Search for Higgs boson pair production in association with a vector boson in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector

The ATLAS Collaboration

This paper reports a search for Higgs boson pair (hh) production in association with a vector boson (W or Z) using 139 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$ recorded with the ATLAS detector at the Large Hadron Collider. The search is performed in final states in which the vector boson decays leptonically ($W \rightarrow \ell\nu$, $Z \rightarrow \ell\ell, \nu\nu$ with $\ell = e, \mu$) and the Higgs bosons each decay into a pair of b-quarks. It targets Vhh signals from both non-resonant hh production, present in the Standard Model (SM), and resonant hh production, as predicted in some SM extensions. A 95% confidence-level upper limit of 183 (87) times the SM cross-section is observed (expected) for non-resonant Vhh production when assuming the kinematics are as expected in the SM. Constraints are also placed on Higgs boson coupling modifiers. For the resonant search, upper limits on the production cross-sections are derived for two specific models: one is the production of a vector boson along with a neutral heavy scalar resonance H, in the mass range 260–1000 GeV, that decays into hh, and the other is the production of a heavier neutral pseudoscalar resonance A that decays into a Z boson and H boson, where the A boson mass is 360–800 GeV and the H boson mass is 260–400 GeV. Constraints are also derived in the parameter space of two-Higgs-doublet models.

Contents

1 Introduction

Since the discovery of the Higgs boson (h) with a mass (m_h) of approximately 125 GeV in 2012 [1, 2], its couplings to vector bosons and fermions have been found to be consistent with Standard Model (SM) predictions within the current measurement precision [3–5], providing strong evidence that the Higgs boson has SM properties. The SM also predicts the Higgs boson self-coupling (or trilinear coupling) as well as quartic couplings with itself and with massive vector bosons. These couplings are direct consequences of electroweak symmetry breaking (EWSB) [6–8] and are yet to be confirmed experimentally.

The Higgs boson self-coupling and quartic coupling to vector bosons can be probed through studies of Higgs boson pair (hh) production. In proton–proton (pp) collisions, SM production of hh is dominated by the gluon–gluon fusion (ggF) process [9, 10]. Extensive searches for this process have led to significant bounds on the Higgs boson self-coupling [11–14]. Production through the vector-boson fusion (VBF) process has the second largest cross-section [9, 10]. Searches for VBF hh production have resulted in additional constraints on the Higgs boson self-coupling and are also sensitive to the SM prediction of the Higgs boson quartic coupling to vector bosons [15–17]. While produced through non-resonant processes in the SM, hh production can also occur in resonant processes in scenarios beyond the Standard Model (BSM) through the decays of heavy resonances, such as the heavy Higgs boson predicted in two-Higgs-doublet models (2HDM) [18–20] or the spin-2 Kaluza–Klein gravitons in Randall–Sundrum models [21–23]. Searches for resonant hh production in ggF and VBF processes have also led to constraints in the parameter spaces of these models [24–26].

This paper reports a search for Higgs boson pairs produced in association with a vector boson, Vhh ($V = W, Z$), a process previously unexplored. The search targets both non-resonant hh production, which occurs in the SM, and BSM-inspired resonant hh production. It is performed on a dataset of pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV collected between 2015 and 2018 with the ATLAS detector at the Large Hadron Collider (LHC), corresponding to an integrated luminosity of 139 ± 2.4 fb^{-1} [27]. The search considers vector bosons decaying into leptons ($W \to \ell\nu$, $Z \to \ell\ell, \nu\nu$) and Higgs bosons decaying into a pair of b-quarks ($h \to bb$), leading to three distinct leptonic channels: $Zhh \to \nu\nu bbbb$ (denoted by 0L), $Whh \to \ell\nu bbbb$ (denoted by 1L), and $Zhh \to \ell\ell bbbb$ (denoted by 2L). Here ℓ denotes either an electron (e) or a muon (μ).[1]

Non-resonant hh production in association with a V boson arises in the SM from three distinct Higgs boson couplings: coupling to vector bosons, self-coupling, and quartic coupling to vector bosons. The leading-order processes are depicted in Figures 1(a), 1(b), and 1(c), respectively. In the SM, the cross-sections of these processes at the LHC are small compared with those of the ggF and VBF processes, 0.50 ± 0.01 fb for Whh (W^+hh: 0.329 ± 0.007 fb, and W^-hh: 0.173 ± 0.005 fb) and 0.36 ± 0.01 fb for Zhh at $\sqrt{s} = 13$ TeV with $m_h = 125$ GeV [9, 10], computed at next-to-next-to-leading-order (NNLO) accuracy in QCD.

Two BSM scenarios are considered for resonant hh production as illustrated in Figures 2(a) and 2(b). The first scenario, labelled as VH, is the 'Higgsstrahlung' production of a generic neutral CP-even scalar H boson which couples directly to vector bosons and decays into hh, i.e. $VH \to Vhh$. Examples of such a scalar resonance are the CP-even heavy Higgs boson predicted in the electroweak singlet model [28] or in the type-II 2HDM [29]. In this search, the scalar H is assumed to be a narrow resonance; i.e. its natural width is much smaller than the expected experimental relative mass resolution of approximately 3%. This scenario was explored previously by ATLAS in a VBF hh search [15]. The VH search is complementary because it is sensitive to the HWW and HZZ couplings separately, while the VBF search is sensitive only

---

[1] Although $W \to \tau\nu$ and $Z \to \tau\tau$ are not considered explicitly, final states with electrons or muons from τ decays are included.
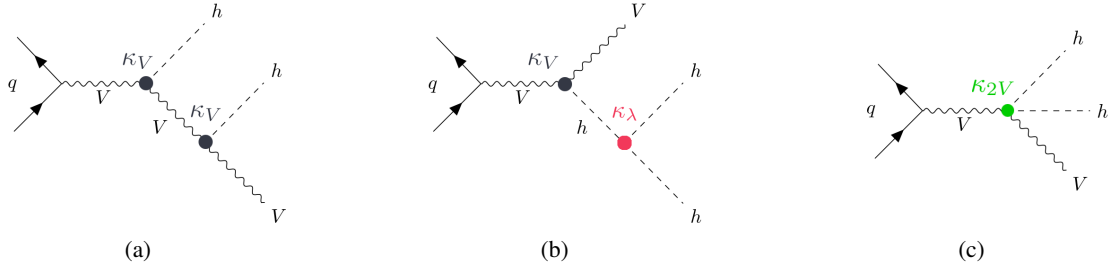
Figure 1: Leading-order Feynman diagrams of non-resonant hh production in association with a vector boson V expected in the SM from (a) Higgs boson coupling to vector bosons, (b) Higgs boson self-coupling, and (c) Higgs boson quartic coupling to vector bosons. The coupling modifiers κ_V, κ_λ, and κ_{2V} are discussed in Section 3.
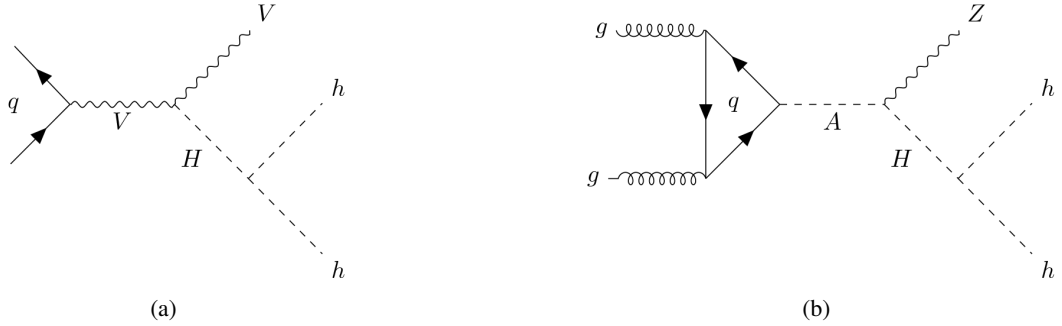

Figure 2: Leading-order Feynman diagrams of resonant hh production in association with a vector boson V predicted in some BSM scenarios from the decay of a heavy scalar H originating from (a) an off-shell vector boson and (b) the decay of a neutral heavy pseudoscalar A.

to their combination. The second scenario, labelled as $A \rightarrow ZH$, is a specific process in the 2HDM which predicts three neutral Higgs bosons: two CP-even scalars h and H (with mass hierarchy $m_H > m_h$), and one CP-odd scalar A. In parts of the 2HDM parameter space where the light Higgs boson h is similar to the SM Higgs boson and has a mass $m_h \sim 125$ GeV favourable for electroweak baryogenesis, the A boson has a mass below about 800 GeV but is heavier than the H boson [30]. If m_H is in the range $2m_h < m_H < 2m_t$, the $H \rightarrow hh$ decay branching ratio could be substantial, leading to a sizeable rate for $gg \rightarrow A \rightarrow ZH \rightarrow Zhh$. Here m_t is the mass of the top quark. For this search, natural widths up to 20% of its mass are considered for A, and a narrow width is assumed for H. Searches for $A \rightarrow ZH$ were performed previously by ATLAS and CMS in the $H \rightarrow bb, \tau\tau$, and WW decay channels [31–34].

2 ATLAS detector

The ATLAS detector [35] at the LHC covers nearly the entire solid angle around the collision point.[2] It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadron calorimeters, and a muon spectrometer incorporating three large superconducting air-core toroidal magnets.

The inner-detector system (ID) is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, the first hit normally being in the insertable B-layer (IBL) installed before Run 2 [36, 37]. It is followed by the silicon microstrip tracker (SCT), which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$. The TRT also provides electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, the electromagnetic calorimeter (ECAL) consists of barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadron calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadron endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic energy measurements respectively.

The muon spectrometer (MS) comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by the superconducting air-core toroidal magnets. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. Three layers of precision chambers, each consisting of layers of monitored drift tubes, covers the region $|\eta| < 2.7$, complemented by cathode-strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

Interesting events are selected by the first-level trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger [38]. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the high-level trigger further reduces in order to record events to disk at about 1 kHz.

An extensive software suite [39] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

---

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

3 Data and Monte Carlo samples

The data used in this analysis were collected using unprescaled single-lepton, missing transverse momentum (E_T^{miss}), or single-photon triggers. The single-lepton trigger requirements are applied as a logical OR of single-electron or single-muon triggers [40, 41], with transverse momentum (p_T) thresholds that started at 20 GeV or 24 GeV in 2015 for muons or electrons, respectively, and increased to 26 GeV in 2016–2018. The E_T^{miss} trigger [42] threshold was raised from 70 GeV to 110 GeV between the 2015 and 2018 data-taking periods. The single-photon trigger [40] was only used in this analysis for background estimation and had a threshold of 140 GeV for the full data-taking period. Events are selected for analysis only if they are of good quality and if all the relevant detector components are known to have been in good operating condition [43], which corresponds to a total integrated luminosity of $139.0\,\mathrm{fb}^{-1}$. The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [44], obtained using the LUCID-2 detector [27] for the primary luminosity measurements. The recorded events contain an average of 34 inelastic pp collisions per bunch-crossing.

Monte Carlo (MC) simulations are used to optimise the search sensitivities and to estimate background contributions. MC samples were produced with various event generators, interfaced to different programs for parton showers, hadronisation, and underlying-event simulations. No potential interference between signal and background processes is considered. Table 1 summarises the simulation of the signal and background processes relevant for the searches described in this paper. All MC samples are passed through the ATLAS detector simulation program [45] based on GEANT4 [46]. Simulated processes are normalised to the most accurate theoretical cross-section predictions currently available. The effects of multiple interactions in the same or nearby bunch crossings (pile-up) were modelled by overlaying minimum-bias events, simulated using the soft QCD processes of PYTHIA 8.186 [47] with the A3 [48] set of tuned parameters and the NNPDF2.3LO [49] PDF. For all samples of simulated events, except for those generated using SHERPA [50], the EVTGEN 1.6.0 program [51] was used to describe the decays of bottom and charm hadrons.

3.1 Non-resonant signal samples

As shown in Figure 1, three types of the Higgs boson couplings are responsible for non-resonant hh production in the SM. Although these couplings are predicted in the SM once the Higgs boson mass m_h is known, their values should be determined experimentally. Deviations from their SM values are traditionally parameterised using the coupling modifiers [71] denoted, in this paper, by κ_V, κ_λ, and κ_{2V} for the hVV, hhh, and $hhVV$ vertices, respectively, assuming the same coupling modifiers for the W and Z bosons.[3]

While it is impractical to produce signal samples for arbitrary values of the coupling modifiers probed by this analysis, such samples can be constructed from six independent samples composed of different combinations of the leading-order (LO) processes shown in Figure 1. These component samples were produced with $\kappa_V = \kappa_\lambda = \kappa_{2V} = 1$ for the following couplings and their combinations:

(1) κ_V: diagram without the hhh and $hhVV$ vertices, Figure 1(a);
(2) κ_λ: diagram with the hhh vertex, Figure 1(b);
(3) κ_{2V}: diagram with the $hhVV$ vertex, Figure 1(c);
(4) $\kappa_V + \kappa_\lambda$: all diagrams except those with an $hhVV$ vertex, Figures 1(a) and 1(b);

---

[3] In the SM, $\kappa_V = 1$, $\kappa_\lambda = 1$, and $\kappa_{2V} = 1$.

Table 1: List of MC event generators, PDFs, and parton shower, hadronisation and underlying event (UE) models used to simulate signal and background processes. Different versions are the results of matching different calculations and models at the time the samples were produced. Here V_ℓ and V_h, respectively, denote V decaying leptonically or hadronically. The last column shows the calculation orders of cross-sections used. Cross-section orders in the last column can be leading order (LO), next-to-leading order (NLO), or next-to-next-to-leading order (NNLO) in terms of QCD or electroweak (EW) accuracy. The mass of the Higgs boson h is set to 125 GeV in the simulation.

