Search for Higgs boson pair production in the bbbb final state from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector


DOI
10.1140/epjc/s10052-015-3628-x

Publication date
2015

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).
Search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state from $pp$ collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector

ATLAS Collaboration
CERN, 1211 Geneva 23, Switzerland

Abstract A search for Higgs boson pair production $pp \to hh$ has been performed with $19.5 \, fb^{-1}$ of proton–proton collision data at $\sqrt{s} = 8$ TeV, which were recorded by the ATLAS detector at the Large Hadron Collider in 2012. The decay products of each Higgs boson are reconstructed as a high-momentum $b\bar{b}$ system with either a pair of small-radius jets or a single large-radius jet, the latter exploiting jet substructure techniques and associated $b$-tagged track-jets. No evidence for resonant or non-resonant Higgs boson pair production is observed. The data are interpreted in the context of the Randall–Sundrum model with a warped extra dimension as well as the two-Higgs-doublet model. An upper limit on the cross-section for $pp \to G^*_KK \to hh \to b\bar{b}b\bar{b}$ of $3.2 \,(2.3) \, fb$ is set for a Kaluza–Klein graviton $G^*_KK$ mass of $1.0 \,(1.5) \, TeV$, at the $95\%$ confidence level. The search for non-resonant Standard Model $hh$ production sets an observed $95\%$ confidence level upper limit on the production cross-section $\sigma(pp \to hh \to b\bar{b}b\bar{b})$ of $202 \, fb$, compared to a Standard Model prediction of $\sigma(pp \to hh \to b\bar{b}b\bar{b}) = 3.6 \pm 0.5 \, fb$.

1 Introduction

The discovery of a Higgs boson ($h$) [1,2] at the Large Hadron Collider (LHC) consistent with the predictions of the Standard Model (SM) [3,4] motivates an enhanced effort to search for new physics via the Higgs sector. Many new physics models predict rates of Higgs boson pair production significantly higher than the SM rate [5–7]. For example, TeV-scale resonances such as the first Kaluza–Klein (KK) excitation of the graviton, $G^*_KK$, predicted in the bulk Randall–Sundrum (RS) model [8,9] or the heavy neutral scalar, $H$, of two-Higgs-doublet models (2HDM) [10] can decay into pairs of Higgs bosons, $hh$. Enhanced non-resonant $pp \to hh$ production can arise in models such as those with new, light, coloured scalars [11], or direct $t\bar{t}hh$ vertices [12,13].

ATLAS has carried out a search in the $b\bar{b} \gamma\gamma$ final state [14], setting limits on both resonant (masses between 260 GeV and 500 GeV) and non-resonant Higgs boson pair production. CMS has searched in the multi-lepton and multi-lepton + photons final-states in the context of 2HDM extensions of the Higgs sector [15]. CMS has also searched for narrow resonances in the $b\bar{b}b\bar{b}$ channel [16].

Recent phenomenological studies have demonstrated that despite the fully hadronic final state being subject to a large multijet background, searches for new physics in the $pp \to hh \to b\bar{b}b\bar{b}$ process have good sensitivity for both resonant [17,18] and non-resonant signals [19]. One contributing factor to this sensitivity is the high expected branching ratio for $h \to b\bar{b}$. The analysis presented in this paper is designed to search for two high-momentum $b\bar{b}$ systems with masses consistent with $m_h$, where each $b\bar{b}$ system contains two jets identified as containing $b$-hadrons (the jets are “$b$-tagged”). Compared to a more inclusive $b\bar{b}b\bar{b}$ final-state analysis, this topology has many benefits due to the large required momentum and angular separation between the two $b\bar{b}$ systems: (i) excellent rejection of all backgrounds; (ii) highly efficient triggering using $b$-tagged multijet triggers; and (iii) negligible combinatorial ambiguity in forming Higgs boson candidates.

Two Higgs boson reconstruction techniques, which are complementary in their acceptance, are presented. The first – “resolved” – technique reconstructs Higgs boson candidates from pairs of nearby anti-$k_t$ jets [20] with radius parameter $R = 0.4$, each $b$-tagged with a multivariate $b$-tagging algorithm [21]. This resolved technique offers good efficiency over a wide range of Higgs boson momenta and so can be used to reconstruct di-Higgs-boson resonances with mass $m_X$ in the range between 500 and 1500 GeV. The sensitivity is best for this technique in the range $500 \leq m_X \lesssim 1100$ GeV. It can be seen in Fig. 1 however, that the acceptance for four $b$-tagged anti-$k_t$ jets decreases for $m_X \gtrsim 1200$ GeV. This loss of acceptance is due to the increased boost of the Higgs boson, which reduces the average separation between the $b$- and $\bar{b}$-quarks from the Higgs boson.
jet clustering. The requirements are defined in Sects. 4.2 and 5.2
between the $b/\Delta 1$ decay, boosted analyses. The shapes of the curves are driven by the separation between the $b$-quarks from the Higgs boson decays and the impact on jet clustering. The requirements are defined in Sects. 4.2 and 5.2.

\begin{equation}
\Delta R = \sqrt{\left(\Delta \eta\right)^2 + \left(\Delta \phi\right)^2},
\end{equation}

Fig. 1 Acceptance times reconstruction and selection efficiency as a function of graviton mass for the bulk RS model, for the resolved and boosted analyses. The shapes of the curves are driven by the separation between the $b$-quarks from the Higgs boson decays and the impact on jet clustering. The requirements are defined in Sects. 4.2 and 5.2.

decay, $\Delta R = \sqrt{\left(\Delta \eta\right)^2 + \left(\Delta \phi\right)^2}$, to values below 0.4. This motivates the use of a second – “boosted” – Higgs boson reconstruction technique that maintains acceptance for these higher-mass resonances through the use of jet substructure techniques. The Higgs boson candidate is reconstructed as a single, trimmed [22] anti-$k_t$ $R = 1.0$ jet which must be associated with two $b$-tagged anti-$k_t$ $R = 0.3$ track-jets [23]. The use of track-jets with a smaller $R$ parameter allows Higgs bosons with higher transverse momentum ($p_T$) to be reconstructed.

The analysis is performed with the dataset recorded by ATLAS in 2012 at $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 19.5 fb$^{-1}$. For the non-resonant search, a counting experiment is performed and the results are interpreted in the context of SM non-resonant Higgs boson pair production. This interpretation is only carried out for the resolved analysis due to its higher sensitivity to such a signal. For the resonant search, a fit to the reconstructed mass spectrum of $hh$ candidates is carried out and the results are interpreted in the context of both bulk RS $G_{kk}^s$ (spin-2) and 2HDM CP-even H boson (spin-0) production. In the bulk RS model, the fermion and boson fields can propagate in a warped extra dimension, which has a curvature parameter $k$. This benchmark model is investigated with three coupling constant values, $k/M_\text{Pl} = 1.0, 1.5$ and $2.0 (M_\text{Pl} = M_\text{Pl}/\sqrt{8\pi}$ is the reduced Planck mass), which cover much of the possible parameter space [8]. The 2HDMs considered have CP-conserving scalar potentials (Type-I, Type-II, Lepton-specific and Flipped) [10], in the regime $m_H = m_A = m_{H^\pm}$, with the potential parameter that mixes the two Higgs doublets $m_{12}^2 = m_A^2 \tan \beta/(1 + \tan^2 \beta)$. Interpretations are made as a function of $\tan \beta$ and $\cos(\beta - \alpha)$. The parameter $\tan \beta$ is the ratio of vacuum expectation values of the two Higgs doublets and $\alpha$ is the mixing angle between the two neutral CP-even scalars.

