Search for the $HH \rightarrow b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

DOI
10.1007/JHEP07(2020)108

Publication date
2020

Document Version
Final published version

Published in
Journal of High Energy Physics

License
CC BY

Citation for published version (APA):
Search for the $HH \rightarrow b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: A search for Higgs boson pair production via vector-boson fusion (VBF) in the $b\bar{b}b\bar{b}$ final state is carried out with the ATLAS experiment using 126 fb$^{-1}$ of proton-proton collision data delivered at $\sqrt{s} = 13$ TeV by the Large Hadron Collider. This search is sensitive to VBF production of additional heavy bosons that may decay into Higgs boson pairs, and in a non-resonant topology it can constrain the quartic coupling between the Higgs bosons and vector bosons. No significant excess relative to the Standard Model expectation is observed, and limits on the production cross-section are set at the 95% confidence level for a heavy scalar resonance in the context of an extended Higgs sector, and for non-resonant Higgs boson pair production. Interpretation in terms of the coupling between a Higgs boson pair and two vector bosons is also provided: coupling values normalised to the Standard Model expectation of $\kappa_{2V} < -0.76$ and $\kappa_{2V} > 2.90$ are excluded at the 95% confidence level in data.

KEYWORDS: Hadron-Hadron scattering (experiments), Higgs physics

ArXiv ePrint: 2001.05178
1 Introduction

The Higgs boson \( (H) \) was discovered by the ATLAS and CMS collaborations in 2012 \cite{1, 2} using proton-proton \( (pp) \) collisions at the Large Hadron Collider (LHC). The measured properties have so far been found to be in agreement with the Standard Model (SM) predictions. The production of a pair of Higgs bosons \( (HH) \) is a rare process in the SM with a cross-section about 1000 times smaller than the single Higgs boson production cross-section, but various theories beyond the SM (BSM) predict cross-sections for \( HH \) production that are significantly higher than the SM prediction. Spin-0 resonances, with narrow or broad width, that decay into Higgs boson pairs, appear in BSM scenarios \cite{3, 4}. Enhanced non-resonant Higgs boson pair production is predicted by many models, for example those featuring light coloured scalars \cite{5} or new contact interactions, such as direct \( t\bar{t}HH \) vertices \cite{6, 7}.
\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures/fig1.png}
\caption{Tree-level Feynman diagrams contributing to Higgs boson pair production via VBF. Diagrams (a), (b) and (c) illustrate the non-resonant production modes scaling with $\kappa_V \kappa_\lambda$, $\kappa_V^2$ and $\kappa_{2V}$, respectively. Diagram (d) illustrates the resonant production mode.}
\end{figure}

Previous searches for Higgs boson pair production in the $b\bar{b}b\bar{b}$ channel were carried out in the gluon-gluon fusion (ggF) production mode by the ATLAS and CMS collaborations [8–12], and limits were set for resonant and non-resonant production. Statistical combinations of search results for $HH$ in various decay channels were also performed by the two experiments [13, 14], profiting from the sensitivity of several final states.

This paper focuses on searches for Higgs boson pair production via vector-boson fusion (VBF), through diagrams such as those presented in figure 1, and using the dominant $H \to b\bar{b}$ decay mode [15]. The VBF process ($pp \to HHjj$) is characterised by the presence of two jets ($j$) with a large rapidity gap resulting from quarks from which a vector boson ($V$) is radiated. In the SM, three different types of couplings are involved in $HH$ production via VBF: the Higgs boson self-coupling ($HHH$), the Higgs-boson-vector-boson coupling ($VVH$) and the quartic (di-vector-boson-di-Higgs-boson, or $VVHH$) coupling. The coupling modifiers $\kappa_\lambda$, $\kappa_V$ and $\kappa_{2V}$ control the strength of the $HHH$, $VVH$ and $VVHH$ couplings with respect to the SM value, respectively, and are normalised so that they are equal 1 in the SM. A deviation of these coupling modifiers from their SM expectations could lead to enhanced $HH$ production. While searches in the ggF mode are more sensitive to deviations in $\kappa_\lambda$, the VBF topology has unique sensitivity to $\kappa_{2V}$ [15] because the ggF mode does not involve the $VVHH$ interaction. For resonant production, two classes of signals are tested to perform a generic inclusive search for resonances with masses $m_X$ in the range 260–1000 GeV. The first signal class is representative of a broad resonance with width typically 10-20% of the signal mass; it corresponds to a heavy scalar of the 2HDM Type II model [16] and is obtained by setting the ratio of vacuum expectation values of the two Higgs doublets $\tan(\beta) = 2.0$ and $\sin(\beta - \alpha) = 0.6$, where $\alpha$ is the mixing angle between the two CP-even Higgs bosons. The second class features a narrow resonance with a fixed width of 4 MeV.

