Combination of searches for Higgs boson pairs in \( pp \) collisions at \( \sqrt{s} = 13 \) TeV with the ATLAS detector

The ATLAS Collaboration

1. Introduction

The discovery of the Higgs boson (H) \([1,2]\) at the Large Hadron Collider (LHC) \([3]\) in 2012 has experimentally confirmed the Brout–Englert–Higgs (BEH) mechanism of electroweak symmetry breaking and mass generation \([4–6]\). The BEH mechanism not only predicts the existence of a massive scalar particle, but also requires this scalar particle to couple to itself. Therefore, observing the production of Higgs boson pairs (HH) and measuring the Higgs boson self-coupling \( \lambda_{HHH} \) is a crucial validation of the BEH mechanism. Any deviation from the Standard Model (SM) predictions would open a window to new physics. Moreover, the form of the Higgs field potential, which generates the Higgs boson self-coupling after electroweak symmetry breaking, can have important cosmological implications, involving, for example, predictions for vacuum stability or models in which the Higgs boson acts as the inflaton field \([7–10]\).

In the SM, the gluon–gluon fusion \( pp \rightarrow HH \) process (ggF) accounts for more than 90% of the Higgs boson pair production cross-section, and only this production mode is considered here. It proceeds via two amplitudes: the first (A1) represented by the diagrams (a) and (b), and the second (A2) represented by the diagram (c) in Fig. 1. The interference between these two amplitudes is destructive and yields an overall cross-section of \( \sigma_{ggF}^{SM}(pp \rightarrow HH) = 33.5^{+2.4}_{-2.2} \) fb at \( \sqrt{s} = 13 \) TeV \([11]\), calculated first at next-to-leading order (NLO) in QCD with the heavy top-quark approximation \([12]\), then numerically with full top-quark mass dependence \([13]\) (confirmed later in Ref. \([14]\)) and analytically computed with some approximation in Ref. \([15]\)) corrected at next-to-next-to-leading order (NNLO) \([16]\) in QCD matched with next-to-next-to-leading logarithmic (NNLL) resummation in the heavy top-quark limit \([17,18]\). The Higgs boson mass used in these calculations and for all results in this paper is \( m_H = 125.09 \) GeV \([19]\). Beyond-the-Standard-Model (BSM) scenarios can bring substantial enhancement of this cross-section by modifying the relative sign of \( A_1 \) and \( A_2 \), and by increasing \( A_2 \). The \( A_2 \) amplitude is proportional to the Higgs boson self-coupling \( \lambda_{HHH} \). The Higgs boson self-coupling modifier due to BSM scenarios is defined as \( k_2 = \lambda_{HHH}/\lambda_{SM} \). In this analysis, all other Higgs boson couplings are assumed to have SM values. Indirect limits on \( k_2 \) have been obtained using the measurements of single Higgs boson production and decay \([20]\) and electroweak precision observables \([21,22]\), constraining \( k_2 \) to the range \(-8 < k_2 < 14\) at 95% confidence level (CL). The Higgs boson self-coupling is discussed in the context of BSM models in Refs. \([22,23]\).

Several BSM models also predict the existence of heavy particles decaying into a pair of Higgs bosons. Two-Higgs-Doublet Models \([24]\), models inspired by the Minimal Supersymmetric Standard Model (MSSM) like habemus MSSM (hMSSM) \([25–28]\), and Elec-
troweak Singlet Models (EWK-singlet) [11,29–31] predict, in addition to the Higgs boson, a second, heavier, CP-even scalar that can decay into two SM Higgs bosons. In the EWK-singlet model, the scalar states are mixed, with a mixing angle $\alpha$. The ratio of the vacuum expectation value of the additional singlet to that of the SM Higgs doublet, $\tan\beta$, is a free parameter. In the hMSSM, the CP-even states also mix, and the model’s phenomenology can be described by the mass $(m_t)$ of a third, CP-odd, resonance and the ratio of the vacuum expectation values of the two Higgs doublets, $\tan\beta$. Alternatively, the Higgs boson pair can be produced resonantly through the decay of a spin-2 Kaluza–Klein (KK) graviton, as predicted in the Randall–Sundrum (RS) model of warped extra dimensions [32]. A schematic diagram for production of a heavy resonance followed by its decay into a Higgs boson pair is shown in Fig. 1(d).

This letter presents a combination of results from searches for both non-resonant and resonant Higgs boson pair production in proton–proton (pp) collisions at $\sqrt{s} = 13$ TeV. The data were collected with the ATLAS detector [33–35] and correspond to an integrated luminosity of up to 36.1 fb$^{-1}$. The combination includes all published ATLAS $HH$ search analyses using $\sqrt{s} = 13$ TeV data, namely those studying the final states $bbbb$ [36], $bbW^+W^-$ [37], $b\bar{b} \tau^+\tau^-$ [38], $W^+W^-W^+W^-$ [39], $b\bar{b}\gamma\gamma$ [40] and $W^+W^-\gamma\gamma$ [41].

