Search for Resonant and Nonresonant Higgs Boson Pair Production in the $b\bar{b}\tau^+\tau^-$ Decay Channel in $pp$ Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector

M. Aaboud et al. *
(ATLAS Collaboration)

(Received 2 August 2018; published 7 November 2018)

A search for resonant and nonresonant pair production of Higgs bosons in the $b\bar{b}\tau^+\tau^-$ final state is presented. The search uses 36.1 fb$^{-1}$ of $pp$ collision data with $\sqrt{s} = 13$ TeV recorded by the ATLAS experiment at the LHC in 2015 and 2016. Decays of the $\tau$-lepton pairs with at least one $\tau$ lepton decaying to final states with hadrons and a neutrino are considered. No significant excess above the expected background is observed in the data. The cross-section times branching ratio for nonresonant Higgs boson pair production is constrained to be less than 30.9 fb, 12.7 times the standard model expectation, at 95% confidence level. The data are also analyzed to probe resonant Higgs boson pair production, constraining a model with an extended Higgs sector based on two doublets and a Randall-Sundrum bulk graviton model. Upper limits are placed on the resonant Higgs boson pair production cross-section times branching ratio, excluding resonances $X$ in the mass range 305 GeV $< m_X < 402$ GeV in the simplified hMSSM minimal supersymmetric model for $\tan\beta = 2$ and excluding bulk Randall-Sundrum gravitons $G_{KK}$ in the mass range 325 GeV $< m_{G_{KK}} < 885$ GeV for $k/\bar{M}_W = 1$.

DOI: 10.1103/PhysRevLett.121.191801

In 2012, the ATLAS and CMS Collaborations at the LHC discovered a new particle with a mass of approximately 125 GeV [1–3]. According to all current measurements it is compatible with the standard model (SM) Higgs boson ($H$) [4–8]. An important pending test of the Brout-Englert-Higgs mechanism is the measurement of Higgs boson pair production. At the LHC, pairs of SM Higgs bosons can be produced via the Higgs self-interaction (“triangle diagram”) and the destructively interfering top-quark loop (“box diagram”) [9,10]. Nonresonant Higgs boson pair production (NR $HH$) can be significantly enhanced relative to the SM prediction by modifications to the top-quark Yukawa coupling, the trilinear Higgs boson coupling $\lambda_{HHH}$, or by introducing production mechanisms with new intermediate particles. Many theories beyond the SM predict heavy resonances that could decay into a pair of SM Higgs bosons, such as a heavy $CP$-even scalar $X$ in two-Higgs-doublet models [11], or spin-2 Kaluza-Klein (KK) excitations of the graviton, $G_{KK}$, in the bulk Randall-Sundrum (RS) model [12–14].

This Letter describes a search for resonant and nonresonant Higgs boson pair production in a final state with two $b$ quarks and two $\tau$ leptons using 36.1 fb$^{-1}$ of $pp$ collision data recorded with the ATLAS detector [15,16] in 2015 and 2016. The $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ decay channels are considered, where the subscripts (lep = electron or muon, had = hadrons) indicate the decay mode of the $\tau$ lepton. Previous searches for Higgs boson pair production were performed at center-of-mass energies $\sqrt{s} = 8$ TeV [17–19] and $\sqrt{s} = 13$ TeV [20–22] by the ATLAS and CMS Collaborations. The ATLAS search in the $4b$ channel constitutes the most sensitive result to date and the observed (expected) limit excludes a cross section greater than 13.0 (20.7) times the SM prediction at 95% confidence level (C.L.).

