeV}$</td>
<td></td>
</tr>
</tbody>
</table>

$p_T^V$ regions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta R(b_1, b_2)$</td>
<td>&lt;3.0</td>
<td>&lt;1.8</td>
<td>&lt;1.2</td>
</tr>
</tbody>
</table>

4 Data set, simulated event samples and multi-jet estimate

The data used in this analysis were collected at a centre-of-mass energy of 13 TeV during the 2015–2017 running periods, and correspond to a total integrated luminosity of $79.8 \pm 1.7$ fb$^{-1}$ [60]. They were collected using missing transverse momentum ($E^\text{miss}_T$) triggers for the 0- and 1-lepton channels and single-lepton triggers for the 1- and 2-lepton channels. Events are selected for analysis only if they are of good quality and if all the relevant detector components are known to be in good operating condition. In the combined dataset, the recorded events contain an average of 32 inelastic pp collisions.

Monte Carlo (MC) simulated events are used to model the SM background and $VH$, $H \rightarrow b\bar{b}$ signal processes. All simulated processes are normalised using the most accurate theoretical predictions currently available for their cross sections. All samples of simulated events were passed through the ATLAS detector simulation [61] based on GEANT 4 [62] and were reconstructed with the standard ATLAS reconstruction software. The effects of pile-up from multiple interactions in the same and nearby bunch crossings were modelled by overlaying minimum-bias events, simulated using the soft QCD processes of PYTHIA 8.186 [63] with the A2 [64] set of tuned parameters (tune) and MSTW2008LO [65] parton distribution functions (PDF). For all samples of simulated events, except for those generated using SHERPA [66], the EVGEN v1.2.0 program [67] was used to describe the decays of bottom and charm hadrons. A summary of all the generators used for the simulation of the signal and background processes is shown in Table 3. Samples produced with alternative generators are used to estimate systematic uncertainties in the event modelling, as described in Section 5.

The MC samples used to model background processes with $W$ or $Z$ boson decays into leptons$^5$, are defined as electroweak (EW) backgrounds. The multi-jet background from strong interactions (QCD) is instead estimated in all three channels using data-driven methods. In both the 0 and 2-lepton channels, the multi-jet contribution is estimated from template fits to data, using the simulated samples to model the EW backgrounds and a functional form to model the multi-jet background. The template fit is performed using a variable which provides significant discrimination between the multi-jet and EW processes, with any selection on that variable removed. In the 0-lepton channel, $\min[\Delta\phi(E^\text{miss}_T, \text{jets})]$ is used and in the 2-lepton channel, the dilepton mass distribution for the case where the two lepton candidates have the

$^5$ This encompasses the $t\bar{t}$ and single-top processes, as well as the diboson and vector-boson plus jet backgrounds.
Table 3: The generators used for the simulation of the signal and background processes. If not specified, the order of the cross-section calculation refers to the expansion in the strong coupling constant ($\alpha_S$). The acronyms ME, PS and UE stand for matrix element, parton shower and underlying event, respectively. (*) The events were generated using the first PDF in the NNPDF3.0NLO set and subsequently reweighted to PDF4LHC15NLO set [68] using the internal algorithm in Powheg-Box v2. (†) The NNLO(QCD)+NLO(EW) cross-section calculation for the $pp \to ZH$ process already includes the $gg \to ZH$ contribution. The $q\bar{q} \to ZH$ process is normalised using the cross-section for the $pp \to ZH$ process, after subtracting the $gg \to ZH$ contribution. An additional scale factor is applied to the $q\bar{q} \to VH$ processes as a function of the transverse momentum of the vector boson, to account for electroweak (EW) corrections at NLO. This makes use of the $VH$ differential cross-section computed with Hawk [69, 70].

<table>
<thead>
<tr>
<th>Process</th>
<th>ME generator</th>
<th>ME PDF</th>
<th>PS and UE</th>
<th>UE model</th>
<th>Cross-section order</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q\bar{q} \to VH$</td>
<td>Powheg-Box v2 [71] + GoSam [74] + MiNLO [75, 76]</td>
<td>NNPDF3.0NLO** [72]</td>
<td>Pythia8.212 [63]</td>
<td>AZNLO [73]</td>
<td>NNLO(QCD)+NLO(EW) [77–83]</td>
</tr>
<tr>
<td>$q\bar{q} \to ZH$</td>
<td>Powheg-Box v2 + GoSam + MiNLO</td>
<td>NNPDF3.0NLO*</td>
<td>Pythia8.212</td>
<td>AZNLO</td>
<td>NNLO(QCD)†+NLO(EW)</td>
</tr>
<tr>
<td>$gg \to ZH$</td>
<td>Powheg-Box v2 + GoSam &amp; MiNLO</td>
<td>NNPDF3.0NLO*</td>
<td>Pythia8.212</td>
<td>AZNLO</td>
<td>NLO*+NLL [84–88]</td>
</tr>
</tbody>
</table>

same charge, assuming the multi-jet contribution is symmetric for opposite and same-charge events, is used. In both cases, it is found that the multi-jet contribution is sufficiently small that it can be neglected in the global likelihood fit without having any impact on the signal extraction.

The multi-jet background is found to be non-negligible in the 1-lepton channel and is estimated separately in the electron and muon sub-channels, and in the 2- and 3-jet categories. A template fit to the transverse mass ($m_T^W$) distributions of the $W$ boson candidate is performed, which offers the clearest discrimination between the multi-jet and EW processes, in order to extract the multi-jet yield. The template used for the multi-jet contribution is obtained from data in a control region after subtraction of the residual EW contribution, based on MC predictions, while the template for the EW contribution in the signal region is obtained directly from MC predictions. The control region is enriched in multi-jet events that are kinematically close to the corresponding signal region but not overlapping with it, and is defined by applying the nominal selection but inverting the tight isolation requirement. To increase the statistical precision of the data-driven estimate, the number of required $b$-tags is reduced from two to one in the multi-jet enriched control region. The template fit applied in the signal region determines the normalisation of the multi-jet contribution, while the shape of the BDT discriminant (or of other relevant observables) is obtained analogously to the $m_T^W$ template. Both the normalisation and shape derived for the BDT discriminant are then used in the global likelihood fit. The multi-jet contribution in the 2-jet region is found to be 1.9% (2.8%) of the total background contribution in the electron (muon) sub-channel, while
in the 3-jet region it is found to be 0.2% (0.4%). These estimates are subject to sizeable systematic uncertainties, which are described in Section 5.

5 Systematic uncertainties

The sources of systematic uncertainty can be broadly divided into four groups: those of experimental nature, those related to the modelling of the simulated backgrounds, those related to the multi-jet background estimation, and those associated with the Higgs boson signal simulation. The uncertainties are all derived closely following the methodology outlined in Ref. [28], which is briefly summarised below.

5.1 Experimental uncertainties

The dominant experimental uncertainties originate from the $b$-tagging simulation-to-data efficiency correction factors, from the jet energy scale corrections and the modelling of the jet energy resolution. The $b$-tagging simulation-to-data efficiency correction factors are derived [102–104] separately for $b$-jets, $c$-jets and light-flavour jets. All three correction factors have uncertainties estimated from multiple measurements, which are decomposed into uncorrelated components which are then treated independently, resulting in three uncertainties for $b$-jets and $c$-jets, and five for light-flavour jets. The approximate size of the uncertainty in the tagging efficiency is 2% for $b$-jets, 10% for $c$-jets and 40% for light jets. Additional uncertainties are considered in the extrapolation of the $b$-jet efficiency calibration to jets with $p_T > 300$ GeV and in the misidentification of hadronically decaying $\tau$-leptons as $b$-jets. The uncertainties in the jet energy scale and resolution are based on their respective measurements in data [57, 105]. The many sources of uncertainty in the jet energy scale correction are decomposed into 23 uncorrelated components which are treated as independent. An additional specific uncertainty is considered that affects the energy calibration of $b$- and $c$-jets.