Previous combinations of searches for $HH$ pair production were performed at $\sqrt{s} = 8$ TeV by the ATLAS experiment [42] and at $\sqrt{s} = 13$ TeV by the CMS experiment [43] combining the final states $bb\bar{b}$ [44–47], $bbVV$ [48], $b\bar{b}\tau^+\tau^-$ [49] and $b\bar{b}\gamma\gamma$ [50].

2. Analysis description

The analysis strategies for each of the final states considered in this letter are summarised below.

- The $bb\bar{b}$ analysis is performed using four anti-$k_t$ jets reconstructed with a radius parameter $R = 0.4$ [51,52] (resolved analysis) or two large-$R$ jets with $R = 1.0$ (boosted analysis). The dataset of the resolved analysis is split according to the years 2015 and 2016, and then statistically combined taking into account the different trigger algorithms used in 2015 and 2016. In part of the 2016 data period, inefficiencies in the online vertex reconstruction affected $b$-jet triggers that were used in the resolved analysis, reducing the total available integrated luminosity to 27.5 fb$^{-1}$. The boosted analysis searches for two large-$R$ jets containing the $b$-quark pairs from the decays of the two Higgs bosons. The large-$R$ jets are identified as originating from $b$-quarks using a $b$-tagging algorithm applied to $R = 0.2$ track-jets [53] associated with the large-$R$ jet [54]. The analysis is divided into three categories: the first category selects events in which each of the two large-$R$ jets has one $b$-tagged track-jet; the second category requires that one large-$R$ jet contains two $b$-tagged track-jets and the other large-$R$ jet contains one $b$-tagged track-jet; the third category requires that both large-$R$ jets contain two $b$-tagged track-jets.

- The $bbW^+W^-$ analysis looks for the $WW \rightarrow \ell\nu qq$ decay channel, where $\ell$ is an electron or muon, and $q$ is a $u, d, s, c$ quark or anti-quark. The $bb$ pair is selected from two $R = 0.4$ jets (resolved analysis) or one $R = 1.0$ large-$R$ jet (boosted analysis), while the jets from the $W$ decay are reconstructed with $R = 0.4$ jets. The resolved analysis is used in the SM $HH$ search, in the search for a scalar resonance with a mass between 500 and 1400 GeV, and in the search for a KK graviton in the mass range 500 to 800 GeV. The boosted analysis looks for scalar resonances in the mass range 1400 to 3000 GeV and for KK gravitons between 800 and 3000 GeV. The resolved and boosted analyses each use one category. The two analyses are not statistically combined due to a significant overlap between the two signal regions.

- The $b\bar{b}\tau^+\tau^-$ analysis looks for final states with two $R = 0.4$ $b$-tagged jets and two $\tau$-leptons. One of the two $\tau$-leptons of the $\tau^+\tau^-$ pair is required to decay hadronically, while the other decays either hadronically ($\tau_{had}^{tag}$) or leptonically ($\tau_{lep}^{tag}$). In the $\tau_{lep}^{tag}$ channel, events are triggered by single lepton triggers (SLT), requiring an electron or a muon in the final state, or by the coincidence of a lepton trigger with a hadronic $\tau$ trigger (LTT). In the $\tau_{had}^{tag}$ channel, events are triggered by single hadronic $\tau$ triggers (STT) or double hadronic $\tau$ triggers (DTT). The analysis is divided into three categories: one selects $\tau_{had}^{tag}$ events, a second selects $\tau_{lep}^{tag}$ events triggered by the SLT, and a third selects $\tau_{lep}^{had}$ events triggered by the LTT. The $\tau_{had}^{tag}$ and the SLT $\tau_{lep}^{had}$ categories are used for all model interpretations, while the LTT
Table 1

Summary of the main characteristics of the analyses used in the non-resonant and resonant searches. The resonant analysis characteristics are indicated between square brackets. ‘BS(HH → 4ℓjj)’ indicates the branching fraction of the HH pair where one H decays into x̄ and the other decays into yj. The branching fraction values are taken from Ref. [11] for a Higgs boson mass $m_H = 125.09$ GeV. $\mathcal{L}_{\text{int}}$ indicates the integrated luminosity of the dataset used in the analysis. ‘Categories’ indicates the number of signal categories. ‘Discriminant’ indicates the distribution used in the final limit-setting fit (‘c.e.’ stands for counting events and that a simple event counting was used in the final fit rather than a distribution shape). ‘Model’ indicates which models each analysis tested: NR stands for SM HH signal model, S for a spin-0 scalar model, and G for a KK graviton model. $m_{\text{HH}}$ gives the probed mass range for the resonant search. “Ref.” reports the reference to the individual final state papers.