The SM nonresonant $HH$ process was simulated with MadGraph5_aMC@NLO at next-to-leading order (NLO) [23–27] using the CT10 parton distribution function (PDF) set [28]. Parton showers and hadronization were simulated with Herwig++ [29] using the UEEE5 set of tuned parameters (tune) [30]. The events were reweighted to reproduce the $m_{HH}$ spectrum obtained in Refs. [9,31], which fully accounts for the finite mass of the top quark. The cross-section times branching ratio to the $b\bar{b}\tau\tau$ final state, evaluated at next-to-next-to-leading order (NNLO) and including next-to-next-to-leading logarithm (NNLL) corrections and NLO top-quark mass effects, is $2.44^{+1.01}_{-0.22}$ fb [32]. Events with a generic narrow-width scalar $X$ or $G_{KK}$ decaying into $HH$ were produced in MadGraph5_aMC@NLO at leading order (LO) and interfaced to the Pythia 8 [33] parton shower model using the A14 tune [34] together with the NNPDF23LO PDF set [35]. The cross section and width of the $G_{KK}$ were taken from Ref. [36] and
depend on $k/M_{Pl}$, where $k$ corresponds to the curvature of the warped extra dimension and $M_{Pl} = 2.4 \times 10^{18}$ GeV is the effective four-dimensional Planck scale. Events with $k/M_{Pl} = 1$ and $k/M_{Pl} = 2$ were simulated.

The dominant SM background processes are $t \bar{t}$, QCD multijet and $Z$ bosons produced in association with jets originating from heavy-flavor quarks ($bb, bc, cc$), subsequently referred to as $Z +$ heavy flavor [37]. SM Higgs boson production in association with $Z$ boson, subsequently decaying into a $b \bar{b} \tau \tau$ [38] final state, is an irreducible background in this analysis. The $t \bar{t}$ and single-top-quark background events were simulated using POWHEG-BOX [39], with the CT10 PDF set, and MADGRAPH5_aMC@NLO [40]. The parton showers were simulated using PYTHIA 6 [41] and the Perugia 2012 tune [42]. The $t \bar{t}$ background was scaled to match the NNLO + NNLL cross sections [43], while the single-top samples were corrected to NLO [44,45] (approximate NNLO [46]) predictions for the $t$- and $s$-channel ($Wt$ final state). Events with $W$ or $Z$ bosons and associated jets were simulated with the SHERPA 2.2.1 generator [47–51], using the NNPDF30NNLO PDF set [52] and normalized to the NNLO cross sections [53]. Diboson and Drell–Yan backgrounds were produced with SHERPA 2.2.1 [47] using the CT10NLO PDF set and the generator cross-section predictions. Quark-induced $ZH$ processes were generated with PYTHIA 8, using the A14 tune and the NNPDF23LO PDF set. The samples were normalized to NNLO cross sections for QCD and NLO for electroweak processes [54–60]. The gluon-induced $ZH$ process [61] was generated with POWHEG using the CT10 PDF set and using PYTHIA 8 with the A1NLO tune [62] to simulate parton showers. Cross sections [63–67] were scaled to NLO + NLL in QCD. SM Higgs boson production in association with a top-quark pair was simulated with MADGRAPH5_aMC@NLO; PYTHIA 8 was used to simulate the parton shower, while the cross section was taken from Ref. [10]. In all signal and background samples, the mass of the $H$ bosons was set to 125 GeV. The contributions from other SM Higgs boson processes are negligible. EVTGEN v1.2.0 [68] was used to model the properties of bottom and charm hadron decays for all processes except those simulated in SHERPA. The detector response to the generated events was simulated with GEANT4 [69,70]. Simulated events are reweighted to match the distribution of the number of inelastic collisions per event (pileup) in data.

