Search for a CP-odd Higgs boson decaying to Zh in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector

ATLAS Collaboration

1. Introduction

After the discovery of a Higgs boson at the LHC in 2012 [1,2], one of the most important remaining questions is whether the newly discovered particle is part of an extended scalar sector. A CP-odd Higgs boson, $A$, appears in many models with an extended scalar sector, e.g. in the case of the two-Higgs-doublet model (2HDM) [3].

The addition of a second Higgs doublet leads to five Higgs bosons after the electroweak symmetry breaking. The phenomenology of such a model is very rich and depends on the vacuum expectation values of the Higgs doublets, the CP properties of the Higgs potential and the values of its parameters and the Yukawa couplings of the Higgs doublets with the fermions. In general, it is possible to accommodate in the model a Higgs boson compatible to the one discovered at the LHC. In the case where the Higgs potential of the 2HDM is CP-conserving, the Higgs bosons after electroweak symmetry breaking are two CP-even ($h$ and $H$), one CP-odd ($A$) and two charged ($H^\pm$) Higgs bosons. Many theories beyond the Standard Model (SM) include a second Higgs doublet, such as the minimal supersymmetric SM (MSSM) [4–8], axion models [e.g. Ref. [9]] and baryogenesis models [e.g. Ref. [10]]. Searches for a CP-odd Higgs boson are reported in Refs. [11–14].

In this Letter, a search for a heavy CP-odd Higgs boson decaying into a $Z$ boson and the $\sim$125 GeV Higgs boson, $h$, is described. The $A \rightarrow Zh$ decay rate can be dominant for part of the 2HDM parameter space, especially for an $A$ boson mass, $m_A$, below the $t\bar{t}$ threshold. In this case, the $A$ boson is produced mainly via gluon fusion and its natural width is typically small: $\Gamma_A/m_A \ll \mathcal{O}(1\%)$.

The search is performed for $m_A$ in the range 220 to 1000 GeV, reconstructing $Z \rightarrow \ell\ell$ decays (where $\ell = e, \mu$) with $h \rightarrow b\bar{b}$ or $h \rightarrow t\bar{t}$, as well as $Z \rightarrow \nu\bar{\nu}$ with $h \rightarrow b\bar{b}$. The selected $h$ boson decay modes provide high branching ratios and the possibility to fully reconstruct the Higgs boson decay kinematics. The reconstructed invariant mass (or transverse mass) of the $Zh$ pair, employing the measured value of the $h$ boson mass, $m_h$, to improve its resolution, is used to search for a signal.

2. Data and simulated samples

The data used in this search were recorded with the ATLAS detector in proton–proton collisions at a centre-of-mass energy of 8 TeV. The ATLAS detector is described in detail elsewhere [15]. The integrated luminosity of the data sample, selecting only periods where all relevant detector subsystems were operational, is 20.3 ± 0.6 fb$^{-1}$ [16]. The data used in the $\ell\ell\tau\tau$ and $\ell\ell\bar{\nu}\nu$ final states were collected using a combination of single-electron, single-muon, dielectron ($ee$) and dimuon ($\mu\mu$) triggers. Depending

---

* E-mail address: atlas.publications@cern.ch.

http://dx.doi.org/10.1016/j.physletb.2015.03.054

0370-2693/Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/). Funded by SCOAP3.
on the trigger choice, the \( p_T^2 \) thresholds vary from 24 to 60 GeV for the single-electron and single-muon triggers, and from 12 to 13 GeV for the \( e \bar{e} \) and \( \mu \bar{\mu} \) triggers. The data used in the \( v\nu bb \) final state were collected with a missing transverse momentum (\( E_T^{\text{miss}} \)) trigger with a threshold of \( E_T^{\text{miss}} > 80 \) GeV.

Signal events from a narrow-width \( A \) boson produced via gluon fusion are generated with MadGraph5 [17] for all final states considered in this search. The parton showering is performed with PYTHIA8 [18,19].

Production of \( W \) and \( Z \) bosons in association with jets is simulated with SHERPA [20]. Top-quark pair and single top-quark production is simulated with POWHEG [21–23] and AcerMC [24]. Production of \( WW \), \( WZ \), and \( ZZ \) dibosons are simulated using POWHEG. The \( WZ \) and \( ZZ \) processes include the production of off-shell \( Z \) bosons (\( Z^* \)) and photons (\( \gamma^* \)). Triboson production (\( \gamma WW^* \), \( ZZW^* \), \( ZZW^* \)) and top pair production in association with a \( Z \) boson are generated with MadGraph5. Finally, the production of the SM Higgs boson in association with a \( Z \) boson is considered as a background in this search. It is simulated using PYTHIA8. The CT10 [26] set of parton distribution functions was used for samples generated with MadGraph5 and PYTHIA8. The CTEQ6L1 [25] set of parton distribution functions was used for the other samples.

All generated samples are passed through the GEANT4-based [27] detector simulation of the ATLAS detector [28]. The simulated events are overlaid with minimum-bias events, to account for the effect of multiple interactions occurring in the same and neighboring bunch crossings ("pile-up"). The events are reweighted so that the average number of interactions per bunch crossing agrees with the data.