<table>
<thead>
<tr>
<th>Process</th>
<th>$m_{\text{HH}}$ (GeV)</th>
<th>$\mathcal{L}_{\text{int}}$ (fb)$^{-1}$</th>
<th>Categories</th>
<th>Discriminant</th>
<th>Model</th>
<th>$m_{\text{HH}}$ (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bbb</td>
<td>0.34</td>
<td>0.25</td>
<td>0.073</td>
<td>0.046</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td>bbbW+W−</td>
<td>2.6</td>
<td>1.0</td>
<td>3.0</td>
<td>1.0</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td>bbr±τ−</td>
<td>2.6</td>
<td>1.0</td>
<td>3.0</td>
<td>1.0</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td>W+W−</td>
<td>2.6</td>
<td>1.0</td>
<td>3.0</td>
<td>1.0</td>
<td>2.6</td>
<td>1.0</td>
</tr>
<tr>
<td>W+γγ</td>
<td>2.6</td>
<td>1.0</td>
<td>3.0</td>
<td>1.0</td>
<td>2.6</td>
<td>1.0</td>
</tr>
</tbody>
</table>

The $\tau_{\text{lep}}$ category is used in the SM HH search (excluding the $k_s$ analysis) and in resonant searches up to a mass of 800 GeV.

- The $W^+W^−W^+W^−$ analysis looks for channels with leptonic and/or hadronic $W$ decays. Three channels are identified: $\ell\nu\ell\nu q\bar{q}$, $\ell\nu\ell\nu q\bar{q}$, $\ell\nu\ell\nu q\bar{q}$, and $\ell\nu\ell\nu q\bar{q}$, with $\ell$ being an electron or muon, $q$ a quark, and $\nu$ a neutrino. The $q$ momentum is reconstructed from the $R = 0.4$ jets. In order to suppress $Z$ + jets and $t\bar{t}$ background, dilepton events are required to have two leptons of the same charge. Events are categorised according to the lepton flavour ($ee$, $e\mu$ and $\mu\mu$). Three-lepton events are selected if the sum of the lepton charges is $\pm 1$. They are divided into two categories according to the number of same-flavour, opposite-charge (SFOS) lepton pairs; one category selects zero SFOS lepton pairs and a second category selects one or two SFOS lepton pairs. Four-lepton events are categorised according to the number of SFOS lepton pairs and the invariant mass ($m_{4l}$) of the four-lepton system. Four categories are defined, requiring that the number of SFOS lepton pairs is less than two or equal to two, and $m_{4l}$ is smaller or larger than 180 GeV. A total of nine categories are fit simultaneously in the searches for both non-resonant and resonant HH production.

- The $b\bar{b}γγ$ analysis searches for a $HH$ pair decaying into $b\bar{b}$ and $γγ$. Two high-$p_T$ isotopically charged photons are required to have $E_T/m_{γγ} > 0.35$ and 0.25 respectively. The events are then analysed using two selections: a ‘loose selection’ requiring a jet with $p_T > 40$ GeV and a second jet with $p_T > 25$ GeV, and a ‘tight selection’ where the two jets are required to have $p_T > 100$ and 30 GeV. All jets have a radius parameter $R = 0.4$. Both selections are subdivided into two categories requiring one $b$-tagged jet or two $b$-tagged jets. The tight selection is used in the SM HH search and in the search for resonances with masses higher than 500 GeV, while the loose selection is used in the $k_s$ analysis and in the search for resonances with masses smaller than 500 GeV. The analysis is therefore divided into four categories, but only two of them are simultaneously fit to extract each result.

- The $W^+γγ$ analysis searches for a $HH$ pair decaying into $γγ$ and $WW$. The analysis uses the same photon selection as the $b\bar{b}γγ$ channel and looks for one $W$ decaying leptonically and a second $W$ decaying hadronically ($WW → τνq̄q$). The hadronic $W$ decay is reconstructed from $R = 0.4$ jets. Only one category is used in the searches for both non-resonant and resonant HH production.

A summary of the main analysis characteristics is given in Table 1. All analyses impose a series of sequential requirements on kinematic variables to select signal events and suppress backgrounds. The $bbr±τ−$ analysis uses a boosted decision tree (BDT) distribution as the final discriminant for both the non-resonant and resonant searches. For the resonant searches, the $bbW+W^−$ and $bbγγ$ analyses use the HH invariant mass ($m_{\text{HH}}$) as the final discriminant, the $W^+W^−γγ$ analysis uses the $γγ$ invariant mass ($m_{γγ}$), while the $W^+W^−W^+W^−$ uses simple event counting. For the SM HH search, the $bbW+W^−$ and $bbγγ$ analyses use the $m_{\text{HH}}$ distribution as a discriminant, profiling from the difference between the shapes of the signal and the dominant multi-jet background. The $bbrγγ$ and $W^+W^−γγ$ analyses fit the $m_{γγ}$ distribution to extract both the signal yield and the background expectation, while the $bbW^+W^−$ and $W^+W^−W^+W^−$ analyses use event counting.