Events are required to have at least one collision vertex reconstructed from at least two charged-particle tracks with transverse momentum $p_T > 0.4$ GeV. The primary vertex for each event is selected as the vertex with the highest $\sum (p_T)$. Jets are formed using the anti-$k_T$ algorithm [72] with a radius parameter $R = 0.4$ and calorimeter energy clusters as inputs [73–75]. These jets are taken as seeds for the reconstruction of the visible products of hadronically decaying $\tau$ leptons ($\tau_{had-vis}$) [76–78], which are subsequently required to have one or three associated tracks. In order to distinguish $\tau_{had-vis}$ from quark- and gluon-initiated jets, a boosted decision tree (BDT) [79], trained separately for $\tau_{had-vis}$ with one and three charged particles, is employed. Selected $\tau_{had-vis}$ candidates must satisfy the “medium” BDT working point [77]. Electron candidates are identified using a likelihood technique in combination with additional track-hit requirements [80]; the transition region between the barrel and end cap calorimeters is excluded. Information from the tracking and muon systems is used to reconstruct muon candidates [81]. Only isolated electrons and muons are considered, where no nearby tracks or calorimeter energy deposits within a $p_T$-dependent variable-size $\Delta R$ cone around the lepton are allowed. Jets arising from pileup are suppressed using dedicated track and vertex requirements [82]. The missing transverse momentum, with magnitude $E_T^{miss}$, is defined as the negative vectorial sum of all reconstructed and fully calibrated objects in the event, along with an additional track-based soft term [83]. Jets containing $b$ hadrons are identified using the MV2c10 multivariate discriminant [84,85] trained against a light-quark-flavor sample also containing 10% of $c$ hadrons. A working point with 70% efficiency on simulated $t \bar{t}$ events is used. An overlap-removal procedure is applied to the reconstructed electrons, muons, $\tau_{had-vis}$, and jets to prevent double counting of energy deposits in the detector as described in Ref. [86].

The selected final state is characterized by one electron or muon and one $\tau_{had-vis}$ of opposite charge, or two $\tau_{had-vis}$ of opposite charge, plus two $b$-tagged jets and $E_T^{miss}$. In all cases, events with additional electrons or muons above 7 GeV or $\tau_{had-vis}$ above 20 GeV are rejected. The off-line selection criteria for the electron, muon, and $\tau_{had-vis}$ depend on the triggers used. In the $\tau_{lep}\tau_{had}$ channel events are selected with a single-lepton trigger (SLT) and a lepton plus $\tau_{had}$ trigger (LTT), which are analyzed separately and combined with the $\tau_{had}\tau_{had}$ channel in the final fit. Depending on the data period, the electron or muon that passes the SLT trigger is required to have $p_T > 25–27$ GeV. Events which fail this requirement are considered for the LTT category if the electron (muon) has $p_T > 18$ GeV (15 GeV). In all cases, $p_T$ requirements are 1 GeV higher than the trigger thresholds to ensure a nearly constant trigger efficiency relative to the off-line selection. The $\tau_{lep}\tau_{had}$ events are required to have one $\tau_{had-vis}$ candidate with $|\eta| < 2.3$ and $p_T > 20$ GeV for SLT events, raised to 30 GeV for LTT events due to $\tau_{had-vis}$ $p_T$ requirements applied in this category of triggers. In the $\tau_{had}\tau_{had}$ channel a logical OR of single $\tau_{had}$ triggers (STT) and $\tau_{had}$ triggers (DTT) is used. The leading $\tau_{had-vis}$ candidate is required to have a minimum $p_T$ of 40 GeV for DTT and between 100 and 180 GeV for STT events, depending on the data-taking period. The subleading $\tau_{had-vis}$ is required to have a minimum $p_T$ of 20 (30) GeV for STT (DTT) events. The leading jet is required to have
inputs to the final fit. The BDTs for the normalization of the dominant backgrounds, form the separation of signal from background. Their distributions in the three trained against the main backgrounds, \( p_T > 20 \) GeV. The invariant mass of the di-\( \tau \) system, \( m_{\tau\tau}^{\text{MMC}} \), is calculated using the Missing Mass Calculator [87] and is required to be greater than 60 GeV. Signal region (SR) events are defined as those meeting the criteria above, and in addition containing two \( b \)-tagged jets; they are further separated into \( \tau_{\text{lep}}\tau_{\text{had}} \) LTT and \( \tau_{\text{had}}\tau_{\text{had}} \) categories.