Uncertainties in the reconstruction, identification, isolation and trigger efficiencies of muons [48] and electrons [46], along with the uncertainty in their energy scale and resolution, are estimated based upon 13 TeV data. These are found to have only a small impact on the result. The uncertainties in the energy scale and resolution of the jets and leptons are propagated to the calculation of $E_T^{\text{miss}}$, which also has additional uncertainties from the scale, resolution and reconstruction efficiency of the tracks used to define the soft term [58], along with the modelling of the underlying event. An uncertainty is assigned to the simulation-to-data $E_T^{\text{miss}}$ trigger scale factors to account for the statistical uncertainty in the measured scale factors and differences between the scale factors determined from $W +$ jets, $Z +$ jets and $t\bar{t}$ events. The uncertainty in the luminosity is 2.0% for the combined dataset [60]. The average number of interactions per bunch crossing is rescaled by 3% to improve the agreement between simulation and data, based on the measurement of the visible cross section in minimum-bias events [106], and an uncertainty, as large as the correction, is included.

5.2 Simulated sample uncertainties

Modelling uncertainties are derived for the simulated samples and broadly cover three areas: normalisations, acceptance differences that affect the relative normalisations between analysis regions with a common normalisation, and the differential distributions of the most important kinematic variables.
The overall normalisations and associated uncertainties for the background processes are taken from the currently most accurate calculations as detailed in Table 3, apart from those backgrounds whose normalisations are unconstrained (floated) in the global likelihood fit. Additional systematic uncertainties on the acceptance differences and shapes are derived either from particle-level comparisons between nominal and alternative samples, or from comparisons to data in control regions. The particle-level comparisons are cross-checked with detector-level simulations whenever these are available, and good agreement is found. The alternative samples are either produced by alternative generators or by altering parameters of the nominal generator. When acceptance uncertainties are estimated all the nominal and alternative samples are normalised using the same production cross section. Shape uncertainties are considered in each of the analysis regions separately, with the samples scaled to have the same normalisation in each region. In this case, the uncertainty is taken from the alternative sample which has the largest shape difference compared to the nominal sample. Shape uncertainties are only derived for the $m_{bb}$ and $p_T^V$ variables, as it was found that it is sufficient to only consider the changes induced in these variables by an alternative sample to cover the overall shape variation of the BDT$_{VH}$ discriminant. Full details are provided in Ref. [28].

5.2.1 Background uncertainties

The systematic uncertainties affecting the modelling of the background samples are summarised in Tables 4 and 5 and key details on the treatment of the backgrounds are detailed below.

**V + jets production** The $V + \text{jets}$ backgrounds are subdivided into three different components based upon the jet flavour labels of the two $b$-tagged jets in the event. The main background contributions ($V + bb$, $V + bc$, $V + bl$ and $V + cc$) are jointly considered as the $V + \text{HF}$ background. Their overall normalisation, separately in the 2- and 3-jet categories, is free to float in the global likelihood fit. The remaining flavour components, $V + cl$ and $V + ll$, make up less than $\sim 1\%$ of the background in each analysis region, so only uncertainties in the normalisation of these backgrounds are included. Acceptance uncertainties are estimated for the relative normalisations of the different regions that share a common floating normalisation parameter. In the case of the $W + \text{HF}$ background, this includes the uncertainties in the ratio of the event yield in the 0-lepton channel to that in the 1-lepton channel and, in the 1-lepton channel, in the ratio of the event yield in the $W + \text{HF}$ control region to that in the signal region. For the $Z + \text{HF}$ background, there is an uncertainty in the ratio of the event yield in the 0-lepton channel to that in the 2-lepton channel. Uncertainties are also estimated in the relative normalisation of the four heavy-flavour components that make up the $V + \text{HF}$ background. These are taken as uncertainties in the $bc$, $cc$ and $bl$ yields compared to the dominant $bb$ yield and are estimated separately in each channel in a manner similar to the acceptance systematic uncertainties. Uncertainties are also derived on the shapes of the $m_{bb}$ and $p_T^V$ distributions, which are evaluated for $W + \text{HF}$ from comparing to alternative samples and for $Z + \text{HF}$ from comparisons to enriched $m_{bb}$ side-bands in data.

**$t\bar{t}$ production** Due to the significantly different regions of phase space probed, the $t\bar{t}$ background in the 0- and 1-lepton channels (jointly referred to as 0+1 lepton in the following) is considered independently from the $t\bar{t}$ background in the 2-lepton channel; different overall floating normalisation factors are considered, and acceptance uncertainties are derived separately and taken as uncorrelated between the 0+1 and 2-lepton channels. For the 0+1 lepton channels, uncertainties are considered in the normalisation ratios of the 3-jet and 2-jet regions, of the $W + \text{HF}$ control region and signal region, and of the 1-lepton and 0-lepton channels. For the 2-lepton channel, the normalisations in the 2- and 3-jet regions are both left floating, and are effectively determined in their respective $e\mu$ control regions. Uncertainties in the shapes of the
Table 4: Summary of the systematic uncertainties in the background modelling for $Z + \text{jets}$, $W + \text{jets}$, $\bar{t}t$, single top-quark and multi-jet production. An "S" symbol is used when only a shape uncertainty is assessed. The regions for which the normalisations float independently are listed in brackets. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed.