The background estimation in this search for most processes is based on data-driven techniques, but in some cases only simulated samples are used. In that case, the simulated samples are normalized using theoretical cross section calculations. In particular, for diboson production both \( gg \) [29] and \( gg \) [30,31] initiated processes are included. Triboson production follows Ref. [32] and top pair production in association with a \( Z \) boson follows Refs. [33, 34]. SM Higgs boson production in association with a \( Z \) boson uses a calculation described in Ref. [35].

3. Object reconstruction

Electrons are identified from energy clusters in the electromagnetic calorimeter that are matched to tracks in the inner detector [36]. Electrons are required to have \( |\eta| < 2.47 \) and \( p_T > 7 \) GeV. Isolation requirements, defined in terms of the calorimetric energy or the \( p_T \) of tracks within cones around the object, as well as quality requirements are applied to distinguish electrons from jets.

Muons are reconstructed by matching tracks reconstructed in the inner detector to tracks or track segments in the muon spectrometer systems [37]. The muon acceptance is extended to the region \( 1.2 < |\eta| < 2.7 \), which is outside the inner detector coverage, using only tracks reconstructed in the forward part of the muon detector. Muons used for this search must have \( |\eta| < 2.7 \), \( p_T > 6 \) GeV and are also required to pass isolation requirements.

Jets are reconstructed using the anti-\( k_T \) algorithm [38] with radius parameter \( R = 0.4 \) and \( p_T > 20 \) GeV (\( p_T > 30 \) GeV) for \( |\eta| < 2.5 \) (2.5 < \( |\eta| < 4.5 \)). Low-\( p_T \) jets from pile-up are rejected with a requirement on the scalar sum of the \( p_T \) of the tracks associated with the jet: for jets with \( |\eta| < 2.4 \) and \( p_T < 50 \) GeV, tracks associated with the primary vertex\(^2\) must contribute over 50% to the sum. Jets from the decay of long-lived heavy-flavor hadrons are selected using a multivariate tagging algorithm (\( b \)-tagging) [39]. The \( b \)-tagging efficiency is 70% for jets from \( b \)-quarks in a sample of simulated \( \tau \tau \) events.

Hadronic decays of \( \tau \) leptons (\( \tau_{\text{had}} \)) [40] are reconstructed starting from clusters of energy in the calorimeter. A \( \tau_{\text{had}} \) candidate must lie within \( |\eta| < 2.47 \), have a transverse momentum greater than 20 GeV, one or three associated tracks and a total charge of \( \pm 1 \). Information on the collimation, isolation, and shower profile is combined into a multivariate discriminant to reduce backgrounds from quark- or gluon-initiated jets. Dedicated algorithms that reduce the number of electrons and muons misidentified as hadronic \( \tau \) decays are applied. In this analysis, two \( \tau_{\text{had}} \) identification selections are used — "loose" and "medium" — with efficiencies of about 65% and 55%, respectively.

The missing transverse momentum (\( E_T^{\text{miss}} \)) is computed using fully calibrated and reconstructed physics objects, as well as clusters of calorimeter-cell energy deposits that are not associated with any object [41]. In addition, a track-based missing transverse momentum (\( P_T^{\text{miss}} \)) is calculated as the negative vector sum of the transverse momenta of tracks with \( |\eta| < 2.4 \) and associated with the primary vertex.

4. Search for \( A \rightarrow Zh \) with \( h \rightarrow \tau\tau \)

In the search for \( A \rightarrow Zh \rightarrow \ell\ell \tau\tau \), three channels are considered, distinguished by the way the \( \tau\tau \) pair decays: two \( \tau \) leptons decaying hadronically (\( \tau_{\text{had}} \)), one leptonic and one hadronic decay (\( \tau_{\text{lep}}\tau_{\text{had}} \)) and, finally, two leptonic decays (\( \tau_{\text{lep}}\tau_{\text{lep}} \)). Electrons in the \( \tau_{\text{had}} \) and \( \tau_{\text{lep}} \) channels are rejected in the transition region between the barrel and end-cap of the detector (\( 1.37 < |\eta| < 1.52 \)). Muons in the \( \tau_{\text{had}} \) and \( \tau_{\text{lep}} \) channels are considered only for \( |\eta| < 2.5 \).

The resolution of the reconstructed \( A \) boson mass is improved using a mass-difference variable,

\[
m_{\text{rec}} = m_{\ell\ell\tau\tau} - m_{\ell\tau} - m_{\tau\tau} + m_{Z} + m_{h}.
\]

where \( m_{Z} \) is the mass of the \( Z \) boson, \( m_{h} = 125 \) GeV is the mass of the CP-even Higgs boson, \( m_{\ell\tau} \) is the invariant mass of the two leptons associated with the \( Z \) boson decay, and \( m_{\ell\ell\tau\tau} \) denotes the \( \ell\ell\tau\tau \) invariant mass. The value of \( m_{\ell\tau} \), the invariant mass of the \( \tau \)’s, is estimated with the Missing Mass Calculator (MMC) [42]. The mass resolution for all \( \tau\tau \) channels ranges from 3% at \( m_{A} = 220 \) GeV to 5% at \( m_{A} = 1 \) TeV.