2 The ATLAS detector

ATLAS is a multi-purpose particle physics experiment [24] at the LHC. The detector\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 of the LHC ring, and the y-axis points upwards. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ is the azimuthal angle around the beam pipe. The pseudorapidity, $\eta$, is defined in terms of the polar angle $\theta$ as $\eta = -\ln[\tan(\theta/2)]$.} consists of inner tracking devices surrounded by a superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner tracking system provides charged-particle tracking in the pseudorapidity region $|\eta| < 2.5$ and vertex reconstruction. It consists of a silicon pixel detector, a silicon microstrip tracker, and a straw-tube transition radiation tracker. The system is surrounded by a solenoid that produces a 2T axial magnetic field. The central calorimeter system consists of a liquid-argon electromagnetic sampling calorimeter with high granularity covering $|\eta| < 3.2$ and a steel/scintillator-tile calorimeter providing hadronic energy measurements in the central pseudorapidity range $(0 < |\eta| < 1.7)$. The endcap and forward regions are instrumented with liquid-argon calorimeters for both electromagnetic and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer is operated in a magnetic field provided by air-core superconducting toroids and includes tracking chambers for precise muon momentum measurements up to $|\eta| = 2.7$ and trigger chambers covering the range $|\eta| < 2.4$. A three-level trigger system is used to select interesting events [25]. The Level-1 trigger reduces the event rate to below 75 kHz using hardware-based trigger algorithms acting on a subset of detector information. Two software-based trigger levels, referred to collectively as the High-Level Trigger (HLT), further reduce the event rate to about 400 Hz using information from the entire detector.

3 Data and simulation samples

The data sample used in this analysis, after applying data quality requirements that include the availability of $b$-jet triggers, corresponds to an integrated luminosity of 19.5 ± 0.5 fb$^{-1}$. The uncertainty in the integrated luminosity (2.8 %) is derived following the same methodology as that detailed in Ref. [26]. The data sample is selected by a combination of five triggers requiring multiple jets or $b$-jets, where $b$-jets are identified by a dedicated HLT $b$-tagging algorithm. This combination of triggers is >99.5 % efficient for signal events passing the offline selection, across the full mass range considered.

Simulated Monte Carlo (MC) event samples are used to model the different signals, as well as the small back-
Table 1 Computed production cross-sections times branching ratio $\sigma (pp \to X \to hh \to b\bar{b}b\bar{b})$ and total widths for selected resonance pole mass values. The bulk RS model predictions are obtained with $k/M_{Pl} = 1.0$; both cross-section and width grow as the square of $k/M_{Pl}$. The 2HDM predictions are for a Type-II model with $\cos (\beta - \alpha) = -0.2$ and $\tan \beta = 1$.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mass [GeV]</th>
<th>$\sigma \times B$ [fb]</th>
<th>$\Gamma$ [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulk RS</td>
<td>1000</td>
<td>1.47</td>
<td>55</td>
</tr>
<tr>
<td>Bulk RS</td>
<td>1500</td>
<td>0.085</td>
<td>90</td>
</tr>
<tr>
<td>2HDM</td>
<td>1000</td>
<td>5.54</td>
<td>130</td>
</tr>
<tr>
<td>2HDM</td>
<td>1500</td>
<td>0.330</td>
<td>332</td>
</tr>
</tbody>
</table>

Non-resonant SM $pp \to hh \to b\bar{b}b\bar{b}$ events are generated using the exact form factors for the top loop, taken from HPAIR [5,42]. The cross-section is defined as the inclusive cross-section. The gluon-fusion production cross-section at NNLO in QCD from Ref. [7] is used, summed with the NLO predictions for vector-boson-fusion, top-pair-associated and vector-boson-associated production from Ref. [43]. The resulting cross-section is $\sigma (pp \to hh \to b\bar{b}b\bar{b}) = 3.6 \pm 0.5$ fb, where the uncertainty term includes the effects of uncertainties in the renormalization and factorization scale, PDFs, $\alpha_S$ and $Br(H \to bb)$.

The $t\bar{t}$ background sample is generated using POWHEG-BOX v1.0 [44–47] interfaced to PYTHIA v6.426 [48], with the top mass fixed to 172.5 GeV and the CT10 [49] NLO PDF set. The NNLO+NNLL prediction of 253 pb for the $t\bar{t}$ cross-section [50–55] is used for normalization. Single-top background is negligible.

The $Z$+jets sample is generated using SHERPA v1.4.3 [56] with the CT10 PDF set and the Z boson decaying to two $b$-quarks. The $Z$+jets cross-section is taken from an NLO POWHEG-BOX v1.0 [57] plus PYTHIA v8.165 prediction, which is found to agree well with measurements in the boosted regime [58].

The generated MC events are processed with the GEANT4-based [59] ATLAS detector simulation [60]. Effects of multiple proton–proton interactions (pile-up) are simulated using PYTHIA v8.1 with the CTEQ6L1 PDF set and the AU2 tune [61]. The simulated events are weighted so that the distribution of the average number of interactions per bunch-crossing matches that in the data. The same reconstruction software is used to process both the data and the simulated samples.

Table 2 summarizes the various event generators and PDF sets, as well as parton shower and hadronization software used for the analyses presented in this paper.

4 Resolved analysis

4.1 Event reconstruction

Jets are reconstructed from topological clusters of calorimeter cell energy deposits at the electromagnetic scale [63] using the anti-$k_t$ jet clustering algorithm, with a radius parameter of $R = 0.4$. The effects of pile-up on jet energies are accounted for by a jet-area-based correction [64]. The jets are then calibrated using $p_T$- and $\eta$-dependent calibration factors based on MC simulations and the combination of several in situ techniques applied to data [65]. Following this, the jets undergo Global Sequential calibration [63] which reduces flavour-dependent differences in calorimeter response. If a muon with $p_T > 4$ GeV and $|\eta| < 2.5$ is found within a cone of $\Delta R = 0.4$ around the jet axis, the four-momentum
of the muon is added to that of the jet (after correcting for the expected energy deposited by the muon in the calorimeter). Such muons are reconstructed by combining measurements from the inner tracking and muon spectrometer systems, and are required to satisfy tight muon identification quality criteria [66]. Jets with a significant energy contribution from pile-up interactions [67] are removed using tracking information. For jets with $p_T < 50$ GeV and $|\eta| < 2.4$, the $p_T$ sum of tracks matched to the jet is calculated and it is required that at least 50% of this $p_T$ sum is due to tracks originating from the primary vertex.2

Jets with $|\eta| < 2.5$ are $b$-tagged using the properties of the tracks associated with them, the most important being the impact parameter (defined as the track’s distance of closest approach to the primary vertex in the transverse plane) of each track, as well as the presence and properties of displaced vertices. The MV1 $b$-tagging algorithm [21] used in this analysis combines the above information using a neural network and is configured to achieve an efficiency of 70% for tagging $b$-jets,3 with a charm-jet rejection of approximately 5 and a light-quark or gluon jet rejection of around 140, as determined in an MC sample of $t\bar{t}$ events. The $b$-tagging efficiency in the simulation is scaled to reproduce the one measured in data [68].

4.2 Selection

The combined acceptance times efficiency ($A \times \epsilon$) at different stages of the event selection is shown in Fig. 2 as a function of resonance mass for the resonant signal models. The event selection begins with the requirement of at least four $b$-tagged jets, each with $p_T > 40$ GeV (shown in Fig. 2 as “4 $b$-tagged jets”). SM non-resonant Higgs boson pair production has a softer Higgs boson $p_T$ spectrum than the $m_X = 500$ GeV resonances, resulting in a lower $A \times \epsilon = 4.9 \%$ for this requirement. The four highest-$p_T$ $b$-tagged jets are then used to form two dijet systems, demanding that the angular distance, $\Delta R$, between the jets in each of the dijets is smaller than 1.5. The transverse momentum of the leading (in $p_T$) dijet system, $p_T^{\text{lead}}$, is required to be greater than 200 GeV, while the subleading dijet system must have $p_T^{\text{subl}} > 150$ GeV. In the rare case that a jet could be used to create more than one dijet which satisfies the above kinematic requirements, the dijet with the highest mass is chosen. Thus two unique dijet systems, with no jets in common, are selected (shown as “2 dijets” in Fig. 2). For SM non-resonant Higgs boson pair production, after this requirement, $A \times \epsilon = 1.2 \%$. The impact of different decay kinematics can be seen by comparing Fig. 2b to Fig. 2a: the decay of spin-0 $H$ bosons gives a softer Higgs boson $p_T$ spectrum than in the case of the spin-2 $G_{KK}^*$ decay (due to the differing angular distributions of $hh$), resulting in lower acceptance for these kinematic requirements at low resonance mass.