BDTs are used in the analysis to improve the separation of signal from background. Their distributions in the three signal regions, along with control region yields to constrain the normalization of the dominant backgrounds, form the inputs to the final fit. The BDTs for the \( \tau_{\text{had}}\tau_{\text{had}} \) channel are trained against the main backgrounds, \( \tilde{t}\tau, Z \rightarrow \tau\tau \), and multijet events; in the \( \tau_{\text{lep}}\tau_{\text{had}} \) channel they are trained solely against the dominant \( \tilde{t}\tau \) background. For the BDT trainings, the \( \tilde{t}\tau \) and \( Z \rightarrow \tau\tau \) backgrounds are taken purely from simulation, while the multi-jet events are estimated using the data-driven approach described below. Variables which provide good discrimination and are minimally correlated are used as inputs to the BDTs, as summarized in Table I. The variables selected in each channel differ, reflecting the different background compositions. In the resonant search, BDTs are trained separately for each signal mass considered, from 260 to 1000 GeV (800 GeV for LTT), where the signal model combines the target resonance mass and its two neighboring mass points, to be sensitive to masses between the simulated points. For NR \( HH \) production, the BDTs are trained on a signal sample with the SM admixture of the contributions from the box diagram and triangle diagram. The BDTs are more sensitive to the box diagram where the two Higgs bosons are produced at higher \( p_T \) and the selection efficiency is greater.

In both channels, simulated events are used to model background processes containing reconstructed \( \tau_{\text{had-vis}} \) that are matched to generated \( \tau_{\text{had}} \) within \( \Delta R = 0.2 \) (subsequently referred to as true \( \tau_{\text{had}} \)) and other minor background contributions. The rate of events with at least one true \( \tau_{\text{had}} \) and a jet reconstructed as an electron or muon is found to be negligible. For \( \tilde{t}\tau \) background events containing one or more true \( \tau_{\text{had}} \) the normalization is obtained in the final fit, constrained mainly by the low \( \tau_{\text{lep}}\tau_{\text{had}} \) BDT score regions, resulting in a normalization factor of \( 1.06 \pm 0.13 \). The normalization of the \( Z \rightarrow ee/\tau\tau + \text{heavy-flavor background} \) is determined using \( Z \rightarrow \mu\mu + \text{heavy-flavor} \) events. Their selection closely follows the event selection used for signal events. Instead of two \( \tau \)-lepton candidates, two muons with \( p_T > 27 \) GeV and dimuon invariant mass between 81 and 101 GeV are selected. To remove the contribution from SM \( ZH(H \rightarrow bb) \) production, \( m_{bb} \) is required to be lower than 80 GeV or greater than 140 GeV. The normalization is determined by including the \( Z \rightarrow \mu\mu + \text{heavy-flavor} \) control region yield in the final fit, resulting in a normalization factor of \( 1.34 \pm 0.16 \). Normalization factors are not applied to the \( Z + \) light-flavor contributions. The modeling of the BDT score

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \tau_{\text{lep}}\tau_{\text{had-vis}} ) channel (SLT resonant)</th>
<th>( \tau_{\text{lep}}\tau_{\text{had-vis}} ) channel (SLT nonresonant &amp; LTT)</th>
<th>( \tau_{\text{had}}\tau_{\text{had}} ) channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_{HH} )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( m_{\tau\tau}^{\text{MMC}} )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( m_{bb} )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( \Delta R(\tau, \tau) )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( \Delta R(b, b) )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( E_T^{\text{miss}} )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( E_T^{\text{miss}} ) ( \phi ) centrality</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( m_W )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( \Delta \phi(H, H) )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>( \Delta p_T(\text{lep}, \tau_{\text{had-vis}}) )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Subleading ( b )-jet ( p_T )</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
distributions is validated in the 0-\(b\)-tag and 1-\(b\)-tag regions as well as in dedicated \(t\bar{t}\) and Z + heavy-flavor validation regions.

Contributions from processes in which a quark- or gluon-initiated jet is misidentified as a \(\tau_{\text{had-vis}}\) candidate (fake-\(\tau_{\text{had}}\)) are estimated using data-driven methods for major backgrounds. A fake-\(\tau_{\text{had}}\) enriched sample is defined by requiring that a \(\tau_{\text{had-vis}}\) fails the “medium” BDT identification but satisfies a very loose requirement on the BDT score. This selection maintains a composition of quark- and gluon-initiated jets similar to those mimicking \(\tau_{\text{had-vis}}\) in the SR. In the case where the event contains more than one such fake \(\tau_{\text{had}}\), one is chosen randomly. The SR selection, except for the \(\tau_{\text{had-vis}}\) identification, is applied to the fake-\(\tau_{\text{had}}\) enriched sample to extract template distributions for the fake-\(\tau_{\text{had}}\) background after the true-\(\tau_{\text{had}}\) contamination is subtracted using simulation. The templates are scaled with fake factors (FF) defined as the ratio of the number of fake \(\tau_{\text{had}}\) that pass the \(\tau_{\text{had-vis}}\) identification to the number that fail, calculated in dedicated control regions (CR) and parametrized in \(p_{\text{R}}(\tau_{\text{had-vis}})\) and the number of associated tracks.