Process	Matrix element			Parton	UE	Cross-section
	generator	order	PDF	shower	model	order
Signal samples						
Non-resonant Vhh	MADGRAPH5_AMC@NLO [52]	LO	NNPDF2.3LO [49]	PYTHIA 8.244 [53]	A14 [54]	NNLO [9]
Resonant VH	MADGRAPH5_AMC@NLO	LO	NNPDF2.3LO	PYTHIA 8.244	A14	LO
Resonant $A \to ZH$	MADGRAPH5_AMC@NLO	LO	NNPDF2.3LO	PYTHIA 8.244	A14	LO
Top quark processes						
$t\bar{t}$	POWHEG BOX v2 [55]	NLO	NNPDF3.0NLO [56]	PYTHIA 8.230	A14	NNLO +NNLL [57]
Single-t: s-channel	POWHEG BOX v2 [58]	NLO	NNPDF3.0NLO	PYTHIA 8.230	A14	NLO [59]
Single-t: t-channel	POWHEG BOX v2 [58]	NLO	NNPDF3.0NLO	PYTHIA 8.230	A14	NLO [60]
Single-t: Wt	POWHEG BOX v2 [61]	NLO	NNPDF3.0NLO	PYTHIA 8.230	A14	NNLO [62]
ttH	POWHEG BOX v2	NLO	NNPDF3.0NLO	PYTHIA 8.230	A14	NLO (QCD+EW) [10]
ttV	MADGRAPH5_AMC@NLO	NLO	NNPDF3.0NLO	PYTHIA 8.230	A14	NLO (QCD+EW)
$tt\ell\ell$§	MADGRAPH5_AMC@NLO	NLO	NNPDF3.0NLO	PYTHIA 8.230	A14	NLO
$tttt$	MADGRAPH5_AMC@NLO	NLO	NNPDF3.1LO	PYTHIA 8.230	A14	NLO (QCD+EW)
ttt	MADGRAPH5_AMC@NLO	LO	NNPDF2.3LO	PYTHIA 8.230	A14	LO
$ttWW$	MADGRAPH5_AMC@NLO	LO	NNPDF2.3LO	PYTHIA 8.230	A14	LO
tWZ	MADGRAPH5_AMC@NLO	NLO	NNPDF2.3LO	PYTHIA 8.212	A14	NLO
tZ	MADGRAPH5_AMC@NLO	LO	NNPDF2.3LO	PYTHIA 8.212	A14	LO
Single- and multi-boson production						
V + jets	SHERPA 2.2.1 [63–65]	NLO†	NNPDF3.0NNLO [56]	SHERPA 2.2.1 [66, 67]	Default	NNLO [68]
γ + jets	SHERPA 2.2.2	NLO†	NNPDF3.0NNLO	SHERPA 2.2.2	Default	NNLO
$qq \to V_\ell V_{\ell\text{ or }h}$	SHERPA 2.2.1	NLO	NNPDF3.0NNLO	SHERPA 2.2.1	Default	NLO‡
$gg \to V_\ell V_{\ell\text{ or }h}$	SHERPA 2.2.2	NLO	NNPDF3.0NNLO	SHERPA 2.2.2	Default	NLO‡
$V_\ell V_\ell V$	SHERPA 2.2.2	NLO	NNPDF3.0NNLO	SHERPA 2.2.2	Default	LO
$V_\ell V_h V_h$	MADGRAPH5_AMC@NLO	LO	NNPDF2.3LO	PYTHIA 8.243	A14	LO
$qq \to VH$	POWHEG BOX v2 [69]	NLO	NNPDF3.0NLO	PYTHIA 8.212	AZNLO [70]	NNLO(QCD) +NLO(EW)
$gg \to ZH$	POWHEG BOX v2	NLO	NNPDF3.0NLO	PYTHIA 8.212	AZNLO	NLO+NLL

§ Here, the same-flavour $\ell\ell$ pairs are from $Z^*/\gamma^* \to \ell\ell$.
† Produced with up to two extra jets at NLO and up to four extra jets at LO.
‡ For the diboson samples the cross-sections are calculated by the Monte Carlo generator at NLO accuracy in QCD.

(5) $\kappa_V + \kappa_{2V}$: all diagrams except those with an hhh vertex, Figures 1(a) and 1(c);

(6) $\kappa_\lambda + \kappa_{2V}$: all diagrams with either an hhh or an $hhVV$ vertex, Figures 1(b) and 1(c).

Samples 1–3 determine the contributions from their respective diagrams, while samples 4–6 allow the determination of the contributions from interference between the three diagrams. Non-resonant signal samples for any coupling modifier deviations from the SM values can be built from the combinations of these component samples, weighted by the coupling modifiers. A SM Vhh sample with $\kappa_V = \kappa_\lambda = \kappa_{2V} = 1$ was also produced to validate this procedure. Kinematic distributions of the reweighted sample are found

to agree with those of the directly produced sample. Since the hVV coupling has been constrained through the single Higgs boson measurements [4], $\kappa_V = 1$ is assumed in this paper.

In addition to the quark-initiated diagrams shown in Figure 1, Zhh can also be produced via the gluon-initiated $gg \rightarrow Zhh$ diagram. Although $gg \rightarrow Zhh$ is technically a higher-order process for Zhh production, its cross-section is predicted to be approximately 24% of that of the $qq \rightarrow Zhh$ process in the SM [10]. A correction is applied to the non-resonant $qq \rightarrow Zhh$ cross-section, as a function of κ_λ and κ_{2V}, to account for the $gg \rightarrow Zhh$ contribution. The generator-level difference, in both normalisation and shape, between the $qq \rightarrow Zhh$ sample and the sum of the $qq \rightarrow Zhh$ and $gg \rightarrow Zhh$ samples is taken as an uncertainty. For a signal as expected in the SM, the normalisation component of this uncertainty has a magnitude of 24%. It tends to be larger than 24% for large κ_{2V} and smaller than 24% for large κ_λ.

3.2 Resonant signal samples

The production of a generic CP-even scalar resonance H in association with a V boson, where the H decays into a pair of the h Higgs bosons, is modelled using the $qq \rightarrow VH \rightarrow Vhh$ process in the 2HDM as shown in Figure 2(a). The heavy Higgs boson is assumed to have a narrow width, i.e. its total decay width is far smaller than the experimental m_{hh} resolution of approximately 3%, but to decay promptly nevertheless. Ten signal samples were produced for each leptonic channel, corresponding to m_H values of 260, 280, 300, 400, 500, 600, 700, 800, 900, and 1000 GeV. The lower bound is dictated by $2m_h$, while the upper bound is limited by the ability to reconstruct two separate jets from the highly boosted $h \rightarrow bb$ decay.

The $gg \rightarrow A \rightarrow ZH \rightarrow Zhh$ signal samples were produced for mass combinations of $m_A = 360, 400, 500, 600, 700,$ and $800\,\text{GeV}$ and $m_H = 260, 300,$ and $400\,\text{GeV}$ subject to the kinematic bound of $m_A > m_H + m_Z$, leading to a total of 15 (m_A, m_H) grid points for each of the 0L and 2L channels. These mass combinations cover the 2HDM parameter space relevant for this search, but unexplored by previous $gg \rightarrow A \rightarrow ZH$ searches [31–34]. The CP-odd A boson could have a substantial total decay width, depending on the 2HDM parameter values. Two sets of samples were produced, one for a narrow-width (NW) A boson and the other for a large-width (LW) A boson. In both scenarios, the total decay width of the heavy CP-even H boson is assumed to be narrow. For the LW samples, the A boson width is assumed to be 20% of its mass. To perform searches for A bosons with different widths, MC distributions for an A boson with width equal to 5%, 10%, or 15% of its mass are derived through reweighting. The generator-level A boson mass distribution for each of these intermediate widths is compared with that of the LW A boson to derive reweighting factors which are then applied to the distributions of the simulated LW MC samples to obtain the corresponding distributions for the intermediate widths.

4 Object reconstruction and identification

As discussed in Section 1, the signal events are characterised by the products of the targeted decays of the vector and Higgs bosons: electrons or muons, jets (b-jets in particular), and missing transverse momentum. Additionally, events with identified hadronically decaying τ-leptons are vetoed to reduce backgrounds. Events with photons are used for background estimation. The reconstruction and identification of these physics objects are described in this section.

Electrons are reconstructed by matching topological energy clusters in the ECAL with tracks in the ID [72] and are required to have $p_T > 7\,\text{GeV}$ within the pseudorapidity range $|\eta| < 2.47$. They are identified

using likelihood-based identification criteria which combine the requirements of calorimeter shower shape, track-to-cluster matching, and associated track qualities. Electron candidates are required to satisfy the *tight* criterion in the 1L channel and the *loose* criterion for the rest. All candidates must satisfy a loose track- and calorimeter-based isolation criterion to minimise the number of jets misidentified as electrons.

Similarly to electrons, photon reconstruction starts with topological energy clusters in the ECAL [72], but the clusters are required to have either no matching tracks (unconverted photons) or one or two matching tracks consistent with a conversion vertex (converted photons). Photon candidates are required to have $p_T > 10$ GeV within the pseudorapidity range $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$ and satisfy the *tight* identification criterion as well as a calorimeter-based isolation requirement called '*TightCaloOnly*' [72] to suppress backgrounds.

Muons are reconstructed by matching tracks in the ID to either full tracks or track segments in the MS or, for $|\eta| < 0.1$ only, to energy deposits in the calorimeter [73]. They must have $p_T > 7$ GeV and be within $|\eta| < 2.5$, the combined acceptance of the ID and MS. Muons must satisfy the *medium* quality criterion in the 1L channel and the *loose* criterion in others. The same isolation requirement used for the electron selection is applied to all muon candidates to reduce the rate of muons from heavy-flavour decays.

Electrons (muons) are required to have associated tracks satisfying $|d_0/\sigma_{d_0}| < 5$ (3) and $|z_0 \sin \theta| < 0.5$ mm, where d_0 is the transverse impact parameter relative to the beam line, σ_{d_0} is its uncertainty, and z_0 is the distance between the longitudinal position of the track along the beam line at the point where d_0 is measured and the longitudinal position of the primary pp collision vertex.

Jets are reconstructed from particle-flow objects using the anti-k_t algorithm [74, 75] with a radius parameter of $R = 0.4$ and are calibrated as described in Ref. [76]. They are required to have $p_T > 20$ GeV and $|\eta| < 4.5$. Jets with $p_T < 60$ GeV and $|\eta| < 2.4$ originating from pile-up are suppressed with the jet-vertex-tagger [77], a likelihood discriminant based on matching enough of the jet's tracks to the primary vertex.

Jets containing b-hadrons (b-jets) are identified with the DL1r algorithm [78]. The algorithm is based on information such as properties of displaced tracks and reconstructed secondary vertices in the jet. A jet is b-tagged if the DL1r value is above a preset threshold. Four thresholds, referred to as working points (WP), are defined with average efficiencies of 60%, 70%, 77%, and 85% for tagging b-jets from simulated $t\bar{t}$ events. A pseudo-continuous b-tagging score is defined for each jet as the number of WPs it satisfies, with zero being failing, and four passing, all WPs. In some cases, instead of directly applying b-tagging in MC simulation, the b-tagging efficiencies for b-, c-, and light-flavour jets are applied as weights in the generated events [79]. This procedure, referred to as 'truth'-tagging, increases the effective MC sample size for subdominant backgrounds. A momentum correction is applied to b-tagged jets to account for the energy lost to soft radiation and to muons and neutrinos in semileptonic b-hadron decays, following the procedure used in Ref. [80]. Furthermore, correction factors are applied to the simulated samples to compensate for the differences between the b-tagging efficiencies in data and simulation [78].

Hadronically decaying τ-leptons, τ_h, are identified from the reconstructed jets [81]. A multivariate discriminant is used to select jets with energy-deposit profiles consistent with those expected from τ_h decay products and to match tracks in the ID to the τ_h candidates, using the 'medium' RNN-based τ_h criteria. The τ_h candidates must have one or three associated tracks and satisfy the requirements of $p_T > 20$ GeV and $|\eta| < 2.5$, excluding the region $1.37 < |\eta| < 1.52$.

The missing transverse momentum $\boldsymbol{p}_T^{\text{miss}}$, with magnitude E_T^{miss}, is calculated as the negative vectorial sum of the transverse momenta of reconstructed physics objects, namely electrons, muons, photons, hadronically

decaying τ-leptons, and jets [82, 83]. A component, called the 'soft-term', from energy deposits due to the underlying event and other soft radiation not included in the physics objects is added in the $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$ calculation. A $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$ significance, $\mathcal{S}(E^\mathrm{miss})$, is defined in order to test whether the measured $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$ is incompatible with zero real $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$. It is calculated from the measured $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$, its resolution, and the correlation between measurements parallel and perpendicular to the $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$ direction [84].

A sequential overlap-removal procedure is applied to ensure that energy deposits in the calorimeter and tracks in the ID are not included in two or more different reconstructed objects. If two electrons share a track, the electron with lower p_T is removed. If an electron and muon share an inner-detector track, the muon is removed if it is calorimeter-tagged, and the electron is removed otherwise. The closest jet within $\Delta R = 0.2$ of a selected electron is removed. If the nearest jet surviving that selection is within $\Delta R = 0.4$ of the electron, the electron is discarded. Muons are usually removed if they are separated from the nearest jet by $\Delta R < 0.4 + 10\,\mathrm{GeV}/p_\mathrm{T}^\mu$, since this reduces the background from heavy-flavour decays inside jets. However, if this jet has fewer than three associated tracks, the muon is kept and the jet is removed instead; this avoids an inefficiency for high-energy muons undergoing significant energy loss in the calorimeter. A τ_h candidate is rejected if it is separated by $\Delta R < 0.2$ from any selected electron, muon, or jet.

5 Analysis

The non-resonant and resonant signal models targeted in this search share the same Vhh final state but have different event kinematics. The search starts with the selection of the Vhh final states which are characterised by a leptonically decaying V boson and four b-jets from the decays of the two Higgs bosons. Three leptonic channels (0L, 1L, 2L), one for each leptonic decay mode of the W and Z bosons ($Z \to \nu\nu$, $W \to \ell\nu$, $Z \to \ell\ell$), define the signal regions (SRs) of the search. Multivariate techniques based on boosted decision trees (BDT), trained to distinguish between signal and background events in each SR and for each signal model, provide the final discriminants to extract potential signal contributions. This is achieved through simultaneous fits to the BDT distributions observed in data with the hypotheses of signal plus background contributions. For the resonant models, mass requirements are applied to the resonance candidates before the fits to the BDT distributions. Background contributions are estimated using both data control regions (CRs) and MC simulations. The SR event selections, background estimations, the designs and trainings of the BDTs, and the mass requirements for resonant models are described in this section.

5.1 Signal region event selection

The 0L channel is intended for the $Zhh \to \nu\nu\,bbbb$ final state. Candidate events are required to have $E_\mathrm{T}^\mathrm{miss} > 150\,\mathrm{GeV}$ and $\mathcal{S}(E_\mathrm{T}^\mathrm{miss}) > 12$. Events with identified *loose* leptons or τ_h are vetoed. To suppress multi-jet backgrounds with mismeasured $E_\mathrm{T}^\mathrm{miss}$, the minimum azimuthal opening angle between $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$ and the Higgs boson candidates (see below) must satisfy $\min\left[\Delta\phi(\boldsymbol{p}_\mathrm{T}^\mathrm{miss}, \boldsymbol{h})\right] > 1$. The ratio of the $E_\mathrm{T}^\mathrm{miss}$ trigger efficiencies in data and simulation, measured in single-muon events as a function of $\boldsymbol{p}_\mathrm{T}^\mathrm{miss} + \boldsymbol{p}_\mathrm{T}^\mu$, is applied as an event-weight scale factor to simulated events [42], with corresponding uncertainties described in Section 6.1.

The 1L channel is designed for the $Whh \to \ell\nu\,bbbb$ final state. In this channel, candidate events are selected by requiring exactly one *tight* electron with $p_\mathrm{T} > 27\,\mathrm{GeV}$ or one *medium* muon with $p_\mathrm{T} > 25\,\mathrm{GeV}$. In addition, the events must have $E_\mathrm{T}^\mathrm{miss} > 30\,\mathrm{GeV}$. Candidates with additional *loose* leptons or identified

τ_h are removed. The 1L channel is split into two SRs based on the charge of the lepton, 1L+ and 1L−, motivated by the expected large difference between the signal W^+hh and W^-hh cross-sections in contrast to the mostly charge-symmetric backgrounds.

The 2L channel targets the $Zhh \rightarrow \ell\ell\,bbbb$ final state. The $Z \rightarrow \ell\ell$ candidates are selected by requiring exactly two oppositely charged *loose* leptons with the same flavour, e^+e^- or $\mu^+\mu^-$, at least one of which has $p_T > 27$ GeV. The invariant mass of the lepton pair must satisfy $81 < m_{\ell\ell} < 101$ GeV for compatibility with the Z boson mass.

Candidate $hh \rightarrow bbbb$ decays are selected by requiring at least four jets passing the 85% b-tagging WP. Within the four-jet combination with the highest pseudo-continuous b-tagging scores (considering all combinations of four jets when more than one combination satisfies this requirement), the jets are paired to form the two $h \rightarrow bb$ candidates, h_1 and h_2, by minimising the value of $|m_{h_1} - 120\,\text{GeV}| + |m_{h_2} - 120\,\text{GeV}|$. Here m_{h_1} and m_{h_2} are the invariant masses of the two candidates and 120 GeV is their most probable value in simulation. The pair with the higher p_T is labelled as h_1 and the other as h_2. The efficiency for correctly identifying the four b-jets coming from the $hh \rightarrow bbbb$ decay amongst the selected b-jets depends on the signal model and, if applicable, the resonance mass values. For the non-resonant SM Vhh signal, the efficiency is approximately 83%. For the resonant VH signal, the efficiency varies from 72% at $m_H = 260$ GeV to 96% at $m_H = 1000$ GeV. Likewise, the probability for correctly pairing the two Higgs boson decays is also model and mass dependent. The efficiencies for the $A \rightarrow ZH$ signals are similar to those for the VH signals.