The resolved analysis considers a large range of resonance masses, $500 \leq m_X \leq 1500$ GeV. Due to the differing kinematics, the optimal selection for low-mass resonances differs from the optimum for higher masses. To increase the analysis sensitivity, three requirements which vary with the reconstructed resonance mass are used. These selection requirements are optimized simultaneously, by performing a three-dimensional scan of threshold values, using the statistical-only exclusion limit (Sect. 6) as the objective function. There are mass-dependent requirements (shown in Fig. 2 as “MDC”) on the minimum $p_T$ of the leading and sub-leading dijets as well as on the maximum difference in pseudorapidity, $|\Delta \eta_{\text{dijets}}|$, between them. These requirements are written in terms of the four-jet mass $m_{4j}$ expressed in GeV:

2 Proton–proton collision vertices are reconstructed requiring that at least five tracks with $p_T > 0.4$ GeV are associated with a given vertex. The primary vertex is defined as the vertex with the highest summed track $p_T^2$.

3 A jet is labelled as a $b$-jet if a $b$-quark with transverse momentum above 5 GeV exists within a cone of $\Delta R = 0.3$ around the jet axis.

---

**Table 2** Summary of MC event generators, PDF sets, parton shower and hadronization used in the analysis for both signal and background processes. ¹ Madgraph was modified [62] to use the exact top-loop form-factors from HPAIR [5,42] for the gluon-fusion production process.

<table>
<thead>
<tr>
<th>Model/Process</th>
<th>Generator</th>
<th>PDF</th>
<th>Parton Shower/Hadronization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulk RS: $pp \rightarrow G_{KK}^* \rightarrow hh \rightarrow b\bar{b}b\bar{b}$</td>
<td>Madgraph v1.5.1 [27,28]</td>
<td>CTEQ6L1 [30]</td>
<td>Pythia v8.175 [29]</td>
</tr>
<tr>
<td>2HDM: $pp \rightarrow H \rightarrow hh \rightarrow b\bar{b}b\bar{b}$</td>
<td>Madgraph v1.5.1 [27]</td>
<td>CTEQ6L1 [30]</td>
<td>Pythia v8.175 [29]</td>
</tr>
<tr>
<td>SM: $pp \rightarrow hh \rightarrow b\bar{b}b\bar{b}$</td>
<td>Madgraph v1.5.1 [27,62]†</td>
<td>CTEQ6L1 [30]</td>
<td>Pythia v8.175 [29]</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>POWHEG v1.0 [44,45]</td>
<td>CT10 [49]</td>
<td>Pythia v6.426 [48]</td>
</tr>
<tr>
<td>$Z+\text{jets}$</td>
<td>SHERPA v1.4.3 [56]</td>
<td>CT10 [49]</td>
<td>SHERPA v1.4.3 [56]</td>
</tr>
</tbody>
</table>
The decay of the $W$ ("extra jets") in the event are used to reconstruct $\bar{t}t$ background. This $\bar{t}t$ background predominantly comprises events where both top quarks decayed hadronically. These hadronic decays lead to three jets for each top quark: one $b$-jet directly from the top decay and two from the decay of the $W$ boson. Since the probability to mis-tag charm-jets is much higher than the probability to mis-tag light-jets, in the majority of cases the dijet is formed from the $b$-jet and a charm-jet from the decay of the $W$ boson. In order to reduce the $\bar{t}t$ background, jets not already used in the formation of the two dijets and top quark candidates by combining them with each of the dijets. These extra jets are required to have $p_T > 30$ GeV, $|\eta| < 2.5$, and $\Delta R < 1.5$ relative to the dijet. The $W$ boson candidates are reconstructed by adding the four-momentum of each of the possible extra jets to the four-momentum of the jet in the dijet system with the lowest probability of being a $b$-jet according to the multivariate $b$-tagging algorithm. Top quark candidates are then reconstructed by summing the dijets with each of the extra jets. The compatibility with the top quark decay hypothesis is then determined using the variable:

$$X_{\bar{t}t} = \sqrt{\frac{(m_W - \bar{m}_W)^2}{\sigma_{m_W}} + \frac{(m_t - \bar{m}_t)^2}{\sigma_{m_t}}}$$

where $m_W$ and $m_t$ are the invariant masses of the $W$ boson and top quark candidates, $\sigma_{m_W} = 0.1 m_W$, $\sigma_{m_t} = 0.1 m_t$, $\bar{m}_W = 80.4$ GeV and $\bar{m}_t = 172.5$ GeV. The values of $\sigma_{m_W}$ and $\sigma_{m_t}$ reflect the dijet and three-jet system mass resolutions. If either dijet in an event has $X_{\bar{t}t} < 3.2$ for any possible combination with an extra jet, the event is rejected. This requirement reduces the $\bar{t}t$ background by $\sim 60 \%$, whilst retaining $\sim 90 \%$ of signal events (shown as "$\bar{t}t$ Veto" in Fig. 2).

The event selection criteria described above are collectively referred to as the "4-tag" selection requirements. These requirements select 1891 events.

Following the 4-tag selection, a requirement on the leading and subleading dijet masses ($m_{2j}^{\text{lead}}$ and $m_{2j}^{\text{subl}}$, respectively) is used to define the signal region. The central value of this region corresponds to the median values of the narrowest dijet mass intervals that contain 90 $\%$ of the MC signal (these were found to be stable with resonance mass). The definition of the signal region is

$$X_{hh} = \sqrt{\frac{(m_{2j}^{\text{lead}} - 124 \text{ GeV})^2}{0.1 m_{2j}^{\text{lead}}}} + \frac{(m_{2j}^{\text{subl}} - 115 \text{ GeV})^2}{0.1 m_{2j}^{\text{subl}}},$$

where the 0.1 $m_{2j}$ terms represent the widths of the leading and subleading dijet mass distributions. The signal region is defined as $X_{hh} < 1.6$. This corresponds to the kinematical requirements illustrated by the inner region in Fig. 3, albeit with data from the 2-tag sample shown. It is optimized to maximize the expected sensitivity of the search. The acceptance times efficiency of the full selection, including this signal region requirement, is shown in Fig. 2 as "Signal Region". For SM non-resonant Higgs boson pair production, the full selection has an $A \times \epsilon = 0.60 \%$.

The final step of the Higgs boson pair resonant production search is to perform a fit to the four-jet mass $m_{4j}$ in the signal region. The sensitivity of this fit is increased by improving
the $m_{4j}$ resolution in this region, using the constraint that the two dijet masses should equal the Higgs boson mass, i.e. $m_{2j}^{\text{lead}} = m_{2j}^{\text{subl}} = m_h$. To this end, each dijet four-momentum is multiplied by a correction factor $\alpha_{\text{dijet}} = m_h/m_{\text{dijet}}$. This leads to an improvement of $\sim 30\%$ in the signal $m_{4j}$ resolution – with a significant reduction of the low-mass tails caused by energy loss – with little impact on the background.

4.3 Background estimation

After the 4-tag selection described above, about 95% of the remaining background in the signal region is expected to originate from multijet events, which are modelled using data. The remaining $\sim 5\%$ of the background is $t\bar{t}$ events. The $t\bar{t}$ yield is determined from data, while the $m_{4j}$ shape is taken from MC simulation. The $Z+\text{jets}$ contribution is $< 1\%$ of the total background and is modelled using MC simulation. The background from all other sources – including processes featuring Higgs bosons – is negligible.

4.3.1 Multijet background

The multijet background is modelled using an independent data sample selected by the same trigger and selection requirements as described in Sect. 4.2, except for the $b$-tagging requirement: only one of the two selected dijets has to be formed from $b$-tagged jets, while the other dijet can be formed from jets that are not $b$-tagged. This “2-tag” selection yields a data sample comprising 485377 events, 98% of which are multijet events and the remaining 2% are $t\bar{t}$. The predicted contamination by the signal is negligible.