<table>
<thead>
<tr>
<th>Z + jets</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z + ll$ normalisation</td>
<td>18%</td>
</tr>
<tr>
<td>$Z + cl$ normalisation</td>
<td>23%</td>
</tr>
<tr>
<td>$Z + \text{HF normalisation}$</td>
<td>Floating (2-jet, 3-jet)</td>
</tr>
<tr>
<td>$Z + bc$-to-$Z + bb$ ratio</td>
<td>30 – 40%</td>
</tr>
<tr>
<td>$Z + cc$-to-$Z + bb$ ratio</td>
<td>13 – 15%</td>
</tr>
<tr>
<td>$Z + bl$-to-$Z + bb$ ratio</td>
<td>20 – 25%</td>
</tr>
<tr>
<td>0-to-2 lepton ratio</td>
<td>7%</td>
</tr>
<tr>
<td>$m_{bb}, p_T^{V}$</td>
<td>S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>W + jets</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$W + ll$ normalisation</td>
<td>32%</td>
</tr>
<tr>
<td>$W + cl$ normalisation</td>
<td>37%</td>
</tr>
<tr>
<td>$W + \text{HF normalisation}$</td>
<td>Floating (2-jet, 3-jet)</td>
</tr>
<tr>
<td>$W + bl$-to-$W + bb$ ratio</td>
<td>26% (0-lepton) and 23% (1-lepton)</td>
</tr>
<tr>
<td>$W + bc$-to-$W + bb$ ratio</td>
<td>15% (0-lepton) and 30% (1-lepton)</td>
</tr>
<tr>
<td>$W + cc$-to-$W + bb$ ratio</td>
<td>10% (0-lepton) and 30% (1-lepton)</td>
</tr>
<tr>
<td>0-to-1 lepton ratio</td>
<td>5%</td>
</tr>
<tr>
<td>$W + \text{HF CR to SR ratio}$</td>
<td>10% (1-lepton)</td>
</tr>
<tr>
<td>$m_{bb}, p_T^{V}$</td>
<td>S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\bar{t}t$ (all are uncorrelated between the 0+1 and 2-lepton channels)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{t}t$ normalisation</td>
<td>Floating (0+1 lepton, 2-lepton 2-jet, 2-lepton 3-jet)</td>
</tr>
<tr>
<td>0-to-1 lepton ratio</td>
<td>8%</td>
</tr>
<tr>
<td>2-to-3-jet ratio</td>
<td>9% (0+1 lepton only)</td>
</tr>
<tr>
<td>$W + \text{HF CR to SR ratio}$</td>
<td>25%</td>
</tr>
<tr>
<td>$m_{bb}, p_T^{V}$</td>
<td>S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Single top-quark</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross section</td>
<td>4.6% (s-channel), 4.4% (t-channel), 6.2% ($Wt$)</td>
</tr>
<tr>
<td>Acceptance 2-jet</td>
<td>17% (t-channel), 55% ($Wt \rightarrow bb$), 24% ($Wt \rightarrow oth$)</td>
</tr>
<tr>
<td>Acceptance 3-jet</td>
<td>20% (t-channel), 51% ($Wt \rightarrow bb$), 21% ($Wt \rightarrow oth$)</td>
</tr>
<tr>
<td>$m_{bb}, p_T^{V}$</td>
<td>S (t-channel, $Wt \rightarrow bb$, $Wt \rightarrow oth$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multi-jet (1-lepton)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalisation</td>
<td>60 – 100% (2-jet), 90 – 140% (3-jet)</td>
</tr>
<tr>
<td>BDT template</td>
<td>S</td>
</tr>
</tbody>
</table>
Table 5: Summary of the systematic uncertainties in the background modelling for diboson production. “PS/UE” indicates parton shower / underlying event. An “S” symbol is used when only a shape uncertainty is assessed. When determining the $(W/Z)Z$ diboson production signal strength, as the normalisations are unconstrained, the normalisation uncertainties are removed. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed.

| ZZ |  
|---|---|
| Normalisation | 20% |
| 0-to-2 lepton ratio | 6% |
| Acceptance from scale variations (var.) | 10 – 18% |
| Acceptance from PS/UE var. for 2 or more jets | 6% |
| Acceptance from PS/UE var. for 3 jets | 7% (0-lepton), 3% (2-lepton) |
| $m_{bb}, p_T^V$, from scale var. | S (correlated with WZ uncertainties) |
| $m_{bb}, p_T^V$, from PS/UE var. | S (correlated with WZ uncertainties) |
| $m_{bb}$, from matrix-element var. | S (correlated with WZ uncertainties) |

| WZ |  
|---|---|
| Normalisation | 26% |
| 0-to-1 lepton ratio | 11% |
| Acceptance from scale var. | 13 – 21% |
| Acceptance from PS/UE var. for 2 or more jets | 4% |
| Acceptance from PS/UE var. for 3 jets | 11% |
| $m_{bb}, p_T^V$, from scale var. | S (correlated with ZZ uncertainties) |
| $m_{bb}, p_T^V$, from PS/UE var. | S (correlated with ZZ uncertainties) |
| $m_{bb}$, from matrix-element var. | S (correlated with ZZ uncertainties) |

| WW |  
|---|---|
| Normalisation | 25% |

$p_T^V$ and $m_{bb}$ distributions are estimated in the 0+1 and 2-lepton channels separately from comparisons to alternative samples.

**Single top-quark production** In the $Wt$ and $t$-channels, uncertainties are derived in the normalisation, acceptance and shapes of the $m_{bb}$ and $p_T^V$ distributions. For the $Wt$-channel, the modelling uncertainties are evaluated based on the flavour of the two $b$-tagged jets, due to the different regions of phase space being probed when there are two $b$-jets ($bb$) present compared to events where there are one or fewer $b$-jets present ($otb$). The $s$-channel only has a normalisation uncertainty derived as its contribution is negligible overall.

**Diboson production** The diboson backgrounds are composed of three distinct processes, $WZ$, $WW$ and $ZZ$ production. Given the small contribution from $WW$ production ($<0.1\%$ of the total background) only a normalisation uncertainty is assigned. The more important contributions from the $WZ$ and $ZZ$ backgrounds have uncertainties derived for the overall normalisation, the relative acceptance between regions and for the $m_{bb}$ and $p_T^V$ shapes. These are derived following the procedure described in Ref. [28] and are outlined in Table 5.

### 5.2.2 Signal uncertainties

The systematic uncertainties that affect the modelling of the signal are summarised in Table 6; they are derived following the procedure outlined in Ref. [28], with uncertainties in the calculations of the $VH$
production cross sections and the $H \rightarrow b\bar{b}$ branching ratio\(^\text{6}\) assigned following the recommendations of the LHC Higgs Cross Section working group [23, 87, 88, 107, 108].

Table 6: Summary of the systematic uncertainties in the signal modelling. “PS/UE” indicates parton shower / underlying event. An “S” symbol is used when only a shape uncertainty is assessed. Where the size of an acceptance systematic uncertainty varies between regions, a range is displayed.

<table>
<thead>
<tr>
<th>Signal</th>
<th>Cross section (scale)</th>
<th>0.7% ($qq$), 27% ($gg$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H → b\bar{b} branching ratio</td>
<td>Cross section (PDF)</td>
<td>1.9% ($qq \rightarrow WH$), 1.6% ($qq \rightarrow ZH$), 5% ($gg$)</td>
</tr>
<tr>
<td>Acceptance from scale variations (var.)</td>
<td>2.5 – 8.8%</td>
<td></td>
</tr>
<tr>
<td>Acceptance from PS/UE var. for 2 or more jets</td>
<td>2.9 – 6.2% (depending on lepton channel)</td>
<td></td>
</tr>
<tr>
<td>Acceptance from PS/UE var. for 3 jets</td>
<td>1.8 – 11%</td>
<td></td>
</tr>
<tr>
<td>Acceptance from PDF+(\alpha_S) var.</td>
<td>0.5 – 1.3%</td>
<td></td>
</tr>
<tr>
<td>(m_{bb}, p_T^V), from scale var.</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>(m_{bb}, p_T^V), from PS/UE var.</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>(m_{bb}, p_T^V), from PDF+(\alpha_S) var.</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>(p_T^V) from NLO EW correction</td>
<td>S</td>
<td></td>
</tr>
</tbody>
</table>

### 5.3 Multi-jet background uncertainties

Systematic uncertainties can have an impact on the data-driven multi-jet estimate used in the 1-lepton channel in two ways: either changing the \(m_T^W\) distributions used in the multi-jet template fits, therefore impacting the extracted multi-jet normalisations, or directly changing the multi-jet BDT distributions used in the global likelihood fit. Several uncertainties are considered, uncorrelated between the electron and muon sub-channels. The respective variations are added in quadrature for the normalisations, or considered as separate shape uncertainties. The variations are obtained by changing the definition of the multi-jet control region (more stringent isolation requirements, a different single-electron trigger to probe a potential trigger bias in the isolation requirements), and varying the normalisation of the contamination from the top (t\(\bar{t}\) and \(Wt\)) and V + jets processes in the multi-jet control region. In addition, the following systematic uncertainties only have an impact on the multi-jet normalisation: use of another discriminant variable instead of \(m_T^W\) for the multi-jet template fit (the azimuthal separation between the directions of the missing transverse momentum and of the vectorial sum of the momenta of the two or three jets) and, for the electron sub-channel only, the inclusion of the \(E_T^{\text{miss}} < 30\) GeV region, which significantly enhances the multi-jet contribution in the template fit.

