The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/197589

Please be advised that this information was generated on 2019-03-19 and may be subject to change.
Search for the Higgs boson produced in association with a vector boson and decaying into two spin-zero particles in the $H \rightarrow aa \rightarrow 4b$ channel in $pp$ collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: A search for exotic decays of the Higgs boson into a pair of spin-zero particles, $H \rightarrow aa$, where the $a$-boson decays into $b$-quarks promptly or with a mean proper lifetime $c\tau_a$ up to 6 mm and has a mass in the range of 20–60 GeV, is presented. The search is performed in events where the Higgs boson is produced in association with a $W$ or $Z$ boson, giving rise to a signature of one or two charged leptons (electrons or muons) and multiple jets from $b$-quark decays. The analysis is based on the dataset of proton-proton collisions at $\sqrt{s} = 13$ TeV recorded in 2015 and 2016 by the ATLAS detector at the CERN Large Hadron Collider, corresponding to an integrated luminosity of 36.1 fb$^{-1}$. No significant excess of events above the Standard Model background prediction is observed, and 95% confidence-level upper limits are derived for the production cross-sections for $pp \rightarrow WH$, $ZH$ and their combination, times the branching ratio of the decay chain $H \rightarrow aa \rightarrow 4b$. For $a$-bosons which decay promptly, the upper limit on the combination of cross-sections for $WH$ and $ZH$ times the branching ratio of $H \rightarrow aa \rightarrow 4b$ ranges from 3.0 pb for $m_a = 20$ GeV to 1.3 pb for $m_a = 60$ GeV, assuming that the ratio of $WH$ to $ZH$ cross-sections follows the Standard Model prediction. For $a$-bosons with longer proper lifetimes, the most stringent limits are 1.8 pb and 0.68 pb, respectively, at $c\tau_a \sim 0.4$ mm.

KEYWORDS: Beyond Standard Model, Hadron-Hadron scattering (experiments), Higgs physics

ArXiv ePrint: 1806.07355
1 Introduction

The discovery of the Higgs boson by the ATLAS and CMS collaborations [1, 2] at the Large Hadron Collider (LHC) has been a major achievement for the Standard Model (SM). A comprehensive programme to explore the properties of this particle is underway, including measurements of the branching ratios to SM particles and searches for decays into “exotic” or non-SM particles. Exotic Higgs boson decays are a powerful probe for physics beyond the SM (BSM). The Higgs boson has a very narrow decay width, so even a small coupling to a non-SM particle could open up a sizeable decay mode. Measurements at the LHC are in agreement with SM predictions, constraining the non-SM branching ratio of the Higgs boson to less than approximately 30% at 95% confidence level (CL) using the 7 and 8 TeV datasets [3–5]. Despite this experimental triumph, there is still ample room for exotic Higgs boson decays compatible with observations to date.

The Higgs boson has been proposed as a possible “portal” for hidden-sector particles to interact with SM particles [6–8]. Exotic decays in particular are predicted by many BSM theories [9], including those with an extended Higgs sector such as the Next-to-Minimal Supersymmetric Standard Model (NMSSM) [10–14], models with a first-order electroweak phase transition [15, 16], models with neutral naturalness [17–19] and models of dark matter [20–24].

The decay of the Higgs boson into a pair of spin-zero particles $a$, which in turn decay into a pair of SM particles, arises in several scenarios of new physics [9]. In particular, if the $a$-boson mixes with the Higgs boson and inherits its Yukawa couplings, decays of the type $a \rightarrow b\bar{b}$ are expected to be dominant for $m_a > 2m_b$. Models of this type have been
Figure 1. Representative tree-level Feynman diagrams for the (a) $WH$ and (b) $ZH$ production processes with the subsequent decays $W \rightarrow \ell \nu$, $Z \rightarrow \ell\ell$ ($\ell = e, \mu$) and $H \rightarrow aa \rightarrow 4b$.

This paper considers the decay mode $H \rightarrow aa$ with the subsequent decay $a \rightarrow b\bar{b}$, building on the previous work of ref. [27], in which a similar analysis was reported with a subset of the data considered here. The previous result set an upper limit on the production cross-section $\sigma(WH)$ times the branching ratio for $H \rightarrow aa \rightarrow 4b$ ranging from 6.2 pb for an $a$-boson mass of $m_a = 20$ GeV to 1.5 pb for $m_a = 60$ GeV, compared with the SM cross-section $\sigma_{SM}(WH) = 1.37$ pb. This paper includes ten times more data, adds the $ZH$ channel and an improved analysis technique.

This search focuses on the $WH$ and $ZH$ processes, with $W \rightarrow \ell\nu$, $Z \rightarrow \ell\ell$ ($\ell = e, \mu$) and $H \rightarrow aa \rightarrow 4b$, as shown in figure 1. The $a$-boson can be either a scalar or a pseudoscalar under parity transformations, since the decay mode considered in this search is not sensitive to the difference in coupling. The $a$-boson signals considered have masses in the range $20$ GeV $\leq m_a \leq 60$ GeV and mean proper lifetimes, $c\tau_a$, up to 6 mm.

The resulting signature has a single lepton or two leptons accompanied by a high multiplicity of jets originating from $b$-quarks ($b$-jets). Since four $b$-jets are produced from the decay of the Higgs boson, they tend to have a low transverse momentum ($p_T$) compared with $m_H$ and can overlap, especially for light $a$-bosons. Events with one or two electrons or muons, including those produced via leptonically decaying $\tau$-leptons, are considered. The $WH$ and $ZH$ processes are chosen for this search because the presence of at least one charged lepton in the final state provides a powerful signature for triggering and for suppressing background from the high cross-section strong-interaction production of four $b$-jets.

Several kinematic variables, including the reconstructed masses in the decay $H \rightarrow aa \rightarrow 4b$, are combined to identify signal events. The background estimation techniques, systematic uncertainties and statistical treatment closely follow those used in other ATLAS searches with similar signatures [28–32].
2 ATLAS detector

The ATLAS experiment [33] is a multipurpose particle physics detector with forward-backward symmetric cylindrical geometry and nearly 4\pi coverage in solid angle.\textsuperscript{1} The interaction point is surrounded by an inner detector (ID) tracking system, a calorimeter system, and a muon spectrometer (MS). The ID covers |\eta| < 2.5 and consists of a silicon pixel detector, a silicon microstrip detector, and a transition radiation tracker. The ID includes the insertable B-layer [34], a pixel layer close to the interaction point, which provides high-resolution measurements at small radius to improve the tracking performance. A thin superconducting solenoid surrounds the ID and provides a 2 T axial magnetic field.

The calorimeter system features a high-granularity lead/liquid-argon sampling calorimeter that measures the energy and the position of electromagnetic showers within |\eta| < 4.9. Liquid-argon sampling calorimeters are also used to measure hadronic showers in the endcap (1.5 < |\eta| < 3.2) and forward (3.1 < |\eta| < 4.9) regions, while a steel/scintillator tile calorimeter measures hadronic showers in the central region (|\eta| < 1.7). The MS surrounds the calorimeters and consists of three large superconducting air-core toroid magnets, each with eight coils, a system of precision tracking chambers (|\eta| < 2.7), and fast trigger chambers (|\eta| < 2.4). For Run 2, the ATLAS detector has a two-level trigger system. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the rate of accepted events to 100 kHz. This is followed by the software-based high-level trigger that reduces the rate of recorded events to 1 kHz.