For the \(t\bar{t}\) final state, fake-\(\tau_{\text{had}}\) background contributions from \(t\bar{t}\), W + jets and multijet processes are estimated using a combined fake-factor method similar to that described in Refs. [86,89]. In order to account for the different sources of fake \(\tau_{\text{had}}\), the FFs are derived separately for each background contribution. The CR for multijet events is defined by inverting the isolation requirement applied to the electron or muon for events with 0 or 1 \(b\)-tagged jets. The \(t\bar{t}\) (W + jets) control region is defined by requiring two (zero) \(b\)-tagged jets and \(m_{W}^{T} > 40\) GeV, where \(m_{T}^{W} = \sqrt{2p_{T}^{\text{lep}}E_{T}^{\text{miss}}(1 - \cos \Delta \phi_{\text{lep},E_{T}^{\text{miss}}})}\), and \(\Delta \phi_{\text{lep},E_{T}^{\text{miss}}}^{\text{lep}}\) is the azimuthal angle between the electron or muon and the \(E_{T}^{\text{miss}}\). Fake factors for \(t\bar{t}\) and W + jets are found to be consistent for both processes. The individual fake factors are then combined as \(\text{FF(comb)} = \text{FF(QCD)} \times \text{r}_{\text{QCD}} + \text{FF}(t\bar{t}/W + \text{jets}) \times (1 - \text{r}_{\text{QCD}})\), where \(r_{\text{QCD}}\) is defined as the fraction of fake \(\tau_{\text{had}}\) from (predominantly multijet) processes contributing to the data in the fake \(\tau_{\text{had}}\) enriched template region that are not accounted for by simulated background processes, and is less than 5% in the 2-\(b\)-tag region. Because of the different origin of fake \(\tau_{\text{had}}\), the FFs for \(t\bar{t}/W + \text{jets}\) can be up to 30% larger than those for multijet processes. Events with two \(b\)-tagged jets but a same-sign charge (SS) electron or muon and \(\tau_{\text{had-vis}}\) are used for validating the fake-\(\tau_{\text{had}}\) background, showing all distributions are well modeled. Given this, and the small size of the contribution, no transfer factor is applied to correct the multijet estimation from the 1-\(b\)-tag region to the 2-\(b\)-tag region.

In the \(\tau_{\text{had}}\) final state, only the multijet background is estimated from data using the FF method. The differential FFs are derived in a 1-\(b\)-tag SS control region, while the overall normalization is taken from the 2-\(b\)-tag SS control region. The \(t\bar{t}\) background is estimated from simulation, where the fake-\(\tau_{\text{had}}\) \(t\bar{t}\) contribution is corrected in bins of \(n_{\text{r}_{\tau_{\text{had-vis}}}}\) using the probability for a jet from a hadronic W-boson decay to mimic a \(\tau_{\text{had-vis}}\) candidate (fake rate), as measured with data in the \(\tau_{\text{lep}}\) \(t\bar{t}\) control region [86]. Contributions from true \(\tau_{\text{had}}\) are subtracted using simulation.