4.1. \( \ell\ell\tau_{\text{had}} \)

Events in the \( \ell\ell\tau_{\text{had}} \) channel are required to contain exactly two opposite-sign leptons \( \ell\ell \) (\( ee \) or \( \mu\mu \)) and exactly two opposite-sign \( \tau_{\text{had}} \). The \( p_T \) requirements for these objects are \( p_T > 26 \) GeV (15 GeV) for the leading (subleading) electron, \( p_T > 25–36 \) GeV (10 GeV) for the leading (subleading) muon, depending on the trigger, and \( p_T > 35 \) GeV (20 GeV) for the leading (subleading) \( \tau_{\text{had}} \) candidates. The \( \tau_{\text{had}} \) candidates are required to satisfy the "loose" \( \tau_{\text{had}} \) identification criterion. In addition, the \( ee/\mu\mu \) invariant mass and the \( \tau\tau \) invariant mass have to lie in the ranges \( 80 < m_{\ell\tau} < 100 \) GeV and \( 75 < m_{\ell\ell} < 175 \) GeV. Finally, the \( p_T \) of the \( \ell\ell \) pair, \( p_T^{\ell\ell} \), is required to be:

\(^2\) 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 upward. Cylindrical coordinates \(( r, \phi \)) are used in the transverse plane, \( \phi \) being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle \( \theta \) as \( \eta = -\ln(\tan(\theta/2)) \). Transverse momenta are computed from the three-momenta, \( \vec{p} \), as \( p_T = |\vec{p}| \sin \eta \).

\(^{\star}\) The primary vertex is taken to be the reconstructed vertex with the highest \( \Sigma p_T^2 \) of the associated tracks.
This requirement maximizes the sensitivity over the whole explored $A$ mass range. In the region of $p_T^Z > 125$ GeV, if $m_A^2 > 400$ GeV, otherwise.

The dominant background for this channel originates from events where one or both of the $\tau$ had is a misidentified jet (“fake-$\tau$ background”). This background is dominated by $Z +$ jets events, with small contributions from dibosons and events with top quarks, and it is estimated using a template method. The shape of the fake-$\tau$ background is taken from a control region (the “template region”) that contains events satisfying all the $\ell\ell\tau\tau$ selection criteria apart from the requirements for an opposite-sign $\tau$ had pair and the $\tau$ had identification criteria. The fake-$\tau$ background is normalized by using two additional control regions. The first region, “A”, contains events that satisfy the signal selection criteria, with the exception that the $m_{\tau\tau}$ constraint is inverted, i.e. $m_{\tau\tau} < 75$ GeV or $m_{\tau\tau} > 175$ GeV. The second region, “B”, contains events that satisfy all the template selection criteria, with the exception that the $m_{\tau\tau}$ constraint is inverted, as in the region “A” definition. The ratio of the number of events in “A” to the number of events in “B” is used to scale the template region events in order to obtain the normalization of the fake-$\tau$ background.

In addition to the fake-$\tau$ background, there are also contributions from backgrounds with real $\ell\ell\tau\tau$ objects in the event. These backgrounds come primarily from $ZZ^{(*)}$ production. SM Higgs boson production in association with a $Z$ boson is estimated using simulation, and contributes 17% of the total background.

4.2. $\ell\ell\tau\ell\tau$ had

Events in the $\ell\ell\tau\ell\tau$ had channel are required to contain exactly three light leptons, $\mu\mu\mu$, $e\mu\mu$, $e\mu\mu$ or $ee\mu$, and exactly one $\tau$ had. The $p_T$ requirements for these objects are $p_T > 26$ GeV (15 GeV) for the leading (remaining) electron(s), $p_T > 25$–36 GeV (10 GeV) for the leading (remaining) muon(s), depending on the trigger, and $p_T > 20$ GeV for the $\tau$ had. Subsequently, all the possible $\ell\ell$ pairs that are composed of opposite-sign, same-flavor leptons are selected. From these pairs, the pair that has the invariant mass closest to $m_{\tau\tau}$ is considered to be the lepton pair from the $Z$ boson decay. The third light lepton is considered to be the leptonic $\tau$ decay, and it is used along with the $\tau$ had to define the $\tau$ had pair. This light lepton is required to have opposite-sign charge with respect to the $\tau$ had. In addition, the $\tau$ had is required to satisfy the “medium” $\tau$ had identification requirement, and $m_{\ell\ell}$ and $m_{\tau\tau}$ have to lie in the ranges $80 < m_{\ell\ell} < 100$ GeV and $75 < m_{\tau\tau} < 175$ GeV. The total acceptance times selection efficiency varies from 6% for $m_A = 220$ GeV, to around 17% for the highest $A$ boson masses considered.