### 6 Statistical analysis

The statistical procedure is based on a likelihood function \(L(\mu, \theta)\), constructed as the product of Poisson probability terms over the bins of the input distributions. The parameter of interest \(\mu\) is the signal strength, that multiplies the SM Higgs boson production cross section times the branching ratio into $b\bar{b}$.

\(^{6}\) Systematic uncertainties fully degenerate with the signal strength (acting on the overall global normalisation of the signal, such as branching ratio and cross-section uncertainties) do not affect the calculation of the significance with respect to the background only prediction.
Table 7: Factors applied to the nominal normalisations of the $t\bar{t}$, $W + HF$ and $Z + HF$ backgrounds, as obtained from the global fit to the 13 TeV data for the nominal multivariate analysis, used to extract the Higgs boson signal. The errors include the statistical and systematic uncertainties.

<table>
<thead>
<tr>
<th>Process</th>
<th>Normalisation factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$ 0- and 1-lepton</td>
<td>0.98 ± 0.08</td>
</tr>
<tr>
<td>$t\bar{t}$ 2-lepton 2-jet</td>
<td>1.06 ± 0.09</td>
</tr>
<tr>
<td>$t\bar{t}$ 2-lepton 3-jet</td>
<td>0.95 ± 0.06</td>
</tr>
<tr>
<td>$W + HF$ 2-jet</td>
<td>1.19 ± 0.12</td>
</tr>
<tr>
<td>$W + HF$ 3-jet</td>
<td>1.05 ± 0.12</td>
</tr>
<tr>
<td>$Z + HF$ 2-jet</td>
<td>1.37 ± 0.11</td>
</tr>
<tr>
<td>$Z + HF$ 3-jet</td>
<td>1.09 ± 0.09</td>
</tr>
</tbody>
</table>

Systematic uncertainties enter the likelihood as nuisance parameters (NP), $\theta$, of three main types. Most of the uncertainties discussed in Section 5 are constrained with Gaussian or log-normal probability density functions. The normalisations of the largest backgrounds, $t\bar{t}$, $W + HF$ and $Z + HF$, can be reliably determined by the fit, therefore they are left unconstrained in the likelihood. The uncertainty coming from the limited number of events in the simulated samples used for the background predictions are included using the Beeston-Barlow technique [109]. As detailed in Ref. [27], systematic variations which are subject to large statistical fluctuations are smoothed, and systematic uncertainties that have a negligible impact on the final results are pruned away region-by-region.

The measurement of the compatibility of the background-only hypothesis with the observed data is done using the $q_0$ test statistic constructed from the profile-likelihood ratio with the asymptotic approximation, as defined in Ref. [110].

### 6.1 Multivariate analysis

As discussed in Section 3, the global fit comprises eight signal regions, defined as the 2- and 3-jet regions in the high-$p_T$ category for the three channels, and the medium-$p_T$ category in the 2-lepton channel. The BDT$_{VH}$ multivariate discriminant is used in these regions. The event yields are used in the two $W+HF$ control regions of the 1-lepton channel. In the four $e\mu$ control regions from the 2-lepton channel, the $m_{hh}$ distributions are input to the fit, except for the high-$p_T$ category of the 2-jet region, where the event yield is used. The post-fit normalisation factors of the unconstrained backgrounds in the global fit to the 13 TeV data are shown in Table 7.

The effect of systematic uncertainties on the measurement of the signal strength are displayed in Table 8. The impact of a category of systematic uncertainties is defined as the quadratic difference of the uncertainty on $\mu$, computed first when all NPs are fitted, then when the NPs of the category are fixed to their best-fit values. The total statistical uncertainty is defined as the uncertainty on $\mu$ when all the NPs are fixed to their best-fit values. The total systematic uncertainty is then defined as the quadratic difference of the total uncertainty on $\mu$ and the total statistical uncertainty. As shown in the table, the systematic uncertainties for the modelling of the signal play a dominant role, followed by the uncertainty due to the limited size of the simulated samples, the modelling of the backgrounds and the $b$-tagging uncertainty.
Table 8: Breakdown of the contributions to the uncertainties in $\mu$. The sum in quadrature of the systematic uncertainties attached to the categories differs from the total systematic uncertainty due to correlations.

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>$\sigma_\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>0.259</td>
</tr>
<tr>
<td>Statistical</td>
<td>0.161</td>
</tr>
<tr>
<td>Systematic</td>
<td>0.203</td>
</tr>
</tbody>
</table>

**Experimental uncertainties**

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>$\sigma_\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jets</td>
<td>0.035</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$</td>
<td>0.014</td>
</tr>
<tr>
<td>Leptons</td>
<td>0.009</td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>0.061</td>
</tr>
<tr>
<td>$c$-jets</td>
<td>0.042</td>
</tr>
<tr>
<td>light jets</td>
<td>0.009</td>
</tr>
<tr>
<td>extrapolation</td>
<td>0.008</td>
</tr>
<tr>
<td>Pile-up</td>
<td>0.007</td>
</tr>
<tr>
<td>Luminosity</td>
<td>0.023</td>
</tr>
</tbody>
</table>

**Theoretical and modelling uncertainties**

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>$\sigma_\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>0.094</td>
</tr>
<tr>
<td>Floating normalisations</td>
<td>0.035</td>
</tr>
<tr>
<td>$Z$ + jets</td>
<td>0.055</td>
</tr>
<tr>
<td>$W$ + jets</td>
<td>0.060</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>0.050</td>
</tr>
<tr>
<td>Single top quark</td>
<td>0.028</td>
</tr>
<tr>
<td>Diboson</td>
<td>0.054</td>
</tr>
<tr>
<td>Multijet</td>
<td>0.005</td>
</tr>
<tr>
<td>MC statistical</td>
<td>0.070</td>
</tr>
</tbody>
</table>

### 6.2 Dijet-mass analysis

In the dijet-mass analysis, the number of signal regions is increased to fourteen as a consequence of splitting the event categories with $p_T^V > 150$ GeV in two, while the $W$+HF CR are merged into the corresponding SR, as outlined in Section 3.4. The $m_{bb}$ distributions are input to the fit in all categories, except for the 2-jet medium- and high-$p_T^V$ categories for the 2-lepton $e\mu$ control region, where the event yield is used.

### 6.3 Diboson analysis

In the diboson analysis, a measurement of the signal strength of the $ZZ$ and $WZ$ processes is conducted to validate the main multivariate analysis. The setup differs from the global fit only by the use of the $\text{BDT}_{VZ}$ distributions as inputs, instead of $\text{BDT}_{VH}$. The parameter of interest $\mu_{VZ}$ is the signal strength.
of the combined $WZ$ and $ZZ$ diboson processes, and the SM Higgs boson is included as a background process normalised to the predicted SM cross section with an uncertainty of 50%.