Table 2 summarises the selections that define the SRs for the three leptonic channels. Also included in the table are selections for control regions discussed below. The products of the acceptances and efficiencies of the SR selections, $\mathcal{A} \times \epsilon$, are shown in Figure 3 as functions of the model parameters for a few selected signal models.

Table 2: Selections for the 0L, 1L, and 2L signal regions, and $t\bar{t}$ and V+jets control regions. The '—' symbol indicates no selection is applied.

|	Signal regions			Control regions	
	0L	1L (1L+/1L−)	2L	$t\bar{t}$	V+jets
Trigger	E_T^{miss}	single-lepton	single-lepton	single-lepton	single-photon
Lepton or photon	0 *loose* leptons, 0 identified τ_h	= 1 *tight* electron with $p_T > 27$ GeV OR 1 *medium* muon with $p_T > 25$ GeV, 0 additional *loose* leptons, 0 identified τ_h	= 2 *loose* leptons (e^+e^- or $\mu^+\mu^-$), ≥ 1 lepton with $p_T > 27$ GeV, $81 < m_{\ell\ell} < 101$ GeV	= 2 *loose* leptons ($e^\pm\mu^\mp$), ≥ 1 lepton with $p_T > 27$ GeV	= 1 photon with $p_T > 150$ GeV, 0 *loose* leptons, 0 identified τ_h
$\boldsymbol{p}_T^{\text{miss}}$	$E_T^{\text{miss}} > 150$ GeV, $\mathcal{S}(E_T^{\text{miss}}) > 12$, $|\Delta\phi(\boldsymbol{p}_T^{\text{miss}}, h)| > 1$	$E_T^{\text{miss}} > 30$ GeV	—	—	—
Jets	≥ 4 jets with $p_T > 20$ GeV and passing the 85% b-tagging WP				

11
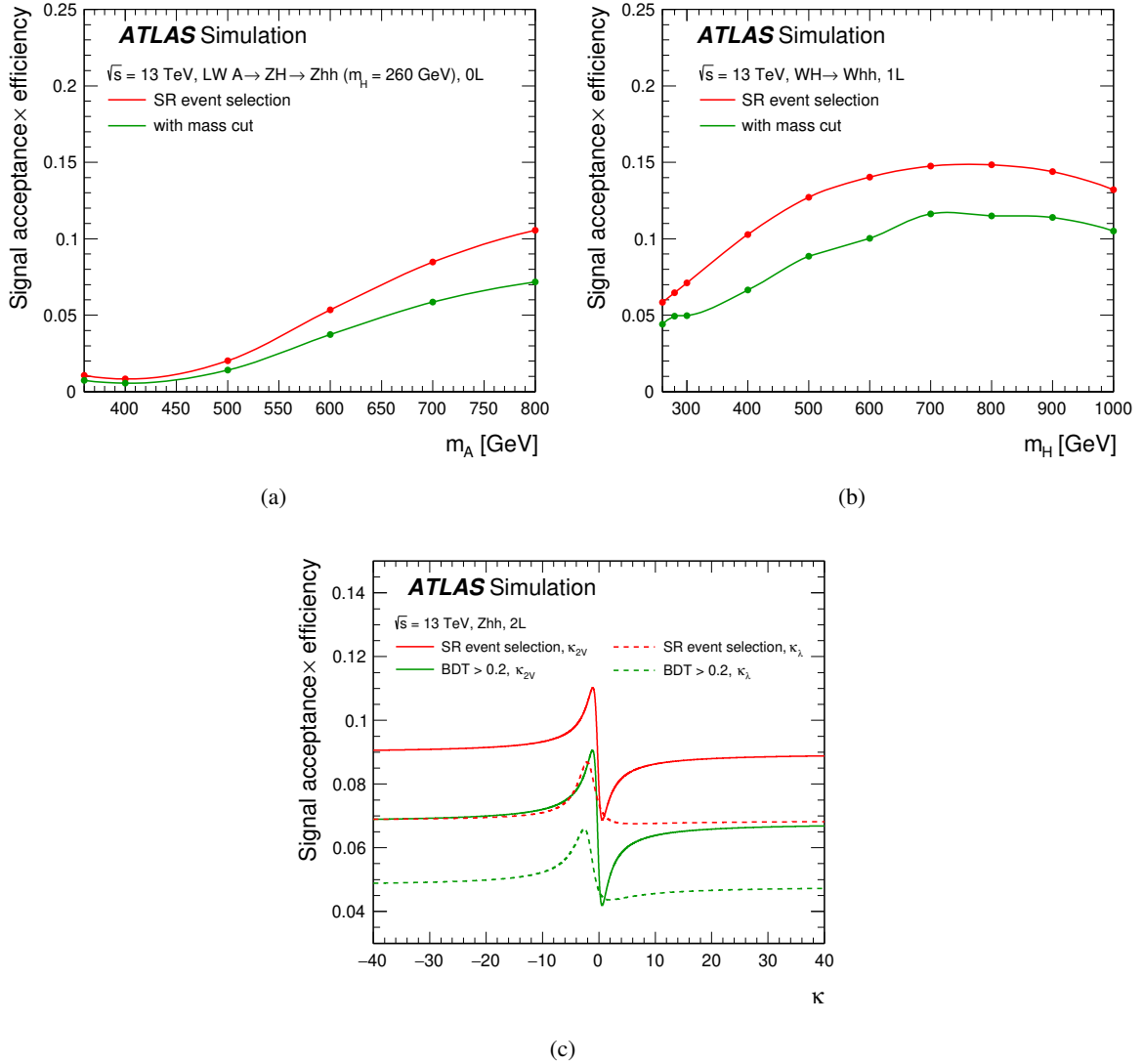
Figure 3: Example dependences of acceptance times efficiency ($\mathcal{A} \times \epsilon$) on signal model parameters: (a) m_A for the $A \to ZH$ search in the 0L channel for the case of $m_H = 260$ GeV and a LW A boson, (b) m_H for the WH search in the 1L channel, and (c) κ_{2V} and κ_λ for the non-resonant Zhh search in the 2L channel. Each $\mathcal{A} \times \epsilon$ value is calculated for the respective $\nu\nu bbbb$, $\ell\nu bbbb$, and $\ell\ell bbbb$ final state, with ℓ including the e, μ, and τ leptons here. For the resonant VH and $A \to ZH$ searches, the $\mathcal{A} \times \epsilon$ value is shown both with and without the mass requirements on the reconstructed resonances. Similarly, for the Zhh search, the $\mathcal{A} \times \epsilon$ value is shown both with and without a BDT requirement. The small decreases in $\mathcal{A} \times \epsilon$ at high m_H values in (b) are due to the merging of b-jets from highly boosted $h \to bb$ decays. The structures in $\mathcal{A} \times \epsilon$ around small values of κ in (c) reflect large changes in the relative contributions of the three production diagrams shown in Figure 1.

5.2 Background estimations

Major background sources in the Vhh search are the production of top-quark pairs ($t\bar{t}$), single top quarks, vector bosons in association with jets (V + jets), diboson and multi-boson events, and multi-jet processes.

Their relative contributions depend on the channel. In the 0L and 1L channels, the $t\bar{t}$ background dominates, with the subleading contribution being from $V+$jets ($Z+$jets for 0L and $W+$jets for 1L). In the 2L channel, $Z+$jets and $t\bar{t}$ are the two leading sources. A mixture of data-driven and simulation-based methods are used to estimate these backgrounds. MC simulations are used to model the kinematics of background processes except for the multi-jet backgrounds. Contributions from $t\bar{t}$ and $V+$jets processes are normalised to the data through the use of CRs, while the rest of the non-multi-jet sources are normalised to their theoretical cross-sections within the estimated uncertainties. For the multi-jet backgrounds, their normalisations and kinematic models are derived from auxiliary data samples. These data-driven procedures are described below.

A $t\bar{t}$ control region ($CR_{t\bar{t}}$) is used to evaluate the $t\bar{t}$ contribution. The control region is defined by using a selection similar to that for the 2L SR, but requiring a different-flavour opposite-sign lepton pair ($e^{\pm}\mu^{\mp}$) instead of a same-flavour opposite-sign lepton pair ($e^{+}e^{-}$ or $\mu^{+}\mu^{-}$) and removing the dilepton mass requirement to increase the sample size. The jet requirements are the same as for the SR. These selections are summarised in Table 2.

Top-quark-pair events have two b-jets from top-quark decays and can mimic signal events if there are two additional genuine b-jets from radiation or if there are misidentified c-jets or light-flavour jets (j) from radiation or hadronic W boson decays. The flavour content of this radiation is difficult to model, particularly for heavy flavours (b or c). To allow for flavour-dependent variations of the $t\bar{t}$ background contribution in the signal extraction, simulated $t\bar{t}$ events are labelled according to the 'truth' flavours of jets from the radiation: events with one or more b-hadrons ($t\bar{t}+\geq 1b$), events with one or more c-hadrons but no b-hadrons ($t\bar{t}+\geq 1c$), and events with no b-hadrons or c-hadrons ($t\bar{t}+j$). The 'truth' flavour of a jet is determined from hadrons with $p_{T} > 5\,\text{GeV}$ found within a cone of size $\Delta R = 0.3$ around the jet axis [78].

The $V+$jets process is a major background source in all three channels ($Z+$jets for 0L and 2L and $W+$jets for 1L). A $V+$jets CR ($CR_{V+\text{jets}}$), defined using $\gamma+$jets events, is used to assess their contributions. As summarised in Table 2, events in $CR_{V+\text{jets}}$ are selected by requiring exactly one photon with $p_{T} > 150\,\text{GeV}$ and no identified e, μ, or τ_{h} from the data collected using a single-photon trigger. They must pass the same jet requirements as events in the SRs. At $p_{T} \gtrsim m_{V}$, $\gamma+$jets events and $V+$jets events are expected to have similar kinematics because they originate from comparable diagrams and the impact of the finite mass of the V bosons becomes small at high p_{T}. Residual differences between $\gamma+$jets and $V+$jets, driven mainly by lower-p_{T} events, are taken into account by extrapolation uncertainties (see Section 6).

The $V+$jets events can be selected as signal events if they have four jets passing the b-tagging requirement, either from genuine heavy-flavour jets or from misidentified light-flavour jets. Like the $t\bar{t}$ events, modelling the jet flavour composition is challenging. The simulated $\gamma+$jets events are categorised according to 'truth' flavour and jet matching: events with three or more b-hadrons ($V+\geq 3b$), events with ≥ 1 c-hadron but with ≤ 2 b-hadrons ($V+\geq 1c$), and events with zero c-hadrons and ≤ 2 b-hadrons ($V+j$).

The contribution from multi-jet processes is negligible in the 0L and 2L SRs as well as in $CR_{t\bar{t}}$, minor in the 1L SRs, but significant in $CR_{V+\text{jets}}$. Both its rate and kinematics are difficult to simulate. Auxiliary data samples rich in multi-jet events are selected to model the corresponding contributions in the 1L SRs and $CR_{V+\text{jets}}$. For $CR_{V+\text{jets}}$, the auxiliary sample is defined by inverting the photon isolation requirement while keeping the rest of the selection the same. The kinematic distributions of multi-jet events in $CR_{V+\text{jets}}$ are modelled by the selected events in this auxiliary sample after subtracting approximately 10% non-multi-jet contributions taken from simulation, and an uncertainty in the shapes of these distributions is defined by varying the non-multi-jet contributions in the auxiliary sample by $\pm100\%$. These distributions are assigned

a pre-fit normalisation equal to the difference between the number of data events and the non-multi-jet contributions in $\text{CR}_{V+\text{jets}}$, with an uncertainty of 100%. The multi-jet modelling in the 1L SRs follows the same approach, except that the auxiliary sample is selected by inverting the lepton isolation requirement and the multi-jet contribution in the 1L SRs is normalised through a fit to the distribution of the transverse mass of the lepton and $E_{\text{T}}^{\text{miss}}$ system,[4] in events selected with the 1L SR requirements but with the $E_{\text{T}}^{\text{miss}}$ criteria reversed to $E_{\text{T}}^{\text{miss}} < 30\,\text{GeV}$, which leads to negligible contamination of potential signals produced with cross-sections comparable to the search sensitivity.

Jet b-tagging scores offer the most sensitive information for separating the different flavour contributions of the $t\bar{t}$ and $V+$ jets backgrounds. Thus the distributions of the sum of the pseudo-continuous b-tagging scores of the four jets with the highest scores, $\sum s_{b\text{-tag}}^{\text{pc}}$, which ranges from 4 to 16, are used to disentangle contributions from the three flavour components. The $\sum s_{b\text{-tag}}^{\text{pc}}$ distributions observed in $\text{CR}_{t\bar{t}}$ and $\text{CR}_{V+\text{jets}}$, divided into three bins (4–9, 10–12, and 13–16), are included in the fits to determine the potential signal contributions as discussed in Section 7. In the fits, the flavour components of the $t\bar{t}$ and $V+$ jets contributions are normalised with their separate normalisation factors. The $\sum s_{b\text{-tag}}^{\text{pc}}$ template, i.e. the shape of the $\sum s_{b\text{-tag}}^{\text{pc}}$ distribution, of each component is taken from simulation except for the multi-jet contribution in $\text{CR}_{V+\text{jets}}$ discussed above. The post-fit $t\bar{t}$ and $V+$ jets contributions comprise 92.4% and 79.7% of the total event yields in their respective CRs.

Validation regions (VRs) are defined for each of the SRs and CRs ($\text{CR}_{t\bar{t}}$ and $\text{CR}_{V+\text{jets}}$) to study the background modelling uncertainties (see Section 6). Events in VRs are selected in the same way as those for SRs and CRs but with different jet requirements. Instead of four b-tagged jets, events are required to have at least four jets, exactly three of which pass the 85% b-tagging WP. The highest-p_{T} non-b-tagged jet in $|\eta| < 2.5$ is taken as the fourth jet to build hh candidates and is assigned a pseudo-continuous b-tagging score of 4, which is found to best reproduce the multivariate discriminants in the signal regions.

5.3 Multivariate discriminant

BDTs are used to exploit the kinematic differences between signal and background events passing the SR event selection. One BDT is constructed for each channel (0L, 1L, 2L) and each signal model (Vhh, VH, $A \rightarrow ZH$), resulting in eight BDTs in total (three each for 0L and 2L, two for 1L). To minimise the complexity of the analysis, these BDTs are built in the same way where possible. They differ in having channel-dependent and signal-model-dependent variables. BDT input variables are chosen through extensive comparisons between their distributions in signal and background events while minimising correlations among the variables. Moreover, for the resonant VH and $A \rightarrow ZH$ searches, only variables with weak correlations with the resonance mass are considered so as to lessen the dependence of the BDTs on any particular hypothesised signal mass value. Requirements on the reconstructed resonance masses are applied separately (Section 5.4). Table 3 summarises the input variables used for the eight BDTs.

A common set of seven input variables are used for all eight BDTs. These variables are

- the sum ($m_{h_1} + m_{h_2}$) and the difference ($m_{h_1} - m_{h_2}$) of the masses of the Higgs boson candidates h_1 and h_2,

- the sum of the b-tagging scores $\sum s_{b\text{-tag}}^{\text{pc}}$,

---

[4] The transverse mass of a system is defined as $\sqrt{m^2 + p_{\text{T}}^2}$, where m and p_{T} are the mass and transverse momentum of the system, respectively.