3 Event and object selection

Events are selected from proton-proton (pp) collisions collected by the ATLAS detector at the LHC at \sqrt{s} = 13 TeV in 2015 and 2016. The dataset corresponds to an integrated luminosity of 3.2±0.1 fb\textsuperscript{-1} recorded in 2015 and 32.9±0.7 fb\textsuperscript{-1} recorded in 2016, for a total of 36.1±0.8 fb\textsuperscript{-1} [35]. The data used for this search were collected using the single-electron or single-muon triggers with the lowest \pT thresholds available, 20 (26) GeV for muons and 24 (26) GeV for electrons in 2015 (2016) [36].

Electrons are reconstructed from energy deposits (clusters) in the electromagnetic calorimeter matched to tracks in the ID [37] and are required to have \pT > 15 GeV and |\eta| < 2.47. Candidates in the transition region between the barrel and endcap calorimeters, 1.37 < |\eta| < 1.52, are excluded. Electrons must satisfy the “tight” identification criterion based on a likelihood discriminant [38]. Muons are reconstructed by combining matching

\textsuperscript{1}ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, \phi) are used in the transverse plane, \phi being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle \theta as \eta = -\ln\tan(\theta/2). The transverse momentum \pT and other transverse variables are defined by projecting these variables into the x-y plane, and the transverse energy \ET is defined as \sqrt{\m^2 + \pT^2}, where m represents the mass of a considered object. The distance in the pseudorapidity-azimuthal-angle space is defined as \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}.
tracks in the ID and the MS, and are required to have $p_T > 10$ GeV and $|\eta| < 2.4$. Muon candidates must satisfy the “medium” identification criterion [39].

In order to distinguish leptons produced in the decays of $W$ and $Z$ bosons from those produced in the decays of heavy-flavour hadrons, all lepton candidates are required to originate from the primary interaction vertex, chosen as the vertex with the highest sum of the $p_T$ of its associated tracks. Furthermore, since lepton candidates arising from background sources, such as decays of hadrons, are typically embedded in jets, all lepton candidates are required to be isolated from other particles in the event. This is achieved by imposing a maximal allowed value on the energy deposited in the calorimeter and/or the momentum of ID tracks within a cone around the direction of the lepton candidate, according to the gradient isolation criteria [38, 39].

Jets are reconstructed from three-dimensional topological energy clusters [40] in the calorimeter using the anti-$k_t$ jet algorithm [41] implemented in the FastJet package [42] with a radius parameter of 0.4. Jets are calibrated using energy- and $\eta$-dependent corrections [43] and are required to have $p_T > 20$ GeV and $|\eta| < 2.5$. Events containing jets arising from non-collision sources or detector noise are removed [44]. Finally, a track-based veto, the Jet Vertex Tagger (JVT), is used to reduce contributions from jets arising from additional $p\bar{p}$ interactions (pile-up) [45].

Jets including $b$-hadrons, referred as $b$-jets, are identified using information from a multivariate $b$-tagging algorithm that combines information from an impact-parameter-based algorithm, from the explicit reconstruction of an inclusive secondary vertex and from a multi-vertex fitter that attempts to reconstruct the $b$- to $c$-hadron decay chain [46, 47]. This $b$-tagging algorithm defines a set of “$b$-tagged” jets. The working point used provides an efficiency to identify jets with $b$-hadrons from the primary vertex of approximately 77%. The rejection factors are 134 against light-quark and gluon jets, about 6 against jets originating from $c$-quarks, and about 22 against hadronically decaying $\tau$-leptons, as determined in a simulated sample of top-quark pair ($t\bar{t}$) events [46, 47]. This $b$-tagging discriminant is used to categorize selected events as discussed in section 5. The $b$-tagging algorithm is also efficient in identifying jets containing $b$-hadrons that do not originate from the primary vertex. The efficiency is largest for proper lifetimes of $\tau_L \sim 0.5$ mm and decreases for longer lifetimes.

The missing transverse momentum, $E_T^{\text{miss}}$, is defined as the magnitude of the transverse momentum imbalance $E_T^{\text{miss}}$, the negative vector sum of the transverse momenta of calibrated selected objects, such as electrons, muons and jets. The transverse momenta of charged-particle tracks compatible with the primary vertex and not matched to any of those objects are also included in $E_T^{\text{miss}}$ [48].

Events are required to have at least one reconstructed electron or muon with $p_T > 27$ GeV which is matched within a cone of size $\Delta R = 0.15$ to the lepton candidate reconstructed by the trigger algorithms. Two event categories (single lepton and dilepton) are used to probe $WH$ and $ZH$, respectively. Events with exactly one lepton are required to satisfy $E_T^{\text{miss}} > 25$ GeV and the transverse mass
\footnote{The transverse mass is defined as $m_T^W \equiv \sqrt{2E_T^{\text{miss}}p_T^L(1 - \cos \Delta \phi)}$, where $p_T^L$ is the transverse momentum of the lepton and $\Delta \phi$ is the azimuthal angle between the lepton and $E_T^{\text{miss}}$ directions.} must fulfil $m_T^W > 50$ GeV, in order to
Table 1. Summary of requirements for the single-lepton and dilepton channels. Here $m_{\ell\ell}$ is the dilepton invariant mass in the $ee$ and $\mu\mu$ channels.

be consistent with $W$ boson decays. Events in the dilepton channel must have exactly two leptons with the same flavour and opposite electric charges. In the $ee$ and $\mu\mu$ channels, the dilepton invariant mass must be consistent with the $Z$ boson mass window 85–100 GeV. Events in the $e\mu$ channel (different flavour) are also used in the analysis to study backgrounds. Finally, events must have at least three jets, of which at least two must be $b$-tagged. The selection requirements are summarized in table 1.

4 Signal and background modelling

Simulated event samples are used to study the characteristics of the signal and to calculate its acceptance, as well as for most aspects of the background estimation. Monte Carlo (MC) samples were produced using the full ATLAS detector simulation [49] based on GEANT 4 [50]. A faster simulation, where the full GEANT 4 simulation of the calorimeter response is replaced by a detailed parameterization of the shower shapes [51], was adopted for some of the samples. To simulate the effects of pile-up, additional interactions were generated using PYTHIA 8.186 [52] with the A2 set of tuned parameters [53] and the MSTW2008LO [54] parton distribution function (PDF) set, and overlaid on the simulated hard-scatter event. Simulated events were reweighted to match the pile-up conditions observed in the data. All simulated events are processed through the same reconstruction algorithms and analysis chain as the data. In the simulation, the top-quark mass is assumed to be $m_t = 172.5$ GeV. Decays of $b$- and $c$-hadrons were performed by EvtGen v1.2.0 [55], except in samples simulated with the SHERPA event generator [56].