2 ATLAS detector

The ATLAS experiment [17–19] at the LHC operates a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and a near $4\pi$ coverage in solid angle.\footnote{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} It consists of an inner tracking detector surrounded by a thin superconducting
solenoid providing a 2 T axial magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip, and transition-radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A hadronic steel/scintillator-tile calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The endcap and forward regions are instrumented with LAr calorimeters for both the EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and incorporates three large air-core toroidal superconducting magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 Tm across most of the detector. The muon spectrometer includes a system of precision tracking chambers and fast detectors for triggering [20]. A two-level trigger system is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most 100 kHz. This is followed by the software-based high-level trigger (HLT), which reduces the accepted event rate to 1 kHz on average.

3 Data and simulated samples

This search is performed using data collected by the ATLAS experiment between 2016 and 2018 in $\sqrt{s} = 13$ TeV LHC $pp$ collisions, which correspond to an integrated luminosity of 126 fb$^{-1}$. Only events recorded during stable beam conditions and when the detector components relevant to the analysis were operating normally are considered [21]. During the 2016 data-taking, a fraction of the data was affected by an inefficiency in the vertex reconstruction in the HLT, which reduced the efficiency of the algorithms used to identify jets originating from $b$-hadron decays; those events were not retained for further analysis. This reduces the integrated luminosity of the 2016 dataset to 24.3 fb$^{-1}$.

Simulated Monte Carlo (MC) event samples are used to model signal production and the backgrounds from top-quark-pair ($t\bar{t}$) production. The dominant process arises from multijet production, and is modelled using data-driven techniques.

Events with a generic scalar resonance produced via VBF and decaying into $HH \rightarrow b\bar{b}b\bar{b}$ were generated with POWHEG-Box v2 [22-24] interfaced to PYTHIA 8.186 [25] for parton showering and hadronisation, with the Higgs boson mass fixed to 125 GeV [26]. The NNPDF23LO parton distribution function (PDF) set [27] and the A14 set of tuned parameters [28] for underlying-event simulation were used. Resonant signal samples were produced with masses ranging from 260 GeV to 1000 GeV.

Non-resonant production of Higgs boson pairs was simulated with MadGRAPH5_AMC@NLO [29]. The Higgs boson self-coupling and the couplings of the Higgs boson to vector bosons were set to their SM values, while the $\kappa_{2V}$ coupling modifier was varied. Samples with $\kappa_{2V}$ equal to 0.0, 0.5, 1.0, 1.5, 2.0, and 4.0 were generated at leading order (LO). Interference between various diagrams contributing to the non-resonant signal of the LHC ring, and the $y$-axis points upwards. 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 $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. 

---
is considered in the simulation. A linear combination of three samples is used to derive distributions for a finer granularity of $\kappa_{2V}$ values, following a technique used to generate $\kappa_3$ distributions [13]. A particular choice of three samples is done based on the $\kappa_{2V}$ range of the generated distributions to avoid large weights and to reduce statistical uncertainties. The cross-section of the VBF $HH$ process, evaluated at next-to-next-to-next-to-leading order (N$^3$LO) in QCD, is $1.73 \pm 0.04$ fb in the SM [30–33]. The N$^3$LO to LO cross-section ratio at the SM value is calculated and this factor is applied to the cross-sections at each $\kappa_{2V}$ point.

To estimate the contribution from ggF $HH$ production with two additional jets that can mimic the VBF $HH$ topology, the SM non-resonant production of Higgs boson pairs via ggF was simulated with MadGraph5_aMC@NLO using the CT10 PDF set [34] and the FTapprox method [35] to include finite top-quark mass effects. In this sample, the generation of $pp \rightarrow H +$ parton is done at next-to-leading order (NLO). Parton showers and hadronisation were simulated with Herwig 7.0.4 [36]. Interference effects with other SM processes are found to be marginal and are ignored. The cross-section is evaluated at next-to-next-to-leading order (NNLO) with the resummation at next-to-next-to-leading-logarithm (NNLL) accuracy and including top-quark mass effects at NLO [35, 37–42]; it is equal to $31.05^{+1.49}_{-1.99}$ fb. The uncertainty includes the variations of the factorisation and renormalisation scales, PDF and $\alpha_S$.

The generation of $t\bar{t}$ events was performed with Powheg-Box v2 [43] using the NNPDF3.0NLO [44] PDF set. The parton showers, hadronisation, and underlying event were simulated using Pythia 8.230 [45] with the NNPDF23LO PDF set and the corresponding A14 set of tuned underlying-event parameters. The predicted $t\bar{t}$ production cross-section is $831.8^{+19.8}_{-29.2} \pm 35.1$ pb as calculated with the Top++ 2.0 program to NNLO in perturbative QCD, including soft-gluon resummation to NNLL accuracy [46], and assuming a top-quark mass of 172.5 GeV. The first uncertainty comes from the independent variations of the factorisation and renormalisation scales, while the second one is associated with variations in the PDF and $\alpha_S$, following the PDF4LHC prescription with the MSTW2008 68% CL NNLO, CT10 NNLO and NNPDF2.3 5f FFN PDF sets [27, 47–49].