- the number of jets (N_{jets}),

- the sum of the transverse energy of jets excluding the four selected b-jets (H_{T}^{ex}),

- $m_{h_1}^{\text{FSR}}$ and $m_{h_2}^{\text{FSR}}$, the masses of the two Higgs boson candidates calculated by adding in any jets, excluding the four b-jets themselves, in a cone of size $\Delta R = 0.8$ around each b-jet's axis. This calculation is intended to recover jet energy lost due to final-state radiation (FSR).

Input variables specific to channels or signal models are the invariant mass m_{hh} and transverse momentum p_{T}^{hh} of the reconstructed Higgs boson pair, $E_{\text{T}}^{\text{miss}}$, the transverse momentum of the vector boson (p_{T}^V), the transverse mass of the lepton–$E_{\text{T}}^{\text{miss}}$ system (m_{T}^W), and four variables exploiting differences between the angular distributions of signal and background events. Because of its strong correlation with the resonance mass, m_{hh} is not used in the BDTs for the resonant searches. Instead, requirements on m_{hh} are applied afterwards (see Section 5.4). The variable p_{T}^{hh} is used in all BDTs except those for the $A \to ZH$ search, due to its correlation with m_A. For the same reason, $E_{\text{T}}^{\text{miss}}$ (p_{T}^V) is not used in the 0L (2L) channel of the $A \to ZH$ search. The variables $\cosh(\Delta\eta)_1 - \cos(\Delta\phi)_1$ and $\cosh(\Delta\eta)_2 - \cos(\Delta\phi)_2$ [5] of the h_1 and h_2 candidates exploit the angular differences between the signal $h \to bb$ decay and the background gluon-splitting $g^* \to bb$ process and are motivated by $m_{h1}^2 \propto \cosh(\Delta\eta)_1 - \cos(\Delta\phi)_1$ and $m_{h2}^2 \propto \cosh(\Delta\eta)_2 - \cos(\Delta\phi)_2$ in the approximation that jets are massless. Similarly, the Higgs bosons and vector bosons are expected to be produced more centrally for signal events and more forwardly for background events. The two rapidity difference variables, between h_1 and h_2 ($|y_{h_1} - y_{h_2}|$) and between V and hh ($|y_V - y_{hh}|$), are designed to take advantage of these subtle differences.

Figure 4 compares the m_{hh}, $m_{h_1} - m_{h_2}$, and H_{T}^{ex} distributions in data with those expected from background processes for events passing the SR selections in the three leptonic channels. The background distributions are obtained from the fit to the background-only hypothesis discussed in Section 7. Also shown are the expected distributions from example signal models.

The BDTs are trained in the TMVA [85] framework. In order to make use of the complete set of simulated MC events for the BDT training and evaluation in an unbiased way, the MC events are split into two samples of equal size. A BDT is trained on each of the two samples and applied to the other, such that the same events are never used for both the BDT training and evaluation. Similarly, half of the events in data are evaluated with one of the two BDTs, and half are evaluated with the other. As a result, while significant overtraining is not observed, the effect of any potential overtraining is minimised by following this method. For each BDT, event weights are included in the training so that the relative importance of each background process is taken into account.

For the search for non-resonant Vhh production, the SM, κ_λ-only, and κ_{2V}-only signal samples are added together, with equal weights, to form a combined signal sample for the BDT training. The inclusion of the κ_λ and κ_{2V} samples in the training improves the sensitivity to their respective couplings. Three different BDTs are trained, one for each leptonic channel. Little degradation in sensitivity is observed with this strategy relative to training BDTs for each signal sample separately.

For the search for resonant Vhh signals, MC samples produced with different values of the resonance mass are combined with equal weights to form the signal sample for the BDT training in each channel. It results in a BDT that is not strongly correlated with the resonance mass and therefore has good sensitivity to signals with different mass values. This strategy is applied to both the VH and $A \to ZH$ searches; the combination for the latter includes samples with different m_A and m_H values as well as A bosons with narrow and large widths.

---

[5] $\Delta\eta$ and $\Delta\phi$ are, respectively, the differences in pseudorapidity and azimuth of the two b-jets of the $h \to bb$ candidate.
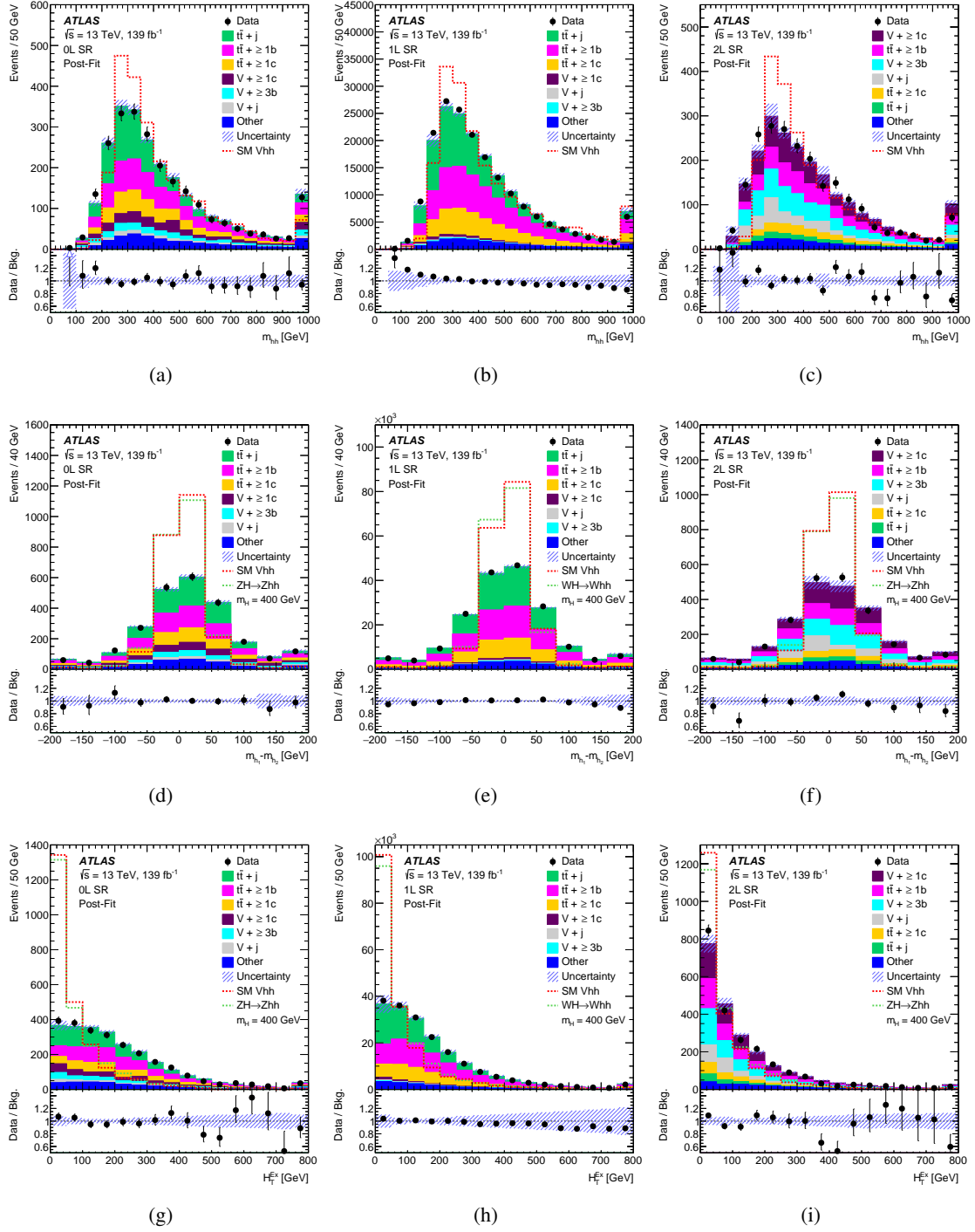
Figure 4: Example distributions of kinematic variables used in the BDT trainings: the invariant mass m_{hh} (a, b, c) and the mass difference $m_{h_1} - m_{h_2}$ (d, e, f) of the two Higgs boson candidates, and H_T^{ex} (g, h, i) of events passing the 0L (a, d, g), 1L (b, e, h), and 2L (c, f, i) SR selections. The 1L SR combines 1L+ and 1L− categories. The expected distributions from example non-resonant and resonant signal models, normalised to the total background expectations, are overlaid. No resonant signal distributions are shown for m_{hh}, which is used in the BDTs for non-resonant production only. The final bins include overflows. The bottom panels show the ratios of the data to the total background expectations. The hatched bands represent the combined statistical and systematic uncertainties in the total background predictions.

16

Table 3: Variables used in the BDT discriminant in each of the channels and for each signal model; see text for the variable definitions. The ✓ symbol indicates the inclusion of the variable. The BDTs for the VH and $A \to ZH$ searches exclude variables strongly correlated with the resonance mass.

Variable	0L			1L		2L		
	Vhh	VH	$A \to ZH$	Vhh	VH	Vhh	VH	$A \to ZH$
$m_{h_1} + m_{h_2}$	✓	✓	✓	✓	✓	✓	✓	✓
$m_{h_1} - m_{h_2}$	✓	✓	✓	✓	✓	✓	✓	✓
N_{jets}	✓	✓	✓	✓	✓	✓	✓	✓
H_T^{ex}	✓	✓	✓	✓	✓	✓	✓	✓
$\sum s_{b\text{-tag}}^{\text{pc}}$	✓	✓	✓	✓	✓	✓	✓	✓
$m_{h_1}^{\text{FSR}}$	✓	✓	✓	✓	✓	✓	✓	✓
$m_{h_2}^{\text{FSR}}$	✓	✓	✓	✓	✓	✓	✓	✓
m_{hh}	✓			✓		✓		
p_T^{hh}	✓	✓		✓	✓	✓	✓	
E_T^{miss}	✓	✓		✓	✓	✓	✓	✓
p_T^V				✓	✓	✓	✓	
m_T^W				✓				
$\cosh(\Delta\eta)_1 - \cos(\Delta\phi)_1$	✓	✓		✓	✓	✓	✓	
$\cosh(\Delta\eta)_2 - \cos(\Delta\phi)_2$	✓	✓		✓	✓	✓	✓	
$|y_{h_1} - y_{h_2}|$	✓	✓		✓	✓	✓	✓	
$|y_V - y_{hh}|$						✓	✓	

5.4 Mass requirements for resonance searches

For the resonant VH and $A \to ZH$ searches, the Higgs boson pair hh is produced from the decay of a new heavy scalar H. Consequently, the signal m_{hh} distributions are expected to peak around m_H with a width equal to the natural width of the new scalar convolved with the detector resolution. Therefore, m_{hh} is a powerful discriminant against the expected continuum SM backgrounds.

Since the lighter Higgs boson h is a narrow resonance with known mass, the m_{hh} resolution can be improved by constraining the measured masses of the two Higgs boson candidates, m_{h_1} and m_{h_2}, to their expected value of 125 GeV. This is achieved by scaling the momenta of the b-jets from each Higgs boson candidate by the ratio of 125 GeV to the measured di-b-jet mass. Figure 5(a) compares the m_{hh} distributions before and after the rescaling for a few selected m_H values in the 2L channel of the ZH search. The rescaling improves the relative m_{hh} resolution from 12.1% (6.1%) to 3.5% (2.6%) at $m_H = 300\,(1000)$ GeV. Similar improvements are obtained in the other leptonic channels of both the VH and $A \to ZH$ searches. Moreover, in the 2L channel of the $A \to ZH$ search the rescaling leads to improvement in the mass resolution of Zhh from the decay of a narrow A resonance, as shown in Figure 5(b). At $m_H = 300$ GeV, the relative Zhh mass resolution improves from 9.4% (5.4%) to 2.8% (2.5%) at $m_A = 400\,(800)$ GeV. For LW A bosons, the rescaling has negligible impact on the width of the Zhh mass distribution. In the following analysis,

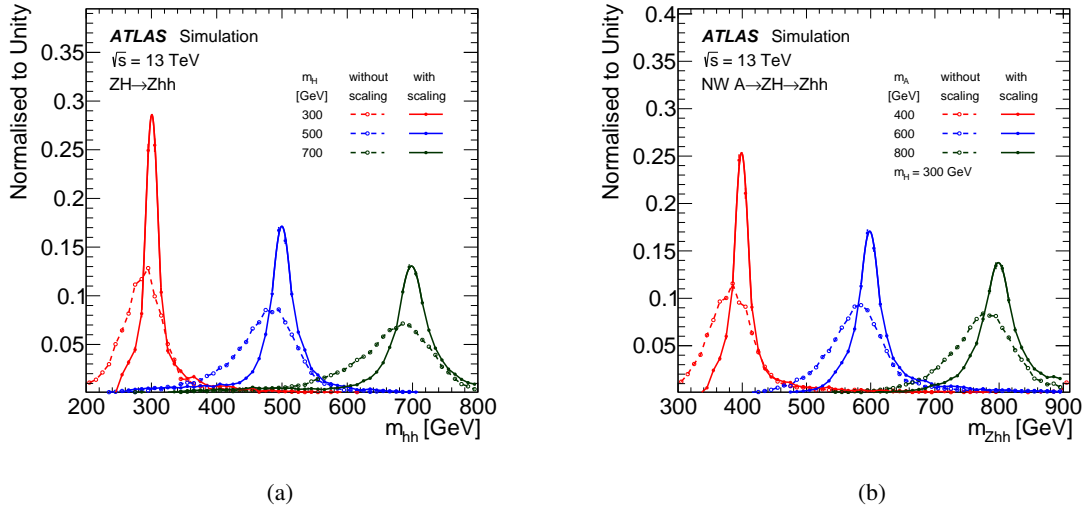the rescaled *b*-jet momenta are used solely to calculate the masses of resonance candidates.

(a) (b)

Figure 5: Reconstructed invariant mass distributions of (a) the hh system in the ZH search and (b) the Zhh system in the $A \to ZH$ search with a NW A boson for a few representative signal mass points in the 2L channel. Distributions before the m_h rescaling are shown as open circles and dashed lines while those after the rescaling are shown as solid circles and solid lines. All distributions are normalised to unity. A NW A boson has a negligible total decay width compared with the experimental mass resolution.

As discussed in Section 5.3, the m_{hh} variable is not used to construct BDTs for the VH and $A \to ZH$ searches. Instead, m_{hh} is required to be in a window around the target m_H value afterwards. The window sizes are optimised to improve the search sensitivities. They vary from 30 GeV at $m_H = 260$ GeV to 220 GeV at $m_H = 1000$ GeV, corresponding to approximately 3–8 times of the expected m_{hh} resolution after the rescaling. At high resonance masses the backgrounds are small, so the mass windows are widened (relative to the resolution) to increase signal efficiencies. The same m_H-dependent m_{hh} windows are used for all channels and for both the VH and $A \to ZH$ searches.

