The following full text is a publisher's version.

For additional information about this publication click this link. http://hdl.handle.net/2066/161327

Please be advised that this information was generated on 2017-09-23 and may be subject to change.
I. INTRODUCTION

The discovery of a Higgs boson ($h$) \cite{1,2} at the Large Hadron Collider (LHC) provides an opportunity to search for physics beyond the Standard Model (SM) in channels involving couplings with the Higgs boson. In particular, the production cross section for Higgs-boson pairs in the SM is significantly smaller than predicted by a host of models, making this channel promising for a search for new phenomena. Examples of such models are the bulk Randall-Sundrum (RS) model \cite{3,4} with a warped extra dimension and the two-Higgs-doublet model (2HDM) \cite{5}. In the RS model, spin-2 Kaluza-Klein (KK) excitations of the graviton $G_{KK}$ are produced via gluon fusion with subsequent decay to the $hh$ final state. Similarly, a heavy spin-0 scalar, $H$, in the 2HDM also gives rise to a resonant $hh$ signature. Enhanced nonresonant $hh$ production is expected in models with light colored scalars \cite{6} or direct $t\bar{t}h$ vertices \cite{7,8}.

Previous searches for $hh$ production have been performed by the ATLAS and CMS collaborations with $pp$ collisions at $\sqrt{s} = 8$ TeV. The final states include $b\bar{b}bb$ \cite{9,10}, $b\bar{b}\tau^+\tau^-$ \cite{11,12}, $b\bar{b}\gamma\gamma$ \cite{13,14}, and $\gamma\gamma WW^*$ \cite{11}. A combination of these different channels has been performed by ATLAS \cite{11}, which shows that for resonance masses above 500 GeV the sensitivity is highest in the $b\bar{b}bb$ channel.

The dominant $h \rightarrow b\bar{b}$ decay mode is exploited in this paper to search for both resonant and nonresonant production of Higgs-boson pairs in $pp$ collisions at $\sqrt{s} = 13$ TeV. Two analyses are presented. The “resolved” analysis is optimized for $hh$ systems that have sufficiently low mass to be resolved into four distinct $b$-jet signatures in the ATLAS detector. The “boosted” analysis focuses on higher-mass $hh$ systems that are characterized by higher-momentum Higgs bosons for which the two $b$-jets cannot be resolved due to the high boost. In this situation, large-radius jets are utilized to capture the by-products of each Higgs-boson decay and small-radius track jets are used to identify the presence of $b$-hadrons. The final results are obtained using the resolved analysis up to resonance masses of 1100 GeV, where its expected sensitivity is higher than that of the boosted analysis, whereas the boosted analysis is used for masses above 1100 GeV.

The two analyses generally follow the same approach as that adopted for the 8 TeV data (see Ref. \cite{9}). The analysis of the 13 TeV data reported in this paper benefits from an enhanced sensitivity to high-mass resonances due to the significant increase in the production cross section in that kinematic region. Furthermore, the boosted analysis includes a channel with only three $b$-tagged track jets, in addition to the channel with four $b$-tagged track jets already included in the previous analysis. This new channel improves sensitivity for resonances with mass above 2000 GeV because the $b$-jet identification efficiency decreases sharply at high transverse momenta. The boosted analysis also operates with smaller track-jet radii to account for the larger boost at 13 TeV.

II. THE ATLAS DETECTOR

The ATLAS experiment \cite{15} at the LHC is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and a near $4\pi$ coverage in solid
angle. 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 (ID) covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip, and transition radiation tracking (TRT) detectors. A new innermost pixel layer [16] inserted at a mean radius of 3.3 cm is used for the first time in the 2015 data taking. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements. A steel/scintillator-tile hadronic 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 (MS) surrounds the calorimeters and includes three large superconducting air-core toroids. The field integral of the toroids ranges between 2.0 and 6.0 T m for most of the detector. The MS includes a system of fast detectors for triggering and precision tracking chambers. A dedicated trigger system is used to select events. The first-level trigger is implemented in hardware and uses the calorimeter and muon detectors to reduce the accepted rate to 100 kHz. This is followed by a software-based high-level trigger (HLT) that reduces the accepted event rate to 1 kHz on average.

### III. DATA AND SIMULATION SAMPLES

The data sample used in this analysis was collected during the 2015 LHC run with $pp$ collisions at $\sqrt{s} = 13$ TeV. After requiring that the data be collected during stable beam conditions and that relevant detector systems be functional, the total integrated luminosity is estimated to be 3.2 fb$^{-1}$ with an uncertainty of 5.0% derived following the methodology detailed in Ref. [17]. In the resolved analysis, events are selected by a combination of three triggers requiring either one or two jets selected by a dedicated HLT $b$-tagging algorithm [18]. These triggers require either one $b$-tagged jet with transverse momentum $p_T > 225$ GeV, two $b$-tagged jets with $p_T > 55$ GeV and an additional jet with $p_T > 100$ GeV, or four jets with $p_T > 35$ GeV, two of which are $b$-tagged. A trigger requiring a single jet of radius 1.0 and $p_T > 360$ GeV is used to select events in the boosted analysis. The $p_T$ thresholds for these single- or multiple-jet triggers are lower at the first level of the trigger system. The combination of all the above triggers has an efficiency rising from 95% to 99% for selecting $bb\bar{b}\bar{b}$ signal events passing the full analysis selection as the resonance mass increases.

Simulated Monte Carlo (MC) event samples are used to model signal production and the background from $tt$ and $Z +$ jets events. A method based on data is used to model the dominant multijet background. Signal $G_{KK}$ events are generated at leading order (LO) with MG5\_AMC@NLO v2.2.2 [19] using the NNPDF2.3 LO parton distribution function (PDF) set [20], and PYTHIA 8.186 [21] to model the parton shower and hadronization process using the A14 set of tuned underlying-event parameters [22]. The Higgs boson mass is set to 125.0 GeV. Values of the signal cross section times branching ratio for $G_{KK}^\pm \rightarrow hh \rightarrow bb\bar{b}\bar{b}$ with the coupling constant $k/M_{Pl} = 1$ are 11.2 fb and 0.185 fb for $G_{KK}$ masses of 1000 GeV and 2000 GeV, respectively. The parameter $k$ corresponds to the curvature of the warped extra dimension and the effective four-dimensional Planck scale $M_{Pl} = 2.4 \times 10^{18}$ GeV. Signal samples are also generated with $k/M_{Pl} = 2$ to study broader resonances. Both the cross section and natural width depend on $(k/M_{Pl})^2$. Generation of the heavy $H$ scalar in a simplified model with a fixed narrow width $\Gamma_H = 1$ GeV is performed with MG5\_AMC@NLO and the CT10 PDF set [23]. With this $\Gamma_H$ choice, the width of the reconstructed $hh$ resonance is dominated by the experimental resolution. For this model, parton showering and hadronization are handled by HERWIG++ [24] with the CTEQ6L1 PDF set [25] and the UEEEE5 underlying-event tune [26]. The scalar interpretation for this search only makes use of the acceptance times efficiency from this model and no interpretation in terms of 2HDM parameters is presented. Nonresonant SM $pp \rightarrow hh \rightarrow bb\bar{b}\bar{b}$ events are generated via the gluon-fusion process with MG5\_AMC@NLO using form factors for the top-quark loop from HPAIR [27,28]. The cross section times branching ratio to the $bb\bar{b}\bar{b}$ final state, evaluated at next-to-next-to-leading order with the summation of logarithms at next-to-next-leading-logarithm accuracy, is $11.3^{+1.0}_{-1.0}$ fb [29]. The uncertainty includes the effects due to renormalization and factorization scales, PDF set, $a_s$, effects of finite top-quark mass in loops, and the $h \rightarrow bb$ branching ratio.