6.4 Run 1, $H \to b\bar{b}$ and $VH$ combinations

The statistical analysis of the 13 TeV data is combined with the results of the data recorded at 7 TeV and 8 TeV [27]. Detailed studies were reported in Ref. [28] on the impact of the correlation of systematic uncertainties between the two analyses. In most cases, the impact of correlations was found to be negligible. Only a $b$-jet-specific jet energy scale, and theory uncertainties on the Higgs boson signal (overall cross section, branching ratio and $p_T$-dependent NLO EW corrections) are correlated across the different datasets.

A second combination is performed with the results of the searches for the $H \to b\bar{b}$ decay in the $t\bar{t}H$ [33, 35] and VBF [30, 32] production modes carried out with the Run 1 and Run 2 data. As the analysis targeting the VBF production mode has a sizeable contribution of gluon fusion events, it is referred to as the VBF+ggF analysis in the following. Assuming the SM cross sections for the various production modes, the combination measures the branching ratio of the Higgs boson into $b$-quarks. The only NP correlated across the six analyses is the $H \to b\bar{b}$ branching ratio. A few other NPs are correlated across some of the analyses, following the studies conducted for the combinations of Run 1 results [15], of analyses on the $t\bar{t}H$ production mode [18], and of Run 2 results [111].

A third combination is also performed combining the Run 2 $VH$, $H \to b\bar{b}$ result with other results for the Higgs boson produced in the $VH$ production mode, but for the case of the Higgs boson decaying to diphotons [40] or via $ZZ^*$ to four leptons [39]. The combination is undertaken as outlined in Ref. [111]. Assuming the SM branching ratios for the $ZZ^*$, diphoton and $b\bar{b}$ decays, this combination measures the signal strength of the $VH$ production mode.

7 Results

7.1 Results of the SM Higgs boson search at $\sqrt{s} = 13$ TeV

Figure 1 shows the BDT output distributions in the most sensitive (high-$p_T$) categories. The background prediction in all post-fit distributions is obtained by normalising the backgrounds and setting the nuisance parameters according to the results of the signal extraction fit. The post-global likelihood fit signal and background yields are shown in Table 9 for all the analysis regions.

For the tested Higgs boson mass of 125 GeV, when all lepton channels are combined, the probability $p_0$ of obtaining from background alone a result at least as signal-like as the observation is $5.3 \cdot 10^{-7}$, whilst the expected value is $7.3 \cdot 10^{-6}$. The observation corresponds to an excess with a significance of 4.9 standard deviations, to be compared to an expectation of 4.3 standard deviations. The fitted value of the signal strength parameter is:

$$\mu_{bb}^{VH} = 1.16^{+0.27}_{-0.25} = 1.16 \pm 0.16 \text{(stat.)}^{+0.21}_{-0.19} \text{(syst.)}. $$
Table 9: The Higgs boson signal, background and data yields for each signal region category in each channel after the full selection of the multivariate analysis. The signal and background yields are normalised to the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties. An entry of “−” indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selection.

<table>
<thead>
<tr>
<th>Signal regions</th>
<th>0-lepton</th>
<th></th>
<th>1-lepton</th>
<th></th>
<th>2-lepton</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2-jet</td>
<td>3-jet</td>
<td>2-jet</td>
<td>3-jet</td>
<td>2-jet</td>
<td>3-jet</td>
</tr>
<tr>
<td>Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z + ll</td>
<td>17 ± 11</td>
<td>27 ± 18</td>
<td>1.5 ± 1.0</td>
<td>3.4 ± 2.3</td>
<td>13.7 ± 8.7</td>
<td>49 ± 32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z + cl</td>
<td>45 ± 18</td>
<td>76 ± 30</td>
<td>3.0 ± 1.2</td>
<td>6.9 ± 2.8</td>
<td>43 ± 17</td>
<td>170 ± 67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z + HF</td>
<td>4770 ± 140</td>
<td>5940 ± 300</td>
<td>179.5 ± 9.1</td>
<td>348 ± 21</td>
<td>7400 ± 120</td>
<td>14160 ± 220</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W + ll</td>
<td>20 ± 13</td>
<td>32 ± 22</td>
<td>31 ± 23</td>
<td>65 ± 48</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W + cl</td>
<td>43 ± 20</td>
<td>83 ± 38</td>
<td>139 ± 67</td>
<td>250 ± 120</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W + HF</td>
<td>1000 ± 87</td>
<td>1990 ± 200</td>
<td>2660 ± 270</td>
<td>5400 ± 670</td>
<td>1.8 ± 0.2</td>
<td>13.2 ± 1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single top quark</td>
<td>368 ± 53</td>
<td>1410 ± 210</td>
<td>2080 ± 290</td>
<td>9400 ± 1400</td>
<td>188 ± 89</td>
<td>440 ± 200</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tt</td>
<td>1333 ± 82</td>
<td>9150 ± 400</td>
<td>6600 ± 320</td>
<td>50200 ± 1400</td>
<td>3170 ± 100</td>
<td>8880 ± 220</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diboson</td>
<td>254 ± 49</td>
<td>318 ± 90</td>
<td>178 ± 47</td>
<td>330 ± 110</td>
<td>152 ± 32</td>
<td>355 ± 68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi-jet e sub-ch.</td>
<td>−</td>
<td>−</td>
<td>100 ± 100</td>
<td>41 ± 35</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi-jet µ sub-ch.</td>
<td>−</td>
<td>−</td>
<td>138 ± 92</td>
<td>260 ± 270</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Total bkg.</td>
<td>7851 ± 90</td>
<td>19020 ± 140</td>
<td>12110 ± 120</td>
<td>66230 ± 270</td>
<td>10964 ± 99</td>
<td>24070 ± 150</td>
</tr>
<tr>
<td>Signal (fit)</td>
<td>128 ± 28</td>
<td>128 ± 29</td>
<td>131 ± 30</td>
<td>125 ± 30</td>
<td>51 ± 11</td>
<td>86 ± 22</td>
</tr>
<tr>
<td>Data</td>
<td>8003</td>
<td>19143</td>
<td>12242</td>
<td>66348</td>
<td>11014</td>
<td>24197</td>
</tr>
</tbody>
</table>
The background contributions after the global likelihood fit are shown as filled histograms. The signal yield extracted from data ($\mu = 1.16$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal ($\mu = 1.16$) and background is shown in the lower panel. The BDT$_{VH}$ output distributions are shown with the binning used to define the global likelihood fit.
Figure 2 shows the data, background and signal yields, where final-discriminant bins in all regions are combined into bins of $\log(S/B)$, with $S$ being the fitted signal and $B$ the fitted background. The Higgs boson signal contribution is shown after rescaling the SM cross section according to the value of the signal strength parameter extracted from data ($\mu = 1.16$). In the lower panel, the pull of the data with respect to the background (the statistical significance of the difference between data and fitted background) is shown with statistical uncertainties only. The full line indicates the pull of the prediction for signal and background with respect to the background prediction.

Table 10 shows the signal strengths, $p_0$ and significance values for the combined global fit, and for a fit where the lepton channels each have their own signal strength parameter. The compatibility of the signal strength parameters measured in the three lepton channels is 80%.

A combined fit is also performed with floating signal strength parameters separately for the $WH$ and $ZH$ production processes. The results of this fit are shown in Figure 3. The $WH$ and $ZH$ production modes have observed (expected) significances of 2.5 (2.3) and 4.0 (3.5) standard deviations, respectively, with a linear correlation between the two signal strengths of $-1\%$.

The compatibility between fits differing only in their number of parameters of interest is evaluated in the asymptotics regime, where the difference between their minimum likelihoods follows a $\chi^2$ distribution whose number of degrees of freedom is equal to the difference between the number of parameters of interest.
Table 10: Measured signal strengths, expected and observed $p_0$ and significance values (in standard deviations) for the lepton channels and their combination using the 13 TeV dataset.