For all simulated events, $c$-hadron and $b$-hadron decays were handled by EvtGen 1.2.0 [50]. To simulate the impact of multiple $pp$ interactions that occur within the same or nearby bunch crossings (pile-up), minimum-bias events generated with Pythia 8.186 using the NNPDF23LO set of PDFs and the A3 set of tuned parameters [51] were overlaid on the hard-scatter process. The detector response was simulated with GEANT 4 [52, 53], and the events were processed with the same reconstruction software as that was used for the data.

## 4 Event reconstruction

Events are required to have at least one reconstructed primary vertex with at least two associated tracks, each with transverse momentum $p_T > 0.4$ GeV. For events with more than one primary-vertex candidate, the one with the largest track $\sum p_T^2$ is chosen as the hard-scatter primary vertex.
Jets are reconstructed from three-dimensional topological clusters of energy deposits in the calorimeter [54] with the anti-$k_t$ algorithm [55] implemented in the FastJet package [56] with radius parameter $R = 0.4$. Clusters are calibrated at the EM scale [57] and their energy is corrected for additional energy deposits from pile-up interactions using an area-based correction [58]. Subsequently, calibration using $p_T$- and $\eta$-dependent factors derived from simulation is applied, followed by the global sequential calibration [57]. The latter reduces the flavour dependence of the calibration and energy leakage effects. The final calibration is based on in situ measurements in collision data [57]. To preferentially reject jets originating from pile-up interactions, a multivariate classification algorithm (jet vertex tagger) based on tracking information [59] is used for jets with $p_T < 60$ GeV and $|\eta| < 2.4$. The selected working point provides an inclusive hard-scatter process efficiency of about 97% in that kinematic region. The efficiency in the simulation is corrected to match that measured in data. Events having jets consistent with noise in the calorimeter or non-collision backgrounds are vetoed [57].

Jets containing $b$-hadrons are identified using a multivariate algorithm (MV2c10) [60, 61], which exploits information about the jet kinematics, the impact parameters of tracks associated with the jet and the presence of displaced vertices to form the decision. The $b$-tagging requirements result in an efficiency of 70% for jets with $p_T > 20$ GeV containing $b$-hadrons, and the misidentification rate is 0.3% (11.2%) for light-flavour (charm) jets. These were determined in a sample of simulated $t\bar{t}$ events. For all simulated events the $b$-tagging efficiencies are corrected to match those measured in data [60, 62, 63].

To further correct the $b$-jet energy for effects that are not considered in the default calibration, a jet energy regression is used. The method uses a boosted decision tree (BDT) algorithm implemented in TMVA [64]. The BDT training is performed using variables the $b$-jet energy resolution is sensitive to: MV2c10 score, energy leakage outside the jet cone, pile-up contamination, hard radiation from the original parton, and energy loss through semileptonic $b$-hadron decays. Both the training and the validation are performed with simulated $t\bar{t}$ samples, resulting in approximately 10% improvement in the jet energy resolution. The performance of the jet energy regression is validated in $Z(\rightarrow \mu\mu) + b/\bar{b}b$ events in data and no mismodelling is found. The $H \rightarrow b\bar{b}$ mass peak is found to be closer to 125 GeV and the standard deviation divided by the mean of the mass distribution is improved by about 25% for the $m_X = 600$ GeV signal sample.

5 Event selection

Events are selected using a combination of $b$-jet triggers, with the lowest jet transverse energy, $E_T$, threshold at 35 GeV, jet $|\eta| < 2.5$ and one or two $b$-tagged jets. The $b$-jet trigger efficiency is measured in data, and the simulated events are corrected to match the measured trigger efficiency.

To select events compatible with VBF production of Higgs boson pairs decaying into four $b$-quarks, exactly four central $b$-tagged jets with $p_T > 40$ GeV and $|\eta| < 2.0$ and at least two forward jets with $p_T > 30$ GeV and $|\eta| > 2.0$ are required. Events with more than four central jets are vetoed in case more than four jets fulfil the $b$-tagging requirement. To
The following pairings. Their invariant mass, $m_{b\bar{b}}$, exceeds 5.0 and $m_{jj}^{VBF}$ is greater than 1000 GeV.

### 5.1 VBF-jets selection

The two highest-$p_T$ forward jets with opposite sign of $\eta$ are considered as remnants of the VBF production process if the absolute value of the pseudorapidity separation between them, $|\Delta\eta_{jj}^{VBF}|$, exceeds 5.0 and their invariant mass, $m_{jj}^{VBF}$, is greater than 1000 GeV.