Generation of $tt$ events is performed with POWHEG-BOX v1 using the CT10 PDF set. The parton shower, hadronization, and the underlying event are simulated using PYTHIA 6.428 [30] with the CTEQ6L1 PDF set and the corresponding Perugia 2012 tune [31]. The top-quark mass is set to 172.5 GeV. Higher-order corrections to $tt$ cross sections are computed with Top++ 2.0 [32]. These incorporate NNLO corrections in QCD, including resummation of next-to-next-to-leading logarithmic soft gluon terms. The overall $tt$ normalization is extracted from the data while the shape of kinematic distributions is taken from.
MC simulation. The $Z + \text{jets}$ sample is generated using PYTHIA 8.186 with the NNPDF2.3 LO PDF set.

For all MC samples, charm-hadron and bottom-hadron decays are handled by EVTGEN 1.2.0 [33]. To simulate the impact of multiple $pp$ interactions that occur within the same or nearby bunch crossings (pileup), minimum-bias events generated with PYTHIA 8 are overlaid on top of the hard scatter event. The detector response is simulated with GEANT 4 [34,35] and the events are processed with the same reconstruction software as that used for the data.

IV. EVENT RECONSTRUCTION

The resolved and boosted analyses rely on the reconstruction of jets with the anti-$k_t$ clustering algorithm [36] but with different values of the radius parameter $R$. Calorimeter jets with $R = 0.4$ (1.0) are used to determine the kinematic properties of Higgs-boson candidates in the resolved (boosted) analysis. Those jets are reconstructed from topological clusters of energy deposits in calorimeter cells. The $R = 0.4$ jet energies are determined from reconstructed cluster energies at the electromagnetic scale with correction factors derived from simulation to account for the response of the calorimeter to hadrons [38]. Jets from pileup are suppressed with the use of tracking information as detailed in Ref. [38]. The $R = 1.0$ jets are built from locally calibrated clusters [37] and are trimmed [39] to minimize the impact of pileup. This trimming proceeds by re-clustering the jet with the $k_t$ algorithm [40] into smaller $R = 0.2$ subjets and removing those subjets with $p_T^{\text{subj}} / p_T^{\text{jet}} < 0.05$, where $p_T^{\text{subj}}$ is the transverse momentum of the subjet and $p_T^{\text{jet}}$ that of the original jet. In addition to the above large-$R$ trimmed jets, the boosted analysis uses track jets with $R = 0.2$ to identify $b$-hadrons from Higgs-boson decays [41]. Such jets are reconstructed from charged-particle tracks with $p_T > 0.4$ GeV and $|\eta| < 2.5$ that satisfy a set of hit and impact parameter criteria to make sure that the tracks originate from the primary vertex, thereby minimizing the impact of pileup. Track jets are associated to large-$R$ jets using ghost association [42]. In this method, the large-$R$ jet algorithm is rerun with both the four-momenta of track jets modified to have infinitesimally small momentum (the “ghosts”) and all topological energy clusters in the event as potential constituents of jets. As a result, the presence of track jets does not alter the large-$R$ jets already found and their association to specific large-$R$ jets is determined by the jet algorithm. Collision vertices are reconstructed requiring a minimum of two tracks with $p_T > 0.4$ GeV in each vertex. The primary vertex is chosen to be the vertex with the largest $\sum p_T^2$, where the sum extends over all tracks associated with the vertex.

The identification of jets containing $b$-hadrons is based on the $R = 0.4$ calorimeter ($R = 0.2$ track) jets in the resolved (boosted) analysis and a multivariate tagging algorithm [43]. This algorithm is applied to a set of tracks with loose impact parameter constraints in a region of interest around each jet axis to enable the reconstruction of the $b$-hadron decay vertex. The $b$-tagging requirements result in an efficiency of 70% (77%) for jets containing $b$-hadrons in the resolved (boosted) analysis, as determined in a sample of simulated $t\bar{t}$ events. The corresponding efficiencies for $c$-hadron jets and light-quark or gluon jets are 12% (29%) and 0.2% (1.4%), respectively. Different $b$-tagging operating points are chosen in the two analyses to maximize their respective sensitivities.

Muons are reconstructed by combining tracks in the ID and MS, and are required to satisfy tight muon identification criteria [44]. The four-momentum of muons with $p_T > 4$ GeV and $|\eta| < 2.5$, that are within $\Delta R$ of 0.4 (0.2) of jets used for $b$-tagging in the resolved (boosted) analysis, is added to the calorimeter jet four-momentum to partially account for the energy lost in semileptonic $b$-hadron decays.

V. EVENT SELECTION

The event selection for the resolved and boosted analyses is described below. These analyses are optimized independently for the reconstruction and selection of $h h \rightarrow b\bar{b}b\bar{b}$ final states, with the resolved analysis aiming at event topologies containing four distinct $b$-jets, whereas the boosted analysis focuses on topologies with higher-momentum Higgs bosons resulting in merged jets.

Different selection and background estimation strategies are adopted for the two analyses. To facilitate the comparison between these different choices, Table I summarizes each of the requirements and approaches described in this section.

A. Resolved analysis

1. Selection

Events selected for the resolved analysis must contain at least four $b$-tagged jets with $|\eta| < 2.5$ and $p_T > 40$ GeV. The four highest-$p_T$ $b$-tagged jets are used to form two dijet systems, requiring an angular distance $\Delta R$ between the jets within the dijet system smaller than 1.5. The transverse momentum of the leading (subleading) dijet system is required to be greater than 200 (150) GeV. These requirements are made to ensure a high trigger efficiency and to avoid ambiguities in forming dijets. In the rare case that a jet is assigned to more than one dijet system, only the combination containing the jets with the highest probability of being $b$-jets according to the multivariate $b$-tagging algorithm is considered.

The resolved analysis considers resonance masses in the range 400–1500 GeV. Event selection that varies as a function of the reconstructed resonance mass ($m_{4j}$) is used to increase the analysis sensitivity across the mass range searched. Mass-dependent selection requirements are made on the leading dijet $p_T$, the subleading dijet $p_T$ and the pseudorapidity difference between the dijets as follows [9]:
TABLE I. Event selection requirements and definition of the different regions used in the resolved and boosted analyses. The methodologies used to estimate the background normalization and shape are also outlined. The variables are defined in the text. Dijet and large-\( R \) jet minimum \( p_T \) values are indicated for leading (subleading) such objects. The functions \( f(m_{dij}) \) and \( f'(m_{dij}) \) represent the mass dependence of the minimum \( p_T \) and maximum \( |\Delta \eta| \) requirements placed on the dijet candidates in the resolved analysis.