Signal samples of associated Higgs boson production with a $W$ or $Z$ boson, $pp \rightarrow WH$ or $ZH$, were generated with POWHEG v2 [57–60] using the CT10 PDF set [61] at next-to-leading order (NLO). The Higgs boson mass is assumed to be $m_H = 125$ GeV. The Higgs boson decay into two spin-zero $a$-bosons and the subsequent decay of each $a$-boson into a pair of $b$-quarks were simulated with PYTHIA 8.186. The $a$-boson decay was performed in the narrow-width approximation and the coupling to the $b$-quarks is assumed to be that of a pseudoscalar. However, since the polarization of the quarks is not observable, this search is insensitive to the chosen parity hypothesis for the $a$-boson. PYTHIA 8.186 was also used for the showering, hadronization, and underlying-event (UE) simulation with the A14 tune [62]. The mass of the $a$-boson was varied for different signal hypotheses in the range $20$ GeV $\leq m_a \leq 60$ GeV, in 10 GeV mass steps.
Signal samples with long-lived $\alpha$-bosons were generated with MADGRAPH5_aMC@NLO [63] at leading order (LO) using the NNPDF2.3LO [64] PDF set and showered with PYTHIA 8.186. The model used is the SM with an additional dark sector that includes a dark vector boson and a dark Higgs boson [9, 65, 66]. In this model, the dark Higgs boson, which plays the role of the $\alpha$-boson, is a scalar under parity transformation and decays promptly. Therefore, the lifetimes of the $\alpha$-bosons were replaced with values sampled randomly from an exponentially falling distribution with the desired mean value. Signal MC samples were produced for $\alpha$-boson mean proper lifetimes of 0.1, 1, and 10 mm. Intermediate $\alpha$-boson lifetimes can be obtained by reweighting these samples. The masses of the $\alpha$-boson are 20, 30, and 60 GeV. The uneven spacing of $\alpha$-boson masses is motivated by the fact that the signal kinematics (and therefore acceptance) change significantly between 20 and 30 GeV, but are quite similar from 30 to 60 GeV. The ATLAS fast detector simulation was used for samples of long-lived $\alpha$-bosons, after verifying that it correctly reproduces the GEANT 4-based simulation for the range of $\alpha$-boson lifetimes under consideration.

The sample used to model the $t\bar{t}$ background was generated using the POWHEG v2 event generator [67], with the NNPDF3.0NLO PDF set. The POWHEG model parameter $h_{\text{damp}}$, which controls matrix element (ME) to parton shower (PS) matching and effectively regulates the high-$p_T$ radiation, was set to $h_{\text{damp}} = 1.5 m_t$ [68]. The PS and the hadronization were modelled by PYTHIA 8.210 [69] with the A14 tune. The renormalization and factorization scales were set to the transverse mass of the top quark, defined as $m_{T,t} = \sqrt{m_t^2 + p_{T,t}^2}$, where $p_{T,t}$ is the transverse momentum of the top quark in the $t\bar{t}$ centre-of-mass reference frame. The $t\bar{t}$ sample is normalized to the next-to-next-to-leading-order (NNLO) theoretical cross-section of $832_{-46}^{+46} \text{ pb}$, obtained with Top++ 2.0 [70]. Alternative $t\bar{t}$ samples used to derive systematic uncertainties are described in section 7.

The simulated $t\bar{t}$ events are categorized depending on the parton-level flavour content of particle jets\(^3\) originating from the decay of the $t\bar{t}$ system, using the procedure described in refs. [28, 29]. Events containing at least one particle jet matched to a $b$-hadron are labelled as $t\bar{t} + b\bar{b}$. Events containing at least one particle jet matched to a $c$-hadron and no $b$-hadron are labelled as $t\bar{t} + c\bar{c}$. The $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$ categories are generically referred to as $t\bar{t}+\text{HF}$ events (with HF standing for “heavy flavour”). Remaining events are labelled as $t\bar{t} + \text{ light-jets}$ (referred to as $t\bar{t} + \text{ light}$) and also include events with no additional particle jets.

To model the dominant $t\bar{t} + b\bar{b}$ background with the highest available precision, the relative contributions of the different heavy-flavour categories in the $t\bar{t}$ sample described above are scaled to match the predictions of an NLO $t\bar{t} + b\bar{b}$ sample including parton showering and hadronization [71], generated with SHERPA+OPENLOOPS [56, 72], using the procedure described in ref. [29]. The sample was produced with SHERPA 2.2.1 and the CT10 four-flavour (4F) scheme PDF set [73, 74]. The renormalization scale for this sample was chosen to be $\mu_R = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$ using the CMMPS prescription [71], while

\(^3\)Particle jets are reconstructed by clustering stable particles, excluding muons and neutrinos, using the anti-$k_t$ algorithm with a radius parameter $R = 0.4$. 

-- 6 --

JHEP10(2018)031
the factorization scale is set to $\mu_F = \frac{1}{2} \sum_{i=t,t,b,b} E_{T,i}$. The resummation scale $\mu_Q$, which sets an upper bound for the hardness of the PS emissions, was also set to $\frac{1}{2} \sum_{i=t,t,b,b} E_{T,i}$.

The production of $W$ and $Z$ bosons in association with jets was simulated with SHERPA 2.2.1 [56] using the NNPDF3.0NNLO PDF set for both the ME calculation and the dedicated PS tuning developed by the SHERPA authors [75]. The ME calculation was performed with COMIX [76] and OPENLOOPS [72], and was matched to the SHERPA PS using the MEPS@NLO prescription [77]. The MEs were calculated for up to two additional partons at NLO and for three and four partons at LO in QCD. The $W/Z$ + jets samples are normalized to the NNLO cross-sections [78, 79].

The diboson + jets samples were generated using SHERPA 2.1.1 as described in ref. [80]. Samples of $t\bar{t}W$ and $t\bar{t}Z$ ($t\bar{t}V$) events were generated with an NLO ME using Madgraph5_aMC@NLO interfaced to PYTHIA 8.210 with the NNPDF3.0NLO PDF set and the A14 tune.

Samples of $Wt$ and $s$-channel single-top-quark backgrounds were generated with Powheg v1 at NLO accuracy using the CT10 PDF set. Overlap between the $t\bar{t}$ and $Wt$ final states was resolved using the “diagram removal” scheme [81]. The $t$-channel single-top-quark events were generated using the Powheg v1 event generator at NLO accuracy with the CT10 4F scheme PDF set. For this process, top quarks were decayed using MadSpin. All single-top-quark samples were interfaced to PYTHIA 6.428 [82] with the Perugia 2012 tune [83]. The single-top quark $t$- and $s$-channel samples are normalized to the NLO theoretical cross-sections [84, 85], while the $Wt$ channel is normalized to the approximate NNLO prediction [86, 87].

Higgs boson production in association with a single top quark is a rare process in the SM, but is included in the analysis and is treated as a background. Samples of single top quarks produced in association with a $W$ boson and a Higgs boson, $tW H$, were produced with Madgraph5_aMC@NLO interfaced to Herwig++ [88] with the CTEQ6L1 PDF set. The other Higgs boson production modes are found to be negligible and are not considered. The production of four top quarks ($t\bar{t}t\bar{t}$) as well as $t\bar{t}WW$ events were generated with Madgraph5_aMC@NLO with LO accuracy and interfaced to PYTHIA 8.186. The $tZW$ production process was also generated with Madgraph5_aMC@NLO interfaced to PYTHIA 8.186, but at NLO accuracy.

In the single-lepton channel, the background from events with a jet or a photon misidentified as a lepton or with non-prompt leptons from hadron decays (hereafter referred to as a fake or non-prompt lepton) is estimated directly from data using a matrix method [89]. A data sample enhanced in fake and non-prompt leptons is selected by removing the lepton isolation requirements and, for electrons, loosening the identification criteria. The efficiency for these “loose” leptons to satisfy the nominal selection (“tight”) criteria is measured in data, separately for real prompt leptons and for fake or non-prompt leptons. For real prompt leptons, the efficiency is measured in $Z$ boson events, while for fake and non-prompt leptons, it is estimated from events with low $E_T^{miss}$ and low values of $m_T^W$.
The number of fake or non-prompt leptons satisfying the tight criteria can then be calculated by inverting the matrix defined by the two equations:

\[ N^l = N^l_r + N^l_f, \quad N^t = \varepsilon_l N^l_r + \varepsilon_l N^l_f, \]

where \( N^l \) (\( N^t \)) is the number of events in data satisfying the loose (tight) lepton selection, \( N^l_r \) (\( N^l_f \)) is the number of events with a real prompt (fake or non-prompt) lepton in the loose lepton sample, and \( \varepsilon_l \) (\( \varepsilon_l \)) is the efficiency for these events to fulfill the tight lepton selection. By generalizing the resulting formula to extract \( \varepsilon_l N^l_f \), a weight is assigned to each event selected in the loose lepton data sample, providing a prediction for both the yields and the kinematic distributions of the fake and non-prompt lepton background.