3. Statistical treatment

The statistical interpretation of the combined search results is based on a simultaneous fit to the data for the cross-section of the signal process and nuisance parameters that encode statistical and systematic uncertainties, using the CLs approach [56]. The asymptotic approximation [57] is used in the analysis of all final states and their combination.

All signal regions considered in the simultaneous fit are either orthogonal by construction or have negligible overlap. The overlap due to object misidentification between $bbrγγ$ and $bbr±τ−$, and between $W^+W^−γγ$ and $W^+W^−W^+W^−$, which are not orthogonal by construction, is evaluated by running the signal and data samples from each channel through the analysis selection of each other channel. Less than 0.1% of simulated signal events overlap between analyses, and no overlap is found in data. There is some irreducible contamination from $bbW^+W^−$ and $bbzz$ events with $τs$ in the final state passing the $bbr±τ−$ selection. This contamination is less than 8% of the $bbr±τ−$ selected events, and it is not taken into account in the $bbr±τ−$ analysis, note that including this contribution would increase the analysis sensitivity therefore the results obtained here are slightly conservative. The detector systematic uncertainties, such as those in jet reconstruction, b-jet tagging, electron, muon and photon reconstruction and identification, as well as the uncertainty on the integrated luminosity [58], are correlated across all final states. Uncertainties on the signal acceptance derived by varying the renormalisation and factorisation scales, the parton distribution functions and the parton shower are correlated too. Theoretical and modelling systematic uncertainties of the backgrounds derived using simulated events are not correlated across different analyses because the overlap among their contributions to the different analyses is negligible.
The ATLAS Collaboration / Physics Letters B 800 (2020) 135103

4. Combination of results on non-resonant Higgs boson pair production

The SM HH analyses use signal samples generated at next-to-leading order (NLO) in QCD with MADGRAPH5_AMC@NLO [59] using the CT10 NLO parton distribution function (PDF) set [60]. Parton showers and hadronisation were simulated with HERWIG++ [61] using parameter values from the UE-EE-5-CTEQ6L1 tune [62]. The so-called FTAprox method [63] is applied in the event generation to include finite top-quark mass effects in the real-radiation NLO corrections. The virtual-loop corrections are realised with Higgs effective field theory (HEFT) assuming infinite top-quark mass. The generated events are then corrected with a generator level bin-by-bin reweighting of the m_{HH} distribution, which is calculated with finite top-quark mass in full NLO corrections [13]. The branching fractions of the Higgs boson are assumed to be equal to the SM predictions [11]. For the SM HH search, the limits are extracted for the cross-section σ_{ggF}(pp → HH) of HH production and are normalised by the SM HH cross-section σ_{SM}^{ggF}(pp → HH). The limits are determined assuming that all kinematic properties of the HH pair are those predicted by the SM, particularly the m_{HH} distribution, and only the total ggF production cross-section, σ_{ggF}(pp → HH), is allowed to deviate from its SM value. The theoretical uncertainties of σ_{SM}^{ggF} are less than 10% [11] and are not included in the fit results.

The upper limits at 95% CL on the cross-section of the ggF Higgs boson pair production normalised to σ_{SM}^{ggF} are shown in Fig. 2 for the individual final states and their combination. The upper limit for each final state is obtained from a fit with minimal changes from previously published results. The changes include an update of the ggF Higgs boson pair production cross-section from 33.4 fb to 33.5 fb for all final states. Additionally, the bb^+bb^- final state included theoretical uncertainties on the ggF inclusive cross-section, σ_{SM}^{ggF}, which are not considered in the present treatment, and the bbyγ final state is updated to use an asymptotic approximation to calculate the observed limit instead of the pseudo-experiment method used for its publication. This results in a 10% change in the observed limit of bbγγ. Moreover, the impact of the asymptotic approximation on all final states combined is found to be 5%.

The combined observed (expected) upper limit on the SM HH production is 6.9 (10) × σ_{SM}^{ggF}(pp → HH). The expected limit is similar to the CMS result of 12.8 × σ_{SM}^{ggF}(pp → HH). The observed limit is more stringent for the ATLAS result than the CMS result of 22.2 × σ_{SM}^{ggF}(pp → HH) because the three leading channels (bbbb, bb^+bb^- and bbyγ) have a data deficit in ATLAS and an excess in CMS [43], remaining however within the two σ uncertainty interval. Detailed comparisons can be found in Ref. [64].

The impact of the systematic uncertainties has been evaluated by recomputing the limit without their inclusion. The limit is then reduced by 13% when removing all of them. The main sources of systematic uncertainty are the modelling of the backgrounds, the statistical uncertainty of simulated events and the τ-lepton reconstruction and identification. When removed the limit reduces by 5%, 3% and 2%, respectively.