The $A \to ZH$ model has an additional resonance, the A boson. In the 2L channel, the $A \to ZH \to \ell\ell hh$ decay can be fully reconstructed. For A bosons with a width significantly smaller than the detector mass resolution, i.e. the NW case, the width of the Zhh mass distribution is dominated by the m_{hh} resolution as a result of the good lepton momentum resolution and the relatively narrow width of the Z boson. Therefore, the invariant mass of the Zhh candidate, m_{Zhh}, is required to fall in a window of the same size as the m_{hh} window for a given m_H, but shifted from the m_H value to the targeted m_A value. For A bosons with a width equal to 20% of the boson mass, i.e. the LW case, such requirements reduce the signal efficiencies substantially as the m_{Zhh} distributions are broadened by the A boson width. In this case, $m_{Zhh} > 475$ GeV is required only for $m_A \geq 500$ GeV to reduce backgrounds which are present mostly at low m_{Zhh} values. In the 0L channel, the $A \to ZH \to \nu\nu hh$ decay cannot be reconstructed fully, due to the escaping neutrinos. Instead, the transverse mass of the hh and E_T^{miss} system, m_T^{Zhh}, is used. However, the m_T^{Zhh} distributions are either too broad in the case of LW A bosons or too severely sculpted by the phase-space limitation for NW A bosons with small mass-splittings between the A and H bosons to be effective in discriminating between signal and backgrounds. Thus a requirement on m_T^{Zhh} is applied only for NW A bosons with $m_A - m_H \geq 200$ GeV. In this case, m_T^{Zhh} must be in the window of $[m_A - 150, m_A + 50]$ GeV for simplicity.

Figure 6 compares the data and expected background distributions of m_{hh} and m_{T}^{Zhh} in the 0L SR, and those of m_{hh} and m_{Zhh} in the 2L SR. Expected distributions from an example signal model are overlaid.

6 Systematic uncertainties

Systematic uncertainties are divided into four categories: experimental uncertainties, theoretical uncertainties of the overall background normalisations, theoretical uncertainties of acceptances and BDT shapes, and data-driven background modelling uncertainties.

6.1 Experimental systematic uncertainties

Jet energy scale (JES) and jet energy resolution (JER) uncertainties comprise the largest group of experimental uncertainties. The JES uncertainties are primarily determined using data-based Z-boson–jet, photon–jet, and multi-jet p_T-balance techniques [86]. Additional uncertainties are applied for the energy scale of jets containing b-quarks. The impact of the JES uncertainties is estimated by scaling the jet energies within their uncertainties. JER uncertainties are also determined from in situ measurements of Z-boson–jet , photon–jet, and dijet p_T balance [86]. The effect of the JER uncertainties is calculated by increasing the resolution within its uncertainties, smearing the jet energy by the resulting change in resolution, and comparing the result with the nominal shape and normalisation in simulation.

Subdominant experimental uncertainties originate from the b-tagging correction factors. The b-tagging correction factors, determined from the difference between the efficiencies measured in data and simulation, are evaluated in five DL1r discriminant bins and are derived separately for b-jets, c-jets, and light-flavour jets [78, 87, 88]. All of the correction factors for the three jet flavours have uncertainties estimated from multiple measurements, which are decomposed into uncorrelated components that are then treated independently. Extra uncertainties are applied when 'truth'-tagging is used (in the highest BDT-score bins for subleading backgrounds), and these are decorrelated between channels and defined as an overall 30% uncertainty in the yield for these backgrounds. The 'truth'-tagging uncertainties are designed to cover any differences in BDT shapes between background samples when 'truth'-tagging is or is not applied.

Uncertainties in the reconstruction, identification, isolation, and trigger efficiencies of electrons [89] and muons [90] are considered, along with the uncertainty in their energy scale and resolution. These are found to have only a small impact on the results. The uncertainties in the energy scale and resolution of the jets and leptons are propagated to the calculation of E_T^{miss}, which also has additional uncertainties from the modelling of the underlying event and the momentum scale, momentum resolution, and reconstruction efficiency of the tracks used to compute the soft-term (see Section 4) [82, 83]. An uncertainty is assigned to the E_T^{miss} trigger correction factors in the 0L channel, defined as the entire difference between the trigger efficiencies for data and simulated events. The uncertainty in the combined 2015–2018 integrated luminosity is 1.7%, as described in Section 3. The average number of interactions per bunch crossing in the simulation is rescaled by 1.03 to improve agreement between the data and the simulation, based on a measurement of the visible cross-section in minimum-bias events [91], and an uncertainty, as large as the correction, is included.
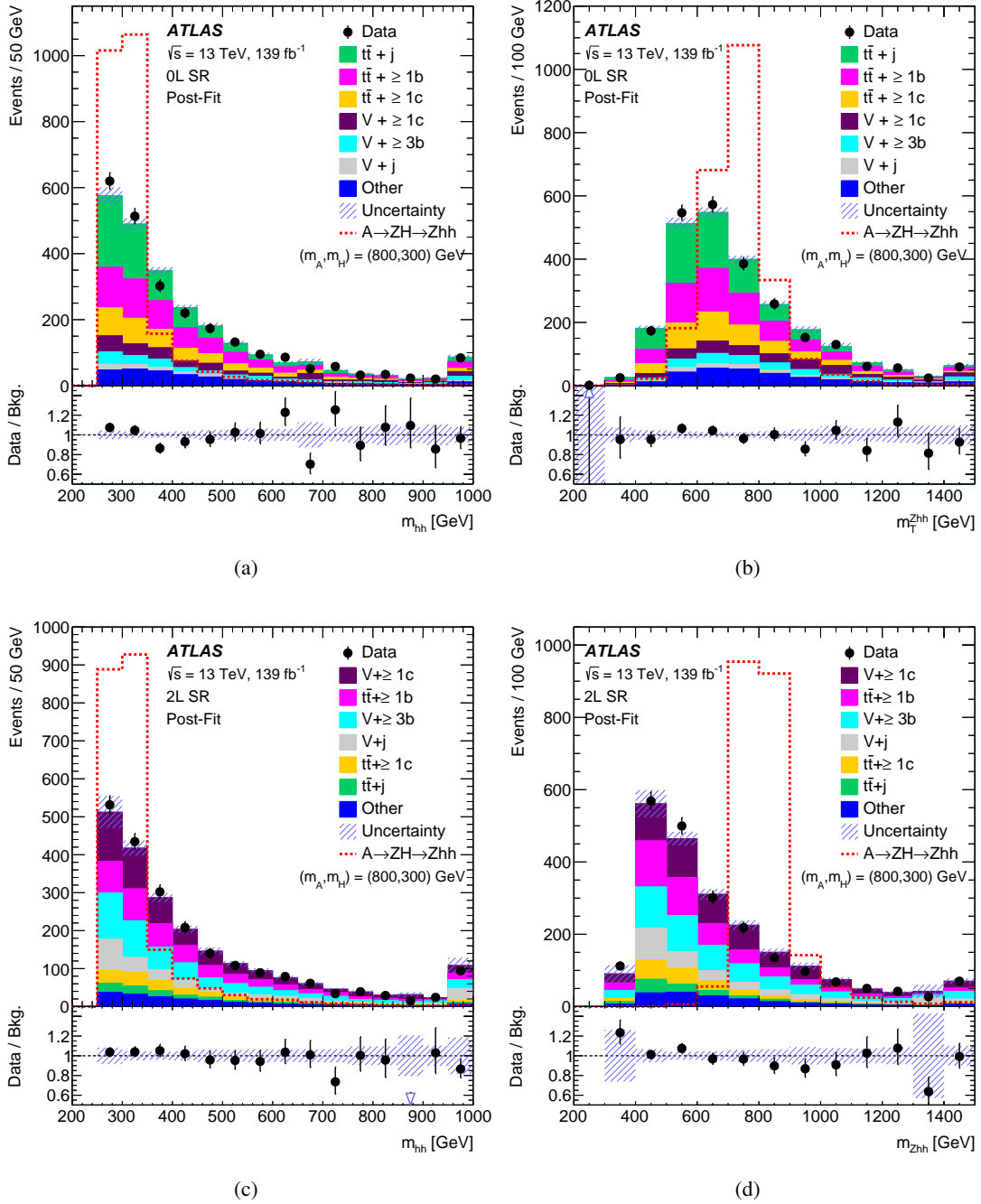
Figure 6: Comparisons between mass distributions of the data and the expected backgrounds: (a) m_{hh} and (b) m_T^{Zhh} of the 0L SR, and (c) m_{hh} and (d) m_{Zhh} of the 2L SR. All masses are calculated after rescaling the measured Higgs boson candidate mass to 125 GeV. The backgrounds are obtained from the background-only fits to the control and signal regions discussed in Section 7. Expected distributions from the $A \rightarrow ZH$ signal at $(m_A, m_H) = (800, 300)$ GeV for a NW A boson, normalised to the total background expectations, are overlaid. The final bins include overflows. The bottom panels show the ratios of the data to the total background expectations. The hatched bands represent the combined statistical and systematic uncertainties in the total background predictions.

6.2 Background normalisation systematic uncertainties

For all background processes, uncertainties are included in the overall normalisations. These uncertainties are correlated across the three leptonic channels. The $t\bar{t} + \geq 1b$ and $V+ \geq 3b$ flavour components are free to float in the fit. The $t\bar{t} + \geq 1c$ and $V+ \geq 1c$ flavour components are assigned a large uncertainty, 100%, intended to be conservative in case there is significant mismodelling of these processes. These uncertainties are always constrained by the fits. The light-flavour $t\bar{t} + j$ and $V + j$ components are assigned a smaller uncertainty, 10%, as these are better-measured processes. In practice, these overall $t\bar{t} + j$ and $V + j$ uncertainties have little impact on the analysis, since the yields from these backgrounds, particularly at high BDT score, are small. These uncertainties are all constrained by the $\sum s_{b\text{-tag}}^{\text{pc}}$ distributions in $\text{CR}_{t\bar{t}}$ and $\text{CR}_{V+\text{jets}}$, as well as by the low BDT-score regions of the SRs.

The above $V +$ jets uncertainties are correlated between $\gamma +$ jets, $Z +$ jets, and $W +$ jets backgrounds. Extra uncertainties are defined for extrapolations between these backgrounds to account for their kinematic differences. Each flavour component of $Z +$ jets is assigned an extra 20% uncertainty, and each component of $W +$ jets is assigned an extra 30% uncertainty. The sizes of the uncertainties are defined so as to cover flavour-composition differences between the $\gamma +$ jets, $Z +$ jets, and $W +$ jets backgrounds.

The remaining normalisation uncertainties are subdominant in the analysis. A 100% uncertainty is applied to the multi-jet backgrounds in both the 1L SRs and $\text{CR}_{V+\text{jets}}$. The 1L SR multi-jet uncertainty has a small impact due to the size of the background, and the $\text{CR}_{V+\text{jets}}$ multi-jet background is constrained by the fits to have an uncertainty smaller than 100%. Uncertainties of 25% are applied to each of the remaining small SM backgrounds. Given the small relative yields from these backgrounds, their uncertainties have little impact on the analysis and are thus chosen to be conservative.

6.3 Data-driven background modelling uncertainties

A set of extra data-driven background uncertainties is included. The relative differences between data and simulation in the validation region BDT distributions define a set of VR non-closure uncertainties, which are determined prior to any fit. In each of the 0L, 1L (1L+ and 1L−), and 2L regions, two uncertainties are defined: one for normalisation and one for shape, each of which is correlated between all backgrounds in the channel. In practice, each signal region tends to be dominated by a single background (either $t\bar{t}$ or $V +$ jets), such that correlating these uncertainties between backgrounds makes little difference compared to decorrelating the uncertainties between backgrounds. In $\text{CR}_{t\bar{t}}$, only the normalisation component of the non-closure is considered, since the individual normalisation factors for each $t\bar{t}$ flavour component can cover any shape difference between data and simulation. In $\text{CR}_{V+\text{jets}}$, no non-closure uncertainty is included, as the normalisation component is redundant because of the multi-jet normalisation uncertainty, which by design covers any non-closure between data and simulation. In total, this defines nine components of the validation region non-closure uncertainty, each of which is taken as 100% of the relative difference in normalisation or shape between data and simulation in the validation regions. These uncertainties are designed to be conservative, and in practice they tend to be constrained by the fits.

For the multi-jet backgrounds in $\text{CR}_{V+\text{jets}}$ and the 1L SRs, shape uncertainties are included in addition to the normalisation uncertainties defined in Section 6.2. The shape uncertainty included in the $\text{CR}_{V+\text{jets}}$ multi-jet template is obtained by increasing or decreasing the $\gamma +$ jets contamination by 100% in this region. Two shape uncertainties are included in the multi-jet template in the 1L SRs: one by increasing or decreasing the prompt-lepton contamination by 30% when evaluating the template, to account for potential

mismodelling of the lepton selection inefficiency, and the other by raising or lowering the m_{T}^{W} threshold used to determine this template, to account for the potential impact of the selection on the extracted BDT shapes.

6.4 Theoretical uncertainties in BDT shapes and acceptances

In addition to the normalisation uncertainties defined in Section 6.2, a number of theoretical uncertainties in background and signal modelling are included. These uncertainties tend to be subdominant relative to the background normalisation and data-driven uncertainties, which are in turn subdominant relative to the statistical uncertainty of the data. Each uncertainty is correlated across all analysis regions, and the effects of each uncertainty on the acceptance and BDT shape are correlated.

For each component of the $t\bar{t}$ and $V+$jets backgrounds and the single-top-quark background, scale uncertainties are included, defined [92] using seven variations of the QCD factorisation and renormalisation scales in the matrix elements by factors of 0.5 and 2, avoiding variations in opposite directions. For each component of the $t\bar{t}$ background and the single-top-quark background, an extra parton-shower uncertainty is defined by comparing the nominal sample with an alternative sample showered with HERWIG 7.04 [93, 94], and an extra matching uncertainty is defined by comparison with an alternative sample produced with MADGRAPH5_AMC@NLO. For each $t\bar{t}$ component, initial-state radiation (ISR) and final-state radiation (FSR) variations are included, following the procedure described in Ref. [95]. For the Wt component of the single-top-quark background, a comparison between the nominal sample, which uses the diagram removal scheme [96], and an alternative sample, which uses the diagram subtraction scheme [97], defines an extra uncertainty. For $V+$jets, variations of the CKKW parameter for merging/matching the matrix element with the parton shower are included, as are variations of the resummation scale.

Modelling of the $t\bar{t}$ background is improved by correcting the top-quark p_{T} distribution to that predicted by calculations of top-quark-pair differential distributions at NNLO QCD and NLO EW accuracy [98]. Previous studies have seen improved agreement between data and prediction in $t\bar{t}$ events, particularly for the top-quark p_{T} distribution, when comparing the data with the NNLO calculations [99]. For each $t\bar{t}$ component, the change in the BDT shape from the NNLO correction is taken as an uncertainty.

For the signal samples, acceptance uncertainties evaluated at the generator level are included to account for scale variations, PDF variations, and variations of the ISR, FSR, multi-parton interaction (MPI), and colour reconnection parameters [54]. For the non-resonant Zhh signal only, an extra uncertainty is included to cover potential mismodelling of $gg \rightarrow Zhh$. An uncertainty is included in the Zhh normalisation and BDT shape, covering the difference between the $qq \rightarrow Zhh$ process alone and the sum of the $gg \rightarrow Zhh$ and $qq \rightarrow Zhh$ processes (the effects on the normalisation and shape are correlated). No other uncertainties are included in the overall cross-section for any signal sample.

6.5 Impact of systematic uncertainties

The effects of the statistical and systematic uncertainties on the search sensitivities are studied for hypothesised signals following the procedure discussed in Section 7. Table 4 lists the leading sources of uncertainty and shows, for a few selected signal models, the expected relative uncertainties in the fitted signal-strength parameter μ, a factor multiplying the predicted cross-section for the hypothesised signal. In the resonant case, the mass values are chosen to be those with the largest excesses observed in the data as discussed in Section 7. The cross-sections used for signal normalisation correspond approximately to their

expected upper limits. In each case, the leading source of uncertainty is the statistical uncertainty of the data.