The uncertainty in the integrated luminosity of the combined 2015 + 2016 data set is 2.1% [90] and is applied to the signal and background components whose normalizations are derived from simulation. An uncertainty related to the pileup reweighting procedure is also applied [91]. Experimental uncertainties in the identification and reconstruction of the electron [92], muon [93], \(\tau_{\text{had-vis}}\) jets [76], and jets [74,94] are accounted for and propagated through the analysis to determine their effect on the final results. These affect the trigger requirements, the identification and reconstruction efficiencies, the isolation, and the reconstructed energies and their resolutions. The uncertainties are propagated to the calculation of the \(E_{T}^{\text{miss}}\) [83], which has an additional uncertainty from the soft term. The uncertainties with the largest impact on the result are those related to the \(\tau_{\text{had-vis}}\) identification efficiency, which correspond to an uncertainty of 16% on the NR signal strength, i.e., the simulated NR \(HH\) yield assuming a cross-section times branching fraction equal to the expected limit and normalized to the SM expectation \((\sigma_{\text{exp}}/\sigma_{\text{SM}})\). Uncertainties in flavor tagging [95,96] also have a significant impact, inducing an uncertainty in the NR signal strength of 8.3%, dominated by those associated with the \(b\)-tagging efficiency.

Theory uncertainties in the modeling of the \(t\bar{t}\) background containing one or more true \(\tau_{\text{had}}\) are assessed by varying the matrix element generator (using aMC@NLO instead of POWHEG-BOX) and the parton shower model (using HERWIG++ instead of PYTHIA 6), and by adjusting the factorization and renormalization scales along with the amount of additional radiation. The resulting variations in the BDT distributions are included as shape uncertainties in the final fit. In order to account for potential acceptance differences between control and signal regions, the normalization of the \(t\bar{t}\) background containing true \(\tau_{\text{had}}\), determined predominantly from the \(\tau_{\text{lep}}\) \(\tau_{\text{had}}\) SR in the final fit, is allowed to vary within a range determined by the acceptance variations associated with the \(t\bar{t}\) modeling uncertainties. This amounts to +30%/-32% for the \(\tau_{\text{had}}\) \(\text{SR}\) and +8.1%/-9.3% for the \(Z \rightarrow \mu\mu + \text{heavy-flavor control region}\). This is the dominant uncertainty in the \(t\bar{t}\) modeling.

For the \(Z + \text{jets}\) background, the theory uncertainties in the modeling of the BDT shapes are derived by comparing the nominal SHERPA sample with an alternative \textsc{MadGraph5}+aMC@NLO + PYTHIA 8 sample and by varying the choice of renormalization and factorization scales, along with the PDF prescription [97]. The normalization of the \(Z \rightarrow \tau\tau + \text{heavy-flavor background in the} \)}
The uncertainty in the modeling of backgrounds due to jets being misidentified as \( r_{\text{had-vis}} \) is estimated by varying the fake factors and fake rates within their statistical uncertainties and varying the amount of true-\( r_{\text{had}} \) background subtracted. Based on studies with simulated \( \bar{t}t \) and \( W + \) jets events, a systematic uncertainty is assigned to cover the difference in the gluon and quark flavor composition of jets misidentified as a \( r_{\text{had-vis}} \) between the signal region and the fake-\( r_{\text{had}} \) enriched sample, parametrized as a function of the \( r_{\text{had-vis}} \) identification BDT score. The uncertainty in the extrapolation of FF(QCD) to the signal region is estimated from the difference between the nominal FFs and alternative ones, calculated either in the SS region for the \( r_{\text{lep},\text{had}} \) channel or a multijet enriched region, where \( \Delta \phi(r_{\text{had-vis}}, r_{\text{had}}) > 2.0 \), in the \( r_{\text{had}} \) case. Similarly, changes in the fake-\( r_{\text{had}} \) determination when varying the \( \bar{t}t \) control region \( m_T^W \) requirement in simulation and data are used to estimate a systematic uncertainty in both the fake factors and fake rates. The overall effect of these uncertainties on the fake-\( r_{\text{had}} \) background estimate leads to an 8.4% variation of the NR signal strength, predominantly due to the true-\( r_{\text{had}} \) subtraction in the \( \bar{t}t \) control region and the composition of the fake \( r_{\text{had}} \).