5. Constraints on the Higgs boson self-coupling

The results in Fig. 2 show that the sensitivity of the SM HH search is driven by the final states bbbb, bb^+bb^- and bbyγ. These final states are used to set constraints on the Higgs boson self-coupling modifer κ λ = g_{HHH}/g_{HHH}^{SM}. After setting all couplings to fermions and bosons to their SM values, a scan of the self-coupling modifier κ λ is performed. The κ λ factor affects both the production cross-section and the kinematic distributions of the Higgs boson pairs, by modifying the A_2 production amplitude. It can also affect the Higgs boson branching fractions due to NLO electroweak corrections [20], but this dependence is neglected in the following.

The signal used in the κ λ fit was simulated according to the following procedure. For each value of κ λ, the m_{HH} spectrum is computed at the generator-level, using the leading-order (LO) version of MADGRAPH5_AMC@NLO [59] with the NNPDF 2.3 LO [65] PDF set, together with PYTHIA 8.2 [66] for the showering model using the A14 tune [67]. Because only one amplitude of Higgs boson pair production depends on κ λ, linear combinations of three LO samples generated with different values of κ λ are sufficient to make predictions for any value of κ λ. Binned ratios of the m_{HH} distributions to the SM distribution are computed for all κ λ values and then used to reweight the events of NLO SM HH signal samples, generated using the full detector simulation. This procedure is validated by comparing kinematic distributions obtained with the reweighting procedure applied to the LO SM sample and LO samples generated with the actual κ λ values set in the event generator. The two sets of distributions are found to be in agreement. This procedure assumes that higher order QCD corrections on the differential cross-section as a function of m_{HH} are independent of κ λ. The reweighted NLO signal sample is used to compute the signal acceptance and the kinematic distributions for different values of κ λ.

This letter presents κ λ results for the first time in the ATLAS bbbb and bb^+bb^- final states and incorporates the previously published result for the bbyγ final state. The κ λ analyses closely follow the SM HH search, with some exceptions which are discussed below for each final state.

- In the bbbb final state, the same analysis selection and final discriminant are used in the κ λ-scan analysis and in the SM HH search. The distribution of the final discriminant m_{HH} is shown in Fig. 3(a), where, with the exception of a small excess in the region below 300 GeV [36] and a small deficit in the 500-600 GeV region, good agreement between data and the
The shape of the $m_{HH}$ distribution has a strong dependence on $\kappa_{\lambda}$, and the signal acceptance varies by a factor 2.5 over the probed range of $\kappa_{\lambda}$-values ($-20 \leq \kappa_{\lambda} \leq 20$) shown in Fig. 4(a). The two effects together determine how the exclusion limits on the cross-section of the $HH$ production vary as a function of $\kappa_{\lambda}$.

- In the $b\bar{b}t^+t^-$ final state, as in the SM $HH$ search, both $t_{\text{lep}}t_{\text{had}}$ and $t_{\text{had}}t_{\text{had}}$ events are used. In contrast with the SM $HH$ search, LTT $t_{\text{lep}}t_{\text{had}}$ events (see Section 2) are not used given their negligible contribution. The SM $HH$ search and the $\kappa_{\lambda}$-scan analysis use the same sets of variables to build BDT discriminants. For the $\kappa_{\lambda}$-scan the BDTs are retrained using the NLO SM signal sample reweighted with $\kappa_{\lambda} = 20$, ensuring good sensitivity over the whole range of probed $\kappa_{\lambda}$-values.

- The BDT score distributions are used in the fit to compute the final results. The shape of the $b\bar{b}t^+t^-$ BDT distributions does not show a $\kappa_{\lambda}$ dependence as strong as in the $b\bar{b}\bar{b}\bar{b}$ final state, as can be seen in Fig. 3. The sensitivity of this analysis is instead affected by a variation in the signal acceptance by up to a factor of three over the probed range of $\kappa_{\lambda}$-values, as shown in Fig. 4(a).

- In the $b\bar{b}\gamma\gamma$ final state, the loose selection is used in the $\kappa_{\lambda}$-scan analysis because the average transverse momentum of the Higgs bosons is lower at large values of $\kappa_{\lambda}$, where $|A_{\lambda}^2|^2$ dominates the production cross-section. As in the SM $HH$ search, the statistical analysis is performed using the $m_{\gamma\gamma}$ distribution, which does not depend on $\kappa_{\lambda}$. The signal acceptance varies by about 30% over the probed range of $\kappa_{\lambda}$-values, as shown in Fig. 4(a). In the previously published analysis [40], LO samples were used for the computation of the signal acceptance, while in this paper the NLO reweighted samples are used, as described above.
The signal acceptance times efficiency as a function of $k_2$, is shown in Fig. 4(a). Given that, for each final state, the same selection was applied over the full scanned $k_2$ range, the shape of the acceptance times efficiency curve is determined by the variation of the event kinematics as a function of $k_2$. For high values of $|k_2|$, the $A_2$ term dominates the total amplitude, causing a softer $m_{HH}$ spectrum, and thus a lower acceptance times efficiency. Around $k_2 = 2.4$ the interference between $A_1$ and $A_2$ amplitudes is maximal, producing the hardest $m_{HH}$ spectrum and, consequently, the highest signal acceptance times efficiency.