<table>
<thead>
<tr>
<th>Region</th>
<th>Resolved</th>
<th>Boosted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event preselection</td>
<td>( \geq 4 ) jets with ( p_T &gt; 40 ) GeV, (</td>
<td>\eta</td>
</tr>
<tr>
<td></td>
<td>( \geq 2 ) dijets with ( p_T &gt; 200(150) ) GeV, ( \Delta R &lt; 1.5 ),</td>
<td>(</td>
</tr>
<tr>
<td></td>
<td>( p_T &gt; f(m_{dij}),</td>
<td>\Delta \eta</td>
</tr>
<tr>
<td>( t\bar{t} ) veto</td>
<td>( X_{\mu} &lt; 3.2 )</td>
<td>3 or 4 ( b )-tagged jets</td>
</tr>
<tr>
<td>Tagging</td>
<td>4 ( b )-tagged jets</td>
<td></td>
</tr>
<tr>
<td>Signal region (SR)</td>
<td>( X_{hh} &lt; 1.6 )</td>
<td></td>
</tr>
<tr>
<td>Sideband region (SB)</td>
<td>Resolved: ( \sqrt{(m_{22}^{lead} - 124 \text{ GeV})^2 + (m_{22}^{subl} - 115 \text{ GeV})^2} &gt; 58 \text{ GeV} )</td>
<td>Boosted: ( \sqrt{(m_{1}^{lead} - 124 \text{ GeV})^2 + (m_{1}^{subl} - 115 \text{ GeV})^2} &gt; 36 \text{ GeV} )</td>
</tr>
<tr>
<td>Control region (CR)</td>
<td>Complementary to SR and SB</td>
<td></td>
</tr>
<tr>
<td>Multijet normalization</td>
<td>Scaled yields from 2-tag SR, scaling derived from</td>
<td>Scaled yields from 2-tag SR, scaling derived from</td>
</tr>
<tr>
<td></td>
<td>4-tag to 2-tag ratio in SB</td>
<td>3(4)-tag to 2-tag fit to leading jet mass in SB</td>
</tr>
<tr>
<td>Multijet shape</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t\bar{t} ) normaliziation</td>
<td>Scaled yields from ( t\bar{t} ) CR, scaling derived from</td>
<td>Scaled yields from MC simulation, scaling derived from</td>
</tr>
<tr>
<td></td>
<td>semileptonic ( t\bar{t} ) events</td>
<td>3(4)-tag to 2-tag fit to leading jet mass in SB</td>
</tr>
<tr>
<td>( t\bar{t} ) shape</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Derived from MC simulation</td>
<td></td>
</tr>
</tbody>
</table>

\[
p_T^{lead} = \begin{cases} 
400 \text{ GeV} & \text{if } m_{dij} > 910 \text{ GeV}, \\
200 \text{ GeV} & \text{if } m_{dij} < 600 \text{ GeV}, \\
0.65m_{dij} - 190 \text{ GeV} & \text{otherwise},
\end{cases} \\
p_T^{subl} = \begin{cases} 
260 \text{ GeV} & \text{if } m_{dij} > 990 \text{ GeV}, \\
150 \text{ GeV} & \text{if } m_{dij} < 520 \text{ GeV}, \\
0.23m_{dij} + 30 \text{ GeV} & \text{otherwise},
\end{cases} \\
|\Delta \eta_{dijets}| = \begin{cases} 
1.0 & \text{if } m_{dij} < 820 \text{ GeV}, \\
1.6 \times 10^{-3}m_{dij} - 0.28 & \text{otherwise}.
\end{cases}
\]

These selection requirements were optimized simultaneously by performing a three-dimensional scan of threshold values, using the expected exclusion limit on the \( G_{KK} \) resonance with \( k/M_{KK} = 1 \) as a metric.

After selecting two dijets that satisfy the mass-dependent criteria, 15% of the total background consists of \( t\bar{t} \) events. This \( t\bar{t} \) background mainly comprises events where both top quarks decay hadronically. These hadronic decays often lead to three jets for each top quark—one \( b \)-jet directly from the top-quark decay and two from the decay of the \( W \) boson. Reduction of the \( t\bar{t} \) background is important as relatively large systematic uncertainties are associated with modeling \( t\bar{t} \) in the signal region. In order to reduce the \( t\bar{t} \) background, jets not already used in the formation of the two dijets (“extra jets”) in the event are used to reconstruct \( W \)-boson and top-quark candidates by combining them with one or both of the jets in a given dijet. 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 candidate is reconstructed by adding the four-momentum of the extra jet 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. The top-quark candidate is reconstructed by summing the dijet and the extra jet. The compatibility with the top-quark decay hypothesis is then determined using the variable

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

where \( m_W \) and \( m_t \) are the invariant masses of the \( W \)-boson and top-quark candidates. The values in the denominator approximate the dijet and three-jet system mass resolutions. If either dijet in an event has \( X_{\mu} < 3.2 \) for any possible
combinations with an extra jet, the event is rejected. This requirement, referred to as the “$t\bar{t}$ veto,” reduces the $t\bar{t}$ background by $\sim 60\%$, while retaining $\sim 90\%$ of signal events. The event selection criteria described above are collectively referred to as the “4-tag” selection requirements.

Following the 4-tag selection, a requirement on the combination of the leading and subleading dijet masses ($m_{2j}^{\text{lead}}$ and $m_{2j}^{\text{subl}}$, respectively) is used to define the signal region. The signal region is defined using the variable

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

(2)

where the $0.1m_{2j}$ terms approximate the widths of the mass distributions. The center of the signal region was optimized using $G_{kk}$ samples with $k/M_H = 1$. On average, the subleading Higgs-boson candidate is reconstructed at lower masses as a result of energy lost from semileptonic hadron decays and final-state radiation. The signal region is defined as $X_{hh} < 1.6$. This corresponds to the kinematic requirements illustrated by the inner region in Fig. 1. The data shown in this figure are derived from a sample of events that satisfy all selection criteria except for having only two jets that pass the $b$-tagging requirements, referred to as the “2-tag” sample.

The acceptance times efficiency for each stage of the resolved-analysis event selection is shown in Fig. 2 for spin-2 and spin-0 resonances. The acceptance times efficiency, $A \times \varepsilon$, of the full selection for the $G_{kk}$ with $k/M_H = 1$ ranges from $0.1\%$ for a $G_{kk}$ of mass 400 GeV to $5.3\%$ for a $G_{kk}$ with a mass of 1000 GeV. The spin of the resonance affects the angular distribution of the decay products, resulting in a lower acceptance in the case of a spin-0 $H$ boson than for the spin-2 $G_{kk}$. As a result, the spin-0 resonance search is performed starting at a mass of 500 GeV. Nonresonant di-Higgs production occurs primarily at the low end of the $m_{4j}$ spectrum, leading to $A \times \varepsilon = 0.64\%$ for the full selection.

The final step in the resonant analysis is to search for an excess in the $m_{4j}$ distribution for events in the signal region. The sensitivity of the search is increased by improving the $m_{4j}$ resolution in this region. This is achieved by scaling the four-momentum of each of the Higgs-boson candidates such that their mass is equal to the Higgs-boson mass. This leads to an improvement of $\sim 30\%$ in the signal $m_{4j}$ resolution with little impact on the background.

2. Background estimation

After the 4-tag selection described above, $\sim 90\%$ of the remaining background in the signal region originates from multijet events, which are modeled using data. The remaining $\sim 10\%$ of the background is expected from $t\bar{t}$ events. The $t\bar{t}$ yield is determined from data, while the $m_{4j}$ shape is taken from MC simulation. The $Z + j$ jets contribution is less than $1\%$ of the total background and is estimated from MC simulation. The background from all other sources—including processes featuring Higgs bosons—is negligible.

Multijet background.—The multijet background is modeled using an independent data sample selected using the same trigger and selection requirements as described above, except for the $b$-tagging requirement: only one of the two selected dijets is formed from $b$-tagged jets, while the other dijet is formed from jets that both fail the $b$-tagging requirements. This “2-tag” selection yields a data sample that consists of $98\%$ multijet events and $2\%$ $t\bar{t}$ events. The predicted signal contamination is negligible.