When applying the matrix method in the case of high jet and \( b \)-tagged jet multiplicities, the number of events in data satisfying the loose and tight lepton selections is significantly reduced, leading to large fluctuations in the background predictions. In order to mitigate this problem, instead of tagging the jets by applying the \( b \)-tagging algorithm, their probabilities to be \( b \)-tagged are parameterized as a function of the jet \( p_T \). This allows all events in the sample before \( b \)-tagging is applied to be used in predicting the normalization and shape of the background from fake or non-prompt leptons after \( b \)-tagging. The tagging probabilities are derived using an inclusive sample of fake or non-prompt leptons and the resulting predictions of this background estimate are in agreement with those obtained by applying the \( b \)-tagging algorithm and have greatly reduced statistical uncertainties.

In the dilepton channel, the background contribution from fake or non-prompt leptons is very small and is estimated from simulation and normalized to data in a control region with two same-charge leptons.

5 Event categorization

Events satisfying the object selection are categorized into analysis regions according to the number of leptons, jets and \( b \)-tagged jets. The regions enhanced in signal \( H \to aa \to 4b \) events relative to the backgrounds are referred to as signal regions (SRs). The other regions, referred to as control regions (CRs), are used to constrain the background predictions and related systematic uncertainties (see section 7) through a profile likelihood fit to the data (see section 8). The signal and backgrounds are derived consistently in the signal and control regions in a combined fit. The discrimination of signal from background is further enhanced in the signal regions by using multivariate techniques, as described in section 6.

The \( H \to aa \to 4b \) decay chain is expected to have multiple \( b \)-tagged jets, often three or four, satisfying the object selection. The dominant background arises from \( t\bar{t} \) events in the single-lepton channel and \( Z + \text{jets} \) events in the dilepton channel, which can also have different jet and \( b \)-tagged jet multiplicities or leptons of different flavour in the case of the dilepton channel. The regions are referred to as \((n_\ell, n_j, n_b)\) indicating \( n_\ell \) leptons, \( n_j \) selected jets and \( n_b \) \( b \)-tagged jets. The SRs contain at least three \( b \)-tagged jets and are \((1\ell, 3j, 3b), (1\ell, 4j, 3b)\) and \((1\ell, 4j, 4b)\) for single-lepton events, and \((2\ell, 3j, 3b), (2\ell, 4j, 3b)\) and \((2\ell, \geq 4j, \geq 4b)\) for same-flavour dilepton events. The CRs are \((1\ell, 3j, 2b), (1\ell, 4j, 2b), (1\ell, \geq 5j, 3b)\) and \((1\ell, \geq 5j, \geq 4b)\) for single-lepton events, \((2\ell, \geq 3j, 2b)\) for same-flavour
Figure 2. Definition of the signal and control regions (SR and CR, respectively) in the single-lepton and dilepton channels. The main background component probed with the CR is indicated. The vertical axis shows the lepton selection, while the horizontal axis shows the jet and $b$-tagged jet multiplicities.

dilepton events, and $(2\ell, \geq 3j, 3b)$ and $(2\ell, \geq 4j, \geq 4b)$ for different-flavour dilepton events. The signal and control regions are summarized in figure 2, indicating the main background sources probed in the CRs. Figure 3 summarizes the background composition in the signal and control regions.

In the single-lepton signal regions, background $t\bar{t}$ events can only satisfy the selection criteria if accompanied by additional $b$-tagged jets. The $t\bar{t}$ + light background is dominant in the sample of events with exactly two or three $b$-tagged jets. The background processes $t\bar{t} + c\bar{c}$ and $t\bar{t} + b\bar{b}$ become more important as the jet and $b$-tagged-jet multiplicities increase. In particular, the $t\bar{t} + b\bar{b}$ background dominates for events with $\geq 5$ jets and $\geq 4$ $b$-tagged jets. In the case of $(1\ell, 3j, 3b)$ or $(1\ell, 4j, 3b)$, the main sources of $t\bar{t}$ background are events with jets mistagged as $b$-tagged jets, particularly from $W \rightarrow cs$ decays, where the $c$-jet is misidentified, and from $W \rightarrow \tau\nu$, where the $\tau$-lepton decays hadronically and is likewise mistagged. In the case of $(1\ell, 4j, 4b)$, the $t\bar{t}$ background includes more events with genuine $b$-quarks from gluon splitting to $b\bar{b}$ pairs.

In the dilepton channel, the background is mainly composed of $t\bar{t}$ and $Z +$ jets events. In the case of the $t\bar{t}$ background, most events contain two prompt leptons from the leptonic decays of the two $W$ bosons, and two $b$-jets from the top-quark decays. Additional jets arise from gluon splitting into $b\bar{b}$ and $c\bar{c}$ and from jets from initial-state radiation and pile-up. In each of these cases, the third and fourth $b$-tags in the event are from additional $b$-tagged jets, or from the mistag of additional $c$- or light-jets. In the samples with exactly three or four jets and exactly three $b$-tagged jets, the contributions of each of these sources is similar. In the case of the sample with exactly four jets and exactly four $b$-tagged jets, the contribution from events with real additional $b$-tagged jets, such as from gluon splitting into $b\bar{b}$, dominates.

In the case of the $Z +$ jets background, the dominant contribution is from $Z$ bosons produced in association with multiple heavy-flavour jets, for example from gluon splitting
Figure 3. Fractional contributions of the various backgrounds to the total background prediction in the (a) single-lepton and (b) dilepton signal and control regions. The predictions for the various background contributions are obtained through the simulation and the data-driven estimates described in section 4. The $t\bar{t}$ background categories are also defined in section 4.

into $bb$. In particular, for events with exactly three jets and three $b$-tagged jets or exactly four jets and four $b$-tagged jets, about half of the events are from $Z + bb$ with a mistagged light-flavoured jet and half are from $Z + b\bar{b}c$ with a mistagged $c$-jet. In the case of the events with exactly four jets and three $b$-tagged jets, approximately a third of the events are from $Z + bb$ with a mistagged light-flavoured jet, a third are from $Z + b\bar{b}c$ with a mistagged $c$-jet and a third are from $Z + b\bar{b}b$.

In the dilepton channel, the control regions are designed to be populated by the two main background processes: $t\bar{t}$ and $Z +$ jets. The control region with two same-flavour leptons, $(2\ell, \geq 3j, 2b)$, is populated by $Z +$ jets and $t\bar{t} +$ light. The control regions with two different-flavour leptons but with the same jet and $b$-tagged jet multiplicities as the signal regions, $(2\ell, \geq 3j, 3b)$ and $(2\ell, \geq 4j, \geq 4b)$, are enhanced in $t\bar{t} + c\bar{c}$ and $t\bar{t} + bb$ processes.