In each analysis, and in their combination, the 95% CL upper limit on the $\sigma_{ggF}(pp \rightarrow HH)$ cross-section is computed for different values of $k_2$. The results are shown in Fig. 4(b). The theoretical $\sigma_{ggF}(pp \rightarrow HH)$ cross-section as a function of $k_2$ is overlaid in the figure. It is computed by multiplying the $HH$ SM cross-section $\sigma_{SM}(pp \rightarrow HH)$ by the ratio $R(k_2)$ of the $pp \rightarrow HH$ cross-section computed at $k_2 = 2.4$, $\sigma_{SM}(pp \rightarrow HH)$ to the same quantity $\sigma_{ggF}(pp \rightarrow HH)$ computed at $k_2 = 1$. The $R(k_2)$ factor is computed at NNLO+NNLL with the infinite top-quark mass approximation [68]. The resulting observed (expected) confidence interval at 95% CL for $k_2$ is: $-5.0 < k_2 < 12.0$ ($-5.8 < k_2 < 12.0$).

In Fig. 4(b) the shape of the upper-limit curves approximately follows the inverse of the signal acceptance shown in Fig. 4(a). In the $bb\bar{b}b$ analysis, the observed limits are more stringent than the expected limits at low values of $k_2$. For these $k_2$ values the signal $m_{HH}$ distributions have significant populations in the region 500-600 GeV, where the data deficit sits, as explained above. For larger values of $k_2$ the $m_{HH}$ distribution is shifted to lower $m_{HH}$ values, and thus the excess in data below 300 GeV leads to the observed limits being less stringent than expected. In the $bb\bar{b}\tau^+\tau^-$ final state the observed limits are more stringent than the expected limits over the whole range of $k_2$, due to a deficit of data relative to the background predictions at high values of the BDT score. The $bb\gamma\gamma$ limit shows a weaker dependence on $k_2$ than the $bb\bar{b}\tau^+\tau^-$ limits because the $bb\gamma\gamma$ acceptance varies less as function of $k_2$.

The 95% CL allowed $k_2$ intervals are given in Table 2. The systematic uncertainties weaken the $k_2$ limits by less than 10% relative to those obtained with only statistical uncertainties. The final state least (most) affected by systematic uncertainties is $bb\gamma\gamma$ ($bb\bar{b}b$). The Higgs boson branching fraction depends on $k_2$, due to $NLO$ electroweak corrections [20]. This dependence is neglected in the present treatment, but its overall impact on the allowed $k_2$ interval is evaluated to be no more than 7%. Theory uncertainties on the signal cross section shown in Fig. 4(b) are not taken into account when computing the $k_2$ limits in Table 2, they affect the limit by less than 8%.

### Table 2

<table>
<thead>
<tr>
<th>Final state</th>
<th>Allowed $k_2$ interval at 95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs.</td>
</tr>
<tr>
<td>$bb\bar{b}b$</td>
<td>$-10.9 - 20.1$</td>
</tr>
<tr>
<td>$bb\tau^+\tau^-$</td>
<td>$-7.4 - 15.7$</td>
</tr>
<tr>
<td>Combination</td>
<td>$-5.0 - 12.0$</td>
</tr>
</tbody>
</table>

### 6. Combination of results for resonant Higgs boson pair production

The resonance decaying into a pair of Higgs bosons is assumed to be either a heavy spin-0 scalar particle, $S$, with a narrow width or a spin-2 KK graviton, $G_{KK}$.

The search for the heavy scalar particle $S$ is performed with all six final states included in this combination. With the exception of $bb\tau^+\tau^-$ and $bb\bar{b}b$, all signal samples were simulated at NLO with MAdGRAPH5_AMC@NLO using the CT10 PDF set. The matrix-element generator was interfaced to HERWIG++ with the UE-EE-5-CTEQ6L1 tune. The $bb\tau^+\tau^-$ final state uses an LO model generated with MAdGRAPH5_AMC@NLO using the NNPDF2.3 LO PDF set interfaced to PYTHIA 8.2 with the A14 tune, while the $bb\bar{b}b$ final state uses the same LO event generator but interfaced to HERWIG++ with the UE-EE-5-CTEQ6L1 tune.