About half of the total background for this channel comes from events where the $\tau$ had and/or the light lepton is a misidentified jet (“fake-$\ell$ background”). This background is dominated by diboson and $Z +$ jets events and it is estimated using a template method. The shape of the fake-$\tau$ background is taken from a control region (the “template region”) that contains events satisfying all $\ell\ell\tau\ell\tau$ had selection criteria, apart from requiring “medium” $\tau$ had identification criteria and opposite-sign charge for the $\tau$ had pair. The fake-$\ell$ background is normalized by using two additional control regions, defined similarly to those in the $\ell\ell\tau\ell\tau$ had channel.

The other half of the background comes from events with real $\ell\ell\tau\ell\tau$ had objects in the event. These backgrounds come primarily from $ZZ^{(*)}$ production. There is also a small (11%) contribution from the SM Higgs boson production in association with a $Z$ boson, which is estimated using simulation.

4.3. $\ell\ell\tau\ell\tau$

Events in the $\ell\ell\tau\ell\tau$ had channel are required to contain at least four leptons, which form one same-flavor and opposite-sign pair consistent with the $Z$ mass ($80 < m_{\ell\ell} < 100$ GeV), and either a same-flavor or different-flavor pair with an invariant mass reconstructed with the MMC algorithm, consistent with a decay from the CP-even Higgs boson ($90 < m_{\tau\tau} < 190$ GeV). One muon is allowed to be reconstructed in the forward region ($2.5 < |\eta| < 2.7$) of the muon spectrometer, or to be identified in the calorimeter with $p_T > 15$ GeV and $|\eta| < 0.1$ [37]. The highest-$p_T$ lepton must satisfy $p_T > 20$ GeV, and the second (third) lepton in $p_T$ order must satisfy $p_T > 15$ GeV ($p_T > 10$ GeV). Among all the possible lepton quadruplets in an event the one minimizing the sum of the mass differences with respect to both the $Z$ and $h$ bosons is chosen.

Two different analysis categories are defined based on the lepton flavors in the Higgs boson decay: $ee$ or $\mu\mu$ (SF), and $e\mu$ (DF). The expected background is very different in the two cases. For the SF channel, the background is dominated by $ZZ^{(*)}$ production with $Z \rightarrow ee/\mu\mu$ decays. For the DF channel, the main background is from the $ZZ^{(*)}$ process through the $Z \rightarrow \ell\ell\tau\ell\tau$ decay chain, but other backgrounds are also important. The signal-to-noise ratio in the SF category is improved by using a set of requirements specifically targeted to suppress the main $ZZ^{(*)}$ background. First, a veto on the on-shell production of $Z$ boson pairs is introduced, requiring the invariant mass of the $h$ boson leptons to lie outside the $Z$ peak; $m_h < 80$ GeV or $m_h > 100$ GeV. Background events are characterized by low missing transverse momentum and are further rejected by requiring $E_{T\text{miss}} > 30$ GeV, and the azimuthal angle between the $E_{T\text{miss}}$ direction and the $Z$ boson transverse momentum to be greater than $\pi/2$. Furthermore, a requirement that the highest-$p_T$ lepton of the $\ell\ell$ pair associated with the $h$ boson has $p_T > 15$ GeV is applied, since it is found to be effective against backgrounds from $Z +$ jets production. The total acceptance times selection efficiency varies from 6.5% (1.5%) for DF (SF) channel for $m_A = 220$ GeV, to around 20% for both channels for the highest $A$ boson masses considered.

The subleading contributions to the background are from diboson and triboson production, $tt$ production in association with a $Z$ boson, and SM Higgs boson production. All these are determined from simulation and amount to about 95% (65%) of the total background in the SF (DF) category. The other background events have at least one lepton which is a misidentified jet or a lepton from a heavy-flavor quark decay and are dominated by $Z +$ jets production, with a smaller contribution from top–quark production. These backgrounds are estimated using a control region where one or both of the leptons in the $\ell\ell$ pair associated with the $h \rightarrow \ell\ell\tau\ell\tau$ decay fail to satisfy the isolation criteria. After subtraction of genuine sources of four-lepton events using simulation, the data are extrapolated to the isolated signal region using normalization factors derived from simulated samples.

4.4. Systematic uncertainties and results

The most important systematic uncertainty for the backgrounds with real $\ell\ell\tau\ell\tau$ objects in the $\tau$ had and $\tau$ had channels comes...
from the uncertainty on the theoretical cross sections used in the normalization. They are due to the parton distribution function choice, the renormalization and factorization scales, as well as the $\alpha_s$ value. This amounts to an uncertainty on the normalization of this background of about 5.0% for the $\tau\ell\tau\ell$ channel and 6.4% for $\ell\tau\ell\tau$ channel. In the $\tau\ell\tau\ell$ channel, the largest contributions come from the $\tau$ identification and energy scale and amounts to 8.9% [40]. The fake-$\tau$/$\ell$ background systematic uncertainty for the $\tau\tau$ channels is dominated by the statistical uncertainty on data in control regions used for the background normalization. It amounts to a normalization uncertainty of 38% and 25% for the $\tau\ell\tau\ell$ and $\tau\ell\tau\ell$ channels, respectively. For the $\ell\ell\tau\ell$ channel, the normalization uncertainty is 65% (25%) for the SF (DF) category.