Table 4: Dominant uncertainties in the best-fit signal-strength parameter μ for hypothesised signals. For this study, the resonance masses are chosen to be those with the largest excesses in the data for illustration: $m_H = 315$ GeV for WH, $m_H = 550$ GeV for ZH, $(m_A, m_H) = (790, 300)$ GeV for NW $A \to ZH$, and $(m_A, m_H) = (420, 320)$ GeV for LW $A \to ZH$. The hypothesised signals are normalised to reference cross-sections approximately equal to their expected upper limits.

Model	Vhh like in SM	WH	ZH	NW $A \to ZH$	LW $A \to ZH$
Systematic uncertainty source			$\Delta\mu/\mu$ [%]		
Background modelling	+20, −15	+14, −11	+4.7, −3.0	+17, −13	+20, −18
MC statistics	+12, −9.1	+13, −7.8	+4.8, −2.2	+7.2, −4.1	+10, −8.3
Objects	+12, −8.6	+8.0, −5.2	+4.5, −2.2	+19, −11	+16, −12
Signal modelling	+10, −4.7	+12, −4.9	+8.6, −3.0	+14, −5.1	+17, −7.6
VR non-closure	+14, −11	+11, −9.4	+4.4, −3.0	+4.9, −3.7	+12, −10
Total systematic uncertainty	+30, −22	+27, −18	+12, −5.8	+30, −18	+33, −24
Statistical uncertainty	+44, −39	+52, −43	+68, −49	+59, −47	+42, −37
Total	+52, −44	+59, −47	+69, −49	+66, −50	+53, −45

7 Results and interpretations

Potential signal contributions in the data are determined through maximum-likelihood fits to the BDT distributions in the SRs and the $\sum s_{b\text{-tag}}^{\text{pc}}$ distributions in $\text{CR}_{t\bar{t}}$ and $\text{CR}_{V+\text{jets}}$. The procedure is based on the framework described in Refs. [100–102]. A profile-likelihood-ratio test statistic is used to test the signal-plus-background hypothesis, with the signal production rate as the parameter-of-interest. All SRs (0L, 1L, and 2L) are included in the fits for the non-resonant Vhh search and the resonant VH search, while only the 0L and 2L SRs are included for the resonant $A \to ZH$ search. The BDT distributions are divided into four bins, determined through optimisations of signal sensitivities while maintaining a reasonable number of background MC events in each bin. The binning boundaries are optimised for each channel separately, but are kept the same for different signal models for simplicity.

The $t\bar{t}$ and $V+$ jets backgrounds in the SRs from MC simulation are decomposed into three jet-flavour categories in the same way as those in the $t\bar{t}$ and $V+$ jets CRs discussed in Section 5.2. These flavour-dependent contributions share the same normalisation factors (NFs) as their corresponding components in the CRs. In the fits, the NFs are unconstrained for the $t\bar{t} + \geq 1b$ and $V+ \geq 3b$ components and are constrained to unity within their estimated uncertainties for other components. Systematic uncertainties, described in Section 6, are incorporated as additional multiplicative terms, parameterised with nuisance parameters, in the likelihood calculations, where each nuisance parameter is given a prior distribution based on individual studies.

To test the overall compatibility of the data with the background expectations, fits are first performed for the background-only hypothesis. Since BDTs are model dependent, a fit to the BDT and $\sum s_{b\text{-tag}}^{\text{pc}}$ distributions for the non-resonant search is chosen to illustrate the background modelling. Figure 4 compares the data with the background expectations from the fit for a few selected kinematic variables used in training the BDTs. Overall, the post-fit backgrounds are found to reproduce the data well. The fitted NF values of the

heavy-flavour components of the $t\bar{t}$ and $V+$jets backgrounds are listed in Table 5, along with those from fits to the different signal-plus-background hypotheses discussed below.

The fits are repeated for the signal-plus-background hypotheses for each signal model and resonance mass assumption. The NF values shown in Table 5 from different fits are stable and consistent. Upper limits on signal production cross-sections are calculated with the CL_s method [103], using the $\tilde{q}_\mu$ test statistic in the asymptotic approximation [104].

Global significances are calculated following the procedure given in Ref. [105]. Pseudo-experiments are generated for the background-only hypothesis. For each pseudo-experiment, a scan over all considered mass points is performed to determine the largest local significance for that pseudo-experiment. The fraction of pseudo-experiments with a local significance greater than the maximum local significance found in the scan over the real data defines the global p-value, which in turn defines the global significance.

Table 5: Values of the normalisation factors of the heavy-flavour background components from the simultaneous fits of the signal and control regions to the background-only and different signal-plus-background hypotheses. For this study, the resonance masses with the largest excesses in the data are chosen: $m_H = 315$ GeV for WH, $m_H = 550$ GeV for ZH, $(m_A, m_H) = (790, 300)$ GeV for NW $A \to ZH$, and $(m_A, m_H) = (420, 320)$ GeV for LW $A \to ZH$.

Background source	Normalisation constraint	Bkg-only fit	Vhh like in SM	VH		$A \to ZH$	
				WH	ZH	NW	LW
$t\bar{t} + \geq 1b$	free-float	$1.21^{+0.19}_{-0.16}$	$1.21^{+0.19}_{-0.17}$	$1.20^{+0.21}_{-0.18}$	$1.18^{+0.22}_{-0.18}$	$1.29^{+0.30}_{-0.23}$	$1.27^{+0.26}_{-0.22}$
$t\bar{t} + \geq 1c$	$\pm 100\%$	$1.54^{+0.62}_{-0.54}$	$1.51^{+0.64}_{-0.53}$	$1.56^{+0.68}_{-0.57}$	$1.44^{+0.61}_{-0.49}$	$1.65^{+0.70}_{-0.63}$	$1.85^{+0.57}_{-0.74}$
$V + \geq 3b$	free-float	$0.98^{+0.28}_{-0.30}$	$0.86^{+0.33}_{-0.30}$	$0.83^{+0.35}_{-0.29}$	$0.90^{+0.31}_{-0.30}$	$0.76^{+0.38}_{-0.30}$	$0.72^{+0.38}_{-0.28}$
$V + \geq 1c$	$\pm 100\%$	$1.29^{+0.55}_{-0.40}$	$1.17^{+0.60}_{-0.35}$	$1.09^{+0.61}_{-0.33}$	$1.24^{+0.59}_{-0.39}$	$1.01^{+0.63}_{-0.31}$	$0.97^{+0.60}_{-0.30}$

7.1 Search for non-resonant Vhh production

The BDTs for non-resonant production are used to search for Vhh signals from three production scenarios: 'SM', κ_λ, and κ_{2V}. The 'SM' scenario assumes SM kinematics but with its cross-section scaled by a signal-strength parameter μ. The κ_λ scenario tests for an anomalous tri-linear hhh coupling, assuming SM couplings for the rest. Similarly, the κ_{2V} scenario tests for an anomalous quartic $hhVV$ coupling. For the κ_λ and κ_{2V} scenarios, both the event kinematics and production cross-section depend on their respective coupling modifier.

Constraints on non-resonant Vhh production are obtained through fits to the signal-plus-background hypotheses described above, assuming the SM value for the $h \to bb$ decay branching ratio [10]. The three non-resonant scenarios have the same data BDT distribution, but differ in the signal BDT distributions. The background BDT distributions are largely the same, barring small variations in the post-fit background contributions. For 'SM' Vhh production, a 95% confidence-level (CL) upper limit of 183 on μ is observed compared with 87^{+41}_{-24} expected. The corresponding post-fit BDT distribution is shown in Figure 7(a). For the κ_λ and κ_{2V} scenarios, Vhh cross-section upper limits are derived from the fits for different values of the coupling modifiers. These limits lead to the observed (expected) 95% CL intervals of $-34.4 < \kappa_\lambda < 33.3$ $(-24.1 < \kappa_\lambda < 22.9)$ and $-8.6 < \kappa_{2V} < 10.0$ $(-5.7 < \kappa_{2V} < 7.1)$ for the two coupling modifiers. The observed bounds are weaker than the expectations largely because of small excesses of data in the highest

BDT bins. These are the first limits derived from the Vhh process, and are considerably weaker than those obtained from the hh searches focused on the ggF and VBF processes [11–16]. In addition, this analysis can search for deviations of the $WWhh$ and $ZZhh$ couplings from their SM values, parameterised by the respective coupling modifiers κ_{2W} and κ_{2Z} (in the SM, $\kappa_{2W} = 1$, and $\kappa_{2Z} = 1$). Cross-section upper limits are derived from the fits for different values of these coupling modifiers, leading to limits for the observed (expected) 95% CL intervals of $-12.3 < \kappa_{2W} < 13.5$ ($-8.6 < \kappa_{2W} < 9.8$) and $-9.9 < \kappa_{2Z} < 11.3$ ($-7.1 < \kappa_{2Z} < 8.5$) for the two coupling modifiers. Higgs boson couplings other than the one being tested are set to their SM values.

7.2 Searches for $VH \rightarrow Vhh$ production

Constraints on the production of a heavy narrow scalar resonance H in association with a V boson are determined through the fits of the BDT distributions for resonant VH production to the signal-plus-background hypothesis in 5 GeV m_H steps. The step size is chosen to be comparable to, or smaller than, the experimental m_{hh} mass resolution. For each tested m_H value, the BDT distributions are obtained after imposing the m_{hh} mass window requirement discussed in Section 5.4. For m_H values with no corresponding MC signal sample, the BDT distributions are linearly interpolated from those of the two closest neighbouring mass points with MC signal samples. To validate this interpolation, the results of fits performed at m_H points with a MC signal sample are compared with the results obtained when the BDT distribution is an interpolation between those of neighbouring points with MC signal samples.

Fits are performed separately for the $WH \rightarrow \ell\nu hh$ and $ZH \rightarrow (\ell\ell/\nu\nu)hh$ searches. All three channels are included in the fits for both searches. The WH signal contributes mostly to the 1L channel, but with a sizeable contribution in the 0L channel due to an inefficiency in lepton identification. Its contribution to the 2L channel is negligible. Inclusion of the 2L channel in the fit effectively makes the channel an additional CR for the $WH \rightarrow \ell\nu hh$ search, further constraining the Z + jets background. Similarly, the ZH signal contributes mostly to the 0L and 2L channels. Inclusion of the 1L channel helps to constrain the W + jets background.

The data are found to be consistent with the estimated background contributions. The largest upward deviations from the background expectations are at $m_H = 315$ GeV with a local (global) significance of 2.5 (1.3) standard deviations (σ) in the WH search and at $m_H = 550$ GeV with a local (global) significance of 2.7σ (1.3σ) in the ZH search. These small excesses are largely correlated with those observed in the non-resonant Vhh search. The post-fit BDT distribution in the WH search at $m_H = 315$ GeV is illustrated in Figure 7(b). The heavy-flavour background NFs from the fits are compared with those from other fits in Table 5.

The observed and expected 95% CL upper limits on the cross-section $\sigma(VH) \times B(H \rightarrow hh \rightarrow bbbb)$ as a function of m_H are shown in Figure 8. The resonant VH search is sensitive to the HWW and HZZ couplings separately. Compared with the search in the VBF channel [15], which is sensitive to the combination of the two couplings, the VH search has better sensitivity for m_H up to ~ 450 GeV, assuming that SU(2) custodial symmetry (like that in the SM) applies to the HWW and HZZ couplings [107].
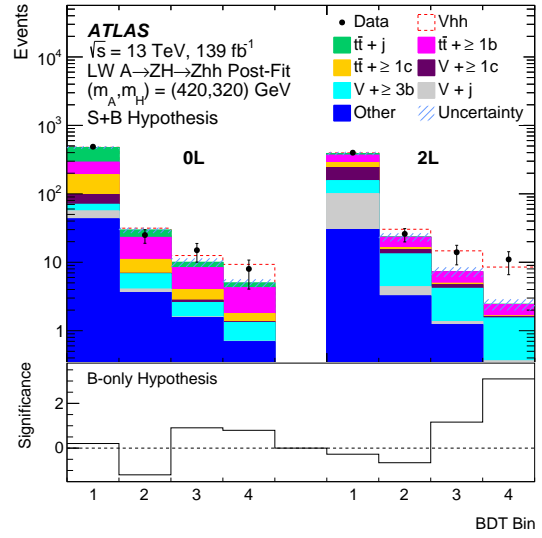
7.3 Search for $A \rightarrow ZH$ production

The $A \rightarrow ZH$ search follows a strategy similar to that used in the VH search. The $A \rightarrow ZH$ BDT distributions in the 0L and 2L channels, after applying the mass requirements discussed in Section 5.4,
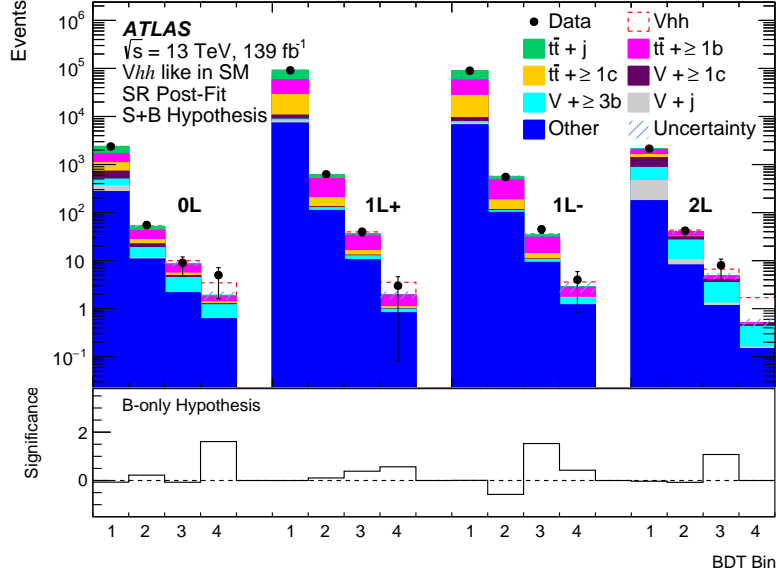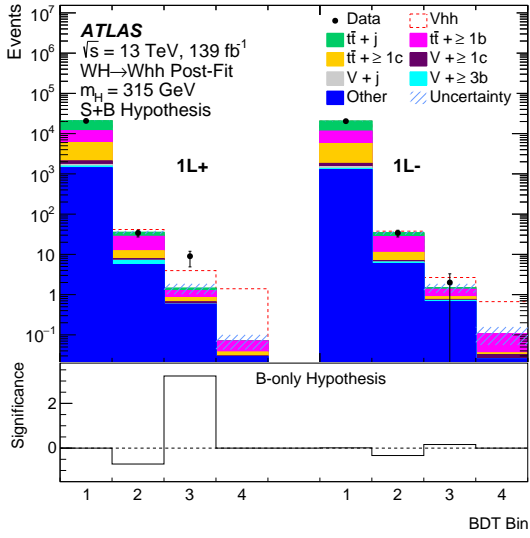
Figure 7: The post-fit BDT distributions of the signal-plus-background hypotheses from the searches for (a) non-resonant Vhh production like in the SM, (b) the resonant WH process at $m_H = 315$ GeV, and (c) the resonant $A \rightarrow ZH$ process at $(m_A, m_H) = (420, 320)$ GeV for a LW A boson. The chosen mass points in (b) and (c) correspond to the most significant excesses of data over the background expectations in their respective search. The hatched bands represent the combined statistical and systematic uncertainties in the total background predictions. The bottom panels show the observed bin-by-bin significances for the background-only hypotheses, calculated following the prescription of Ref. [106]. The LW A boson has a total decay width equal to 20% of its mass.
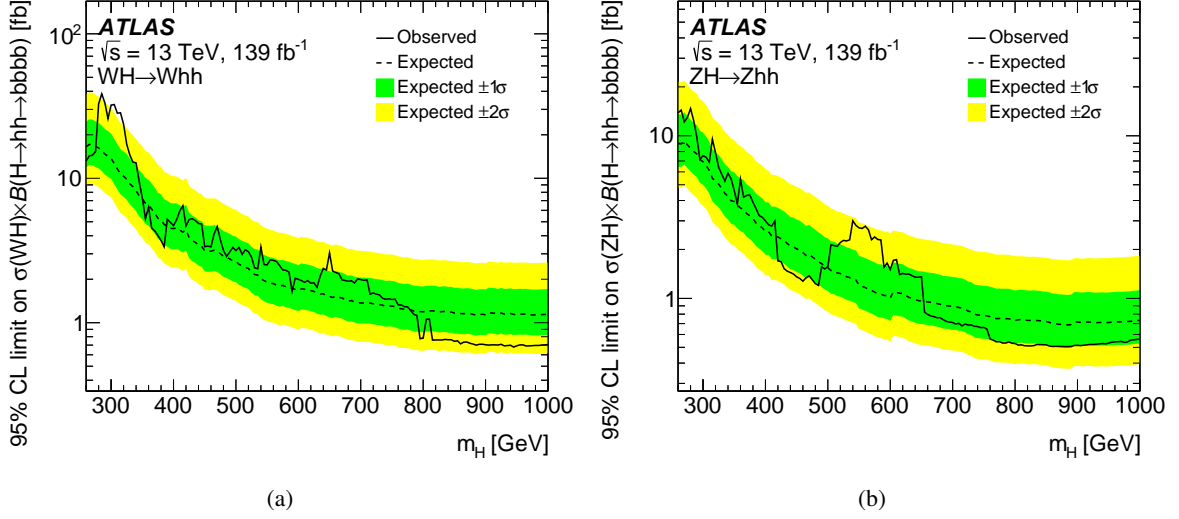
Figure 8: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross-section at $\sqrt{s}$ = 13 TeV of a heavy narrow scalar resonance H in the decay mode $H \rightarrow hh \rightarrow bbbbb$ in association with (a) a W boson and (b) a Z boson as a function of the resonance mass. The green (inner) and yellow (outer) bands represent $\pm1\sigma$ and $\pm2\sigma$ uncertainty in the expected limits.