This 2-tag sample is normalized to the 4-tag sample and its kinematical distributions are corrected for differences introduced by the additional $b$-tagging. These differences arise because the $b$-tagging efficiency as well as the charm- and light-jet rejection vary as a function of jet $p_T$ and $\eta$, the various multijet processes contribute in different fractions, and the fraction of events passed by each trigger path changes. The normalization and kinematic corrections are determined using a signal-free sideband region of the $m_{2j}^{\text{subl}}$ plane, in dedicated samples collected without mass-dependent requirements, which increases the statistical precision of the kinematic corrections. The resulting background model is verified and the associated uncertainties are estimated using a control region. The sideband and control regions are shown in Fig. 3. The sideband region is defined as: $\sqrt{(m_{2j}^{\text{lead}} - 124 \text{ GeV})^2 + (m_{2j}^{\text{subl}} - 115 \text{ GeV})^2} > 58$ GeV, while the control region is defined as the region between the signal and sideband regions. These definitions are chosen to be orthogonal to the signal region and to give approximately equal event yields in the sideband and control regions.

The normalization of the multijet background prediction is set by scaling the number of events in each region of the 2-tag sample by the following factor, $\mu_{\text{QCD}}$, calculated in the sideband region:

$$\mu_{\text{QCD}} = \frac{N_{4\text{-tag}}^{\text{QCD}}}{N_{2\text{-tag}}^{\text{QCD}}} = \frac{N_{4\text{-tag}}^{\text{data}} - N_{t\bar{t}}^{4\text{-tag}} - N_{Z}^{4\text{-tag}}}{N_{2\text{-tag}}^{\text{data}} - N_{t\bar{t}}^{2\text{-tag}} - N_{Z}^{2\text{-tag}}},$$

where $N_{4\text{-tag}}^{2\text{-tag}}$ is the number of events observed in the sideband region in the 2- or 4-tag data sample, respectively. The yields $N_{t\bar{t}}^{2\text{-tag}}$ and $N_{Z}^{2\text{-tag}}$ are the estimated number of $t\bar{t}$ and $Z+\text{jets}$ events in the 2-/4-tag selected sideband region. The $t\bar{t}$ normalization is estimated from data, as described in Sect. 4.3.2, while the $Z+\text{jets}$ is estimated using MC simulation.

To predict the distributions of the multijet background in each region, the predicted $t\bar{t}$ and $Z+\text{jets}$ 2-tag distributions are first subtracted from the 2-tag data sample distribution before the distribution is scaled by $\mu_{\text{QCD}}$.

The correction for the kinematic differences between 2-tag and 4-tag samples is performed by reweighting events in the 2-tag sample. The weights are derived in the sideband region from linear fits to the ratio of the total background model to data for three kinematic distributions which are found to have the largest disagreement between 2-tag and 4-tag events: the leading dijet $p_T$, the $\Delta R$ separation between the jets in the subleading dijet, and the $\Delta R$ separation between the two dijets. The reweighting is done using one-dimensional distri-
butions, but is iterated so that correlations between the three variables are approximately accounted for. Three iterations are found to be sufficient. After the correction process, there is good agreement between the background model and sideband region data in kinematic variables that were not explicitly corrected. Systematic uncertainties in the normalization and shape of the multijet background model in the signal region are assessed using control-region data, as described in Sect. 4.4.

4.3.2 $t\bar{t}$ background

The $t\bar{t}$ background is described using a hybrid model: the normalization is derived from data in a $t\bar{t}$ control sample, while the shape is taken from MC simulation because there are too few events in the $t\bar{t}$ control sample to describe the shape precisely enough.

The $t\bar{t}$ control sample is formed from events which pass the 4-tag selection, except for the top veto, which is reversed: if either of the dijets fails the top veto, the event enters the $t\bar{t}$ control sample. This selection leads to a sample of 41 events within the signal region of the $t\bar{t}$ control sample, of which $\sim 50\%$ are $t\bar{t}$ and $\sim 50\%$ multijet. The multijet background component is estimated using the same methods as used for the nominal selection, but with a wider control region in order to reduce the sideband region $t\bar{t}$ fraction. After subtracting the multijet background, the $t\bar{t}$ control sample yield is then extrapolated to predict the $t\bar{t}$ yield in the nominal signal region, $N_{t\bar{t}}$, using the following equation:

$$N_{t\bar{t}} = \frac{\epsilon_t}{1 - \epsilon_t} \times N_{t\bar{t}}^{\text{CS}},$$  \hspace{1cm} (3)

where $N_{t\bar{t}}^{\text{CS}}$ is the number of events in the signal region, after subtraction of the multijet background, within the $t\bar{t}$ control sample, and $\epsilon_t$ is the efficiency for a selected dijet in a $t\bar{t}$ event to pass the top veto. This equation relies on the assumption that the $\epsilon_t$ of each dijet in the event is uncorrelated, an assumption validated in MC simulation. The $\epsilon_t$ is measured using an independent “semileptonic $t\bar{t}$” data sample that has a high $t\bar{t}$ purity. Events in this sample are selected by requiring one dijet candidate to pass the nominal selection with $p_T > 150$ GeV and one “leptonic top-quark” candidate. The leptonic top quark candidate is defined using a reconstructed muon and one $b$-tagged jet. This $b$-tagged jet is required to be distinct from jets in the dijet candidate, and the muon is required to have $p_T > 25$ GeV, be isolated, and fall within a cone of radius 1.2 around the $b$-tagged jet. The leptonic top quark candidate is required to have $p_T > 180$ GeV, where the leptonic top $p_T$ is defined as the vector sum of the $b$-jet $p_T$, the muon $p_T$, and the missing transverse momentum in the event. The latter is reconstructed from energy deposits in the calorimeter, including corrections for identified jets, electrons and muons. The $t\bar{t}$ veto efficiency is then measured as the fraction of the reconstructed dijet candidates which passed the $t\bar{t}$ veto, yielding $\epsilon_t = 0.53 \pm 0.03 \text{ (stat.)} \pm 0.05 \text{ (syst.)}$. The systematic uncertainty in $\epsilon_t$ is assigned to cover potential differences between $\epsilon_t$ as measured in the semileptonic $t\bar{t}$ sample and $\epsilon_t$ in the full 4-tag selection, where the method is applied in $t\bar{t}$ MC simulation to evaluate such differences. The measured $\epsilon_t$ agrees well with the corresponding semileptonic $t\bar{t}$ MC prediction of 0.54.

Equation (3) gives a data-driven $t\bar{t}$ background prediction of $5.2 \pm 2.6$ events in the signal region after the full selection. The uncertainty is dominated by the statistical uncertainty in $N_{t\bar{t}}^{\text{CS}}$, with a smaller contribution from the uncertainty in $\epsilon_t$.

Due to the limited number of events in the $t\bar{t}$ control sample, the $m_{4j}$ shape of the $t\bar{t}$ background is modelled using MC simulation. However, despite the use of a large $t\bar{t}$ sample, very few events pass the full 4-tag selection. Therefore, the $t\bar{t}$ shape is derived from MC simulation using the “2-tag” selection, with a systematic uncertainty assigned to cover differences between the 2-tag and 4-tag $m_{4j}$ distributions.

4.4 Systematic uncertainties

Two classes of systematic uncertainties are evaluated: those affecting the modelling of the signal and those affecting the background prediction.

The signal modelling uncertainties comprise: theoretical uncertainties in the acceptance, uncertainties in the jet energy scale (JES) and resolution (JER), and uncertainties in the $b$-tagging efficiency.

The theoretical uncertainties considered arise from initial- and final-state radiation modelling (ISR and FSR), PDF uncertainties and uncertainty in the LHC beam energy. These are evaluated using particle-level samples generated using the same generator configurations as the nominal signal samples but with appropriate variations, assessing the difference in yields after the full analysis selection. The ISR and FSR uncertainty is evaluated by varying the relevant parton shower parameters in Pythia 8. The PDF uncertainty is estimated by taking the maximum difference between the predictions when using MSTW2008nlo [69], NNPDF2.3 [70] and CTEQ6L1. The uncertainty due to the beam energy is determined by varying coherently the energy of each beam by $\pm 26.5$ GeV [71] in the simulation. Only FSR has a significant impact on the acceptance, leading to a $\pm 1.0\%$ theoretical modelling acceptance uncertainty.