The reconstructed $A$ boson mass distributions for events passing the $\ell\tau\ell\tau$ and $\ell\ell\tau\ell$ selections are shown in Fig. 1. The number of events passing the $\ell\tau\ell\tau$ channel selections are shown in Table 1. The agreement of the expectation with data is very good.

5. Search for $A \rightarrow Zh$ with $h \rightarrow bb$

This section describes the searches in the $A \rightarrow Zh \rightarrow \ell\ell b\bar{b}$ and $A \rightarrow Zh \rightarrow \nu\nu b\bar{b}$ channels.

5.1. $\ell\ell b\bar{b}$ selection

Events in the $\ell\ell b\bar{b}$ channel are selected by requiring either two electrons or two muons. In the case of muons they are required to be of opposite-sign charge. Leptons must have $p_T > 7 \text{ GeV}$, and electrons are restricted to $|\eta| < 2.47$, while muons must have $|\eta| < 2.7$. Tighter acceptance requirements are placed on one of the leptons in each event in order to select a sample for which the trigger efficiency is high and to reduce the multi-jet background, while keeping a high signal acceptance. These requirements are that the leptons have $p_T > 25 \text{ GeV}$, and, if they are muons, satisfy $|\eta| < 2.5$. A dilepton invariant mass window of $83 < m_{\ell\ell} < 99 \text{ GeV}$ is imposed to reduce top-quark and multi-jet backgrounds.

The $h \rightarrow bb$ decay is reconstructed by requiring two $b$-tagged jets with $p_T > 45 \text{ GeV} (20 \text{ GeV})$ for the leading (subleading) jet. Events with more than two $b$-tagged jets are removed but all events with one or more additional jets failing $b$-tagging are retained. The $h \rightarrow bb$ decay is selected by requiring that the invariant mass of the two $b$-tagged jets lies within the range $105 < m_{bb} < 145 \text{ GeV}$.

The top-quark background, which includes top-quark pair and single top-quark production, is reduced by requiring $E_T^{\text{miss}}/\sqrt{H_T} < 3.5 \text{ GeV}^{1/2}$, where $H_T$ is defined as the scalar sum of the $p_T$ of all jets and leptons in the event.

The reconstructed $A$ boson mass, $m_A^{\text{rec}}$, is the invariant mass of the two leptons and two $b$-tagged jets. In this calculation, the four-momentum of each $b$-tagged jet is scaled by $125 \text{ GeV}/m_{bb}$ in order to improve the resolution. The resulting $m_A^{\text{rec}}$ resolution ranges from 2% at $m_A = 220 \text{ GeV}$ to 3% at $m_A = 1 \text{ TeV}$.

In order to reduce the dominant $Z +$jets background, a requirement is imposed on the transverse momentum of the $Z$ boson, $p_T^Z$, reconstructed from the two leptons: $p_T^Z > 0.44 \times m_A^{\text{rec}} - 106 \text{ GeV}$, where $m_A$ is in units of GeV. The requirement depends on $m_A^{\text{rec}}$ since the background is generally produced at low $p_T^Z$, whereas the mean $p_T^Z$ increases with $m_A$ for the signal. The total acceptance times selection efficiency varies from 7% for $m_A = 220 \text{ GeV}$, to around 16% for the highest $A$ boson masses considered.

5.2. $\nu\nu b\bar{b}$ selection

The event selection in the $\nu\nu b\bar{b}$ channel follows closely the SM $h \rightarrow bb$ analysis in Ref. [43]. Events are selected with $E_T^{\text{miss}} > 120 \text{ GeV}$, $p_T^{\nu} > 30 \text{ GeV}$ and no electrons or muons with $p_T > 7 \text{ GeV}$. In addition to the jet selection of the $\ell\ell b\bar{b}$ analysis, additional restrictions are applied. In order to suppress top-quark background, which is larger than in the $\ell\ell b\bar{b}$ channel, events are rejected if any of the following conditions is satisfied: there is a jet with $|\eta| > 2.5$; there are four or more jets; one of the $b$-tagged jets is the third-highest-$p_T$ jet. In order to select a sample for which the trigger efficiency is high, $H_T$ is required to be above 120 GeV (150 GeV) for events with two (three) jets. There are also requirements on the separation between the two $b$-jets in the $\eta-\phi$ space, $\Delta R_{bb}$, to suppress $Z +$jets and $W +$jets backgrounds as described in Ref. [43]. As in the $\ell\ell b\bar{b}$ channel, the $h$ boson is selected by requiring $105 < m_{bb} < 145 \text{ GeV}$.
Additional requirements are imposed on angular quantities sensitive to the presence of neutrinos in order to suppress the multijet background: the azimuthal angle between $E_{\text{miss}}^{\ell \ell}$ and $p_{T}^{\ell \ell}$ for the $\ell\ell$ jets background: the lowest momentum angle between $E_{\text{miss}}^{\ell \ell}$ and any jet $\min(\Delta \phi(E_{\text{miss}}^{\ell \ell}, \text{jet})) > 1.5$; and the azimuthal angle between $E_{\text{miss}}^{\ell \ell}$ and the b-jet pair $\Delta \phi(E_{\text{miss}}^{\ell \ell}, \text{bb}) > 2.8$. The total acceptance times selection efficiency varies from 4%, for $m_{tA} = 400$ GeV, to around 7% for the highest $A$ boson masses considered.