The scalar resonance search is performed in the mass range 260–3000 GeV, and within this range no statistically significant excess is observed. In the combination, the largest observed deviation from the background expectation is $1\sigma$ for the search mass range. The combined upper limit on the cross-section is shown as a func-
tion of the resonance mass in Fig. 5(a). Systematic uncertainties have a sizeable effect on the upper limits depending on the probed resonance mass. The total impact of systematics or the impact of a single systematic uncertainty has been evaluated by computing the percentage reduction of the upper limit obtained by removing all systematic uncertainties or a particular source. Overall the systematic uncertainties affect the limit by 12% (11%) for a resonance mass of 1 (3) TeV. Among them, the largest systematic uncertainties are due to the modelling of the backgrounds, impacting the upper limit by 7% (9%) at 1 (3) TeV. The second leading systematic uncertainty comes from $b$-tagging, which affects the upper limit by 2% at 1 TeV, but its impact is negligible at 3 TeV where relative background and statistical uncertainties increase significantly. At 3 TeV the second leading systematic uncertainty is related to the jet energy scale and resolution, changing the limit by 2%. Interpretations in specific spin-0 BSM models are provided in Section 7.

The search for a spin-2 KK graviton is performed with the $bb\bar{b}$, $bbW^+W^-$ and $bb\tau^+\tau^-$ final states only. Gravitons were simulated using an LO model in MADGRAPH5_aMC@NLO with the NNPDF 2.3 LO PDF set interfaced to PYTHIA 8.2 with the A14 tune. The resonance width changes with the graviton mass and depends on the parameter $k/\sqrt{M_{\text{Pl}}}$. The curvature of the warped extra dimension in the bulk RS model and $M_{\text{Pl}} = 2.4 \times 10^{18}$ GeV is the effective four-dimensional Planck mass. The search is performed for models with $k/\sqrt{M_{\text{Pl}}}$ equal to 1 and 2. For $k/\sqrt{M_{\text{Pl}}} = 1$ (2), the width ranges from 3% (11%) for a 0.3 TeV graviton mass to 6% (25%) for a 3 TeV graviton mass.

The upper limits in the $G_{KK}$ search are shown as a function of the resonance mass in Figs. 5(b) and 5(c) for $k/\sqrt{M_{\text{Pl}}}$ equal to 1 and 2, respectively. In the combination, the largest observed deviation from the background expectation is $1.5\sigma$ (0.7$\sigma$) for the search mass range with $k/\sqrt{M_{\text{Pl}}} = 1$ (2). Exclusion ranges on the KK graviton mass are obtained by comparing the upper limit with the production cross section calculated at LO. In the case of $k/\sqrt{M_{\text{Pl}}} = 1$, the bulk RS model is excluded at 95% CL in the graviton mass range from 310 GeV to 1380 GeV. In the case of $k/\sqrt{M_{\text{Pl}}} = 2$, the model is excluded at 95% CL for graviton masses from 260 GeV, where the scan starts, to 1760 GeV.

The impact of the systematic uncertainties on the upper limits on $G_{KK}$ has a small dependence on the resonance mass. It is $\sim 20\%$ over the whole mass range for $k/\sqrt{M_{\text{Pl}}} = 1$, and 29% (25%) at a mass of 1 TeV (3 TeV) for $k/\sqrt{M_{\text{Pl}}} = 2$. The largest systematic uncertainties are from the modelling of the backgrounds, affecting the limit by 11% (15%) at 1 TeV (3 TeV) for $k/\sqrt{M_{\text{Pl}}} = 1$ and 16% (21%) at 1 TeV (3 TeV) for $k/\sqrt{M_{\text{Pl}}} = 2$. For $k/\sqrt{M_{\text{Pl}}} = 1$, the subleading systematic uncertainties come from $b$-tagging at low $G_{KK}$ mass, that affect the limit by 3%, and from jet energy scale and resolution at high mass, that affect the upper limit by 2% (3%) at 1 TeV (3 TeV). For...
The calculated states are presented in Fig. 7. Direct constraints from SM Higgs coupling measurements [11,26–28,69] are shown with horizontal lines. The dotted lines indicate the separation between regions where the resonance width is larger than 2% and 5% of the resonance mass. The hatch-marked region corresponds to regions that cannot be excluded because the width of the resonance exceeds 10% of the resonance mass, corresponding to the maximum of the experimental mass resolution among all analysed final states.

**Fig. 7.** Excluded regions in (a) the EWK-singlet model for \( m_S = 260 \text{ GeV} \) and (b) the hMSSM model using the experimental upper limits obtained in the spin-0 resonance searches. In the EWK-singlet exclusion, the indirect constraints from SM Higgs coupling measurements [74] are shown with vertical lines.

\( k/M_H = 2 \), subleading systematic uncertainties are from jet energy scale and resolution, impacting the upper limits by 5% at 1 TeV and 4% at 3 TeV. The systematic uncertainties affect upper limits more for \( k/M_H = 2 \) than for \( k/M_H = 1 \), because the natural width of the signal graviton is four times larger with \( k/M_H = 2 \).

### 7. Constraints on the hMSSM and EWK-singlet models

Exclusion limits are also presented for two specific models, namely the EWK-singlet model [11,29–31] and the hMSSM model [11,26–28,69]. The sensitivity of the \( b\bar{b}W^+W^- \), \( W^+W^−W^+W^- \) and \( W^+W^−γγ \) final states to these models is negligible, so the presented results combine only the \( b\bar{b}b\bar{b} \), \( b\bar{b}τ⁺τ⁻ \) and \( b\bar{b}γγ \) final states.