The JES systematic uncertainty is evaluated using 15 separate and orthogonal uncertainty components, which allow for the correct treatment of correlations across the kinematic bins [65]. The JER uncertainty is evaluated by smearing jet energies according to the systematic uncertainties of the resolution measurement performed with data [65]. For $b$-jets with $p_T < 300$ GeV the uncertainty in the $b$-tagging efficiency is
Table 3 The number of events in data and predicted background events after applying the mass-dependent requirements in preselection and in the sideband and control regions for the resolved analysis. The uncertainties are purely statistical. The $t\bar{t}$ yield in this table, in contrast to the final result, is estimated using MC simulation.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sideband region</th>
<th>Control region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>907 ± 3</td>
<td>789 ± 3</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>25.5 ± 0.3</td>
<td>57.5 ± 0.4</td>
</tr>
<tr>
<td>Z+jets</td>
<td>14 ± 1</td>
<td>20 ± 1</td>
</tr>
<tr>
<td>Total</td>
<td>947 ± 3</td>
<td>867 ± 3</td>
</tr>
<tr>
<td>Data</td>
<td>952</td>
<td>852</td>
</tr>
</tbody>
</table>

evaluated by propagating the systematic uncertainty in the measured tagging efficiency for $b$-jets [68], which ranges from 2% to 8% depending on $b$-jet $p_T$ and $\eta$. However, for the higher resonance masses considered in this analysis, there are a significant number of events containing at least one $b$-jet with $p_T > 300$ GeV. The systematic uncertainties in the tagging efficiencies of these jets are derived from MC simulation and are larger, reaching 24% for $p_T > 800$ GeV.

Systematic uncertainties in the normalization and shape of the multijet background model are assessed in the control region. Table 3 shows the estimated background yields in the control and sideband regions. The control region background prediction agrees with the observed data within the data statistical uncertainty of ±3.5%. To further test the robustness of the background estimation and the assumptions behind it, predictions are made with different sideband and control region definitions and different $b$-tagging requirements on the 2-tag sample. Redefinitions of the sideband and control region changed the kinematic composition of these regions, enhancing the sideband region in either high mass or low mass dijets and therefore altering the kinematic corrections that are applied. These variations induce a maximum change of ±6% in the estimated multijet yield and so the uncertainty is set to this value. Different $b$-tagging requirements on the $b$-tagged dijet in the 2-tag sample are used in order to change the composition of the sample and to vary the degree of $b$-tagging-related kinematic bias. No additional uncertainty is required.

The uncertainty in the description of the multijet $m_{4j}$ distribution is determined by comparing the total background prediction to data in the control region, as shown in Fig. 4. Good agreement in the shape is observed and a straight line fit to the ratio of the distributions gives a slope consistent with zero. This fit, along with its uncertainties, is shown in the bottom panel of Fig. 4. The uncertainty in the multijet background shape is defined using the uncertainty in the fitted slope.

The uncertainty in the $t\bar{t}$ normalization is dominated by the statistical uncertainty of the yield in the $t\bar{t}$ control sample, with a subdominant contribution from the uncertainties in the top veto efficiency, $\epsilon_t$, leading to a total uncertainty of ±50%. The uncertainty in the MC-derived $t\bar{t}$ $m_{4j}$ distribution is dominated by the uncertainty associated with using the shape after the 2-tag selection, rather than the 4-tag selection. This uncertainty is assessed by comparing the 2-tag to 4-tag MC predictions in the signal region. A straight line fit to the ratio of the normalized distributions is made and used to define a shape uncertainty in the same way as the multijet background. Due to the large statistical uncertainties of the 4-tag $t\bar{t}$ sample, the assigned shape uncertainty is large: ~30% and ~100% in the event yield at $m_{4j} = 400$ GeV and 1500 GeV, respectively.

Table 4 shows the relative impact of the uncertainties in the event yields. Figure 5 shows the relative impact on the expected limit for $\sigma$ ($pp \rightarrow t\bar{t} \rightarrow b\bar{b}KK \rightarrow hh \rightarrow b\bar{b}b\bar{b}$). The calculation of the expected limit is described in Sect. 6. It can be seen that for resonance masses below 700 GeV, the effect on the limit is dominated by the multijet description, with a small contribution from the $t\bar{t}$ background since both backgrounds are predominately at low mass. Above $m_X = 700$ GeV, the uncertainty associated with the modelling of the $b$-tagging efficiency has the largest impact, since the larger high-$p_T$ uncertainties have an increasingly important effect with mass.

4.5 Results of the resolved analysis

Table 5 shows the predicted number of background events in the signal region, the number of events observed in the data, and the predicted yield for two potential signals. The
numbers of predicted background events and observed events are in excellent agreement. Figure 5 shows a comparison of the predicted $m_{4j}$ background distribution to that observed in the data. The predicted background agrees with the observed distributions, with no significant deviation.

5 Boosted analysis

5.1 Event reconstruction

The boosted analysis differs from the resolved analysis primarily by the use of large-radius jets designed to contain the decay products of a single $h \rightarrow b\bar{b}$ decay. Those large-radius jets, denoted by the subscript J in the remainder of this paper, are reconstructed from locally calibrated topological clusters of calorimeter cells [63] using the anti–$k_t$ jet clustering algorithm with a radius parameter of $R = 1.0$. To minimize the impact of energy depositions due to pile-up and the underlying event, the jets are trimmed [22]. This trimming algorithm reconstructs subjets within the large-$R$ jet using the $k_t$ algorithm with radius parameter $R_{sub} = 0.3$, then removes any subjet with $p_T$ less than 5% of the large-$R$ jet $p_T$. Further calibration of both the energy and mass scales is applied as a function of $p_T$ and $\eta$ as determined from simulation and in situ measurements [65].

A novel aspect of the boosted technique presented here is the use of track-jets [23] to identify the presence of $b$-quarks inside the large-$R$ jet. Such track-jets are built solely from
tracks with $p_T > 0.5$ GeV and $|\eta| < 2.5$, satisfying a set of hit and impact parameter criteria to make sure that those tracks are consistent with originating from the primary vertex, thereby reducing the effects of pile-up. Track jets are reconstructed using the anti-$k_T$ algorithm with $R = 0.3$. Flavour-tagging of those track-jets proceeds in the same way as for the $R = 0.4$ calorimeter jets used in the resolved analysis described in the previous section, except for a slightly looser requirement on the output of the MV1 neural network for a track-jet to be $b$-tagged. This leads to $b$-jets being $b$-tagged with an efficiency of 74%, with a charm-jet rejection factor of approximately 4 and a light-quark or gluon jet rejection factor of around 65, as determined in an MC sample of $t\bar{t}$ events. The $b$-tagging efficiency for track-jets in the MC simulation is adjusted based on studies of $t\bar{t}$ events in the data (Sect. 5.4).

5.2 Selection

The combined acceptance times efficiency at different stages of the event selection for the boosted analysis is shown in Fig. 7.

Events are required to contain at least two large-$R$ jets with $p_T > 250$ GeV and $|\eta| < 2.0$. To suppress contamination from $t\bar{t}$ events, the leading jet is additionally required to have $p_T > 350$ GeV. This ensures that the top-quark decay products are typically fully contained in a single large-$R$ jet with mass close to that of the top quark. These requirements are shown in Fig. 7 as “2 large-$R$ jets”. Only the leading and subleading large-$R$ jets are retained for further consideration.

Track jets are associated with large-$R$ jets using “ghost association” [64,72,73]. Each of the leading and subleading large-$R$ jets must have at least two track-jets ghost-associated with their respective untrimmed parents, where the track-jets must have $p_T > 20$ GeV and $|\eta| < 2.5$, as well as be consistent with originating from the primary vertex of the event (shown in Fig. 7 as “4 track-jets”). The drop in the $\Lambda \times \epsilon$ value at masses above 1500 GeV is due to the decrease in the angular separation between the two track-jets from the $h \rightarrow b\bar{b}$ decay to below $\Delta R = 0.3$.

To suppress contamination from multijet events, the two selected large-$R$ jets in the event are required to have a separation $|\Delta \eta| < 1.7$. This requirement (shown in Fig. 7 as “$\Delta \eta$”) has only a small impact on the signal acceptance since high-mass resonances tend to produce jets that are more central than those from multijet background processes.

Selection of $h \rightarrow b\bar{b}$ candidates proceeds by requiring that both the leading and subleading track-jets associated with each of the two large-$R$ jets satisfy the $b$-tagging selection (shown in Fig. 7 as “4 $b$-tagged jets”).