<table>
<thead>
<tr>
<th>Signal strength parameter</th>
<th>Signal strength $\mu_{bb}^{VH}$</th>
<th>$p_0$</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-lepton</td>
<td>1.04^{+0.34}_{-0.32}</td>
<td>9.5 $\cdot 10^{-4}$</td>
<td>5.1 $\cdot 10^{-4}$</td>
</tr>
<tr>
<td>1-lepton</td>
<td>1.09^{+0.46}_{-0.42}</td>
<td>8.7 $\cdot 10^{-3}$</td>
<td>4.9 $\cdot 10^{-3}$</td>
</tr>
<tr>
<td>2-lepton</td>
<td>1.38^{+0.46}_{-0.42}</td>
<td>4.0 $\cdot 10^{-3}$</td>
<td>3.3 $\cdot 10^{-4}$</td>
</tr>
<tr>
<td>$VH, H \to bb$ combination</td>
<td>1.16^{+0.27}_{-0.25}</td>
<td>7.3 $\cdot 10^{-6}$</td>
<td>5.3 $\cdot 10^{-7}$</td>
</tr>
</tbody>
</table>

Figure 3: The fitted values of the Higgs boson signal strength parameter $\mu_{bb}^{VH}$ for $m_H = 125$ GeV for the $WH$ and $ZH$ processes and their combination. The individual $\mu_{bb}^{VH}$ values for the ($W/Z$)$H$ processes are obtained from a simultaneous fit with the signal strength for each of the $WH$ and $ZH$ processes floating independently. The compatibility of the individual signal strengths is $84\%$. 
7.2 Results of the dijet-mass analysis

The $m_{bb}$ distribution for all channels and regions summed, weighted by their respective value of the ratio of fitted Higgs boson signal and background yields, and after subtraction of all backgrounds except for the $(W/Z)Z$ diboson processes, is shown in Figure 4. The data and the sum of expected signal and backgrounds are found to be in good agreement. For all channels combined the fitted value of the signal strength parameter is

$$\mu_{bb}^{VH} = 1.06^{+0.36}_{-0.33} = 1.06 \pm 0.20 \text{(stat.)}^{+0.30}_{-0.26} \text{(syst.)},$$

in good agreement with the result of the multivariate analysis. The observed excess has a significance of 3.6 standard deviations, in comparison to an expectation of 3.5 standard deviations. Good agreement is also found in the values of signal strength parameters in the individual channels for the dijet-mass analysis compared to those for the multivariate analysis.

Figure 4: The distribution of $m_{bb}$ in data after subtraction of all backgrounds except for the $WZ$ and $ZZ$ diboson processes, as obtained with the dijet-mass analysis. The contributions from all lepton channels, $p_T^V$ intervals and number-of-jets categories are summed weighted by $S/B$, with $S$ being the total fitted signal and $B$ the total fitted background in that region. The expected contribution of the associated $WH$ and $ZH$ production of a SM Higgs boson with $m_H = 125$ GeV is shown scaled by the measured combined signal strength ($\mu = 1.06$). The size of the combined statistical and systematic uncertainty for the fitted background is indicated by the hatched band.

7.3 Results of the diboson analysis

As a validation of the Higgs analysis, the measurement of $VZ$ production based on the multivariate analysis described in Section 6.3 returns a value of signal strength

$$\mu_{VZ}^{bb} = 1.20^{+0.20}_{-0.18} = 1.20 \pm 0.08 \text{(stat.)}^{+0.19}_{-0.16} \text{(syst.)},$$

in good agreement with the Standard Model prediction. The $VZ$ signal is observed with a significance of 9.6 standard deviations, to be compared to an expected significance of 8.7 standard deviations. Analogously
to the $VH$ signal, fits are also performed with separate signal strength parameters for the $WZ$ and $ZZ$ production modes, and the results are shown in Figure 5.

![Figure 5: The fitted values of the $VZ$ signal strength parameter $\mu_{bb}^{VZ}$ for the $WZ$ and $ZZ$ processes and their combination. The individual $\mu_{bb}^{VZ}$ values for the $(W/Z)Z$ processes are obtained from a simultaneous fit with the signal strength parameters for each of the $WZ$ and $ZZ$ processes floating independently. The compatibility of the individual signal strengths is 47%.](image)

7.4 Results of combination

7.4.1 Run-1 and Run-2 combination for $VH$, $H \rightarrow b\bar{b}$

The Run-2 analysis is first combined with the Run 1 $VH$, $H \rightarrow b\bar{b}$ result following the methodology described in Section 6.4. The observed $p_0$ value is $5.5 \cdot 10^{-7}$, corresponding to an excess with a significance of 4.9 standard deviations, compared to an expectation of 5.1 standard deviations. The measured signal strength is:

$$\mu_{VH}^{bb} = 0.98^{+0.22}_{-0.21} = 0.98 \pm 0.14(\text{stat.})^{+0.17}_{-0.16}(\text{syst.}).$$

Fits are also performed with the signal strength parameters floated independently for the $WH$ and $ZH$ production processes. The compatibility of the signal strengths for the $WH$ and $ZH$ production processes is 72%, and the results of this fit are shown in Figure 6.

7.4.2 Observation of $H \rightarrow b\bar{b}$ decays

The $VH$ result is further combined with results of the searches for the Standard Model Higgs boson decaying into a $b\bar{b}$ pair produced in association with a $t\bar{t}$ pair and in vector boson fusion for both Run 1
and Run 2, to perform a search for the $H \rightarrow b\bar{b}$ decay. For the tested Higgs boson mass of 125 GeV, and assuming the SM production cross sections, the observed significance for the $H \rightarrow b\bar{b}$ decay is of 5.4 standard deviations, to be compared to an expectation of 5.5 standard deviations. For all channels combined the fitted value of the signal strength parameter is

$$\mu_{VH \rightarrow bb} = 1.01 \pm 0.20 = 1.01 \pm 0.12(\text{stat.})^{+0.16}_{-0.16}(\text{syst.}).$$

Table 11 shows the significance values independently for the VBF+ggF, $t\bar{t}H$ and $VH$ channels in the combination of the Run-1 and Run-2 datasets, and for the combined global fit. The signal strengths obtained from a fit where individual signal strength parameters are fitted for the three production modes are displayed in Figure 7. Fits are also performed with the signal strength parameters floated independently for each of the production processes in both Run 1 and Run 2. The compatibility of the six individual measurements is 54%.