are used to constrain $gg \rightarrow A \rightarrow ZH \rightarrow Zhh \rightarrow Zbbbb$ production for each (m_A, m_H) hypothesis with 10 GeV steps in both masses, separately for NW and LW A bosons. The BDT distributions for (m_A, m_H) hypotheses without MC samples are linearly interpolated from those of the four closest neighbouring mass points with MC samples. For a NW A boson, the expected and observed upper limits in the (m_A, m_H) plane are shown in Figure 9(a) and Figure 9(b) respectively. The upper limits for a LW A boson, with a width of 20%, are shown in Figure 9(c) (expected) and Figure 9(d) (observed).

The most significant excesses are observed at (m_A, m_H) = (790, 300) GeV with a local (global) significance of 3.9σ (2.1σ) in the NW scenario and at (m_A, m_H) = (420, 320) GeV with a local (global) significance of 3.8σ (2.8σ) in the LW scenario. Figure 7(c) shows the post-fit BDT distributions in the LW scenario from the search at (m_A, m_H) = (420, 320) GeV. Due to the selection for the LW scenario, there is little sensitivity to either the A boson mass or width, and a real LW signal would likely produce a higher-than-otherwise-expected cross-section upper limit over a wide range of probed A boson mass values. Thus the broad excess in the A boson mass distribution in the LW scenario is consistent with the expectation for a signal.

The 2HDM benchmark used for interpretations has four free parameters: m_A, m_H, $\tan\beta$, and $\cos(\beta - \alpha)$, where $\tan\beta$ is the ratio of the vacuum expectation values of the two doublets and α is the mixing angle of the CP-even Higgs bosons. The limiting case of $\cos(\beta - \alpha) \rightarrow 0$ corresponds to the 2HDM weak decoupling limit [108] in which the lightest CP-even Higgs boson has the same couplings as the SM Higgs boson at the lowest order. The $A \rightarrow ZH$ search results are interpreted as constraints in the plane defined by $\cos(\beta - \alpha)$ and m_A for given m_H and $\tan\beta$ values. For the remaining 2HDM parameters, the mass of the charged Higgs boson is set to be equal to m_A and the potential parameter m_{12}^2 is set to $m_A^2 \tan\beta/(1 + \tan^2\beta)$.

For regions relevant to the sensitivity of this search, the natural width of the H boson remains narrow, especially where $\cos(\beta - \alpha)$ is close to zero. It is estimated that the cross-section upper limits should be

valid as long as the natural width of the H boson is less than 1% and, therefore, the search constrains only parts of the 2HDM parameter space that conform with this requirement. On the other hand, the natural width of the A boson varies from narrow to about 20%. This means that cross-section limits have to be calculated for a range of A boson natural widths in the (m_A, m_H) plane, which are subsequently interpreted in the 2HDM planes discussed previously. The signal hypothesis is tested for several values of the A boson natural width, and cross-section upper limits for those widths are derived as a function of m_A. Linear interpolation is used to derive the limit for any natural width between a pair of tested width values.

(a)

(b)

(c)

(d)
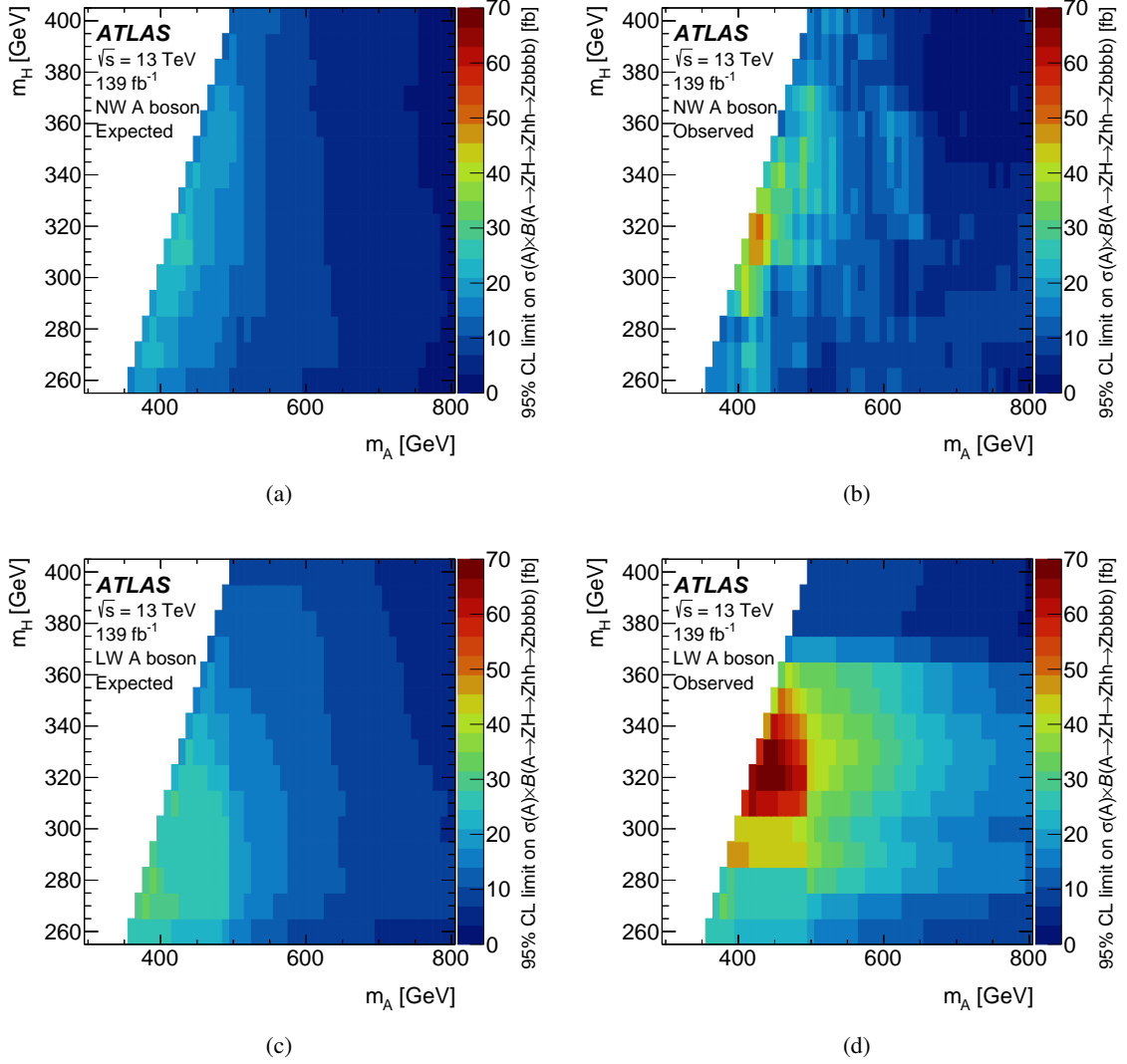
Figure 9: Upper bounds at 95% CL on $\sigma(A) \times B(A \to ZH \to Zhh \to Zbbbb)$ in the (m_A, m_H) plane for (a, b) a NW A boson and (c, d) a LW A boson. The expected upper limits are shown in (a) and (c) while the observed limits are shown in (b) and (d). The A boson has a total decay width that is negligible compared to the experimental mass resolution in the NW case and is 20% of its mass in the LW case.

To interpret these results in the 2HDM, the upper limits on the cross-section are compared with the theoretical predictions of the model. In the type-I and lepton-specific 2HDMs, only gluon–gluon fusion

production of the A boson, $gg \to A$, is relevant, and it is calculated with corrections at up to NNLO in QCD as implemented in SusHi [109–112]. The widths and branching ratios of the Higgs bosons (A, H, and h) are calculated using the 2HDMC code [113]. The procedure used to calculate the cross-sections and branching ratios, as well as to choose the 2HDM parameter values, follows Ref. [10]. The upper limits are shown for some representative values in Figure 10 for the type-I 2HDM and in Figure 11 for the lepton-specific 2HDM. The Hhh coupling vanishes at $\cos(\beta - \alpha) = 0$, a feature which is reflected by the inability of this analysis to exclude this region of the $(\cos(\beta - \alpha), m_A)$ plane. For $\tan\beta = 1$ the sensitivity is the same for both 2HDM scenarios and therefore the corresponding results are omitted from Figure 11. The sensitivity of this search is complementary to that of the $A \to ZH \to \ell\ell bb / \ell\ell WW$ search [34] and to the constraints from the Higgs boson coupling measurements [4].
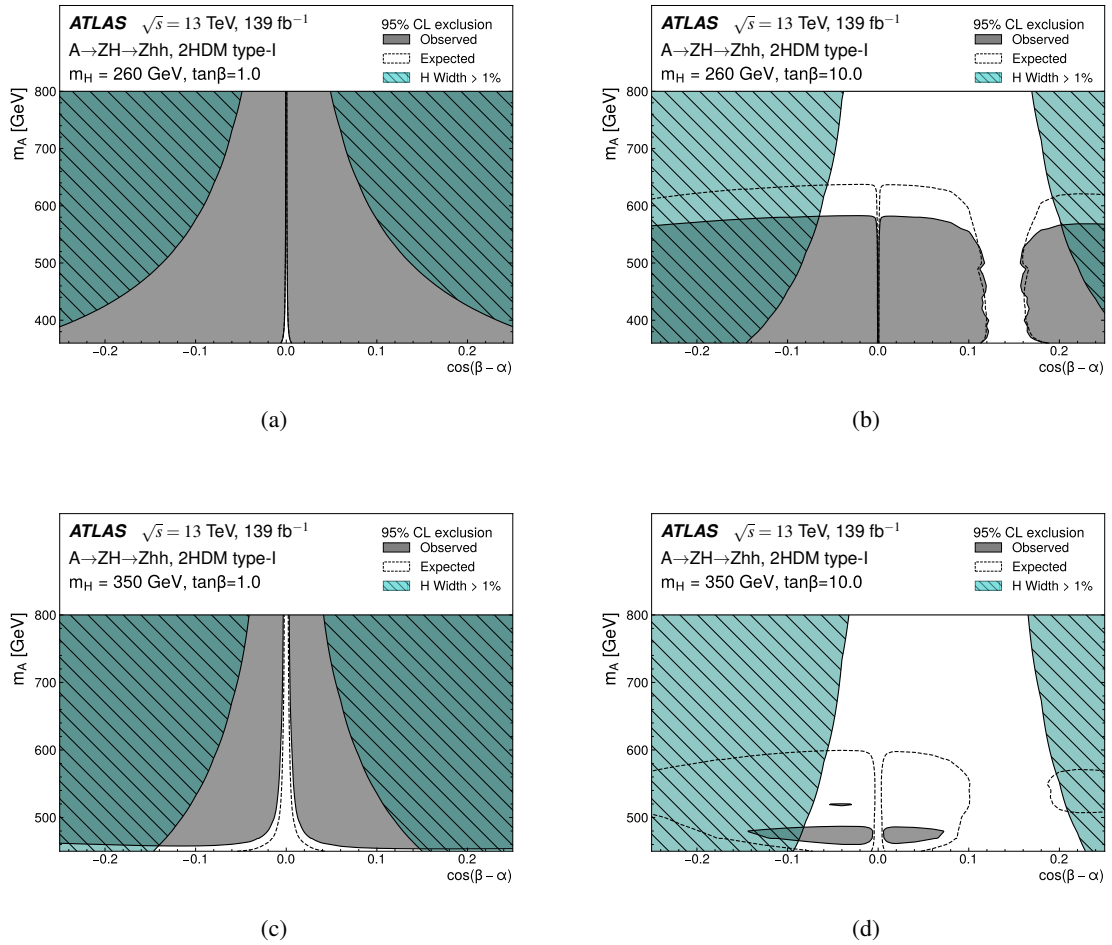
Figure 10: Interpretation of the upper limits on $\sigma(A) \times B(A \to ZH \to Zhh \to Zbbbb)$ in the parameter space of the type-I 2HDM. The shaded areas correspond to 95% CL exclusion regions in the $(\cos(\beta - \alpha), m_A)$ plane for a given m_H and $\tan\beta$. The hatched area corresponds to natural widths of the H boson for which the upper limits are not valid. Panels (a) and (b) refer to $m_H = 260$ GeV and to $\tan\beta = 1$ and 10, respectively; (c) and (d) refer to $m_H = 350$ GeV and to $\tan\beta = 1$ and 10, respectively.
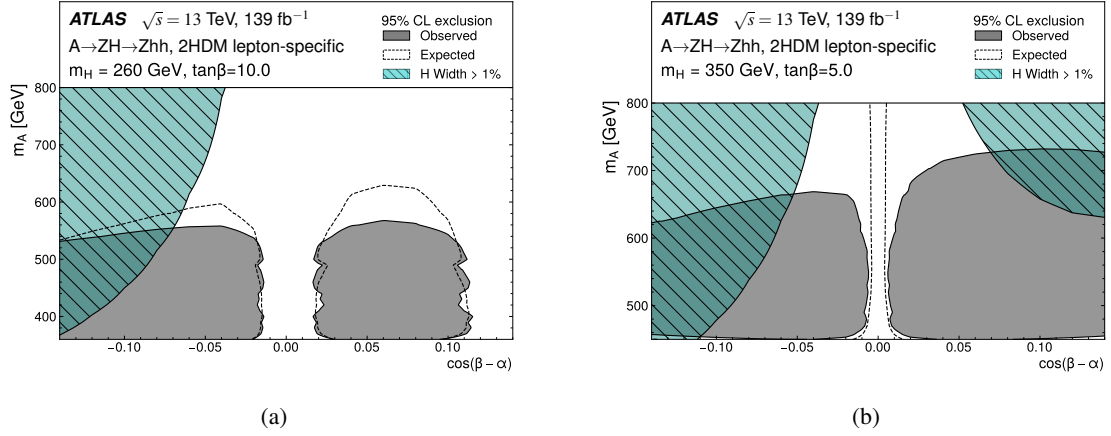
Figure 11: Interpretation of the upper limits on $\sigma(A) \times B(A \to ZH \to Zhh \to Zbbbb)$ in the parameter space of the lepton-specific 2HDM. The same notation as in Figure 10 is used. Panel (a) refers to $m_H = 260$ GeV, $\tan \beta = 10$ and (b) refers to $m_H = 350$ GeV, $\tan \beta = 5$.