A final correction to the large-$R$ jet four-momentum is applied to account for semileptonic $b$-hadron decays. If a muon passing the requirements outlined in Sect. 4.1 is ghost-associated with any of the selected $b$-tagged track-jets, its four-momentum is added to that of the large-$R$ jet. If more than one muon is associated with a given track-jet, the muon closest to the track-jet axis is used. This correction improves the mass resolution for large-$R$ jets in signal MC simulation, especially for the subleading jet.

The last requirement used to select signal event candidates is to require that the large-$R$ jet mass is consistent with the Higgs boson mass. This requirement is defined identically to that for the resolved analysis in Eq. (1), except for the replacement of the small-$R$ dijet mass with the large-$R$ jet mass. The signal region is defined by the requirement $X_{hh} < 1.6$. This final selection is shown in Fig. 7 as “Signal Region”.

5.3 Background estimation

After the 4-tag selection described in Sect. 5.2, the background composition is similar to that of the resolved anal-
ysis. Multijet events comprise approximately 90% of the total background and are modelled entirely using data. The remaining ∼10% of the background is \( t\bar{t} \) events. The \( t\bar{t} \) yield is determined using data, while the \( m_{3J} \) shape is taken from MC simulation. The \( Z+jet \) contribution is <1% of the total background and is modelled using MC simulation. The background from all other sources – including processes featuring Higgs bosons – is negligible.

Estimation and validation of the background described below relies on two data samples defined as follows.

- The “4-tag” sample corresponds to the set of events that satisfy all the requirements detailed in Sect. 5.2, except that the final requirement on the mass of the leading and subleading large-\( R \) jets is not applied.
- The “2+3-tag” sample is identical to the 4-tag sample except for having only two or three of the four track-jets \( b \)-tagged. For events with only two \( b \)-tagged track-jets, both are required to be associated with the same large-\( R \) jet.

Both samples are further subdivided based on the large-\( R \) jet masses, with each subsample having a sideband region to determine the multijet background kinematics and a control region to validate the background estimate. The control region is defined by requirements on the mass of the leading and subleading large-\( R \) jets of \( 95 < m_{\text{lead}} < 160 \) GeV and \( 85 < m_{\text{subl}} < 155 \) GeV respectively, while excluding the signal region defined by \( X_{hh} < 1.6 \). The sideband region is complementary to the signal and control regions. Figure 8 illustrates the sideband and control regions with data from the 2 + 3-tag sample.

The choice of control region (and consequently sideband region) ensures that the multijet background can be estimated by extrapolation of event yields and kinematic properties from the 2 + 3-tag sample to the 4-tag sample with a normalization given by the relative event yields in the sideband region. Furthermore, the control region is chosen such that event kinematics in that region are representative of the kinematics in the signal region.

The estimated background yield in the 4-tag sample, \( N_{bkg}^{4\text{-tag}} \), is computed according to

\[
N_{bkg}^{4\text{-tag}} = \mu_{QCD} N_{QCD}^{2+3\text{-tag}} + \alpha_{t\bar{t}} N_{t\bar{t}}^{4\text{-tag}} + N_{Z}^{4\text{-tag}},
\]

where \( N_{QCD}^{2+3\text{-tag}} \) is the number of multijet events in the 2 + 3-tag data sample, \( N_{t\bar{t}}^{4\text{-tag}} \) and \( N_{Z}^{4\text{-tag}} \) are the numbers of events in the 4-tag \( t\bar{t} \) and \( Z+jets \) MC samples. The parameter \( \mu_{QCD} \) corresponds to the ratio of multijet event yields in the 4-tag and 2 + 3-tag data samples, as defined in Eq. (2), except for including both 2- and 3-tag events in the denominator. Finally, the parameter \( \alpha_{t\bar{t}} \) is a scale factor designed to adjust the \( t\bar{t} \) event yield from the MC simulation. Both \( \mu_{QCD} \) and \( \alpha_{t\bar{t}} \) are extracted from a binned likelihood fit to the leading large-\( R \) jet mass distribution obtained in the sideband region of the 4-tag data sample, as depicted in Fig. 9. Due to the large minimum \( p_T \) requirement for the leading large-\( R \) jet, much of the \( t\bar{t} \) contribution is concentrated at high mass close to the top-quark mass. In this fit, the multijet distribution is extracted from the 2+3-tag data sample, after subtraction of the \( t\bar{t} \) and \( Z+jets \) contributions predicted by the MC simulation. The \( t\bar{t} \) and \( Z+jets \) distributions in the sideband region of the 4-tag data sample are taken from the MC simulation, but the \( Z+jets \) contribution is very small and its distribution is added to the multijet distribution for the fit. The resulting fit values are \( \mu_{QCD} = 0.0071 \pm 0.0007 \) and \( \alpha_{t\bar{t}} = 1.44 \pm 0.50 \) with a correlation coefficient of \(-0.67\) between these two parameters.

Figure 10a shows the dijet mass distribution for the 4-tag data sample in the sideband region with the background estimated using the above method. This figure indicates that the 2 + 3-tag sample provides a valid description of the background kinematics in the 4-tag sample. The modelling of the background yield and kinematics is further validated by testing in the control region of the 4-tag data sample. Good agreement is observed between the data and the predicted background in various kinematical distributions for leading and subleading large-\( R \) jets, as well as for the dijet mass, as shown in Fig. 10b. The shapes of the \( t\bar{t} \) kinematical distributions in the signal region are determined from the MC simulation requiring only three \( b \)-tagged track-jets instead of four due to the limited MC sample size. The number of \( t\bar{t} \) events is then normalized to the expected yield in the 4-tag sample.
Fig. 9 Leading large-\( R \) jet mass distribution for 4-tag events in the sideband region for data (points) and the two dominant sources of background for the boosted analysis. The normalization for each of those two background components is obtained with a fit to the data as described in the text.

Fig. 10 Dijet mass distributions for 4-tag events in the boosted analysis. a shows the sideband region and b the control region for data (points) and the expected background (histograms). The filled blocks represent the statistical uncertainty in the total background estimate.
The shape of the $t\bar{t}$ background used in the fit shown in Fig. 9 is varied by extracting the shape from MC samples with zero, one, two or three $b$-tagged track-jets. Similarly, the uncertainty in the shape of other $t\bar{t}$ kinematical distributions is extracted from those samples. The uncertainty in the shape of the multijet background extracted from the sideband region of the $2+3$-tag sample is constrained by the level of agreement between the background prediction and the observed data in the control region following the procedure described in Sect. 4.4. Good agreement is observed between the data and the predicted background in both the sideband and control regions of the $4$-tag sample as shown in Table 6.

Systematic uncertainties in both the background and signal event yields are summarized in Table 7. A 2.8 % luminosity uncertainty is applied to the $Z$+jets background and to the signal samples. The JER/JES/JMR/JMS uncertainties are applied to signal, $t\bar{t}$ and $Z$+jets samples. The track-jet $b$-tag uncertainty is applied only to the signal samples as the normalization and shape differences in the $t\bar{t}$ sample are accounted for through other sources of systematic uncertainty.

Theoretical uncertainties affecting the signal acceptance are also considered, as described in Sect. 4.4. These sources do not have significant dependence on the assumed resonance mass and the largest contribution is found to be due to the ISR modelling.

The uncertainty in the multijet event yield is derived from the difference between the predicted and observed multijet yields in the control region. This source of uncertainty is dominated by the statistical uncertainty in that region. The $t\bar{t}$ entry in Table 7 accounts for the shape uncertainty in the simulated $t\bar{t}$ leading-jet mass distribution in the sideband region used to fit for $\mu_{QCD}$ and $\alpha_{t\bar{t}}$. This uncertainty is determined by comparing the shape of the 4-tag and 2-tag $t\bar{t}$ distributions. Finally, the “Bkgd stat” accounts for the statistical uncertainties in the extraction of $\mu_{QCD}$ and $\alpha_{t\bar{t}}$. Uncertainties in the $m_{Z\bar{Z}}$ shape of the multijet and $t\bar{t}$ backgrounds are not listed in Table 7, as they do not affect the event yield, but are accounted for in the statistical analysis.