The 2-tag sample is normalized to the 4-tag sample and its kinematics corrected for differences introduced by the additional $b$-tagging requirement on the 4-tag sample. These kinematic differences arise because the $b$-tagging efficiency varies 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 $m_{2j}^{\text{lead}} - m_{2j}^{\text{subl}}$ plane. 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. 1. The sideband region is defined as $\sqrt{(m_{2j}^{\text{lead}} - 124 \text{ GeV})^2 + (m_{2j}^{\text{subl}} - 115 \text{ GeV})^2} > 58 \text{ GeV}$.
while the control region is defined as the region in the $m^{\text{lead}}_Z - m^{\text{subl}}_Z$ plane 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 both 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 $\mu_{\text{Multijet}}$ calculated in the sideband region:

$$
\mu_{\text{Multijet}} = \frac{N^4_{\text{Sideband}}}{N^2_{\text{Multijet}}} = \frac{N^4_{\text{data}}}{N^2_{\text{data}}} - \frac{N^4_{\text{Z+jets}}}{N^2_{\text{Z+jets}}},
$$

where $N^{2/4\text{-tag}}_{\text{data}}$ is the number of events observed in the sideband region in the 2- or 4-tag data sample, respectively.

The yields $N^{2/4\text{-tag}}_{\text{data}}$ are the estimated number of $\bar{t}t$ events in the 2- or 4-tag selected sideband region estimated from MC simulation. To predict the distributions of the multijet background in each region, the predicted $\bar{t}t$ 2-tag distributions are first subtracted from the 2-tag data sample before the distribution is scaled by $\mu_{\text{Multijet}}$.

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

The multijet background model is validated in the control region. Table II compares the observed data yield in the control region with the corresponding background estimate. The modeling of the $m_4$ distribution in the control region is shown in Fig. 3. The 4-tag events in the control region, from linear fits to the ratio of the total background model to data for three kinematic distributions that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

The multijet background model is validated in the control region. Table II compares the observed data yield in the control region with the corresponding background estimate. The modeling of the $m_4$ distribution in the control region is shown in Fig. 3. The 4-tag events in the control region, from linear fits to the ratio of the total background model to data for three kinematic distributions that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

The multijet background model is validated in the control region. Table II compares the observed data yield in the control region with the corresponding background estimate. The modeling of the $m_4$ distribution in the control region is shown in Fig. 3. The 4-tag events in the control region, from linear fits to the ratio of the total background model to data for three kinematic distributions that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data. 

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.

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 that are found to have the largest disagreement between 2-tag and 4-tag: the leading dijet $p_T$, the angular separation between the jets in the subleading dijet, and the angular separation between the two dijets. The reweighting is performed using one-dimensional distributions but is iterated so that correlations between the three variables are taken into account. After the correction process, there is agreement between the background model and sideband region data.
region are well described by the background model in both normalization and $m_{dij}$ shape.

tt background.—The normalization of the $t\bar{t}$ background is derived from data in a $t\bar{t}$ control region. Due to the limited yield in this control region, the shape of the $t\bar{t}$ background is taken from MC simulation. To further decrease statistical uncertainties, the $t\bar{t}$ shape is derived from MC simulation using the “2-tag” selection, with a systematic uncertainty assigned to cover the differences between the 2-tag and 4-tag $m_{dij}$ distributions.

The $t\bar{t}$ control region is formed from events which pass the 4-tag selection, except for the $t\bar{t}$ veto, which is reversed: if either of the dijets fails the $X_{mtt}$ requirement, the event enters the $t\bar{t}$ control region. This selection leads to a sample of 21 events, of which 13.3 are estimated to be multijet events using the 2-tag sample described previously. After subtracting the multijet background, the $t\bar{t}$ control region yield is extrapolated to predict the $t\bar{t}$ yield in the signal region, $N_{t\bar{t}}$, using the following equation:

$$N_{t\bar{t}} = \frac{e_i^2}{1 - e_i^2} \times N_{t\bar{t}}^{CR},$$

where $N_{t\bar{t}}^{CR}$ is the number of events in the $t\bar{t}$ control region, after subtraction of multijet background, and $e_i$ is the efficiency for a selected dijet in a $t\bar{t}$ event to pass the $t\bar{t}$ veto. This equation relies on the assumption that the $e_i$ of each dijet in the event is uncorrelated, an assumption validated in $t\bar{t}$ MC simulation. The $e_i$ is measured using an independent, high-purity “semileptonic $t\bar{t}$” data sample. Events in this sample are selected by requiring one dijet candidate passing 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 a distance $\Delta R < 1.5$ of the $b$-tagged jet. The leptonic top-quark candidate is required to have $p_T > 150$ GeV, where the leptonic top $p_T$ is defined as the vector sum of the $b$-jet $p_T$ and the muon $p_T$. 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 $e_{vtt} = 0.60 \pm 0.04$ (stat) $\pm 0.06$ (syst). A 10% systematic uncertainty is assigned to cover potential differences between $e_i$ as measured in the semileptonic $t\bar{t}$ sample and $e_i$ in the full 4-tag selection, where the method is applied in $t\bar{t}$ MC simulation to evaluate such differences.

The measured $e_i$ agrees well with the corresponding semileptonic $t\bar{t}$ MC prediction of 0.58.

Equation (4) gives a data-driven $t\bar{t}$ background prediction of $4.2 \pm 3.8$ events. The uncertainty is dominated by the statistical uncertainty in the yield in the $t\bar{t}$ control region, with a smaller contribution from the uncertainty in the measured $t\bar{t}$ veto efficiency.

### 3. Systematic uncertainties

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

The signal modeling 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 following sources of theoretical uncertainty in the acceptance are evaluated: missing higher-order terms in the matrix elements and PDF set, as well as modeling of the underlying event, hadronic showers, initial- and final-state radiation. The total theoretical uncertainty is dominated by the uncertainties associated with the modeling of the initial- and final-state radiation.

The jet energy uncertainties are derived based on in situ measurements performed during Run 1 and from MC simulation extrapolations from Run-1 to Run-2 conditions [45]. The JES systematic uncertainty is evaluated using three separate and orthogonal uncertainty components [46]. The JER uncertainty is evaluated by smearing jet energies according to the systematic uncertainties of the resolution measurement [46]. The uncertainty in the $b$-tagging efficiency is evaluated by propagating the systematic uncertainty in the measured tagging efficiency for $b$-jets [47]. The efficiencies are measured as a function of $b$-jet $p_T$ and $\eta$. For $b$-jets with $p_T > 300$ GeV, systematic uncertainties in the tagging efficiencies are extrapolated with MC simulation and are consequently larger [18].

Systematic uncertainties in the normalization and shape of the multijet background model are assessed in the control region. The background prediction in the control region agrees with the observed data to within $\pm 5\%$, which is taken as the uncertainty in the predicted multijet yield. To further test the robustness of the background estimation, the background model is re-evaluated using different sideband and control region definitions and different $b$-tagging requirements on the “2-tag” sample. These changes affect the kinematic and flavor compositions of the various regions used in the background prediction. The control region agreement and signal region predictions of all variations considered are all consistent to within the assigned $\pm 5\%$ uncertainty in the multijet background prediction.

The uncertainty in the description of the multijet $m_{dij}$ distribution is determined by comparing the background prediction to the data in the control region as shown in Fig. 3. To evaluate the level of agreement, a linear fit is performed on the ratio of the distributions. This fit, along with its uncertainties, shown in the bottom panel of Fig. 3, gives a slope consistent with zero. 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 described above. The uncertainty in the MC-derived $t\bar{t}$ $m_{dij}$
distribution is dominated by the uncertainty associated with using the 2-tag selection to model the 4-tag selection. This uncertainty is assessed by comparing the 2-tag and 4-tag MC predictions in the signal region.