It is not possible to accurately reconstruct the invariant mass of the $A$ boson due to the presence of neutrinos in the final state. Therefore, the transverse mass is used as the final discriminant:

$$m_{T}^{\text{rec}} = \sqrt{E_{\text{miss}}^{\ell \ell}^{2} - (p_{T}^{a} + E_{\text{miss}}^{\ell \ell})^{2}},$$

where $p_{T}^{a}$ and $E_{\text{miss}}^{\ell \ell}$ are the transverse energy and transverse momentum of the $A$ boson.

### 5.3. Backgrounds

All backgrounds in $\ell\ell bb$ and $\nu\nu bb$ final states are determined from simulation, apart from the multijet background, which is determined from data. The multijet background in the $\mu\ell\ell$ final state is found to be negligible. In the $e\ell\ell bb$ final state, the background is determined by selecting a sample of events with the electron isolation requirement inverted. The sample is normalized by fitting the $m_{\ell\ell}$ distribution. In the $\nu\nu\ell\ell bb$ final state, the multijet background is determined by inverting the $\Delta \phi(E_{\text{miss}}^{\ell \ell}, \text{bb})$ requirement. The sample is normalized using the region with $\min(\Delta \phi(E_{\text{miss}}^{\ell \ell}, \text{jet})) < 0.4$.

The $Z + \text{jets}$ simulated sample is split into different components according to the true flavor of the jets, i.e., $Z + \ell \ell$, $Z + c\ell$, $Z + c c$, $Z + b \ell$, $Z + b c$, and $Z + b b$, where $l$ denotes a light quark ($u, d, s$) or a gluon. These components are constrained by defining control samples which have the same selection as the $\ell\ell bb$ final state, but with the requirements on the number of b-tagged jets changed to either zero or one. The samples are further divided into events with two or at least three jets. In order to improve the description of the data, corrections are applied to the simulation as a function of the azimuthal angle between the two leading jets, $\Delta \phi_{bb}$, for $Z + \ell \ell$ events and a function of $p_{T}^{\ell}$ for the other components, as described in detail in Ref. [43].

The $W + \ell \ell$ background, which contributes significantly only in the $\nu\nu\ell\ell$ final state, is split into its components in the same way as the $Z + \ell \ell$ sample. It is constrained by defining a sample of events that are selected using the $E_{\text{miss}}^{\ell \ell}$ triggers and contain exactly one lepton with $p_{T} > 25$ GeV and a tightened isolation requirement. The transverse momentum of the lepton and the $E_{\text{miss}}^{\ell \ell}$ system ($p_{T}^{W}$) is required to be above 120 GeV to approximately match the phase space of the signal region. The sample is split into events with zero, one or two b-tagged jets and into events with 2 and 3 jets. A correction depending on $\Delta \phi_{bb}$ is applied to $W + \ell \ell$ and $W + \ell \ell$ events, following studies similar to those performed for the $Z + \ell \ell$ background [43].

A correction is made to the $p_{T}$ distribution of $t\bar{t}$ production in the simulation to account for an observed discrepancy with the data [44]. The normalization of top-quark pair production in the $\ell\ell bb$ channel is measured by defining a sample of events with exactly one electron and one muon, one of which has $p_{T} > 25$ GeV, and two b-tagged jets with $50 < m_{bb} < 180$ GeV.

### 5.4. Systematic uncertainties and results

The most important experimental systematic uncertainties in the $\ell\ell bb$ and $\nu\nu bb$ final states come from the jet energy scale uncertainty and the b-tagging efficiency.

### Table 2

Predicted and observed number of events for the $\ell\ell bb$ and $\nu\nu bb$ final states shown after the profile likelihood fit to the data.

<table>
<thead>
<tr>
<th></th>
<th>$\ell\ell bb$</th>
<th>$\nu\nu bb$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z + \text{jets}$</td>
<td>$1443 \pm 60$</td>
<td>$225 \pm 11$</td>
</tr>
<tr>
<td>$W + \ell \ell$</td>
<td>$-$</td>
<td>$55 \pm 8$</td>
</tr>
<tr>
<td>Top</td>
<td>$317 \pm 28$</td>
<td>$203 \pm 15$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$30 \pm 5$</td>
<td>$10.8 \pm 1.6$</td>
</tr>
<tr>
<td>SM $Zh$, $Wh$</td>
<td>$317 \pm 1.8$</td>
<td>$22.5 \pm 1.2$</td>
</tr>
<tr>
<td>Multi-jet</td>
<td>$20 \pm 1.6$</td>
<td>$3.2 \pm 3.1$</td>
</tr>
<tr>
<td>Total background</td>
<td>$1843 \pm 34$</td>
<td>$521 \pm 12$</td>
</tr>
<tr>
<td>Data</td>
<td>$1857$</td>
<td>$511$</td>
</tr>
</tbody>
</table>