6 Analysis strategy

In each of the six signal regions, a boosted decision tree (BDT) discriminant that combines information from several variables provides additional discrimination between signal and background. In the control regions, kinematic variables are used to provide additional discrimination between distinct sources of background. The distribution of $H_T$, defined as the scalar sum of the $p_T$ of the jets, is used in the control regions with two $b$-tagged jets, while the invariant mass of the $b$-tagged jets is used in the control regions with three or
four $b$-tagged jets. A statistical analysis based on a binned likelihood function constructed as a product of Poisson probability terms over all regions and all bins considered in the search is used to derive the background predictions and uncertainties, and to test for the presence of signal. The statistical procedure and the results are described in section 8.

The toolkit for multivariate analysis (TMVA) [90] is used to train the BDT discriminant. In the single-lepton channel, dedicated BDTs are trained to discriminate each of the signals with $a$-boson masses of 20, 30 and 50 GeV from $t\bar{t}$ events. The discriminant trained at 50 GeV is found to have good sensitivity for the $a$-boson mass range 40–60 GeV. In the dilepton channel, dedicated BDTs are trained to discriminate each of the signals from both the $t\bar{t}$ and $Z +$ jets events. The discriminant trained at 30 GeV is found to have good sensitivity over the full mass range (20–60 GeV) for each of the signal regions.

The choice of inputs used in the BDT discriminants was based on the following considerations. Signal events are characterized by the presence of a resonance resulting from the Higgs boson decay $H \rightarrow aa \rightarrow 4b$. Several variables are used to reconstruct the particles from the signal decay chain. The first is the reconstructed invariant mass of the $b$-tagged jets, $m_{bb}$ or $m_{bbbb}$, defined for events with three or four $b$-tagged jets respectively, which peaks around the Higgs boson mass for signal events. In the case of three $b$-tagged jets, the peak is due to events where two $b$-quarks are merged in a single jet or one of the $b$-quarks is very soft in an asymmetric decay and has a small impact on the kinematics. In the case of events with four $b$-tagged jets, the invariant masses of the two $b$-tagged jet pairs are discriminating variables between signal and background. The pairings are chosen to minimize the difference between the invariant masses of the $b$-tagged jet pairs, and are labelled $m_{bb1}$ and $m_{bb2}$, such that $m_{bb1} > m_{bb2}$.

Additional kinematic variables exhibit differences between signal and background. In both channels, the average angular separation between $b$-tagged jets, referred to as average $\Delta R(b,b)$, is typically larger for background events where the $b$-tagged jets originate from the decays of different particles, such as the two top quarks in $t\bar{t}$ events. In the single-lepton channel, several additional kinematic variables are included in the BDT discriminant. The $H_T$ variable is a measure of the total hadronic energy in the event, which is typically larger for $t\bar{t}$ than for $WH$ events. The transverse momentum of the $W$ boson, $p_T^W$, constructed from the vector sum of the $E_T^{\text{miss}}$ and the lepton $p_T$, is slightly higher for signal $WH$ events, where the $W$ boson recoils against the Higgs boson, than for background $t\bar{t}$ events.

Finally, two other variables are used to identify particles from the dominant $t\bar{t}$ background decay chain in the single-lepton channel. The first variable is used in the $(1\ell, 4j, 3b)$ channel to select $t\bar{t}$ events with two $b$-tagged jets from the top-quark decays and a third $b$-tagged jet from a misidentified $c$- or light-jet from the hadronically decaying $W$ boson. This variable is the invariant mass of two $b$-tagged jets (selected as the pair with the smallest $\Delta R$ separation) and the non-$b$-tagged jet, $m_{bbj}$, which reconstructs the hadronically decaying top quark, peaking around the top-quark mass for these background events. The second variable is an $m_{T2}$ observable, defined as the minimum “mother” particle mass compatible with all transverse momenta and mass-shell constraints [91], that identifies events with several invisible particles. In the case of the $t\bar{t}$ background events, in addition to the $E_T^{\text{miss}}$ from the neutrino from a leptonic $W$ boson decay, invisible particles may arise from
Table 2. List of variables used to train the BDT multivariate discriminant for each signal region.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$(1\ell, 3j, 3b)$</th>
<th>$(1\ell, 4j, 3b)$</th>
<th>$(1\ell, 4j, 4b)$</th>
<th>$(2\ell, 3j, 3b)$</th>
<th>$(2\ell, 4j, 3b)$</th>
<th>$(2\ell, 4j, 4b)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{bb}$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$m_{bbbj}$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$m_{bbl}$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$m_{bbj}$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Average $\Delta R(b, b)$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$H_T$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$p_T^V$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$m_{bb}$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$m_{T2}$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\Delta R(\ell, \ell)$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\Delta R(Z, H)$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\cos \theta^*$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$E_{miss}^T$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 2 summarizes the variables used to train each of the BDT discriminants for the six signal regions. Figures 4 and 5 show the expected distributions of the kinematical variables inclusively in number of jets and $b$-tagged jets. The jets with the largest values of the $b$-tagging discriminant are used to define the variables shown. The distributions are obtained “post-fit”, after accounting for the systematic uncertainties and applying the statistical procedure described in sections 7 and 8, respectively.

7 Systematic uncertainties

Many sources of systematic uncertainties affect this search, including those related to the integrated luminosity, to the reconstruction and identification of leptons and jets, and to the modelling of signal and background processes. Some uncertainties affect only the overall normalization of the samples, while others also impact the shapes of the distributions used to categorize events and build the final discriminants.

A single nuisance parameter is assigned to each source of systematic uncertainty, as described in section 8. Some of the systematic uncertainties, in particular most of the
Figure 4. Comparison of data with the post-fit background estimates for (a) $m_{bb}$, (b) $m_{b\bar{b}b\bar{b}}$, (c) average $\Delta R(b,b)$, (d) $H_T$, (e) $p_T^W$ and (f) $m_{bbj}$ in the single-lepton sample inclusive in number of jets and $b$-tagged jets. Comparisons use events with $\geq 3$ jets, except when $\geq 4$ jets are necessary to define the variable, in which case events with $\geq 4$ jets are used. Distributions for the signal model ($WH$, $H \rightarrow aa \rightarrow 4b$), with $m_a = 60$ GeV, normalized to the SM $pp \rightarrow WH$ cross-section, assuming $B(H \rightarrow aa \rightarrow 4b) = 1$ and scaled by a factor as indicated in the figure, are overlaid. The hashed area represents the total uncertainty in the background. The last bin contains the overflow.
Figure 5. Comparison of data with the post-fit background estimates for (a) $m_{bb}$, (b) $m_{bb+}$, (c) $m_{bb1}$, (d) average $R(b,b)$, (e) $\Delta R(\ell,\ell)$ and (f) $\cos \theta^*$ in the dilepton sample inclusive in number of jets and $b$-tagged jets. Comparisons use events with $\geq 3$ jets, except when $\geq 4$ jets are necessary to define the variable, in which case events with $\geq 4$ jets are used. Distributions for the signal model ($ZH, H \rightarrow aa \rightarrow 4b$), with $m_a = 60$ GeV, normalized to the SM $pp \rightarrow ZH$ cross-section, assuming $B(H \rightarrow aa \rightarrow 4b) = 1$ and scaled by a factor as indicated in the figure, are overlaid. The hashed area represents the total uncertainty in the background. The last bin contains the overflow.
experimental uncertainties, are decomposed into several independent sources, as specified in the following. Each individual source then has a correlated effect across all channels, analysis categories, signal and background samples. For modelling uncertainties, especially the $t\bar{t}$ and $Z + \text{jets}$ modelling, additional nuisance parameters are included to split some uncertainties into several sources affecting different subcomponents of a particular process independently.