Table III summarizes the relative impact of the uncertainties in the event yields.

### 4. Event yields

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 are presented in Table IV. The numbers of predicted background events and observed events are in agreement.

Figure 4 shows a comparison of the predicted $m_{4j}$ background distribution to that observed in the data. The predicted background and observed distributions are in agreement, with no significant local excesses.

### B. Boosted analysis

#### 1. Selection

The boosted analysis selects events with at least two large-$R$ jets with $250 < p_T < 1500$ GeV, $|\eta| < 2.0$, and mass $m_{1j} > 50$ GeV. The upper bound on the transverse momentum and the mass requirement correspond to the kinematic region where jet calibration uncertainties are available from Refs. [41] and [48]. Only the two large-$R$ jets with highest $p_T$ are retained for further selection. In order to reduce the contamination from $t\bar{t}$ events, the leading jet is additionally required to have $p_T > 350$ GeV, thus ensuring that all top-quark decay products are contained in a single large-$R$ jet with mass close to that of the top quark.

![Figure 4](ATLAS.png)

**Figure 4.** Distribution of $m_{4j}$ in the signal region of the resolved analysis for data compared to the predicted background. The hatched bands represent the combined statistical and systematic uncertainty in the total background estimate. The expected signal distribution for a $G_{KK}$ resonance with mass of 800 GeV is also shown.

### TABLE III. Summary of systematic uncertainties (expressed in percentage yield) in the total background and signal event yields in the signal region of the resolved analysis. Uncertainties are provided for nonresonant SM Higgs pair production, for a $G_{KK}$ resonance with $k/M_{Pl} = 1$ and $m = 500$ GeV, and for three resonances with $m = 800$ GeV: a $G_{KK}$ resonance with $k/M_{Pl} = 1$, a $G_{KK}$ resonance with $k/M_{Pl} = 2$, and a spin-0 narrow-width $H$ boson.

<table>
<thead>
<tr>
<th>Source</th>
<th>Background</th>
<th>SM $hh$</th>
<th>$k/M_{Pl} = 1$</th>
<th>$\vec{k}/\vec{M}_{Pl} = 1$</th>
<th>$\vec{k}/\vec{M}_{Pl} = 2$</th>
<th>$H$ (800 GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>JER</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>JES</td>
<td>12</td>
<td>14</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>$b$-tagging</td>
<td>18</td>
<td>15</td>
<td>26</td>
<td>27</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Theoretical</td>
<td>9</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Multijet</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>24</td>
<td>21</td>
<td>28</td>
<td>28</td>
<td>28</td>
</tr>
</tbody>
</table>

### TABLE IV. The number of predicted background events in the $hh$ signal region for the resolved analysis, compared to the data. The yield for two potential signals, SM nonresonant Higgs pair production and an 800 GeV $G_{KK}$ resonance with $k/M_{Pl} = 1$ are shown. The quoted errors include both the statistical and systematic uncertainties.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Signal region yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>43.3 ± 2.3</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>4.3 ± 3.0</td>
</tr>
<tr>
<td>$Z +$ jets</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>47.6 ± 3.8</td>
</tr>
<tr>
<td>Data</td>
<td>46</td>
</tr>
<tr>
<td>SM $hh$</td>
<td>0.22 ± 0.05</td>
</tr>
<tr>
<td>$G_{KK}$ (800 GeV), $k/M_{Pl} = 1$</td>
<td>5.7 ± 1.5</td>
</tr>
</tbody>
</table>
At least two track jets must be found by the ghost method [42] to be associated with each large-\( R \) jet. They are required to be consistent with the primary vertex of the event as well as to satisfy \( p_T > 10 \text{ GeV} \) and \( |\eta| < 2.5 \).

Since high-mass resonances tend to produce jets that are more central than multijet background processes, the two large-\( R \) jets are required to have a separation \( |\Delta \eta| < 1.7 \).

Signal event candidates are selected if each of the large-\( R \) jets has a mass consistent with that of the Higgs boson. This is defined as for the resolved analysis in Eq. (2), where the small-\( R \) dijet mass is replaced by the large-\( R \) jet mass, requiring \( X_{hh} < 1.6 \). This requirement defines the signal region in the leading-subleading large-\( R \) jet mass plane.

Two samples of events are selected based on the number of \( b \)-tagged leading and subleading track jets associated with each large-\( R \) jet. They are referred to as the “3-tag” and the “4-tag” samples, and require exactly three or at least four track jets passing the \( b \)-tagging selection, respectively. In the 3-tag sample, the fourth jet is explicitly required to fail the \( b \)-tagging requirements to define orthogonal samples.

The signal region corresponds to the kinematic requirements illustrated by the inner region in Fig. 5. The data shown in this figure are derived from a sample of events that satisfy all selection criteria except for having only two track jets that pass the \( b \)-tagging requirements, referred to as the “2-tag” sample. This sample is used to estimate the background contribution as described below.

The acceptance times efficiency for each stage of the boosted-analysis event selection is shown in Fig. 6 for the \( G_{KK} \) and heavy scalar models. The requirement that at least two individual track jets be associated to the large-\( R \) jets becomes less efficient at high mass due to merging. The full selection for a \( G_{KK} \) resonance with a mass of 1000 GeV (2000 GeV) and \( k/M_D = 1 \) has an acceptance times efficiency of 9% (11%) in the 3-tag sample and 8% (5%) in the 4-tag sample.

2. Background estimation

As in the resolved analysis, the dominant source of background stems from multijet (80%–90%) events and the rest is primarily due to \( t\bar{t} \) production. The background estimation method generally follows the same approach as that described in Sec. VA2. Differences are highlighted below.

The shape of the multijet background in both the 3-tag and 4-tag samples is derived from the 2-tag sample. Due to the large statistical uncertainty in the background prediction for dijet masses \( (m_{jj}) \) above 1500 GeV, an exponential fit to the data in the range between 900 and 2000 GeV is used to model the high-mass tail of the dijet distribution in
the signal region. The estimated background yield in each signal region, \( N_{\text{bkg}}^{3(4)-\text{tag}} \), is computed according to

\[
N_{\text{bkg}}^{3(4)-\text{tag}} = \frac{3(4)-\text{tag}}{\text{multijet}}N_{\text{multijet}}^{2(3)-\text{tag}} + \alpha_{t\bar{t}}^{3(4)-\text{tag}}N_{t\bar{t}}^{3(4)-\text{tag}} + N_{Z\text{-jets}}^{3(4)-\text{tag}},
\]

where \( N_{\text{multijet}}^{2(3)-\text{tag}} \) is the number of multijet events in the 2-tag sample, \( N_{t\bar{t}}^{3(4)-\text{tag}} \) and \( N_{Z\text{-jets}}^{3(4)-\text{tag}} \) are the numbers of events predicted by the 3(4)-tag \( t\bar{t} \) and \( Z + \text{jets} \) MC samples. The parameter \( \mu_{\text{multijet}}^{3(4)-\text{tag}} \) corresponds to the ratio of multijet event yields in the 3(4)-tag and 2-tag samples, as defined in Eq. (3), except for considering 3- or 4-tag events in the numerator. Finally, the parameter \( \alpha_{t\bar{t}}^{3(4)-\text{tag}} \) is a scale factor designed to correct the \( t\bar{t} \) event yield estimated from the MC simulation.