### 5.2 Signal kinematics selection

The four central $b$-tagged jets are considered in three possible combinations of two-jet pairings. Their invariant mass, $m_{4b}$, is used to define criteria to select signal-like events. The following $\Delta R$ requirements must be satisfied; these reflect the correlation of $m_{4b}$ with
the Lorentz boost of the Higgs bosons and the angle between their decay products in the laborator frame:

\[
\begin{aligned}
\frac{360 \text{ GeV}}{m_{4b}} - 0.5 &< \Delta R_{bb}^{\text{lead}} < \frac{653 \text{ GeV}}{m_{4b}} + 0.475 \\
\frac{235 \text{ GeV}}{m_{4b}} &< \Delta R_{bb}^{\text{subl}} < \frac{875 \text{ GeV}}{m_{4b}} + 0.35
\end{aligned}
\]

if \( m_{4b} < 1250 \text{ GeV} \),

\[
\Delta R_{bb}^{\text{lead}} < 1
\]

\[
\Delta R_{bb}^{\text{subl}} < 1
\]

if \( m_{4b} \geq 1250 \text{ GeV} \),

where \( \Delta R_{bb}^{\text{lead}} \) and \( \Delta R_{bb}^{\text{subl}} \) are the angular distances between the jets that form, respectively, the leading and subleading Higgs boson candidates (ordered in \( p_T \)). These criteria are optimised for both non-resonant and resonant Higgs boson pair production, and the numerical values are chosen to maximise the signal sensitivity.

Out of the possible pairings fulfilling the previous selection, the combination that leads to pairs with a dijet mass closest to the SM Higgs boson mass should be the optimal choice. However, due to energy loss through semileptonic \( b \)-hadron decays, this criterion is relaxed. The mass values of 123.7 GeV for the leading Higgs boson candidate and 116.5 GeV for the subleading Higgs boson candidate are found to maximise the signal significance for a resonance with a mass of 600 GeV, which lies in the middle of the covered mass range. The same target values are used for all other signal hypotheses. For a given pairing, the quantity \( D_{HH} \) that corresponds to the distance of the leading and subleading Higgs boson candidate masses, in the \( (m_{2b}^{\text{lead}}, m_{2b}^{\text{subl}}) \) plane, from the line connecting \((0 \text{ GeV}, 0 \text{ GeV})\) and \((123.7 \text{ GeV}, 116.5 \text{ GeV})\), can be computed as:

\[
D_{HH} = \sqrt{(m_{2b}^{\text{lead}})^2 + (m_{2b}^{\text{subl}})^2} \sin \left( \tan^{-1} \left( \frac{m_{2b}^{\text{subl}}}{m_{2b}^{\text{lead}}} \right) - \tan^{-1} \left( \frac{116.5 \text{ GeV}}{123.7 \text{ GeV}} \right) \right).
\]

The pairing with the smallest value of \( D_{HH} \) is chosen. Studies based on simulation indicate that for SM non-resonant \( HH \) production the correct pairs are identified in at least 83% of the signal events, while for broad resonances the corresponding fraction is greater than 91%.

### 5.3 Selection for background suppression

In order to enhance the sensitivity to signal, various requirements are applied to suppress the background. The magnitude of the vector sum of the transverse momenta of the selected four \( b \)-jets and the two VBF-jets tends to peak at lower values for signal events than for multijet events. Consequently, it is required to be less than 60 GeV. The pseudorapidity difference between the reconstructed Higgs boson candidates, \(|\Delta \eta_{HH}|\), is required to be below 1.5, and mass-dependent requirements on the transverse momenta of the leading and subleading Higgs boson candidates, respectively \( p_{T,H}^{\text{lead}} \) and \( p_{T,H}^{\text{subl}} \), are:

\[
p_{T,H}^{\text{lead}} > 0.5 m_{4b} - 103 \text{ GeV},
\]

\[
p_{T,H}^{\text{subl}} > 0.33 m_{4b} - 73 \text{ GeV}.
\]
The resulting dijet pairs are still dominated by multijet events. To further increase the search sensitivity, the dijet masses are required to fulfill:

\[ X_{HH} = \sqrt{\left( \frac{m_{2b}^{\text{lead}} - 123.7 \text{ GeV}}{11.6 \text{ GeV}} \right)^2 + \left( \frac{m_{2b}^{\text{subl}} - 116.5 \text{ GeV}}{18.1 \text{ GeV}} \right)^2} < 1.6, \]

where 11.6 GeV and 18.1 GeV are the experimental widths of the simulated leading and subleading Higgs boson candidates, respectively. The mass resolution of the subleading Higgs boson candidate is worse because it is composed of the lower \( p_T \) jets. These values are derived using a 600 GeV resonant signal sample and are similar for other signal samples.