The jet energy scale systematic uncertainty arises from several sources including uncertainties from the in situ calibration, pile-up dependent corrections and the jet flavor composition [45]. In addition, an uncertainty on the jet energy resolution is applied. The jet energy scale and resolution uncertainties are propagated to the $E_{\text{miss}}^{\ell \ell}$. The uncertainty on $E_{\text{miss}}^{\ell \ell}$ also has a contribution from hadronic energy that is not associated with jets [41]. The $b$-tagging efficiency uncertainty depends on jet $p_{T}$ and comes mainly from the uncertainty on the measurement of the efficiency in $t\bar{t}$ events [39]. Similar uncertainties are derived for the $c$-tagging and light-flavor jet tagging [46].

Other experimental systematic uncertainties that are included but have a smaller impact are uncertainties from lepton energy scale and identification efficiency, the efficiency of the $E_{\text{miss}}^{\ell \ell}$ trigger and the uncertainty on the multi-jet background estimate, which is taken to be 100% of the estimated number of events.

In addition to the experimental systematic uncertainties, modeling systematic uncertainties are applied, accounting for possible differences between the data and the simulation model used for each process. For the background samples, the procedure described in Ref. [43] is followed. The $Z + \text{jets}$ and $W + \text{jets}$ backgrounds include uncertainties on the relative fraction of the different flavor components, and on the $m_{bb}$, $\Delta \phi_{bb}$ and $p_{T}^{\ell}$ distributions. For $t\bar{t}$ production, uncertainties on the top-quark transverse momentum, $m_{bb}$, $E_{\text{miss}}^{\ell \ell}$ and $p_{T}^{\ell}$ distributions are included. Uncertainties on the ratio of two-jet to three-jet events are also included for each background.

The $m_{T}^{\text{rec}}$ and $m_{A}^{\text{rec}}$ distributions for events passing the $\ell\ell bb$ and $\nu\nu bb$ final-state selections, respectively, are shown in Fig. 2. The distributions are shown after a profile-likelihood fit, which constrains simultaneously the signal yield and the background normalization and shape, which is performed in the same manner as in Ref. [43]. The overall background is more constrained than the individual components, causing the errors of individual components to be anti-correlated. The number of events passing the $\ell\ell bb$ and $\nu\nu bb$ final state selections are shown in Table 2, where the values for the expectations and uncertainties are obtained from the profile-likelihood fit.

### 6. Results

In all channels, no significant excess of events is observed in the data compared to the prediction from SM background sources. The significance of local excesses is estimated using $p$-values calculated with a test statistic based on the profile likelihood [47]. The largest data excesses are at $m_{A} = 220$ GeV ($p$-value = 0.014) and $m_{A} = 260$ GeV ($p$-value = 0.14) in the combined final states with $h \rightarrow b \bar{b}$ and $h \rightarrow \tau\tau$, respectively. Exclusion limits at the 95% confidence level (CL) are set on the production cross section times the branching ratio $BR(A \rightarrow Zh)$ as a function of the $A$ boson mass. The exclusion limits are calculated with a modified frequentist method [48], also known as CLs, and the profile likelihood method.
using the binned $m^\text{rec,T}_A$ mass distributions for $\ell\ell\tau\tau$ and $\ell\ell bb$ final states and the binned $m^\text{rec,T}_A$ distribution for the $\nu\nu bb$ final state.

Fig. 3 shows the 95% CL limits on the production cross section times the branching ratio, $\sigma (gg \to A) \times \text{BR}(A \to Zh) \times \text{BR}(h \to bb/\tau\tau)$, as well as the expected limits for each individual subchannel. The limit on the production times the branching ratio is in the range 0.098–0.013 pb and 0.57–0.014 pb for $m_A$ in the range 220–1000 GeV for the $\tau\tau$ and $bb$ channels, respectively. The $\tau\tau$ channels use few signal mass points beyond $m_A = 500$ GeV, since a coarse binning in $m^\text{rec,T}_A$ is adopted in view of the very small predicted number of background events.