### 7.4.3 Observation of $VH$ production

The Run 2 $VH$, $H \rightarrow b\bar{b}$ result is further combined with other Run-2 searches for the Higgs boson produced in the $VH$ production mode, but decaying to either diphotons or to four leptons via $ZZ^*$ decays. For the tested Higgs boson mass of 125 GeV, and assuming the SM branching ratios in the three decay modes considered, the observed significance for $VH$ production is 5.3 standard deviations, to be compared to an expectation of 4.8 standard deviations.
Table 11: Expected and observed significance values (in standard deviations) for the $H \rightarrow b\bar{b}$ channels fitted independently and their combination using the 7 TeV, 8 TeV and 13 TeV dataset.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp.</td>
</tr>
<tr>
<td>VBF+ggF</td>
<td>0.9</td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>1.9</td>
</tr>
<tr>
<td>VH</td>
<td>5.1</td>
</tr>
<tr>
<td>$H \rightarrow b\bar{b}$ Combination</td>
<td>5.5</td>
</tr>
</tbody>
</table>

Figure 7: The fitted values of the Higgs boson signal strength parameter $\mu_{H\rightarrow bb}$ for $m_H = 125$ GeV separately for the $VH$, $t\bar{t}H$ and VBF+ggF analyses, along with their combination. The individual $\mu_{H\rightarrow bb}$ values for the different production modes are obtained from a simultaneous fit with the signal strength parameters for each of the processes floating independently. The compatibility of the individual signal strengths is 83%.
8 Conclusion

A search for the Standard Model Higgs boson decaying into a $b\bar{b}$ pair and produced in association with a $W$ or $Z$ boson is presented, using data collected by the ATLAS experiment in proton–proton collisions from Run 2 of the Large Hadron Collider. This dataset corresponds to an integrated luminosity of $79.8\,\text{fb}^{-1}$ collected at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. An excess over the expected background is observed, with a significance of 4.9 standard deviations compared to an expectation of 4.3. The measured signal strength with respect to the SM prediction for $m_H = 125\,\text{GeV}$ is found to be $\mu_{VH}^{bb} = 1.16 \pm 0.16^{+0.21}_{-0.19}\,\text{(syst.)}$. This result is combined with previous results based on the full Run-1 dataset collected at centre-of-mass energies of 7 TeV and 8 TeV. An excess over the expected Standard Model background is observed, with a significance of 4.9 standard deviations compared to an expectation of 5.1. The measured signal strength with respect to the SM expectation is found to be $\mu_{VH}^{bb} = 0.98 \pm 0.14^{+0.17}_{-0.15}\,\text{(syst.)}$. Results for the Standard Model Higgs boson decaying into a $b\bar{b}$ pair in the $VH$, $t\bar{t}H$ and VBF+ggF production modes at centre-of-mass energies of 7 TeV, 8 TeV and 13 TeV are also combined, assuming the relative production cross sections of these processes to be as predicted by the Standard Model. An excess over the expected Standard Model background is observed, with a significance of 5.4 standard deviations compared to an expectation of 5.5, providing an observation of the $H \rightarrow b\bar{b}$ decay mode. The measured signal strength with respect to the SM expectation is found to be $\mu_{H \rightarrow bb} = 1.01 \pm 0.12^{+0.17}_{-0.15}\,\text{(syst.)}$. Assuming the Standard Model production strength, the result is consistent with the value of the Yukawa coupling to bottom quarks in the Standard Model.

In addition, the Run 2 $VH$, $H \rightarrow b\bar{b}$ result is further combined with other Run 2 searches for the Higgs boson decaying to either four leptons (via $ZZ^*$) or diphotons in the $VH$ production mode, under the assumption that the relative branching fractions of the three decay modes are as predicted by the Standard Model. The result is an observed significance of 5.3 standard deviations, to be compared to an expectation of 4.8 standard deviations. This provides a direct observation of the Higgs boson being produced in association with a vector boson.
References


[34] ATLAS Collaboration, Search for the Standard Model Higgs boson decaying into \( b\bar{b} \) in \( pp \) collisions at \( \sqrt{s} = 8 \text{ TeV} \) with the ATLAS detector, JHEP 05 (2016) 160, arXiv: 1604.03812 [hep-ex].

[35] ATLAS Collaboration, Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a \( b\bar{b} \) pair in \( pp \) collisions at \( \sqrt{s} = 13 \text{ TeV} \) with the ATLAS detector, Phys. Rev. D (2017), arXiv: 1712.08895 [hep-ex].


[37] CMS Collaboration, Search for \( t\bar{t}H \) production in the \( H \rightarrow b\bar{b} \) decay channel with lepton-\( t\bar{t} \) decays in proton-proton collisions at \( \sqrt{s} = 13 \text{ TeV} \), (2018), arXiv: 1804.03682 [hep-ex].


[40] ATLAS Collaboration, Measurement of Higgs boson properties in the diphoton decay channel using 80 \( fb^{-1} \) of \( pp \) collision data at \( \sqrt{s} = 13 \text{ TeV} \) with the ATLAS detector, ATLAS-CONF-2018-028, 2018, url: http://cdsweb.cern.ch/record/2628771.


[79] G. Ferrera, M. Grazzini and F. Tramontano, 
Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO, 

[80] O. Brein, R. Harlander, M. Wiesemann and T. Zirke, 
Top-Quark Mediated Effects in Hadronic Higgs-Strahlung, 

[81] G. Ferrera, M. Grazzini and F. Tramontano, 
Higher-order QCD effects for associated WH production and decay at the LHC, 

[82] G. Ferrera, M. Grazzini and F. Tramontano, 
Associated ZH production at hadron colliders: the fully differential NNLO QCD calculation, 

[83] J. M. Campbell, R. K. Ellis and C. Williams, 
Associated production of a Higgs boson at NNLO, 

[84] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak and T. J. E. Zirke, 
Gluon-induced Higgs-strahlung at next-to-leading order QCD, 

[85] R. V. Harlander, S. Liebler and T. Zirke, 
Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-Doublet Model, 

[86] S. Alioli, P. Nason and G. Ridolfi, 
A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction, 


Appendix
Figure 8: The post-fit distributions for $E_{T}^{miss}$ (top left), $m_{W}$ (middle left), $m_{\ell\ell}$ (bottom left) and $m_{bb}$ (right) 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 normalised to the signal yield extracted from data ($\mu = 1.16$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The entries in overflow are included in the last bin. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.
Figure 9: Distributions of the $p_T^V$ for all 2-jet signal and control regions, except for the 0-lepton channel which is shown in Fig. 8. Shown are the data (points with error bars) and expectation (histograms). The background contributions after the global 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 normalised to the signal yield extracted from data ($\mu = 1.16$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend for the signal regions. In the $W + HF$ and $e\mu$ CRs, the unstacked unfilled histograms for the signal are not shown. The dashed histogram shows the total pre-fit background. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty for 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: Distributions used as input to the global likelihood fit for the MVA analysis: BDT_{VH} distributions in the medium-\(p_T\) categories of the 2 lepton channel, event yields in the W + HF control regions. These distributions complete the set shown in Figures 1 and 11. Shown are the data (points with error bars) and expectation (histograms). The background contributions after the global 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 after rescaling the SM cross section according to the value of signal strength extracted from data (\(\mu = 1.16\)), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend for the signal regions. In the W + HF CRs the unstacked unfilled histograms for the signal are not shown. The dashed histogram shows the total pre-fit background. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty for 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 11: Distributions used as input to the global likelihood fit for the MVA analysis: $m_{bb}$ distributions in the $e\mu$ control regions. These distributions complete the set shown in Figures 1 and 10. Shown are the data (points with error bars) and expectation (histograms). The background contributions after the global fit are shown as filled histograms. The dashed histogram shows the total pre-fit background. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty for 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.
<table>
<thead>
<tr>
<th>Control regions</th>
<th>1-lepton</th>
<th>2-lepton</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p_T^V &gt; 150$ GeV, 2-$b$-tag</td>
<td>$75$ GeV &lt; $p_T^V$ &lt; $150$ GeV, 2-$b$-tag</td>
</tr>
<tr>
<td>Sample</td>
<td>2-jet</td>
<td>3-jet</td>
</tr>
<tr>
<td>$Z + HF$</td>
<td>$15.1 \pm 1.4$</td>
<td>$33 \pm 2.5$</td>
</tr>
<tr>
<td>$W + ll$</td>
<td>$2.1 \pm 1.5$</td>
<td>$3.8 \pm 2.6$</td>
</tr>
<tr>
<td>$W + cl$</td>
<td>$8.4 \pm 4.1$</td>
<td>$13.5 \pm 6.6$</td>
</tr>
<tr>
<td>$W + HF$</td>
<td>$498 \pm 34$</td>
<td>$1044 \pm 92$</td>
</tr>
<tr>
<td>Single top quark</td>
<td>$23.8 \pm 5.4$</td>
<td>$122 \pm 23$</td>
</tr>
<tr>
<td>$\bar{t}t$</td>
<td>$68 \pm 18$</td>
<td>$307 \pm 77$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$13.4 \pm 3.7$</td>
<td>$22.6 \pm 7.5$</td>
</tr>
<tr>
<td>Multi-jet e sub-ch.</td>
<td>$8.3 \pm 8.5$</td>
<td>$3.6 \pm 2.9$</td>
</tr>
<tr>
<td>Multi-jet $\mu$ sub-ch.</td>
<td>$6.9 \pm 4.6$</td>
<td>$13 \pm 13$</td>
</tr>
<tr>
<td>Total bkg.</td>
<td>$644 \pm 23$</td>
<td>$1563 \pm 39$</td>
</tr>
<tr>
<td>Signal (fit)</td>
<td>&lt; 1</td>
<td>$2.3 \pm 0.6$</td>
</tr>
<tr>
<td>Data</td>
<td>$642$</td>
<td>$1567$</td>
</tr>
</tbody>
</table>