Additional requirements are imposed to reduce the number of hadronically decaying \( t\bar{t} \) events by vetoing candidate events compatible with a top-quark decay. The jet with the highest \( b \)-tagging score is considered as the \( b \)-jet originating from a top-quark decay and the remaining central jets are considered to stem from the \( W \)-boson decay. Since the top quarks are expected to be produced centrally, only central jets are tested. All possible two-jet combinations in the event are tested and the selected combination is the one with the smallest value of \( X_{Wt} \), defined as:

\[ X_{Wt} = \sqrt{\left( \frac{m_W - 80.4 \text{ GeV}}{0.1 m_W} \right)^2 + \left( \frac{m_t - 172.5 \text{ GeV}}{0.1 m_t} \right)^2} \]

where \( m_W \) and \( m_t \) are the reconstructed invariant masses of the \( W \)-boson and top-quark candidates, respectively. The event is vetoed if \( X_{Wt} \leq 1.5 \). This requirement reduces the \( t\bar{t} \) contamination by about 50% with negligible impact on the signal efficiency.

All requirements listed above define the signal region (SR). The number of selected signal events divided by the number of generated events after each selection step (cumulative acceptance times efficiency) is shown in figure 2 for the non-resonant signal as a function of the \( \kappa_{2V} \) coupling modifier and for the resonant signal models as a function of the generated mass. The acceptance times efficiency increases as a function of the resonance mass, while for the non-resonant signal a significant drop is observed at \( \kappa_{2V} = 1 \). The trigger and jet selection requirements cause the drop for \( \kappa_{2V} \) values around 1, while the smaller acceptance times efficiency for low-mass resonances is due to the softer \( p_T \) spectrum of \( b \)-jets.

To estimate the background and to validate the background estimation technique, two regions orthogonal to the SR are used: the sideband region (SB) and the validation region (VR). The events in the SB and VR must fail the requirement defined in eq. (5.1) and fulfill

\[ \sqrt{\left( m_{2b}^{\text{lead}} - 123.7 \text{ GeV} \right)^2 + \left( m_{2b}^{\text{subl}} - 116.5 \text{ GeV} \right)^2} < 30 \text{ GeV} \]

requirement in the VR and

\[ 30 \text{ GeV} < \sqrt{\left( m_{2b}^{\text{lead}} - 123.7 \text{ GeV} \right)^2 + \left( m_{2b}^{\text{subl}} - 116.5 \text{ GeV} \right)^2} < 45 \text{ GeV} \]

requirement in the SB. The \( m_{2b} \) distributions of the leading versus subleading Higgs boson candidates for the non-resonant signal and the multijet background are shown in figure 3, together with the contours of the SR, VR and SB.
Figure 2. Cumulative acceptance times efficiency at each stage of the event selection, as detailed in section 5. The number of events surviving the selection divided by the number of generated events is reported separately for the non-resonant signal as a function of the $\kappa_{2\nu}$ coupling modifier and for the narrow- and broad-width resonance production hypotheses as a function of the generated mass.

Figure 3. Two-dimensional mass regions used in the analysis. The signal region is inside the inner (red) dashed curve, the validation region is outside the signal region and within the intermediate (orange) circle, and the sideband is outside the validation region and within the outer (yellow) circle. The regions are shown for (a) simulated events from the SM non-resonant $HH$ process and (b) the estimated multijet background.

6 Background estimation

After the event selection described in section 5, the background is dominated by multijet and $tt$ events. The multijet events constitute about 95% of the total background and are modelled using data. The remaining 5% are $tt$ events, which are modelled using simulation. The normalisation of the all-hadronic $tt$ background is determined from data, whereas the non-all-hadronic $tt$ background is normalised to the SM prediction. In the SM, the contribution of the $HH$ pairs produced via ggF is small compared to other backgrounds.
and three times larger than for the VBF production. Thus $HH$ production via ggF is treated as a background in this analysis and is fixed to the SM prediction. Other minor backgrounds with contributions below 0.5% are neglected. The background estimation technique is the same as in ref. [8].

6.1 Multijet background

The data-driven multijet background estimation uses data events with lower $b$-jet multiplicity and reweights them to model events with higher $b$-jet multiplicity. The multijet events are selected using the same trigger and selection requirements as those used in the SR, except for the $b$-tagging requirement. In particular, the SR requires at least four $b$-jets (“four-tag sample”). To derive a background estimate for this region, events with at least four central jets, but with only two of them $b$-tagged (“two-tag sample”), are used. The events in the two-tag sample are reweighted by applying a product of two event weights. The first event weight corrects for the additional $b$-tagged jet activity and the second event weight corrects for the kinematic differences caused by requiring additional $b$-tagged jets. These differences can arise for a variety of reasons: the $b$-tagging efficiency varies as a function of jet $p_T$ and $\eta$; the various multijet processes contribute with different fractions in each sample; and the fraction of events accepted by each trigger path changes. The reweighting is performed using one-dimensional distributions and is iterated until the weights converge to stable values. Details of the reweighting procedure can be found in ref. [8]. The weights are derived in the SB using the procedure described above and validated in the VR.