Figure 11 presents the impact of each source of systematic uncertainty on the expected cross-section limit for the production of $G_{KK}$ as a function of resonance mass with the choice $k/M_{Pl} = 1$. These values are obtained following the statistical analysis described below while neglecting each source of uncertainty in turn. The multijet background uncertainty dominates for resonance masses below $1000 \text{ GeV}$, with $b$-tagging, large-$R$ jet mass and the number of sideband data events for the background estimate becoming the most important at higher mass.
5.5 Results of the boosted analysis

A total of 34 events is observed in the data whereas the background expectation is estimated to be $25.7 \pm 4.2$, see Table 8 for a breakdown of the various sources of background. The significance of this excess of events in the data is evaluated below.

The dijet mass distribution in the signal region is shown in Fig. 12. For this distribution and the statistical analysis, the estimated background prediction from multijet ($t\bar{t}$) events is fit to an exponential function at masses above 900 (800) GeV and the associated uncertainty is propagated to the statistical analysis.

6 Results

The results from the analyses in Sects. 4 and 5 are interpreted separately using the statistical procedure described in Ref. [1].

Table 8  The number of predicted background events in the $hh$ signal region, compared to the data for the boosted analysis. Errors correspond to the total uncertainties in the predicted event yields. The yield for a 1000 GeV $G_{KK}^*$ in the bulk RS model with $k/\sqrt{s} = 1$ is also given.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Signal region yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>$23.5 \pm 4.1$</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>$2.2 \pm 0.9$</td>
</tr>
<tr>
<td>$Z + $jets</td>
<td>$0.14 \pm 0.06$</td>
</tr>
<tr>
<td>Total</td>
<td>$25.7 \pm 4.2$</td>
</tr>
<tr>
<td>Data</td>
<td>$34$</td>
</tr>
<tr>
<td>$G_{KK}^*(1000 \text{ GeV}), k/\sqrt{s} = 1$</td>
<td>$2.1 \pm 0.6$</td>
</tr>
</tbody>
</table>

Fig. 12  Dijet mass distributions for data (points) as well as expected background (solid histograms) in the signal region of the boosted analysis. The filled blocks represent the combined statistical and systematic uncertainty in the total background estimate. Two simulated $G_{KK}^*$ signal peaks predicted by the bulk RS model with $k/\sqrt{s} = 1$ are also shown as dashed lines.

and references therein. Hypothesized values of $\mu$, the global signal strength factor, are tested with a test statistic based on the profile likelihood ratio [74,75]. In the profile likelihoods, the maximum likelihood values are obtained with the systematic uncertainties treated as independent, Gaussian or log-normal constraint terms. The statistical analysis described below is performed using data from the signal region solely. In the case of the search for non-resonant $hh$ production, only the number of events passing the final selection is used whereas the $m_{4j}$ or $m_{2J}$ distributions are used in the case of the search for $hh$ resonances.

6.1 Background-only hypothesis tests

Tests of the background-only hypothesis ($\mu = 0$) are carried out to determine if there are any statistically significant local excesses in the data. The significance of an excess is quantified using the local $p_0$, the probability that the background could produce a fluctuation greater than or equal to the excess observed in data. A global $p_0$ is also calculated for the most significant discrepancy, using background-only pseudo-experiments to derive a correction for the look-elsewhere effect across the mass range tested.

In the case of the resolved analysis, the largest deviation from the background-only hypothesis is found to be $2.1 \sigma$ for a $pp \rightarrow H \rightarrow hh \rightarrow b\bar{b}b\bar{b}$ signal with fixed $\Gamma_H = 1$ GeV at $m_{4j} = 1200$ GeV. This corresponds to a global significance of $0.42 \sigma$. The significance of any deviation for a $G_{KK}^*$ signal with $k/\sqrt{s} = 1$ is very similar, albeit with slightly smaller local discrepancies as a result of the larger signal $m_{4j}$ width.

In the case of the boosted analysis, the largest local deviation corresponds to the data excess at $m_{2J} \sim 900$ GeV apparent in Fig. 12, with a local significance of $2.6 \sigma$ for $pp \rightarrow G_{KK}^* \rightarrow hh \rightarrow b\bar{b}b\bar{b}$ with $k/\sqrt{s} = 1$. The global significance of this deviation corresponds to $0.78 \sigma$.

Given these low significance values, the results of both analyses are consistent with the background-only hypothesis. Of the 117 events selected in the data by either the resolved or boosted analysis, only four events are common to both.

6.2 Exclusion limits

The data are used to set upper limits on the cross-sections for the different benchmark signal processes. Exclusion limits are based on the value of the statistic $CL_s$ [76], with a value of $\mu$ regarded as excluded at $95\%$ confidence level (CL) when $CL_s$ is less than $5\%$.

The non-resonant search is performed using the resolved analysis, since it has better sensitivity than the boosted analysis. Using the SM $hh$ non-resonant production as the signal model, the observed $95\%$ CL upper limit is $\sigma(pp \rightarrow hh \rightarrow b\bar{b}b\bar{b}) = 202$ fb. This can be compared to the inclusive SM
resolved analysis gives better sensitivity for each analysis for the expected and observed cross-section upper limits from complementary resonance mass regions. Figure 13 shows the 95% CL upper limits for $p\bar{p} \to \gamma_{KK}^\ast \to hh \to b\bar{b}b\bar{b}$ in the bulk RS model with $k/M_{Pl} = 1$ and $k/M_{Pl} = 2$, as well as $p\bar{p} \to H \to hh \to b\bar{b}b\bar{b}$ with fixed $\Gamma_H = 1$ GeV. The prediction (as defined in Sect. 3) of $\sigma(p\bar{p} \to hh \to b\bar{b}b\bar{b}) = 3.6 \pm 0.5$ fb.

For the resonant Higgs boson pair production search, the resolved and boosted analyses offer their best sensitivity in complementary resonance mass regions. Figure 13 shows the expected and observed cross-section upper limits from each analysis for $p\bar{p} \to \gamma_{KK}^\ast \to hh \to b\bar{b}b\bar{b}$ within the bulk RS model with $k/M_{Pl} = 1$. The resolved analysis can be seen to give a more stringent expected exclusion limit for resonance masses up to 1100 GeV, while the boosted analysis offers better sensitivity beyond that mass. This motivates a simple combination of the separate exclusion limits from the resolved and boosted analyses. For each of the signal models, the limit for each mass point is taken from the analysis which offers the most stringent expected exclusion.

Figure 14 shows the combined 95 % CL upper limits for three signal models: $p\bar{p} \to \gamma_{KK}^\ast \to hh \to b\bar{b}b\bar{b}$ within the bulk RS model with $k/M_{Pl} = 1$ and 2, and the $p\bar{p} \to H \to hh \to b\bar{b}b\bar{b}$ with a fixed $\Gamma_H = 1$ GeV. The most stringent limits of $\sigma(p\bar{p} \to X \to hh \to b\bar{b}b\bar{b}) \sim 3$ fb are set in the range $900 < m_X < 1600$ GeV, where there is little expected background and either the resolved or boosted analysis provides good signal acceptance. The excluded mass ranges for the bulk RS KK graviton are shown in Table 9.

The excluded mass range for the 2HDM is parameter dependent, principally because the production cross-section varies, but also because the exclusion limit depends on the parameter-dependent $H$ boson width, $\Gamma_H$. The theoretical cross-section used to determine the 95 % CL excluded regions is the sum of the cross-sections of gluon-fusion production, vector-boson-fusion production and $b$-associated production.

The effects of $\Gamma_H$ are accounted for by creating $m_H$ distributions with a range of widths, $0 < \Gamma_H/m_H \leq 0.5$, for each $m_H$ considered. These distributions are based on parameterizations which include resolution and acceptance effects combined with a Breit–Wigner line-shape. A grid of limits are calculated with each of these mass distributions. Then, for each point in $m_H$, $\cos(\beta - \alpha)$, and $\tan \beta$ space, the cross-section limit is determined by interpolating between the appropriate limits, based on the $\Gamma_H$ given by the model for that point. For the widest signals considered, the exclusion limits worsen by up to a factor of three. The exclusion regions determined through this process are shown as a function of $\cos(\beta - \alpha)$.
Fig. 15 Excluded regions of the \((\cos(\beta-\alpha), \tan \beta)\) parameter space for (a) the Type-I 2HDM signal model and (b) the Type-II 2HDM signal model. The grey areas demarcate the phase-space regions where \(\Gamma_H / m_H > 0.15\), for which the cross-section limits have not been demonstrated to be reliable.