Table 12: The fitted signal and background yields for each control region category in each channel ($W + HF$ in the 1-lepton channel, $e\mu$ events in the 2-lepton channel), corresponding to the selection applied to the control regions for the multivariate analysis. The yields are normalised by the results of the global likelihood fit. All systematic uncertainties are included in the indicated uncertainties. An entry of “–” indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selection.
Figure 12: Impact of systematic uncertainties for the fitted Higgs boson signal-strength parameter $\mu$ for the nominal MVA analysis applied to the 13 TeV data. The systematic uncertainties are listed in decreasing order of their impact on $\mu$. The boxes show the variations of $\hat{\mu}$, referring to the top $x$-axis, when fixing the corresponding individual nuisance parameter $\theta$ to its fitted value $\hat{\theta}$ modified upwards or downwards by its fitted uncertainty, and performing the fit again, with all the other parameters allowed to vary, so as to take correlations between systematic uncertainties properly into account. 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 fitted uncertainties of the nuisance parameters, relative to their nominal uncertainties. As explained in section 5.1, the $b$-tagging uncertainties are decomposed into uncorrelated components; the labels 0 and 1 refer to the leading and second-leading components. The large data statistics in the 0- and 2-lepton mass sidebands allow to pull and constrain the nuisance parameter on the $m_{bb}$ shape of the $Z+HF$ background. The pull corrects a mismodelling, observed in $Z+HF$ enriched side-band regions, of the $m_{bb}$ distribution by the simulation.
Figure 13: Distributions of $m_{bb}$ used as input to the global fit of the di-jet mass analysis. The distributions refer to the signal regions of the 0-lepton channel. Shown are the data (points with error bars) and expectation (histograms). The background contributions after the global 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 normalised to the signal yield extracted from data ($\mu = 1.06$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty for 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 14: Distributions of $m_{bb}$ used as input to the global fit of the di-jet mass analysis. The distributions refer to the signal regions of the 1-lepton channel. Shown are the data (points with error bars) and expectation (histograms). The background contributions after the global 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 normalised to the signal yield extracted from data ($\mu = 1.06$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty for 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 15: Distributions of $m_{bb}$ used as input to the global fit of the di-jet mass analysis. The distributions refer to the signal regions of the 2-lepton channel. Shown are the data (points with error bars) and expectation (histograms). The background contributions after the global 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 normalised to the signal yield extracted from data ($\mu = 1.06$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The entries in overflow are included in the last bin. The size of the combined statistical and systematic uncertainty for 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 16: The distribution of $m_{bb}$ in data, as obtained with the dijet-mass analysis (left), and after subtraction of all backgrounds (bottom panel of left plot and also shown in standalone plot on right). The contributions from all lepton channels, $p_T^V$ intervals and number-of-jets categories are summed weighted by their respective value of the ratio of fitted Higgs boson signal and background. The expected contribution of the associated $WH$ and $ZH$ production of a SM Higgs boson with $m_H = 125$ GeV is shown scaled by the measured combined signal strength ($\mu = 1.06$). The size of the combined statistical and systematic uncertainty for the fitted background is indicated by the hatched band.
Figure 17: The fitted values of the Higgs boson signal strength parameter $\mu_{bb}^{VH}$ for $m_{H} = 125$ GeV for the 0-, 1- and 2-lepton channels and their combination with the 7 TeV, 8 TeV and 13 TeV datasets combined. The results are shown both for the nominal multivariate analysis (MVA) and for the di-jet mass analysis (CBA). The individual $\mu_{bb}^{VH}$ values for the lepton channels are obtained from a simultaneous fit with the signal strength parameter for each of the lepton channels floating independently.
Figure 18: The BDT\textsubscript{VZ} output post-fit distributions in the 0-lepton (top) and 1-lepton (bottom) channels for 2 b-tag events, for all jet-multiplicity categories and \( p_T^V \) regions used in the fit. Only the distributions in the signal regions are shown. The background contributions after the global likelihood fit are shown as filled histograms. The VZ diboson signal is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data (\( \mu = 1.20 \)), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.
Figure 19: The BDT_{VZ} output post-fit distributions in the 2-lepton channel for 2 $b$-tag events, for all jet-multiplicity categories and $p_T^V$ regions used in the fit. Only the distributions in the signal regions are shown. The background contributions after the global likelihood fit are shown as filled histograms. The VZ diboson signal is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data ($\mu = 1.20$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel.
Figure 20: Event yields as a function of $\log(S/B)$ for data, background and a diboson signal ($VZ$) with $m_H = 125$ GeV. Final-discriminant bins in all regions are combined into bins of $\log(S/B)$, with $S$ being the fitted signal and $B$ the fitted background. The diboson signal contribution is shown after rescaling the SM cross section according to the value of the signal strength parameter extracted from data ($\mu = 1.20$). In the lower panel, the pull of the data with respect to the background (the statistical significance of the difference between data and fitted background) is shown with statistical uncertainties only. The full line indicates the pull of the prediction for signal and background with respect to the background prediction.
Figure 21: The fitted values of the $VZ$ signal-strength parameter $\mu_{VZ}^{bb}$ for the 0-, 1- and 2-lepton channels and for their combination. The individual $\mu_{VZ}^{bb}$ values for the lepton channels are obtained from a simultaneous fit with the signal-strength parameters for each of the lepton channels floating independently. The compatibility between the individual signal strengths is 64%.

Figure 22: The fitted values of the Higgs boson signal strength parameter $\mu_{H \rightarrow bb}$ for $m_H = 125$ GeV separately for the $VH$, $ttH$ and ggF+VBF analyses in both Run 1 and Run 2, along with their combination. The individual $\mu_{H \rightarrow bb}$ values for the different production modes are obtained from a simultaneous fit with the signal strength parameters for each of the processes floating independently. The compatibility of the individual signal strengths is 54%.