The uncertainty of the combined integrated luminosity for 2015 and 2016 is 2.1%. It is determined using a methodology similar to that detailed in ref. [35]. Uncertainties in the modelling of pile-up are also estimated, and cover the differences between the predicted and measured inelastic cross-sections [92].

Uncertainties associated with leptons arise from the trigger, reconstruction, identification, and isolation efficiencies, as well as the momentum scale and resolution. These are measured in data using leptons in $Z \to \ell^+\ell^-$, $J/\psi \to \ell^+\ell^-$ and $W \to e\nu$ events [38, 39] and have only a small impact on the result.

Uncertainties associated with jets arise from their reconstruction and identification efficiencies. These are due to the uncertainty in the jet energy scale (JES), resolution and the efficiency of the JVT requirement that is meant to remove jets from pile-up. The JES and its uncertainty are derived by combining information from test-beam data, LHC collision data and simulation [43]. Additional uncertainties are also considered, associated with the jet flavour and pile-up corrections. The total per-jet uncertainties are 1–6%, although the effects are amplified by the large number of jets in the final state.

The efficiency to correctly tag $b$-jets is measured in data using dilepton $t\bar{t}$ events [94]. The mistag rate for $c$-jets is measured in events with $W$ bosons decays into $q\bar{q}$. For light jets, it is measured in multi-jet events using jets containing secondary vertices and tracks with impact parameters consistent with a negative lifetime [46]. The uncertainty associated with the $b$-tagging efficiency ranges between 2% and 10% depending on the jet $p_T$. The size of the uncertainties associated with the mistag rates is 5–20% for $c$-jets depending on the jet $p_T$ and 10–50% for light jets depending on the jet $p_T$ and $\eta$.

For the long-lived $a$-boson signals, the secondary vertices of $b$-jets are, on average, further displaced from the primary vertex than those of $b$-jets from $t\bar{t}$ events. An additional “displaced $b$-tagging” systematic uncertainty is applied to long-lived signal samples to account for a displacement-dependent mismodelling of the $b$-tagging efficiency. The uncertainty is determined using the “adjusted MC” method [95], which was originally developed for the calibration of the mistag rate for light-flavour jets. The resulting uncertainty increases approximately linearly with the $a$-boson proper lifetime, from $\sim 2\%$ for prompt $a$-bosons to $\sim 10\%$ for proper lifetimes of 10 mm. It is applied in addition to the standard $b$-tagging uncertainties.

Uncertainties associated with energy scales and resolutions of leptons and jets are propagated to $E_T^{\text{miss}}$. An uncertainty in the contribution from charged-particle tracks not associated with reconstructed leptons and jets is also included in the $E_T^{\text{miss}}$ uncertainty [48].

Several sources of systematic uncertainty affecting the modelling of the main backgrounds, $t\bar{t}$ and $Z + \text{jets}$ are considered. For the $t\bar{t}$ background, the procedure closely follows that described in ref. [28]. An uncertainty of 6% is assumed for the $t\bar{t}$ production
cross-section [70], including contributions from variations of the factorization and renormalization scales, and uncertainties arising from the PDFs, $\alpha_S$, and the top-quark mass.

Systematic uncertainties affecting the shape of the $t\bar{t}$ background account for the choice of generator, the choice of PS and hadronization models, and the effects of initial- and final-state radiation. The uncertainties are derived from comparisons between the nominal simulation (POWHEG+PYTHIA) and alternative samples produced with SHERPA+OPENLOOPS (varying the NLO event generator, PS and hadronization models) or POWHEG+HERWIG 7 [96] (varying only the PS and hadronization models).

Additional uncertainties are evaluated to account for the use of SHERPA+OPENLOOPS NLO to model the $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$ backgrounds. Uncertainties are also assessed for the choice of scheme to treat massive quarks and the choice of PDF sets, as well as the choice of shower recoil model and scale. Uncertainties are also included to account for differences in the relative contributions of the $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$ and $t\bar{t}$+light processes. All uncertainties are treated as uncorrelated across the $t\bar{t}$ flavour components. The normalization of the $t\bar{t} + b\bar{b}$ process is included as an independent free-floating factor, while the $t\bar{t} + c\bar{c}$ component is assigned a 50% normalization uncertainty, derived from studies of alternative background samples [29].

In the case of the $W + \text{jets}$ and $Z + \text{jets}$ backgrounds, all normalizations are included as independent free-floating factors. In the case of the $Z + \text{jets}$ background in the dilepton channel, a separate normalization factor is considered for each jet and $b$-tagged jet multiplicity bin: $(2\ell, \geq 3j, 2b)$, $(2\ell, 3j, 3b)$, $(2\ell, \geq 4j, 3b)$ and $(2\ell, \geq 4j, \geq 4b)$. Additional assigned uncertainties are based on variations of the factorization and renormalization scales and of the matching parameters in the SHERPA simulation.

A cross-section uncertainty of $+5\% -4\%$ is assigned to the three single-top-quark production modes [86, 97, 98]. For the $Wt$ and $t\bar{t}$-channel production modes, uncertainties associated with the choice of PS and hadronization model and with initial- and final-state radiation are evaluated by using a set of alternative samples. The uncertainty in modelling of the interference between $Wt$ and $t\bar{t}$ production at NLO is assessed by comparing the default simulation to an alternative one that resolves the interference at the cross-section level (“diagram subtraction” scheme) instead of the amplitude level (“diagram removal” scheme) [81].

A 50% normalization uncertainty in the diboson background is assumed, which includes uncertainties in the inclusive cross-section and the production of additional jets [80]. The uncertainties in the $t\bar{t}W$ and $t\bar{t}Z$ NLO cross-section predictions are 13% and 12%, respectively [66, 100], due to PDF and scale uncertainties, and are treated as uncorrelated between the two processes. An additional modelling uncertainty for $t\bar{t}W$ and $t\bar{t}Z$, related to the choice of event generator, PS and hadronization models, is derived from comparisons of the nominal samples with alternative ones generated with SHERPA.

In the single-lepton channel, uncertainties in the estimation of the background with fake or non-prompt leptons come from the limited number of events in the data sample without the lepton isolation requirement and from uncertainties in the measured non-prompt and prompt lepton efficiencies. The normalization uncertainty assigned to this background is 50%, as derived by comparing the background prediction with data in control
regions obtained by inverting the requirements on $E_{\text{T}}^\text{miss}$ and on $m_{T}^{W}$. An uncertainty in the shape of the predicted background distribution covers the difference between the prediction obtained using an inclusive sample before $b$-tagging is applied and the prediction after $b$-tagging. In the dilepton channel, the simulated non-prompt lepton background is assigned a 25% uncertainty, correlated across lepton flavours and all analysis categories.

Several sources of systematic uncertainty affect the theoretical modelling of the signal. Uncertainties originate from the choice of PDFs, the factorization and renormalization scales, and the PS, hadronization and UE models.

8 Results

The distributions of the discriminant for each of the analysis categories are combined in a profile likelihood fit to test for the presence of signal, while simultaneously determining the normalizations and constraining the differential distributions of the most important background components. As described in section 6, in the signal regions, the output of the BDT classifier is used as the discriminant, while $H_{T}$ or the invariant mass of the $b$-jets is used in the control regions.

The likelihood function, $L(\mu, \theta)$, is constructed as a product of Poisson probability terms over all bins in each distribution. The Poisson probability depends on the predicted number of events in each bin, which in turn is a function of the signal-strength parameter $\mu = \sigma \times B(H \rightarrow aa \rightarrow 4b)$, where $\sigma$ are the $pp \rightarrow WH$ and $ZH$ cross-sections. The nuisance parameters, $\theta$, encode the effects of systematic uncertainties. The nuisance parameters are implemented in the likelihood function as Gaussian, log-normal or Poisson priors. The statistical uncertainty of the prediction, which incorporates the statistical uncertainty of the simulated events and of the data-driven fake and non-prompt lepton background estimate, is included in the likelihood as a nuisance parameter for each bin.