For the EWK-singlet model, the experimental limits on the spin-0 resonance (as reported in Section 6) are interpreted as constraints in the \( m_3−\sin α \) plane (where \( m_3 \) is the resonance mass) for \( \tan β = 1 \) and \( \tan β = 2 \), shown in Fig. 6(a) and Fig. 6(b) respectively. The expected cross-section for each point in the parameter space is obtained by scaling the heavy Higgs cross-section calculated at NNLO+NNLL [11] with singlet coupling modifiers. The branching fractions are computed with sHDECAY [70]. In this model, the width of the heavy scalar can be large in some regions of the parameter space. Due to the use of narrow-width signal models in the event generation, results presented here are valid only in regions of the model parameter-space where the resonance width (\( Γ_5 \)) is smaller than the experimental resolution at the resonance mass. This holds when \( Γ_5/m_S < 2\% \) for \( b\bar{b}γγ \), \( Γ_5/m_S < 5\% \) for \( b\bar{b}b\bar{b} \) and \( Γ_5/m_S < 10\% \) for \( b\bar{b}τ⁺τ⁻ \). Therefore, the excluded region in the plot is obtained by combining the three final states for \( Γ_5/m_S < 2\% \), by combining the \( b\bar{b}b\bar{b} \) and \( b\bar{b}τ⁺τ⁻ \) final states for \( 2\% < Γ_5/m_S < 5\% \), and using only \( b\bar{b}τ⁺τ⁻ \) for \( 5\% < Γ_5/m_S < 10\% \). The hatched region shows points where \( Γ_5/m_S ≥ 10\% \), where no exclusion can be provided. Fig. 7(a) shows limits for the \( \sin α \), \( \tan β \) parameter space for \( m_3 = 260 \text{ GeV} \) where, due to the limited decay phase space, the resonance width is narrow in a wide region of the parameter space.

The experimental limits on a spin-0 resonance are also interpreted as constraints in the \( m_A−\tan β \) plane of the hMSSM model in Fig. 7(b). The expected cross-section for each point in the parameter space is obtained using the gluon-gluon fusion cross-section from SUSHI 1.5.0 [71,72] and the branching fractions computed with HDECAY 6.4.2 [73]. The excluded region is more than doubled along \( \tan β \) relative to the previous combined results in Ref. [42] at 8 TeV, and excludes values of \( m_A \) from 190 GeV to 560 GeV depending on \( \tan β \).
kink at low tan β and high mχ values is caused by removing the bbγγ final state from the combination in the region where the predicted width of the heavy CP-even Higgs boson is larger than the experimental resolution on mχ in the bbγγ analysis.

8. Conclusion

A statistical combination of six final states bbbb, bbW+W−, bbτ+τ−, W+W−W−, bbγγ and W+W−γγγ, is presented for the search for non-resonant and resonant production of Higgs boson pairs. These searches use up to 36.1 fb−1 of proton–proton collision data at 13 TeV recorded with the ATLAS detector at the LHC.

In both resonant and non-resonant searches, no statistically significant excess of events above the Standard Model predictions is found. For the Standard Model HH production mode, the observed (expected) 95% confidence level upper limit on the gluon–gluon fusion pp → HH cross-section is 6.9 (10) times the Standard Model prediction. The expected limit is comparable to the CMS result, while the observed limit is significantly stronger than CMS’s due to a data deficit compared to expected background in ATLAS and an excess in CMS. For the resonant case, upper limits are set on the production cross-section of heavy spin-0 and spin-2 resonances decaying into pairs of Higgs bosons in the mass range 260–3000 GeV.

Upper limits on the pp → HH cross-section are also computed as a function of the Higgs boson self-coupling modifier κS = λHHHH/λHHH, by combining the bbbb, bbτ+τ− and bbγγ final states. The combination excludes κS values outside the range −5.0 < κS < 12.0 (−5.8 < κS < 12.0) at 95% confidence level in observation (expectation). The three final states are also combined to constrain the Electroweak Singlet Model in the (mχ, sin α) and the (sin α, tan β) parameter spaces and the habemus Minimal Supersymmetric Standard Model in the (mχ, tan β) parameter space.

Acknowledgements

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

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MINE/IFA, Romania; MES of Russia and NRCC, 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, Canarie, CRC and Compute Canada, Canada; COST, EEC, ERDF, Horizon 2020, and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, France; DFG and AVH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-funded by EU-ESF and the Greek NRFF, Greece; BSF-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya, 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. [75].

References


1 All results are available in digital format on HEpdata at the following link: https://www.hepdata.net/record/90521.


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

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

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

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

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

Also at TRIUMF, Vancouver, BC; Canada.

Also at Universita di Napoli Parthenope, Napoli; Italy.

* Deceased.