6.2 $t\bar{t}$ background

The shape of the $t\bar{t}$ background is modelled using simulation. The $t\bar{t}$ events are expected to contain two $b$-jets from the decay of two top quarks and additional jets stemming from the hadronic $W$-boson decay or additional quarks or gluons produced together with two top quarks. To reduce the statistical uncertainty, simulated $t\bar{t}$ events in the two-tag region, corrected by the kinematic weights derived for the multijet background, are also used. The procedure is validated in the SB, and good agreement is observed between the corrected two-tag sample and the four-tag sample within the uncertainties. Samples for all-hadronic and non-all-hadronic $t\bar{t}$ decays are handled separately.

6.3 Background normalisation

The normalisations of multijet and all-hadronic $t\bar{t}$ backgrounds are derived simultaneously by fitting the $X_{Wt}$ distribution to data in the SB. The $X_{Wt}$ distribution differs for the two backgrounds: the region of $X_{Wt} < 0.75$ is enriched in all-hadronic $t\bar{t}$ events and the region $X_{Wt} > 0.75$ is enriched in multijet events. The normalisation of the non-all-hadronic $t\bar{t}$ background is fixed to its SM prediction in the fit due to its small contribution to the total yields. Two parameters are used in the normalisation fit: $f_{\text{multijet}}$ and $f_{\text{all-had. } t\bar{t}}$. The $f_{\text{multijet}}$ parameter scales the multijet yield from the two-tag to the four-tag sideband region after the reweighting described in section 6.1. The $f_{\text{all-had. } t\bar{t}}$ parameter corrects the normalisation of all-hadronic $t\bar{t}$ yields in the four-tag sideband region. The fit results are cross-checked in the VR, where the same results are obtained within statistical uncertainties.
7 Systematic uncertainties

Background normalisation uncertainties are propagated from the fit, described in the previous section, which determines the multijet and all-hadronic \( t\bar{t} \) yields. The statistical uncertainty of the multijet and all-hadronic \( t\bar{t} \) normalisation parameters is accounted for, including their correlations. Two nuisance parameters are defined in the final fit described in section 8 by calculating two eigenvectors from the covariance matrix of the normalisation fit. Furthermore, the normalisation of the multijet background estimate is verified in the VR, where agreement with data is found; its statistical uncertainty in the VR is thus applied as a normalisation systematic uncertainty of the multijet background.

Two shape uncertainties for the multijet background modelling are evaluated using data from the VR. The multijet modelling uncertainty is related to the level of agreement between the data and the background model in this region. To evaluate this uncertainty, the \( m_{4\text{b}} \) distribution is split into low and high mass regions at 400 GeV and a linear fit to the ratio of this distribution between data and the sum of all backgrounds is performed in both mass regions. Studies show that the ratio, which appears to be constant, can be sufficiently well described by a straight-line fit, and that the change in the result is marginal for other splitting points. The \(+1\sigma\) variation of the slope of the fitted line and the \(+1\sigma\) variation of the fitted line with inverted slope are used as up and down variations of the uncertainty in the background shape. The yield of the varied multijet templates is fixed to its nominal value. The varied templates are derived separately for the low- and high-\( m_{4\text{b}} \) mass regions, the corresponding systematics are treated as uncorrelated in the final fit. The kinematic reweighting uncertainty is assessed by deriving an alternative multijet template using the same procedure as in the nominal case, but using data from the VR. This difference between the multijet templates derived in the SB and the VR is symmetrised around the nominal template, keeping the yield fixed to its nominal value.

The shape modelling uncertainty of the \( t\bar{t} \) background is evaluated by comparing the nominal \( t\bar{t} \) template derived using the two-tag sample as described in section 6 and a \( t\bar{t} \) template derived using the four-tag sample after the basic preselection requirements defined in section 5. A straight-line fit to the ratio of the two-tag to four-tag \( m_{4\text{b}} \) templates is performed. With a procedure similar to the one applied for the multijet shape uncertainties, the up and down variations around the nominal \( t\bar{t} \) template are extracted using a straight-line fit to the \(+1\sigma\) variation of the slope of the fitted line and the \(+1\sigma\) variation of the fitted line with inverted slope, while keeping the yield fixed to its nominal value. This uncertainty is derived separately for the non-all-hadronic (using MC) and all-hadronic (using data) \( t\bar{t} \) samples and is not correlated in the fit between these two samples.