The likelihood function depends on six free-floating normalization factors for $t\bar{t} + b\bar{b}$, $Z + \text{jets}$ for the four jet and $b$-tagged jet multiplicities and $W + \text{jets}$. No prior knowledge from theory or subsidiary measurements is assumed for the normalization factors, hence they are only constrained by the profile likelihood fit to the data. As shown in table 3, the normalization factors are compatible with SM expectations within the uncertainties. The other main background components, particularly $t\bar{t} + \text{light}$ and $t\bar{t} + c\bar{c}$, are also compatible with SM expectations within the uncertainties described in section 7. In the combination of the single-lepton and dilepton channels, the ratio of $WH$ to $ZH$ cross-sections is assumed to follow the SM prediction.

The test statistic $t_{\mu}$ is defined as the profile likelihood ratio: $t_{\mu} = -2 \ln(L(\mu, \hat{\theta}_\mu)/L(\mu, \hat{\theta}))$, where $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters which maximize the likelihood function, and $\hat{\theta}_\mu$ are the values of the nuisance parameters which maximize the likelihood function for a given value of $\mu$. This test statistic is used to measure the probability that the observed data is compatible with the signal+background hypothesis, and to perform statistical inferences about $\mu$, such as upper limits using the CLs method [101–103]. The uncertainty of the best-fit value of the signal strength, $\hat{\mu}$, is obtained when varying $t_{\mu}$ by one unit.
Table 3. Normalization factors included as independent free-floating factors in the likelihood fit. The uncertainties include statistical and systematic components.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Normalization factor</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single-lepton</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + b\bar{b}$</td>
<td>1.5 ± 0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W + \text{jets}$</td>
<td>0.7 ± 0.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z + \text{jets} (2\ell, \geq 3j, 2b)$</td>
<td>—</td>
<td>1.0 ± 0.1</td>
<td>1.1 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>$Z + \text{jets} (2\ell, 3j, 3b)$</td>
<td>—</td>
<td>1.2 ± 0.2</td>
<td>1.1 ± 0.2</td>
<td></td>
</tr>
<tr>
<td>$Z + \text{jets} (2\ell, \geq 4j, 3b)$</td>
<td>—</td>
<td>1.1 ± 0.2</td>
<td>1.1 ± 0.1</td>
<td></td>
</tr>
<tr>
<td>$Z + \text{jets} (2\ell, \geq 4j, \geq 4b)$</td>
<td>—</td>
<td>1.4 ± 0.5</td>
<td>1.4 ± 0.5</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Summary of the impact of the considered systematic uncertainties (in %) on the yields for the main backgrounds and the signal ($m_a = 60 \text{ GeV}$) for the single-lepton and dilepton regions ($1\ell, 4j, 4b$) and ($2\ell, \geq 4j, \geq 4b$) after the fit. The total uncertainty can differ from the sum in quadrature of individual sources due to correlations.

After performing the fit, the leading sources of systematic uncertainty are the modelling of the $t\bar{t}$ and $Z + \text{jets}$ backgrounds and the $b$-, $c$- and light-jet-tagging efficiencies. Table 4 summarizes the main systematic uncertainties by indicating their impact on the normalization of the main backgrounds and the signal ($\hat{m}$) with $m_a = 60 \text{ GeV}$ in the most sensitive signal regions ($1\ell, 4j, 4b$) and ($2\ell, \geq 4j, \geq 4b$). The uncertainties in the normalization are obtained by varying the parameter corresponding to the source of uncertainty under consideration up and down by one standard deviation, while keeping the other nuisance parameters fixed at their central value.
Figure 6. Comparison between data and prediction in the single-lepton control regions of the $H_T$ variable for (a) ($1\ell$, 3j, 2b) and (b) ($1\ell$, 4j, 2b) and the invariant mass of $b$-tagged jets for (c) ($1\ell$, $\geq 5j$, 3b) and (d) ($1\ell$, $\geq 5j$, $\geq 4b$), after the combined single-lepton and dilepton fit to the data. The last bin contains the overflow.

Figures 6 and 7 show the distributions in the control regions of the single-lepton and dilepton channels, respectively, after performing the likelihood fit of these distributions to data assuming $\mu = 0$. The fit produces better agreement between the data and the background predictions, and it reduces the uncertainties as a result of the nuisance-parameter constraints and the correlations generated by the fit. Figure 8 shows the post-fit background distributions for the signal regions of the dilepton and single-lepton channels. Table 5 compares the observed event yield with the SM background prediction in each signal region, as well as the expected number of signal events for a few representative values of $m_a$. The overall acceptance times efficiency for signal is approximately 0.2% and 0.4% for the signal regions ($1\ell$, 4j, 4b) and ($2\ell$, $\geq 4j$, $\geq 4b$), respectively.

No significant deviations from the SM predictions are found and upper limits are set on $\sigma \times B(H \rightarrow aa \rightarrow 4b)$ for a mass hypothesis $m_a$ tested in steps of 10 GeV between...
**Figure 7.** Comparison between data and prediction in the dilepton control regions of (a) the $H_T$ variable for same-flavour dilepton events with $(2\ell, 3j, 2b)$ and the invariant mass of $b$-tagged jets for different-flavour dilepton events with $(2\ell, \geq 3j, 3b)$ and (c) $(2\ell, \geq 4j, \geq 4b)$, after the combined single-lepton and dilepton fit to the data. The last bin contains the overflow.

**Table 5.** Expected event yields of the post-fit background estimates and observed data in three signal regions for the single-lepton and dilepton channels after the combined single-lepton and dilepton fit to the data. The expected number of signal events for a few representative values of $m_a$ are indicated, assuming the SM $pp \rightarrow WH$ and $ZH$ cross-sections, $\sigma_{SM}(WH) = 1.37$ pb and $\sigma_{SM}(ZH) = 0.88$ pb [104], respectively, and $B(H \rightarrow aa \rightarrow 4b) = 1$. The uncertainties include statistical and systematic components. The total uncertainty can differ from the sum in quadrature of individual sources due to correlations.
Figure 8. Comparison between data and prediction for the BDT discriminants in the (left) single-lepton signal regions trained with a signal with $m_a = 50$ GeV: (a) (1$\ell$, 3j, 3b), (c) (1$\ell$, 4j, 3b) and (e) (1$\ell$, 4j, 4b) and in the (right) same-flavour dilepton signal regions trained with a signal with $m_a = 30$ GeV: (b) (2$\ell$, 3j, 3b), (d) (2$\ell$, $\geq$4j, 3b) and (f) (2$\ell$, $\geq$4j, $\geq$4b), after performing the combined single-lepton and dilepton fit of the predictions in all SRs and CRs to the data. The signal yield (solid red) is normalized to $\hat{\mu}$ after the fit for a signal with $m_a = 60$ GeV.
20 and 60 GeV, as shown in figure 9. The observed (expected) upper limits on $\mu$ for the single-lepton channel range from 2.9 (2.4) pb, for $m_a = 20$ GeV, to 1.6 (0.94) pb, for $m_a = 60$ GeV. For the dilepton channel they range from 1.6 (1.2) pb, for $m_a = 20$ GeV, to 0.54 (0.39) pb for $m_a = 60$ GeV. For comparison, the SM NNLO cross-sections for $pp \rightarrow WH$ and $ZH$ are $\sigma_{SM}(WH) = 1.37$ pb and $\sigma_{SM}(ZH) = 0.88$ pb, respectively [104].