The results of the search in the $\tau\tau$ and $bb$ channels are combined in the context of the CP-conserving 2HDM [3], which has seven free parameters and four arrangements of the Yukawa couplings to fermions. In particular, the free parameters are the Higgs boson masses $(m_h, m_A, m_{H^\pm}, m_{H^0})$, the ratio of the vacuum expectation values of the two doublets ($\tan\beta$), the mixing angle between the CP-even Higgs bosons ($\alpha$) and the potential parameter $m^2_{\text{eff}}$ that mixes the two Higgs doublets. The Yukawa coupling arrangements distinguish four different 2HDM models, determining which of the two doublets, $\Phi_1$ and $\Phi_2$, couples to up- and down-type quarks and leptons. In the Type-I model, $\Phi_2$ couples to all quarks and leptons, whereas in the Type-II, $\Phi_1$ couples to down-type fermions and $\Phi_2$ couples to up-type fermions. The lepton-specific model is similar to Type-I apart from the fact that the leptons couple to $\Phi_1$, instead of $\Phi_2$. The Flipped model is similar to Type-II apart from the leptons coupling to $\Phi_2$, instead of $\Phi_1$. In all these models, the limit $\cos(\beta - \alpha) \to 0$ is such that the light CP-even Higgs boson, $h$, has indistinguishable properties from a SM Higgs boson with the same mass. The cross sections for production by gluon fusion are calculated using SusHi [49–54] and the branching ratios are calculated with 2HDMC [55]. For the branching ratio calculations, it is assumed that $m_A = m_H = m_{H^\pm}$, $m_H = 125$ GeV and $m^2_{\text{eff}} = m_A^2 \tan^2 \beta/(1 + \tan^2 \beta)$.

The constraints derived from the combined search in $\tau\tau$ and $bb$ final states are presented as a function of 2HDM parameters. The exclusion region in the $\cos(\beta - \alpha)$ versus $\tan\beta$ plane for $m_A = 300$ GeV are shown in Fig. 4 for the four 2HDM models, while the constraints obtained in the $m_A - \tan\beta$ plane for $\cos(\beta - \alpha) = 0.10$ are shown in Fig. 5. The width of the $A$ boson in the 2HDM may be larger than the experimental mass resolution, and it is taken into account in the 2HDM parameter exclusion regions for widths up to 5% of $m_A$. For Type-II and Flipped models, Higgs boson production
in association with $b$-quarks dominates over gluon fusion for large tan $\beta$ values ($\tan \beta \gtrsim 10$). The cross section for the $b$-associated production uses an empirical matching of the cross sections in the four- and five-flavor schemes [56]. Cross sections in the four-flavor scheme are calculated according to Refs. [57,58] and cross sections in the five-flavor scheme are calculated using SusHi. The relative efficiencies for the $b$-associated and gluon fusion production as well as the predicted cross-section ratio are taken into account when deriving the constraints in the two-dimensional planes shown in Fig. 4. The $b$-associated production efficiencies are estimated using PYTHIA8 and SHERPA samples. The regions of parameter space excluded at 95% CL by the $A \rightarrow \tau\tau$ decay mode are displayed in the same plots, using the results of a search for a heavy Higgs boson decaying into $\tau\tau$ (Ref. [13]), reinterpreted considering only the production of an $A$ boson via gluon fusion and $b$-associated production. For $m_A$ values below the $t\bar{t}$ kinematic threshold, the search presented here can exclude $\cos(\beta - \alpha)$ values down to a few percent for $\tan \beta$ values up to $\approx 3$.

7. Conclusions

Data recorded in 2012 by the ATLAS experiment at the LHC, corresponding to an integrated luminosity of 20.3 fb$^{-1}$ of proton-proton collisions at a centre-of-mass energy 8 TeV, are used to search for a CP-odd Higgs boson, $A$, decaying to $Zh$, where $h$ denotes a light CP-even Higgs boson with a 125 GeV mass. No deviations from the SM background predictions are observed in the three final states considered: $Zh \rightarrow \ell\ell\tau\tau$, $Zh \rightarrow \ell\ellbb$, and $Zh \rightarrow \nu\nu bb$. Upper limits are set at the 95% confidence level for $\sigma(gg \rightarrow A) \times BR(A \rightarrow Zh) \times BR(h \rightarrow \ell\ell)$ of 0.098–0.013 pb for $f = \tau$ and 0.57–0.014 pb for $f = b$ in the range of $m_A = 220$–1000 GeV. This $Zh$ resonance search improves significantly the previously published constraints on CP-odd Higgs boson production in the low $\tan \beta$ region of the 2HDM.

Acknowledgements

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; BMWFW 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; DLR and DFKI, Germany; INFN, Italy; BMBF, DFG, HGF, MPG and AvH Founda-
The interpretation of the cross-section limits in the context of the various 2HDM types as a function of the parameters $\tan \beta$ and $m_A$ for $\cos (\beta - \alpha) = 0.1$: (a) Type-I (a), (b) Type-II, (c) Lepton-specific, and (d) Flipped. Variations of the natural width up to $\Gamma_A/m_A = 5\%$ are taken into account. The grey solid area indicates that the width is larger than 5\% of $m_A$. For Type-II and Flipped 2HDM, the $b$-associated production is included in addition to the gluon fusion. The blue (in the web version) shaded area denotes the area excluded by taking into account the constraints on the CP-odd Higgs boson derived by considering the $A \to t \bar{t}$ decay mode after reinterpreting the results in Ref. [11].

![Diagram](image_url)
Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary.

Also at Department of Physics, Oxford University, Oxford, United Kingdom.

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

Also at Discipline of Physics, University of KwaZulu-Natal, Durban, South Africa.

Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia.

* Deceased.