A sideband region defined by \( \sqrt{(m_{t\bar{t}}^{\text{obs}} - 124 \text{ GeV})^2 + (m_{Z}^{\text{subl}} - 115 \text{ GeV})^2} > 36 \text{ GeV} \) is used to measure \( \mu_{\text{multijet}}^{3(4)-\text{tag}} \) and \( \alpha_{t\bar{t}}^{3(4)-\text{tag}} \) from the data. The background estimate is validated in a control region defined to be complementary to the sideband and signal regions.

Both \( \mu_{\text{multijet}}^{3(4)-\text{tag}} \) and \( \alpha_{t\bar{t}}^{3(4)-\text{tag}} \) are extracted from a binned likelihood fit to the leading large-\( R \) jet mass distribution obtained in the sideband region of the 3(4)-tag sample, as shown in Fig. 7. In this fit, the multijet distribution is extracted from the 2-tag sample, after subtraction of the \( t\bar{t} \) and \( Z + \text{jets} \) contributions predicted by the MC simulation. The \( t\bar{t} \) and \( Z + \text{jets} \) distributions in the sideband region of the 3(4)-tag sample are taken from the MC simulation.

The resulting fit values and their statistical uncertainties for the 3-tag sample are \( \mu_{\text{multijet}}^{3\text{-tag}} = 0.160 \pm 0.003 \) and \( \alpha_{t\bar{t}}^{2\text{-tag}} = 1.02 \pm 0.09 \), with a correlation coefficient of \(-0.60\). A large anticorrelation is observed since the multijet and \( t\bar{t} \) background contributions are constrained to add up to the total number of events in the sideband region of the 3-tag and 4-tag data samples.

The modeling of the background yield and kinematics is validated in the control region of the 3-tag and 4-tag samples. Good agreement is observed between the data and the predicted background in both the sideband and control regions of the 3-tag and 4-tag samples as shown in Table V. The shapes of the \( t\bar{t} \) kinematic distributions in the 4-tag signal region are extracted from the MC simulation in the 3-tag signal region due to the limited size of the 4-tag MC sample.

### TABLE V. The number of events in data and predicted background events in the \( hh \) sideband and control regions of the 3-tag and 4-tag samples for the boosted analysis. The number of multijet and \( t\bar{t} \) background events in the sideband regions are constrained by the number of observed events, as explained in the text. The uncertainties are purely statistical.

<table>
<thead>
<tr>
<th>Sample (3-tag)</th>
<th>Sideband region</th>
<th>Control region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>4328 ( \pm ) 27</td>
<td>607 ( \pm ) 10</td>
</tr>
<tr>
<td>( t\bar{t} )</td>
<td>683.5 ( \pm ) 8.1</td>
<td>99.6 ( \pm ) 3.1</td>
</tr>
<tr>
<td>( Z + \text{jets} )</td>
<td>31.8 ( \pm ) 3.7</td>
<td>7.7 ( \pm ) 1.8</td>
</tr>
<tr>
<td>Total</td>
<td>5043 ( \pm ) 28</td>
<td>715 ( \pm ) 11</td>
</tr>
<tr>
<td>Data</td>
<td>5043</td>
<td>724</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample (4-tag)</th>
<th>Sideband region</th>
<th>Control region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>247.4 ( \pm ) 1.5</td>
<td>34.7 ( \pm ) 0.6</td>
</tr>
<tr>
<td>( t\bar{t} )</td>
<td>28.4 ( \pm ) 1.5</td>
<td>5.1 ( \pm ) 0.7</td>
</tr>
<tr>
<td>( Z + \text{jets} )</td>
<td>3.4 ( \pm ) 1.2</td>
<td>0.6 ( \pm ) 0.5</td>
</tr>
<tr>
<td>Total</td>
<td>279.2 ( \pm ) 2.5</td>
<td>40.3 ( \pm ) 1.0</td>
</tr>
<tr>
<td>Data</td>
<td>279</td>
<td>45</td>
</tr>
</tbody>
</table>
SEARCH FOR PAIR PRODUCTION OF HIGGS BOSONS IN ...

FIG. 8. Dijet mass distribution in the control region for data (points) and background estimate (histograms) in the boosted analysis for events in the (left) 3-tag and (right) 4-tag categories. The hatched bands represent the statistical uncertainty in the total background estimate.

3. Systematic uncertainties

Evaluation of systematic uncertainties in the boosted analysis generally follows the same approach as that described in Sec. VA 3. Differences are highlighted here.

The large-\( R \) jet energy resolution and scale uncertainties as well as the jet mass resolution (JMR) and scale (JMS) uncertainties are derived \textit{in situ} from 8 TeV \( pp \) collisions, taking into account MC simulation extrapolations for the different detector and beam conditions present in 8 and 13 TeV data-taking periods [49]. The uncertainty in the \( b \)-tagging efficiency for track jets is evaluated with the same method used for \( R = 0.4 \) calorimeter jets.

Systematic uncertainties in the normalization and shape of the background model are assessed in the control region. The background predictions in both the 3-tag and 4-tag control regions agree with the observed data to within statistical uncertainties. The statistical uncertainties in the control region yields are assigned as systematic uncertainties in the multijet background normalization. The uncertainty in the shape of the multijet background is assessed in the control region via a linear fit to the ratio of the control region yields to the background prediction is assigned by fitting the 2-tag dijet mass distribution shown in Fig. 8.

An additional uncertainty in the shape of the tail of the background prediction is assigned by fitting the 2-tag dijet mass distribution with a variety of empirical functions designed to model power-law behavior, as described in Ref. [50]. The largest difference between the exponential function predictions and those from alternative fit functions, considering the variation of the fitted parameters within their statistical uncertainties, is taken as a systematic uncertainty.

Relative systematic uncertainties in both the background and signal event yields are summarized in Table VI for the 3-tag and 4-tag selections. For the background, the entry labeled “Statistical” corresponds to the statistical uncertainty from the fit to the leading large-\( R \) jet mass (see Sec. VB 2) used to extract the multijet and \( t \bar{t} \) background yields, taking the correlation between these yields into account. It also includes the \( t \bar{t} \) modeling uncertainties and the statistical uncertainty associated with the data yield in the 2-tag sample. Uncertainties in the \( m_{2j} \) shape of the multijet and \( t \bar{t} \) backgrounds are not listed in Table VI, as they do not affect the event yields, but are accounted for in the statistical analysis.