Theoretical uncertainties in the ggF background yield are evaluated by varying the renormalisation and factorisation scales and from the uncertainty associated with the choice of PDF set. The resulting variation of the expected ggF background yield is about 10%. When considering the same sources of theoretical uncertainty for the VBF \( HH \) signal, its acceptance times efficiency varies by 3%.

The experimental uncertainties listed below affect only MC samples. The uncertainties in the jet energy resolution and scale are evaluated at \( \sqrt{s} = 13 \) TeV using in situ measure-
ment techniques described in ref. [57]. The sources of uncertainty in these measurements are treated as fully correlated between the $p_T$ and mass scales. The resolution uncertainty is evaluated in measurements documented in ref. [65] and is assessed by applying an additional smearing to these observables. The flavour tagging efficiency and its uncertainty for $b$- and $c$-jets is estimated in $t\bar{t}$ events, while the light-jet misidentification rate and uncertainty is determined using dijet events [60, 62, 63]. In addition, an uncertainty in the $b$-jet trigger efficiency is derived from the per-jet online $b$-tagging measurements [66]. The uncertainty in the integrated luminosity is 1.7% [67], obtained using the LUCID-2 detector [68] for the primary luminosity measurements.

8 Results

Following the statistical procedures outlined in ref. [1], a test statistic based on the profile likelihood ratio [69] is used to test hypothesised values $\sigma_{VBF}$ of the cross-section of the signal model in units of fb. This test statistic extracts the information about the signal cross-section from a likelihood fit to the data. The likelihood function includes all parameters which describe the systematic uncertainties and their correlations discussed in section 7. As no significant excess over the background prediction is observed, exclusion limits are computed using the asymptotic formula [69]. The exclusion limits are based on the CLs method [70], where a value of $\sigma_{VBF}$ is regarded as excluded at the 95% confidence level (CL) when CLs is smaller than 5%. The accuracy of the asymptotic approximation is verified with sampling distributions generated using pseudo-experiments.

The mass of the four selected $b$-jets, $m_{4b}$, in the SR is used as the final discriminant for limit setting. Figure 4 shows the distribution of data and the SM background after the background-only fit in the SR and VR. In addition, the signal prediction for the narrow-width resonance hypothesis with $m_X = 800$ GeV and the non-resonant signal at $\kappa_{2V} = 3$ are shown in the SR.

Upper limits on the cross-section are set for all tested models. Figure 5 shows the 95% CL upper limits for resonant $HH$ production via VBF as a function of the resonance mass $m_X$ for narrow- and broad-width resonance hypotheses. The significance of the excess over the background-only prediction is quantified using the local $p_0$-value, defined as the probability of the background-only model to produce a signal-like fluctuation at least as large as that observed in the data. The most extreme $p_0$-value corresponds to a local significance of 1.5 standard deviations at 550 GeV.

The expected and observed limits on SM non-resonant $HH$ production via VBF are given in table 2. Limits are also calculated as a function of $\kappa_{2V}$, as presented in figure 6. The observed excluded region corresponds to $\kappa_{2V} < -0.76$ and $\kappa_{2V} > 2.90$, while the expected exclusion is $\kappa_{2V} < -0.91$ and $\kappa_{2V} > 3.14$. For $\kappa_{2V}$ values deviating from the SM prediction, growing non-cancellation effects result in a harder $m_{HH}$ spectrum, and thereby higher-$p_T$ $b$-jets, which in turn lead to increased signal acceptance times efficiency as shown in figure 2. This search is therefore not sensitive to the region close to the SM prediction, corresponding to $\kappa_{2V} = 1$.

Table 3 summarises the relative impact of the uncertainties on the best-fit signal cross-section for two different narrow-width resonance production hypotheses, with masses equal
Figure 4. Post-fit mass distribution of the $HH$ candidates in the (a) signal and (b) validation regions. The expected background is shown after the profile-likelihood fit to data with the background-only hypothesis; the narrow-width resonant signal at 800 GeV and the non-resonant signal at $\kappa_{2V} = 3$ are overlaid in the signal region, both normalised to the corresponding observed upper limits on the cross-section. The lower panel shows the ratio of the observed data to the estimated SM background. The distribution of events is shown per mass interval corresponding to the bin width of 40 GeV, while the overflow events are included in the last bin.

Figure 5. Observed and expected 95% CL upper limits on the production cross-section for resonant $HH$ production via VBF as a function of the mass $m_X$. The (a) narrow- and (b) broad-width resonance hypotheses are presented.

to 300 GeV and 800 GeV. Only major sources of systematic uncertainty are quoted along with the impact of the statistical uncertainty. The uncertainties of similar nature are grouped into unique categories and the fit is performed independently for the two hypothesised signals. The systematic uncertainties related to the multijet background estimate have the largest impact on the result.
Table 2. Upper limits at 95% CL for SM non-resonant \( HH \) production via VBF in \( \text{fb} \) (first row) and normalised to its SM expectation, \( \sigma_\text{VBF}^{\text{SM}} \) (second row). Uncertainties related to the branching ratio of the \( H \to bb \) decay are not considered.