8 Summary

Searches for Higgs boson pair production in association with a vector boson in pp collisions at $\sqrt{s} = 13$ TeV are performed using a data sample corresponding to an integrated luminosity of 139 fb^{-1}, recorded by the ATLAS experiment between 2015 and 2018 at the LHC. The Higgs bosons are identified via their decays into a pair of b-quarks and the vector bosons are required to decay into leptons, leading to final states with zero, one or two charged leptons along with four b-jets. The searches target both SM-inspired non-resonant hh production and BSM-motivated resonant hh production. The non-resonant Vhh search is carried out for scenarios with either SM kinematics but an enhanced production cross-section or modified Higgs boson couplings to vector bosons or itself. The resonant Vhh searches are designed for the production of a vector boson along with a heavy neutral scalar Higgs boson H decaying into hh, either directly or indirectly from the decay of another heavier neutral pseudoscalar Higgs boson A.

In general, the data are found to be in good agreement with the estimated background contributions, except for a few notable excesses. The most significant global excess is observed in the $gg \to A \to ZH \to Zhh$ search for a large-width A boson at $(m_A, m_H) = (420, 320)$ GeV, where the local (global) significance is 3.8 (2.8) standard deviations. More data are needed to ascertain the nature of this excess. Upper bounds on the Vhh production cross-sections are derived. For non-resonant production with SM kinematics, a 95% CL upper limit of 183 (87) is observed (expected) for the Vhh cross-section relative to its SM prediction. For resonant production, the observed (expected) upper limits are presented as a function of m_H in the range 260–1000 GeV for WH and ZH separately, and in the (m_A, m_H) plane for $A \to ZH$, covering the m_A range 360–800 GeV and m_H range 260–400 GeV. The constraints on $A \to ZH$ production are also interpreted in the $(\cos(\beta - \alpha), m_A)$ parameter space of type-I and lepton-specific two-Higgs-doublet models.

References

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

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

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

[4] ATLAS Collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, Nature **607** (2022) 52, arXiv: 2207.00092 [hep-ex].

[5] CMS Collaboration, *A portrait of the Higgs boson by the CMS experiment ten years after the discovery*, Nature **607** (2022) 60, arXiv: 2207.00043 [hep-ex].

[6] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[7] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[8] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.

[9] J. Baglio et al., *The measurement of the Higgs self-coupling at the LHC: theoretical status*, JHEP **04** (2013) 151, arXiv: 1212.5581 [hep-ph].

[10] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph].

[11] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **800** (2020) 135103, arXiv: 1906.02025 [hep-ex].

[12] CMS Collaboration, *Combination of Searches for Higgs Boson Pair Production in Proton–Proton Collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. Lett. **122** (2019) 121803, arXiv: 1811.09689 [hep-ex].

[13] ATLAS Collaboration, *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2021), arXiv: 2112.11876 [hep-ex].

[14] CMS Collaboration, *Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2021) 257, arXiv: 2011.12373 [hep-ex].

[15] ATLAS Collaboration, *Search for the $HH \to b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07** (2020) 108, arXiv: 2001.05178 [hep-ex], Erratum: JHEP **01** (2021) 145.

[16] CMS Collaboration, *Search for Higgs boson pair production in the four b quark final state in proton–proton collisions at* $\sqrt{s}$ = 13 *TeV*, (2022), arXiv: `2202.09617 [hep-ex]`.

[17] CMS Collaboration, *Search for nonresonant pair production of highly energetic Higgs bosons decaying to bottom quarks*, (2022), arXiv: `2205.06667 [hep-ex]`.

[18] T. D. Lee, *A Theory of Spontaneous T Violation*, Phys. Rev. D **8** (1973) 1226.

[19] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1, arXiv: `1106.0034 [hep-ph]`.

[20] F. Kling, J. M. No and S. Su, *Anatomy of exotic Higgs decays in 2HDM*, JHEP **09** (2016) 093, arXiv: `1604.01406 [hep-ph]`.

[21] L. Randall and R. Sundrum, *Large Mass Hierarchy from a Small Extra Dimension*, Phys. Rev. Lett. **83** (1999) 3370, arXiv: `hep-ph/9905221`.

[22] A. L. Fitzpatrick, J. Kaplan, L. Randall and L.-T. Wang, *Searching for the Kaluza-Klein graviton in bulk RS models*, JHEP **09** (2007) 013, arXiv: `hep-ph/0701150`.

[23] K. Agashe, H. Davoudiasl, G. Perez and A. Soni, *Warped gravitons at the CERN LHC and beyond*, Phys. Rev. D **76** (2007) 036006, arXiv: `hep-ph/0701186`.

[24] ATLAS Collaboration, *Search for resonant pair production of Higgs bosons in the* $b\bar{b}b\bar{b}$ *final state using pp collisions at* $\sqrt{s}$ = 13 *TeV with the ATLAS detector*, Phys. Rev. D **105** (2022) 092002, arXiv: `2202.07288 [hep-ex]`.

[25] CMS Collaboration, *Search for resonant pair production of Higgs bosons decaying to bottom quark-antiquark pairs in proton–proton collisions at* 13 *TeV*, JHEP **08** (2018) 152, arXiv: `1806.03548 [hep-ex]`.

[26] CMS Collaboration, *Search for a heavy Higgs boson decaying into two lighter Higgs bosons in the* $\tau\tau bb$ *final state at 13 TeV*, JHEP **11** (2021) 057, arXiv: `2106.10361 [hep-ex]`.

[27] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017.

[28] M. Mühlleitner, M. O. P. Sampaio, R. Santos and J. Wittbrodt, *Phenomenological comparison of models with extended Higgs sectors*, JHEP **08** (2017) 132, arXiv: `1703.07750 [hep-ph]`.

[29] B. Coleppa, F. Kling and S. Su, *Constraining type II 2HDM in light of LHC Higgs searches*, JHEP **01** (2014) 161, arXiv: `1305.0002 [hep-ph]`.

[30] G. C. Dorsch, S. J. Huber, K. Mimasu and J. M. No, *Echoes of the Electroweak Phase Transition: Discovering a Second Higgs Doublet through* $A_0 \rightarrow ZH_0$, Phys. Rev. Lett. **113** (2014) 211802, arXiv: `1405.5537 [hep-ph]`.

[31] CMS Collaboration, *Search for neutral resonances decaying into a Z boson and a pair of b jets or τ leptons*, Phys. Lett. B **759** (2016) 369, arXiv: `1603.02991 [hep-ex]`.

[32] ATLAS Collaboration, *Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the* $\ell\ell bb$ *final state in pp collisions at* $\sqrt{s}$ = 13 *TeV with the ATLAS detector*, Phys. Lett. B **783** (2018) 392, arXiv: `1804.01126 [hep-ex]`.

[33] CMS Collaboration, *Search for new neutral Higgs bosons through the* $H \rightarrow ZA \rightarrow \ell^+\ell^- b\bar{b}$ *process in pp collisions at* $\sqrt{s}$ = 13 *TeV*, JHEP **03** (2020) 055, arXiv: `1911.03781 [hep-ex]`.

[34] ATLAS Collaboration,
Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the $\ell\ell bb$ and $\ell\ell WW$ final states in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector,
Eur. Phys. J. C **81** (2020) 396, arXiv: 2011.05639 [hep-ex].

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

[36] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*,
ATLAS-TDR-19; CERN-LHCC-2010-013, 2010,
URL: https://cds.cern.ch/record/1291633, Addendum: ATLAS-TDR-19-ADD-1;
CERN-LHCC-2012-009, 2012, URL: https://cds.cern.ch/record/1451888.

[37] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*,
JINST **13** (2018) T05008, arXiv: 1803.00844 [physics.ins-det].

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

[39] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*,
ATL-SOFT-PUB-2021-001, 2021, URL: https://cds.cern.ch/record/2767187.

[40] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*,
Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex].

[41] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*,
JINST **15** (2020) P09015, arXiv: 2004.13447 [hep-ex].

[42] ATLAS Collaboration, *Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking*, JHEP **08** (2020) 080, arXiv: 2005.09554 [hep-ex].

[43] ATLAS Collaboration,
ATLAS data quality operations and performance for 2015–2018 data-taking,
JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det].

[44] ATLAS Collaboration,
Luminosity determination in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC,
ATLAS-CONF-2019-021, 2019, URL: https://cds.cern.ch/record/2677054.

[45] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823,
arXiv: 1005.4568 [physics.ins-det].

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

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

[48] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*,
ATL-PHYS-PUB-2016-017, 2016, URL: https://cds.cern.ch/record/2206965.

[49] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244,
arXiv: 1207.1303 [hep-ph].

[50] E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034,
arXiv: 1905.09127 [hep-ph].

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

[52] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079,
arXiv: 1405.0301 [hep-ph].

[53] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159,
arXiv: 1410.3012 [hep-ph].

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

[55] S. Frixione, G. Ridolfi and P. Nason,
A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction,
JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph].

[56] R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040,
arXiv: 1410.8849 [hep-ph].

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

[58] S. Alioli, P. Nason, C. Oleari and E. Re,
NLO single-top production matched with shower in POWHEG: s- and t-channel contributions,
JHEP **09** (2009) 111, arXiv: 0907.4076 [hep-ph], Erratum: JHEP **02** (2010) 011.

[59] N. Kidonakis, *NNLL resummation for s-channel single top quark production*,
Phys. Rev. D **81** (2010) 054028, arXiv: 1001.5034 [hep-ph].

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

[61] E. Re,
Single-top Wt-channel production matched with parton showers using the POWHEG method,
Eur. Phys. J. C **71** (2011) 1547, arXiv: 1009.2450 [hep-ph].

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

[63] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034,
arXiv: 1905.09127 [hep-ph].

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

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

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

[67] S. Höche, F. Krauss, M. Schönherr and F. Siegert,
QCD matrix elements + parton showers. The NLO case, JHEP **04** (2013) 027,
arXiv: 1207.5030 [hep-ph].

[68] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini, *Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order*, Phys. Rev. Lett. **103** (2009) 082001, arXiv: `0903.2120 [hep-ph]`.

[69] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: `1002.2581 [hep-ph]`.

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

[71] J. R. Andersen et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, (2013), ed. by S. Heinemeyer, C. Mariotti, G. Passarino and R. Tanaka, arXiv: `1307.1347 [hep-ph]`.

[72] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: `1908.00005 [hep-ex]`.

[73] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: `2012.00578 [hep-ex]`.

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

[75] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: `1111.6097 [hep-ph]`.

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

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

[78] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: `1907.05120 [hep-ex]`.

[79] ATLAS Collaboration, *Measurements of WH and ZH production in the $H \rightarrow b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 178, arXiv: `2007.02873 [hep-ex]`.

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

[81] ATLAS Collaboration, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 295, arXiv: `1512.05955 [hep-ex]`.

[82] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: `1802.08168 [hep-ex]`.

[83] ATLAS Collaboration,
E_T^{miss} *performance in the ATLAS detector using 2015–2016 LHC pp collisions*,
ATLAS-CONF-2018-023, 2018, URL: https://cds.cern.ch/record/2625233.

[84] ATLAS Collaboration,
Object-based missing transverse momentum significance in the ATLAS Detector,
ATLAS-CONF-2018-038, 2018, URL: https://cds.cern.ch/record/2630948.

[85] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007,
arXiv: physics/0703039 [physics.data-an].

[86] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at*
$\sqrt{s} = 13$ *TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2020) 689,
arXiv: 2007.02645 [hep-ex].

[87] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in tt̄ events using a*
likelihood approach with the ATLAS detector, ATLAS-CONF-2018-001, 2018,
URL: https://cds.cern.ch/record/2306649.

[88] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS*
proton–proton collision data at $\sqrt{s} = 13$ *TeV*, ATLAS-CONF-2018-006, 2018,
URL: https://cds.cern.ch/record/2314418.

[89] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using*
the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ *TeV*,
Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [hep-ex].

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

[91] ATLAS Collaboration, *Measurement of the Inelastic Proton–Proton Cross Section at* $\sqrt{s} = 13$ *TeV*
with the ATLAS Detector at the LHC, Phys. Rev. Lett. **117** (2016) 182002,
arXiv: 1606.02625 [hep-ex].

[92] E. Bothmann, M. Schönherr and S. Schumann,
Reweighting QCD matrix-element and parton-shower calculations, Eur. Phys. J. C **76** (2016) 590,
arXiv: 1606.08753 [hep-ph].

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

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

[95] ATLAS Collaboration,
Study of top-quark pair modelling and uncertainties using ATLAS measurements at $\sqrt{s} = 13$ *TeV*,
ATL-PHYS-PUB-2020-023, 2020, URL: https://cds.cern.ch/record/2730443.

[96] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber,
Single-top hadroproduction in association with a W boson, JHEP **07** (2008) 029,
arXiv: 0805.3067 [hep-ph].

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

[98] M. Czakon et al., *Top-pair production at the LHC through NNLO QCD and NLO EW*, JHEP **10** (2017) 186, arXiv: 1705.04105 [hep-ph].

[99] ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 538, arXiv: 1511.04716 [hep-ex].

[100] ATLAS Collaboration, *Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **86** (2012) 032003, arXiv: 1207.0319 [hep-ex].

[101] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, ed. by T. Speer et al., arXiv: 1009.1003 [physics.data-an].

[102] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, ed. by L. Lyons and M. Karagoz, arXiv: physics/0306116.

[103] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693.

[104] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an], Erratum: Eur. Phys. J. C **73** (2013) 2501.

[105] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, Eur. Phys. J. C **70** (2010) 525, arXiv: 1005.1891 [physics.data-an].

[106] ATLAS Collaboration, *Formulae for Estimating Significance*, ATL-PHYS-PUB-2020-025, 2020, URL: https://cds.cern.ch/record/2736148.

[107] P. Sikivie, L. Susskind, M. Voloshin and V. Zakharov, *Isospin breaking in technicolor models*, Nucl. Phys. B **173** (1980) 189.

[108] J. F. Gunion and H. E. Haber, *CP-conserving two-Higgs-doublet model: the approach to the decoupling limit*, Phys. Rev. D **67** (2003) 075019, arXiv: hep-ph/0207010.

[109] R. V. Harlander, S. Liebler and H. Mantler, *SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM*, Comput. Phys. Commun. **184** (2013) 1605, arXiv: 1212.3249 [hep-ph].

[110] R. V. Harlander and P. Kant, *Higgs production and decay: analytic results at next-to-leading order QCD*, JHEP **12** (2005) 015, arXiv: hep-ph/0509189.

[111] R. V. Harlander and W. B. Kilgore, *Higgs boson production in bottom quark fusion at next-to-next-to-leading order*, Phys. Rev. D **68** (2003) 013001, arXiv: hep-ph/0304035.

[112] R. V. Harlander and W. B. Kilgore, *Next-to-Next-to-Leading Order Higgs Production at Hadron Colliders*, Phys. Rev. Lett. **88** (2002) 201801, arXiv: hep-ph/0201206.

[113] D. Eriksson, J. Rathsman and O. Stål, *2HDMC - two-Higgs-doublet model calculator physics and manual*, Comput. Phys. Commun. **181** (2010) 189, arXiv: 0902.0851 [hep-ph].