<table>
<thead>
<tr>
<th>Source</th>
<th>Background</th>
<th>( G_{KK}^* (1500 \text{ GeV}) )</th>
<th>( k/M_{Pl} = 1 )</th>
<th>( k/M_{Pl} = 2 )</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Luminosity</th>
<th>3-tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>JER</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>JES</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>JMR</td>
<td>1</td>
</tr>
<tr>
<td>JMS</td>
<td>5</td>
</tr>
<tr>
<td>( b )-tagging</td>
<td>1</td>
</tr>
<tr>
<td>Theoretical</td>
<td>\ldots</td>
</tr>
<tr>
<td>Multijet</td>
<td>\ldots</td>
</tr>
<tr>
<td>Statistical</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4-tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>JER</td>
</tr>
<tr>
<td>JES</td>
</tr>
<tr>
<td>JMR</td>
</tr>
<tr>
<td>JMS</td>
</tr>
<tr>
<td>( b )-tagging</td>
</tr>
<tr>
<td>Theoretical</td>
</tr>
<tr>
<td>Multijet</td>
</tr>
<tr>
<td>Statistical</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
TABLE VII. The number of predicted background events in the \( hh \) 3-tag and 4-tag signal regions, compared to the data for the boosted analysis. Errors correspond to the total uncertainties in the background estimate. The uncertainty band includes both the statistical and systematic uncertainties in the background estimate.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Signal region (3-tag)</th>
<th>Signal region (4-tag)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>235 ± 14</td>
<td>13.5 ± 2.4</td>
</tr>
<tr>
<td>( t\bar{t} )</td>
<td>48 ± 22</td>
<td>1.2 ± 1.0</td>
</tr>
<tr>
<td>Z + jets</td>
<td>2.0 ± 2.2</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>Total</td>
<td>285 ± 19</td>
<td>14.6 ± 2.4</td>
</tr>
<tr>
<td>Data</td>
<td>316</td>
<td>20</td>
</tr>
</tbody>
</table>

\( G_{\text{KK}}^{*}(1000 \text{ GeV}), \ k/M_{\text{Pl}} = 1 \) is also given.

4. Event yields

The predicted number of background events in the 3-tag and 4-tag signal regions, the number of events observed in the data, and the predicted yield for a potential signal are reported in Table VII. One event in the 4-tag signal region, with a mass of 852 GeV, is in common with the resolved analysis. The dijet mass distribution in the signal region is shown in Fig. 9. An excess of data is observed in the 3-tag signal region for \( m_{3\ell} \sim 900 \text{ GeV} \) and in the range between 1600 and 2000 GeV. The significance of these excesses is evaluated below.

VI. RESULTS

The results from the resolved and boosted analyses are interpreted separately using the statistical procedure described in Ref. [1] and references therein. A test statistic based on the profile likelihood ratio [51] is used to test hypothesized values of \( \mu \), the global signal strength factor, separately for each model tested. The statistical analysis described below is performed using the data observed in the signal regions. The systematic uncertainties are treated as independent within each signal region using Gaussian or log-normal constraint terms in the definition of the likelihood function. In the boosted analysis, the data from the 3-tag and 4-tag signal regions are fitted simultaneously treating data-derived systematic uncertainties related to the multijet background estimate as uncorrelated and all other systematic uncertainties as fully correlated. In the case of the search for nonresonant \( hh \) production, only the number of events passing the final selection is used whereas the \( m_{4\ell} \) or \( m_{3\ell} \) distributions are used in the case of the search for \( hh \) resonances.

A. Background-only hypothesis tests

In order to determine if there are any statistically significant local excesses in the data, a test of the background-only hypothesis (\( \mu = 0 \)) is performed. 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 pseudoexperiments to derive a correction for the look-elsewhere effect across the mass range tested [52].

In the case of the resolved analysis, the largest deviation from the background-only hypothesis occurs around 900 GeV and is found to have a local significance less than \( 2\sigma \).

In the case of the boosted analysis, the largest local deviation corresponds to a broad data excess in the 3-tag signal region starting at \( m_{3\ell} \sim 1700 \text{ GeV} \). The local significance of this excess is 2.0\( \sigma \) assuming a \( G_{\text{KK}}^{*} \) resonance with \( k/M_{\text{Pl}} = 1 \).

B. 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 \( CL_s \) [53], with a

![Graph](image-url)

FIG. 9. Dijet mass distribution in the \( hh \) signal region for data (points) and background estimate (histograms) in the boosted analysis for events in the (left) 3-tag and (right) 4-tag categories. The expected signal distributions for \( G_{\text{KK}}^{*} \) masses of 1000, 1500 and 1800 GeV are also shown. The uncertainty band includes both the statistical and systematic uncertainties in the background estimate.
The nonresonant search is performed using the resolved or boosted analysis. Using the SM $hh$ nonresonant production as the signal model, the observed 95% C.L. upper limit is $\sigma(pp \to hh \to b\bar{b}b\bar{b}) < 1.22$ pb, a value to be compared with the inclusive SM prediction (as defined in Sec. III) of $\sigma(pp \to hh \to b\bar{b}b\bar{b}) = 11.3^{+0.9}_{-1.0}$ fb.

For the resonant Higgs-boson pair production search, the resolved and boosted analyses offer their best sensitivity in complementary resonance mass regions. The resolved analysis gives a more stringent expected exclusion limit for resonance masses up to (and including) 1100 GeV, while the boosted analysis offers better sensitivity beyond that mass. A simple combination of the separate exclusion limits from the resolved and boosted analyses is used. This is achieved by taking the limit from the analysis with the more stringent expected exclusion at each mass point for each of the signal models.

Figure 10 shows the combined 95% C.L. upper limits for three different resonances: a spin-2 $G_{KK}$ in the bulk RS model with $k/M_{Pl} = 1$ and 2, and a spin-0 narrow-width $H$ boson. For the spin-2 $G_{KK}$ with $k/M_{Pl} = 1$, limits on $\sigma(pp \to G_{KK} \to hh \to b\bar{b}b\bar{b})$ are set in the range between 21 and 73 fb for masses between 600 and 3000 GeV. The corresponding range of limits for the $G_{KK}$ resonance with $k/M_{Pl} = 2$ is 34 to 86 fb. Although no events are observed at masses near 3000 GeV, the observed limit remains about $1\sigma$ weaker than the expected limit due to a substantial low-mass tail in the shape of high-mass resonance signals and the slight data excess observed at high mass. The cross-section limits for resonance masses below 600 GeV weaken substantially due to the lower acceptance times efficiency (see Fig. 2) and the increased level of background. These cross-section upper limits translate into observed (expected) excluded mass ranges of 480–770 (470–735) GeV for $k/M_{Pl} = 1$ and < 965 (< 995) GeV for $k/M_{Pl} = 2$. The cross-section upper limits for the spin-0 narrow-width $H$ boson are similar, with 95% C.L. exclusion limits ranging from 30 to 300 fb in the mass range between 500 and 3000 GeV.

The search sensitivity of this analysis is similar to that achieved at $\sqrt{s} = 8$ TeV with 19.5 fb$^{-1}$ for resonance masses below 1350 GeV but exceeds it above that mass by factors of 1.4 at 1500 GeV and 12 at 2000 GeV. The search has also been extended to resonance masses beyond 2000 GeV, up to 3000 GeV.

FIG. 10. The expected and observed upper limit for $pp \to G_{KK} \to hh \to b\bar{b}b\bar{b}$ in the bulk RS model with (a) $k/M_{Pl} = 1$ and (b) $k/M_{Pl} = 2$, as well as (c) $pp \to H \to hh \to b\bar{b}b\bar{b}$ with fixed $\Gamma_H = 1$ GeV, at the 95% confidence level. The results of the resolved analysis are used up to a mass of 1100 GeV and those of the boosted analysis are used at higher mass where its expected sensitivity is higher. The red curves show the predicted cross sections as a function of resonance mass for the models considered.

The nonresonant search is performed using the resolved analysis, since it has better sensitivity than the boosted analysis. Using the SM $hh$ nonresonant production as the signal model, the observed 95% C.L. upper limit is $\sigma(pp \to hh \to b\bar{b}b\bar{b}) < 1.22$ pb, a value to be compared with the inclusive SM prediction (as defined in Sec. III) of $\sigma(pp \to hh \to b\bar{b}b\bar{b}) = 11.3^{+0.9}_{-1.0}$ fb.