<table>
<thead>
<tr>
<th>( \sigma_\text{VBF} ) [( \text{fb} )]</th>
<th>Observed</th>
<th>(-2\sigma)</th>
<th>(-1\sigma)</th>
<th>Expected</th>
<th>(+1\sigma)</th>
<th>(+2\sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_\text{VBF} / \sigma_\text{VBF}^{\text{SM}} )</td>
<td>1460</td>
<td>510</td>
<td>690</td>
<td>950</td>
<td>1330</td>
<td>1780</td>
</tr>
<tr>
<td>840</td>
<td>290</td>
<td>400</td>
<td>550</td>
<td>770</td>
<td>1030</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. Observed and expected 95% CL upper limits on the production cross-section for non-resonant \( HH \) production via VBF as a function of the di-vector-boson-di-Higgs-boson coupling modifier \( \kappa_{2V} \). The theory prediction of the cross-section as a function of \( \kappa_{2V} \) is also shown. More details on the predicted cross-section can be found in section 3.

<table>
<thead>
<tr>
<th>Source</th>
<th>( m_X = 300 \text{ GeV} )</th>
<th>( m_X = 800 \text{ GeV} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet normalisation</td>
<td>46%</td>
<td>Multijet modelling</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>26%</td>
<td>Jet energy resolution</td>
</tr>
<tr>
<td>Multijet modelling</td>
<td>18%</td>
<td>Jet energy scale</td>
</tr>
<tr>
<td>Multijet kinematic reweighting</td>
<td>17%</td>
<td>Multijet kinematic reweighting</td>
</tr>
<tr>
<td>( tt ) modelling</td>
<td>11%</td>
<td>Multijet normalisation</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>10%</td>
<td>( tt ) modelling</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>64%</td>
<td>Total systematic uncertainty</td>
</tr>
<tr>
<td>Statistical uncertainty</td>
<td>77%</td>
<td>Statistical uncertainty</td>
</tr>
</tbody>
</table>

Table 3. Dominant relative uncertainties in the best-fit signal cross-section \( \sigma_{\text{VBF}}^{\text{best fit}}(pp \to X jj \to HHjj) \) of hypothesised resonant \( HH \) signal production. The leading sources of systematic uncertainty, the total systematic uncertainty and the data statistical uncertainty are provided. Two mass points are selected: \( m_X = 300 \text{ GeV} \) with the best-fit cross-section of 140 \( \text{fb} \) and \( m_X = 800 \text{ GeV} \) with 4.7 \( \text{fb} \), which correspond to the low and high mass regions. The groups of uncertainties do not add up in quadrature to the total uncertainty, because only the dominant uncertainties are shown and also due to correlations between the uncertainties.
9 Conclusion

A search for both resonant and non-resonant production of pairs of Standard Model Higgs bosons via vector-boson fusion has been carried out in the $b\bar{b}b\bar{b}$ channel. The analysed data were collected from $\sqrt{s} = 13$ TeV proton-proton collisions by the ATLAS detector at the LHC in 2016-2018 and correspond to an integrated luminosity of 126 fb$^{-1}$. Results for resonant $HH$ production are reported in the mass range 260–1000 GeV. The largest deviation from the background-only hypothesis is observed at 550 GeV with a local significance of 1.5 standard deviations. Upper limits on the production cross-section are set for narrow- and broad-width scalar resonances at 95% CL. Limits are also set on the cross-section for non-resonant $HH$ production, and as a function of the di-vector-boson-di-Higgs-boson coupling modifier, $\kappa_{2V}$. The observed 95% CL upper limit on the SM non-resonant $HH$ production cross-section is 1460 fb, compatible with the expected limit at a level below two standard deviations. The observed excluded region corresponds to $\kappa_{2V} < -0.76$ and $\kappa_{2V} > 2.90$, while the expected exclusion is $\kappa_{2V} < -0.91$ and $\kappa_{2V} > 3.11$.

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; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRT, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRS, Morocco; NWO, Netherlands; RCN, Norway; MNI SW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russia Federation; JINR; MESTD, 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, CANARIE, Compute Canada and CRC, Canada; ERC, ERDF, Horizon 2020, Marie Sklodowska-Curie Actions and COST, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya and PROMETEO Programme Generalitat Valenciana, Spain; Göran Gustafsson Stiftelse, Sweden; 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. [71].

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

[8] ATLAS collaboration, Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s}=13$ TeV with the ATLAS detector, JHEP 01 (2019) 030 [arXiv:1901.06174] [inSPIRE].


[70] A.L. Read, Presentation of search results: the CLs technique, [inSPIRE].

The ATLAS collaboration