Fig. 16 Excluded regions of the \((\cos(\beta-\alpha), \tan \beta)\) parameter space for (a) the Lepton-specific 2HDM signal model and (b) the Flipped 2HDM signal model. The grey areas demarcate the phase-space regions where \(\Gamma_H / m_H > 0.15\), for which the cross-section limits have not been demonstrated to be reliable.

and \(\tan \beta\) for \(m_H = 500\) GeV in Figs. 15 and 16, and as a function of \(m_H\) and \(\tan \beta\) for \(\cos(\beta-\alpha) = -0.2\) in Figs. 17 and 18. The validity of the process has been tested using the widest available signals, gravitons in the bulk RS model with \(k/M_{Pl} = 2\). Phase-space regions with \(\Gamma_H\) greater than these graviton widths are considered unvalidated and are shown in the figures as grey areas.

7 Conclusions

Two searches for Higgs boson pair production with the ATLAS detector at the LHC using the \(b\bar{b}b\bar{b}\) final state have been presented: one reconstructs Higgs boson candidates from pairs of nearby \(-k_t\) \(b\)-tagged jets with \(R = 0.4\); the other reconstructs Higgs boson candidates using trimmed anti-\(-k_t\) jets with \(R = 1.0\) matched to two \(b\)-tagged anti-\(-k_t\) track-jets with \(R = 0.3\). Thanks to the high expected \(h \to bb\) branching ratio and the large background rejection factors offered by the boosted dijet topology, the sensitivity for Higgs boson pair production is high, with a mass reach spanning the range between 500 and 2000 GeV. There is no evidence for any signal in 19.5 fb\(^{-1}\) of \(pp\) collision data with \(\sqrt{s} = 8\) TeV. The largest deviation from the background-only hypothesis has a global significance of only 0.78 \(\sigma\). The observed 95 % CL upper limit on \(\sigma(pp \to X \to hh \to b\bar{b}b\bar{b})\) is 3.2(2.3) fb for narrow resonances with a mass of 1.0(1.5) TeV.

Constraints are placed on several benchmark models. For the bulk RS model with \(k/M_{Pl} = 1\), KK gravitons in the
Fig. 17 Excluded regions of the \((m_H, \tan \beta)\) parameter space for a the Type-I 2HDM signal model and b the Type-II 2HDM signal model. The grey areas demarcate the phase-space regions where \(\Gamma_H/m_H\) is large (\(\Gamma_H/m_H > 0.15\) for \(m_H = 500\) GeV, increasing to \(\Gamma_H/m_H > 0.23\) for \(m_H = 1100\) GeV) and the limits have not been demonstrated to be reliable.

Fig. 18 Excluded regions of the \((m_H, \tan \beta)\) parameter space for a the Lepton-specific 2HDM signal model and b the Flipped 2HDM signal model. The grey areas demarcate the phase-space regions where \(\Gamma_H/m_H\) is large (\(\Gamma_H/m_H > 0.15\) for \(m_H = 500\) GeV, increasing to \(\Gamma_H/m_H > 0.23\) for \(m_H = 1100\) GeV) and the limits have not been demonstrated to be reliable.

mass range \(500 \leq m_{G^*} \leq 720\) GeV are excluded at the 95 % CL. For non-resonant signals, using Standard Model \(hh\) non-resonant production as the benchmark, the observed 95 % CL upper limit on \(\sigma(pp \to hh \to b\bar{b}b\bar{b})\) is 202 fb, in good agreement with the expected exclusion. This is to be compared to a SM prediction of \(3.6 \pm 0.5\) fb.

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; BMWFW and FWF, Austria; ANAS, Azerbaijan; STIC, 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 DSM, Germany; INFN, Italy; INFN, Italy; IN2P3-CNRS, CEA-DSM/IRFU, France; GSRT and NSRF, Greece; RGC, Hong Kong SAR, China; ISF, MISTRAL, GIF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRS, Morocco; FOM and NWO, Netherlands; BRF and RCN, Norway; MNiSW and NCN, Poland; GRICES and FCT, Portugal; MINEFF, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, UK; DOE and NSF, USA. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France),
KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (The Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

Open Access  This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Funded by SCOAP3.

References

16. CMS Collaboration, Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 8 TeV. arXiv:1503.04114
49. H.-L. Lai et al., New parton distributions for collider physics at the Tevatron: first genuine NNLO QCD corrections to $q\bar{q}\to t\bar{t}+X$. Phys. Rev. D 82, 074024 (2010), arXiv:1007.2241
50. M. Cacciari et al., Top-pair production at hadron colliders: the all-fermionic scattering channels. JHEP 04, 095 (2013), arXiv:1009.4568
International Center for Elementary Particle Physics and Department of Physics, The 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
Department of Physics, University of Toronto, Toronto, ON, Canada
Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba, Japan
Department of Physics and Astronomy, Tufts University, Medford, MA, USA
Centro de Investigaciones, Universidad Antonio Narino, Bogotá, Colombia
Department of Physics and Astronomy, University of California Irvine, Irvine, CA, USA
Department of Physics and Astronomy, University of Illinois, Urbana, IL, USA
Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
Instituto de Física Corpuscular (IFIC) and Departamento de Física Atómica, Molecular y Nuclear and Departamento de Ingeniería Electrónica and Instituto de Microelectrónica de Barcelona (IMB-CNMI), University of Valencia and CSIC, Valencia, Spain
Department of Physics, University of British Columbia, Vancouver, BC, Canada
Department of Physics and Astronomy, University of Victoria, Victoria, BC, Canada
Department of Physics, University of Warwick, Coventry, UK
Waseda University, Tokyo, Japan
Department of Particle Physics, The Weizmann Institute of Science, Rehovot, Israel
Department of Physics, University of Wisconsin, Madison, WI, USA
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität, Würzburg, Germany
Fachbereich C Physik, Bergische Universität Wuppertal, Wuppertal, Germany
Department of Physics, Yale University, New Haven, CT, USA
Yerevan Physics Institute, Yerevan, Armenia
Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France

a Also at Department of Physics, King’s College London, London, UK
b Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
c Also at Novosibirsk State University, Novosibirsk, Russia
d Also at TRIUMF, Vancouver, BC, Canada
e Also at Department of Physics, California State University, Fresno, CA, USA
f Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
g Also at Departamento de Física e Astronomia, Faculdade de Ciencias, Universidade do Porto, Porto, Portugal
h Also at Tomsk State University, Tomsk, Russia
i Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
j Also at Universita di Napoli Parthenope, Naples, Italy
k Also at Institute of Particle Physics (IPP), Victoria, Canada
l Also at Particle Physics Department, Rutherford Appleton Laboratory, Didcot, UK
m Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia
n Also at Louisiana Tech University, Ruston, LA, USA
o Also at Institucio Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain
p Also at Department of Physics, National Tsing Hua University, Hsinchu, Taiwan
q Also at Department of Physics, The University of Texas at Austin, Austin, TX, USA
f Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia
s Also at CERN, Geneva, Switzerland
t Also at Georgian Technical University (GTU), Tbilisi, Georgia
u Also at Ochadai Academic Production, Ochanomizu University, Tokyo, Japan
v Also at Manhattan College, New York NY, USA
w Also at Institute of Physics, Academia Sinica, Taipei, Taiwan
x Also at LAL, Université Paris-Sud and CNRS/IN2P3, Orsay, France
Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan

Also at School of Physics, Shandong University, Shandong, China

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

Also at Section de Physique, Université de Genève, Geneva, Switzerland

Also at International School for Advanced Studies (SISSA), Trieste, Italy

Also at Department of Physics and Astronomy, University of South Carolina, Columbia, SC, USA

Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China

Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia

Also at National Research Nuclear University MEPhI, Moscow, Russia

Also at Department of Physics, Stanford University, Stanford CA, USA

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

Also at Department of Physics, The University of Michigan, Ann Arbor, MI, USA

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

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

* Deceased