Upper limits are derived for the combination of the single-lepton and dilepton channels assuming that the ratio of $WH$ to $ZH$ cross-sections follows the SM prediction. The observed (expected) 95% CL upper limits on $\mu$ range from 3.0 (2.2) pb, for $m_a = 20$ GeV, to 1.3 (0.74) pb, for $m_a = 60$ GeV. The combined result excludes branching ratios as low as 0.45 in the mass range $m_a$ between about 25 GeV and 60 GeV, assuming the SM $WH$ and $ZH$ cross-sections. The reduced sensitivity for the lighter $a$-boson hypotheses, where it is not possible to set limits on the branching ratio, is due to a lower acceptance caused by overlapping $b$-jets.

**Figure 9.** Summary of the 95% CL upper limits on $\sigma \times B(H \rightarrow aa \rightarrow 4b)$ for (a) the single-lepton channel and (b) the dilepton channel, and (c) the combination of both channels. The observed limits are shown, together with the expected limits (dotted black lines). In the case of the expected limits, one- and two-standard-deviation uncertainty bands are also displayed.
9 Conclusion

This paper presents a search for exotic decays of the Higgs boson into a pair of spin-zero particles, $H \rightarrow aa$, where the $a$-boson decays into two $b$-quarks. The search focuses on processes in which the Higgs boson is produced in association with a $W$ or $Z$ boson that decays leptonically. The analysis uses the $pp$ collision dataset at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector at the LHC in 2015 and 2016, corresponding to an integrated luminosity of $36.1 \pm 0.8$ fb$^{-1}$. The search for $H \rightarrow aa \rightarrow 4b$ is performed in the mass range $20 \text{ GeV} \leq m_a \leq 60 \text{ GeV}$, considering prompt as well as long-lived $a$-bosons with mean proper lifetimes $c\tau_a$ up to 6 mm. The analysis uses several kinematic variables combined in a multivariate discriminant in signal regions and uses control regions to reduce uncertainties in the background estimates. No significant excess of data is observed relative to the SM predictions. Upper limits at 95% CL are derived for the $WH$ and $ZH$ production cross-section times branching ratio of the decay $H \rightarrow aa \rightarrow 4b$. The combined observed upper limit for promptly decaying $a$-bosons ranges from 3.0 pb for $m_a = 20$ GeV to 1.3 pb for $m_a = 60$ GeV, assuming that the ratio of $WH$ to $ZH$ cross-sections follows the SM prediction. For $a$-bosons with longer proper lifetimes, the best limits are 1.8 pb and 0.68 pb, for $m_a = 20$ GeV and 60 GeV respectively, at $c\tau_a \sim 0.4$ mm.
Acknowledgments

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNR and DSSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MUR (Russia) and NRC KI, Russian Federation; JINR; MEQ, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partage le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [105].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


The ATLAS collaboration

International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo; Japan.

Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo; Japan.

Department of Physics, Tokyo Institute of Technology, Tokyo; Japan.

Tomsk State University, Tomsk; Russia.

Department of Physics, University of Toronto, Toronto ON; Canada.

(a) TRIUMF, Vancouver BC; (b) Department of Physics and Astronomy, York University, Toronto ON; Canada.

Division of Physics and Tomonaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba; Japan.

Department of Physics and Astronomy, Tufts University, Medford MA; United States of America.

Department of Physics and Astronomy, University of California Irvine, Irvine CA; United States of America.

Department of Physics and Astronomy, University of Uppsala, Uppsala; Sweden.

Department of Physics, University of British Columbia, Vancouver BC; Canada.

Department of Physics, University of Illinois, Urbana IL; United States of America.

Instituto de Física Corpuscular (IFIC), Centro Mixto Universidad de Valencia - CSIC, Valencia; Spain.

Department of Physics, University of Victoria, Victoria BC; Canada.

Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg, Würzburg; Germany.

Department of Physics, University of Warwick, Coventry; United Kingdom.

Waseda University, Tokyo; Japan.

Department of Particle Physics, Weizmann Institute of Science, Rehovot; Israel.

Department of Physics, University of Wisconsin, Madison WI; United States of America.

Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal; Germany.

Department of Physics, Yale University, New Haven CT; United States of America.

Yerevan Physics Institute, Yerevan; Armenia.

a Also at Borough of Manhattan Community College, City University of New York, NY; United States of America.

b Also at California State University, East Bay; United States of America.

c Also at Centre for High Performance Computing, CSIR Campus, Rosebank, Cape Town; South Africa.

d Also at CERN, Geneva; Switzerland.

e Also at CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille; France.

f Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland.

g Also at Departament de Física de la Universitat Autònoma de Barcelona, Barcelona; Spain.

h Also at Departamento de Física Teorica y del Cosmos, Universidad de Granada, Granada (Spain); Spain.

i Also at Department of Applied Physics and Astronomy, University of Sharjah, Sharjah; United Arab Emirates.

j Also at Department of Financial and Management Engineering, University of the Aegean, Chios; Greece.

k Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY; United States of America.

l Also at Department of Physics and Astronomy, University of Sheffield, Sheffield; United Kingdom.

m Also at Department of Physics, California State University, Fresno CA; United States of America.

n Also at Department of Physics, California State University, Sacramento CA; United States of America.

o Also at Department of Physics, King’s College London, London; United Kingdom.
Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg; Russia.

Also at Department of Physics, Stanford University; United States of America.

Also at Department of Physics, University of Fribourg, Fribourg; Switzerland.

Also at Department of Physics, University of Michigan, Ann Arbor MI; United States of America.

Also at Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa; Italy.

Also at Giresun University, Faculty of Engineering, Giresun; Turkey.

Also at Graduate School of Science, Osaka University, Osaka; Japan.

Also at Hellenic Open University, Patras; Greece.

Also at Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest; Romania.

Also at H. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen; Germany.

Also at Institut Catalana de Recerca i Estudis Avançats, ICREA, Barcelona; Spain.

Also at Institut für Experimentalphysik, Universität Hamburg, Hamburg; Germany.

Also at Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen; Netherlands.

Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest; Hungary.

Also at Institute of Particle Physics (IPP); Canada.

Also at Institute of Physics, Academia Sinica, Taipei; Taiwan.

Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku; Azerbaijan.

Also at Institute of Theoretical Physics, Ria State University, Tbilisi; Georgia.

Also at Istanbul University, Dept. of Physics, Istanbul; Turkey.

Also at LAL, Université Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay; France.

Also at Louisiana Tech University, Ruston LA; United States of America.

Also at LPNHE, Sorbonne Université, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris; France.

Also at Manhattan College, New York NY; United States of America.

Also at Moscow Institute of Physics and Technology State University, Dolgoprudny; Russia.

Also at National Research Nuclear University MEPhI, Moscow; Russia.

Also at Near East University, Nicosia, North Cyprus, Mersin; Turkey.

Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg; Germany.

Also at School of Physics, Sun Yat-sen University, Guangzhou; China.

Also at The City College of New York, New York NY; United States of America.

Also at The Collaborative Innovation Center of Quantum Matter (CICQM), Beijing; China.

Also at Tomsk State University, Tomsk, and Moscow Institute of Physics and Technology State University, Dolgoprudny; Russia.

Also at TRIUMF, Vancouver BC; Canada.

Also at Universita di Napoli Parthenope, Napoli; Italy.

* Deceased