For the resonant Higgs-boson pair production search, the resolved and boosted analyses offer their best sensitivity in complementary resonance mass regions. The resolved analysis gives a more stringent expected exclusion limit for resonance masses up to (and including) 1100 GeV, while the boosted analysis offers better sensitivity beyond that mass. A simple combination of the separate exclusion limits from the resolved and boosted analyses is used. This is achieved by taking the limit from the analysis with the more stringent expected exclusion at each mass point for each of the signal models.

Figure 10 shows the combined 95% C.L. upper limits for three different resonances: a spin-2 $G_{KK}$ in the bulk RS model with $k/M_{Pl} = 1$ and 2, and a spin-0 narrow-width $H$ boson. For the spin-2 $G_{KK}$ with $k/M_{Pl} = 1$, limits on $\sigma(pp \to G_{KK} \to hh \to b\bar{b}b\bar{b})$ are set in the range between 21 and 73 fb for masses between 600 and 3000 GeV. The corresponding range of limits for the $G_{KK}$ resonance with $k/M_{Pl} = 2$ is 34 to 86 fb. Although no events are observed at masses near 3000 GeV, the observed limit remains about $1\sigma$ weaker than the expected limit due to a substantial low-mass tail in the shape of high-mass resonance signals and the slight data excess observed at high mass. The cross-section limits for resonance masses below 600 GeV weaken substantially due to the lower acceptance times efficiency (see Fig. 2) and the increased level of background. These cross-section upper limits translate into observed (expected) excluded mass ranges of 480–770 (470–735) GeV for $k/M_{Pl} = 1$ and < 965 (< 995) GeV for $k/M_{Pl} = 2$. The cross-section upper limits for the spin-0 narrow-width $H$ boson are similar, with 95% C.L. exclusion limits ranging from 30 to 300 fb in the mass range between 500 and 3000 GeV.

The search sensitivity of this analysis is similar to that achieved at $\sqrt{s} = 8$ TeV with 19.5 fb$^{-1}$ for resonance masses below 1350 GeV but exceeds it above that mass by factors of 1.4 at 1500 GeV and 12 at 2000 GeV. The search has also been extended to resonance masses beyond 2000 GeV, up to 3000 GeV.

VII. CONCLUSIONS

A search for both resonant and nonresonant production of pairs of Standard Model Higgs bosons has been carried out in the dominant $b\bar{b}b\bar{b}$ channel with 3.2 fb$^{-1}$ of $pp$ collision data collected by ATLAS during the 2015 run of the LHC at $\sqrt{s} = 13$ TeV. Results are reported for the resolved analysis with each $h \to b\bar{b}$ decay reconstructed as two separate $b$-tagged jets and for the boosted analysis with each $h \to b\bar{b}$ decay reconstructed as a single large-radius jet associated with two small-radius track jets and a
minimum of three $b$-tags for the $hh$ system. No significant data excess is observed above the estimated background consisting mainly of multijet and $t\bar{t}$ events. Upper limits on the production cross section times branching ratio to the $bb\bar{b}\bar{b}$ final state are set for spin-0 and spin-2 resonances with values ranging between 24 and 113 fb (at 95% C.L.) for resonance masses in the range between 600 and 3000 GeV. For nonresonant production, the upper limit is 1.22 pb (at 95% C.L.). The search sensitivity of this analysis exceeds that achieved at $\sqrt{s} = 8$ TeV with 19.5 fb$^{-1}$ for resonance masses above 1350 GeV. Furthermore, the search has been extended to cover the mass range between 2000 and 3000 GeV.

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 DNRSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC NI, Russian 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, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [54].
Search for Pair Production of Higgs Bosons at the ATLAS Experiment at the CERN Large Hadron Collider.


LHC Higgs Cross Section Working Group, Current recommendations for di-Higgs cross-sections, http://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCXSWG.


SEARCH FOR PAIR PRODUCTION OF HIGGS BOSONS IN …

PHYSICAL REVIEW D 94, 052002 (2016)

052002-23
SEARCH FOR PAIR PRODUCTION OF HIGGS BOSONS IN …

PHYSICAL REVIEW D 94, 052002 (2016)

131 State Research Center Institute for High Energy Physics (Protvino), NRC KI, Russia
132 Particle Physics Department, Rutherford Appleton Laboratory, Didcot, United Kingdom
133 INFN Sezione di Roma, Italy
133a Dipartimento di Fisica, Sapienza Università di Roma, Roma, Italy
133b INFN Sezione di Roma Tor Vergata, Italy
134 INFN Sezione di Roma Tor Vergata, Roma, Italy
134a Dipartimento di Fisica, Università di Roma Tor Vergata, Roma, Italy
134b INFN Sezione di Roma Tre, Italy
135 INFN Sezione di Roma Tre, Roma, Italy
135a Facoltà des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies—Université Hassan II, Casablanca, Morocco
135b Centre National de l’Energie des Sciences Techniques Nucleaires, Rabat, Morocco
136 INFN Sezione di Roma Tre, Roma, Italy
136a D Dipartimento di Fisica, Università di Roma Tre, Roma, Italy
137 DSM/IRFU (Institut de Recherches sur les Lois Fondamentales de l’Univers), CEA Saclay (Commissariat à l’Energie Atomique et aux Energies Alternatives), Gif-sur-Yvette, France
138 Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz, California, USA
139 Department of Physics, University of Washington, Seattle, Washington, USA
140 Department of Physics and Astronomy, University of Sheffield, Sheffield, United Kingdom
140a Department of Physics, Shinshu University, Nagano, Japan
141 Fachbereich Physik, Universität Siegen, Siegen, Germany
142 Department of Physics, Simon Fraser University, Burnaby, BC, Canada
143 SLAC National Accelerator Laboratory, Stanford, California, USA
144 Faculty of Mathematics, Physics & Informatics, Comenius University, Bratislava, Slovak Republic
145 Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
145a Department of Physics, University of Cape Town, Cape Town, South Africa
145b Department of Physics, University of Johannesburg, Johannesburg, South Africa
146 School of Physics, University of the Witwatersrand, Johannesburg, South Africa
146a Department of Physics, Stockholm University, Sweden
146b The Oskar Klein Centre, Stockholm, Sweden
147 Physics Department, Royal Institute of Technology, Stockholm, Sweden
148 Departments of Physics & Astronomy and Chemistry, Stony Brook University, Stony Brook, New York, USA
149 Department of Physics and Astronomy, University of Sussex, Brighton, United Kingdom
150 School of Physics, University of Sydney, Sydney, Australia
150a Institute of Physics, Academia Sinica, Taipei, Taiwan
151 Department of Physics, Technion: Israel Institute of Technology, Haifa, Israel
152 Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
153 Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece
154 International Center for Elementary Particle Physics and Department of Physics, The University of Tokyo, Tokyo, Japan
155 Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo, Japan
156 Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
157 Department of Physics, University of Toronto, Toronto, ON, Canada
158 TRIUMF, Vancouver, BC, Canada
159 Department of Physics and Astronomy, York University, Toronto, ON, Canada
160 Faculty of Pure and Applied Sciences, and Center for Integrated Research in Fundamental Science and Engineering, University of Tsukuba, Tsukuba, Japan
161 Department of Physics and Astronomy, Tufts University, Medford, Massachusetts, USA
162 Department of Physics and Astronomy, University of California Irvine, Irvine, California, USA
162a INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine, Italy
163 Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
164 ICTP, Trieste, Italy
164a Dipartimento di Chimica, Fisica e Ambiente, Università di Udine, Udine, Italy
165 Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
166 Department of Physics, University of Illinois, Urbana, Illinois, USA

052002-27