Theory uncertainties in the signal acceptance are calculated by independently varying the renormalization and factorization scales, the choice of PDF and each PDF set by its uncertainties. The uncertainty in the parton shower is taken into account by comparing the default HERWIG++ with PYTHIA 8. Uncertainties in the underlying event, initial-state radiation and final-state radiation are accounted for by changing the PYTHIA tune, but are small. The effects of various categories of uncertainty on the measured nonresonant signal strength corresponding to the expected upper limit at 95% C.L. are summarized in Table II. The individual sources of uncertainty making up the categories listed in the table are grouped together in the final fit to determine their correlated combined effect on the signal strength. For all signal hypotheses, the statistical uncertainties dominate.

For each signal model considered, a profile-likelihood fit [100] is applied to the BDT score distributions simultaneously in the three SRs to extract the signal cross section, along with the \( \bar{t}t \) and \( Z + \) heavy-flavor normalizations. The lattermost is constrained by including the dedicated control region in the fit. All sources of systematic and statistical uncertainty in the signal and background models are implemented as deviations from the nominal model, scaled by nuisance parameters that are profiled in the fit. None of the dominant nuisance parameters are significantly constrained or pulled relative to their input value by the fit. The BDT score distributions for the nonresonant search and the \( G_{KK} \) signal are shown in Fig. 1 after performing the fit and assuming a background-only hypothesis. The acceptance times efficiency for the \( \bar{t}t \) control region over the full BDT spectrum to \( Z \rightarrow \tau\tau \) is 4.2% (2.9%) in the combined SLT and LTT \( r_{\text{lep},\text{had}} \) \( (r_{\text{had}}, r_{\text{had}}) \) channel over the full BDT distribution, decreasing to 3.3% (2.4%) for the two most sensitive BDT bins. As no significant excess over the expected background is observed, upper limits are set on nonresonant and resonant Higgs boson pair production at 95% C.L. using the \( CL_s \) method [101].

Table III presents the upper limits on the cross section for nonresonant \( HH \) production times the \( HH \rightarrow b\tau\tau \) branching ratio, and comparisons with the SM prediction. The observed (expected) limit is 30.9 fb (36.0 fb), 12.7 (14.8) times the SM prediction. In order to compare with previous results, the BDTs are trained and applied to the signal sample without reweighting the \( m_{HH} \) spectrum to Refs. [9,31], giving an observed (expected) limit of 37.4 fb (33.5 fb), 15.4 (13.8) times the SM prediction.

The results of searches for resonant \( HH \) production are presented as exclusion limits on the cross-section times branching fraction equal to the 95% C.L. expected limit of 14.8 times the SM expectation.
FIG. 1. Distributions of the BDT score for NR the combined statistical and systematic uncertainty in the background. The ratio of the data to the sum of the backgrounds is shown after the fit to the background-only hypothesis and the signal is scaled to approximately the expected limit. The hatched band indicates
(a),(b) \( \tau^{\text{lep}} \) single-lepton trigger (SLT), (c),(d) lepton + \( \tau^{\text{had}} \) trigger (LTT) and (e),(f) \( \tau^{\text{had}} \) channels. Distributions are shown after the fit to the background-only hypothesis and the signal is scaled to approximately the expected limit. The hatched band indicates the combined statistical and systematic uncertainty in the background. The ratio of the data to the sum of the backgrounds is shown in the lower panel.
Scalar resonances $X$ and $G_{KK}$ signal models are shown in Fig. 2. For scalar resonances, the results are interpreted in a simplified minimal supersymmetric model, the hMSSM [102,103], where the mass of the light CP-even Higgs boson is fixed to 125 GeV. The mass range $305 \text{ GeV} < m_X < 402 \text{ GeV}$ is excluded at 95% C.L. for $\tan\beta = 2$, where $\tan\beta$ is the ratio of the vacuum expectation values of the scalar doublets. Gravitons are excluded at 95% C.L. in the mass range $325 \text{ GeV} < m_{G_{KK}} < 85 \text{ GeV}$ assuming $k/M_{Pl} = 1$. Above $\sim 600 \text{ GeV}$, the limits are largely insensitive to the value of $k/M_{Pl}$, while at low $m_{HH}$ they improve significantly with increasing $k$ due to the larger natural width. The limits on resonant $HH$ production are significantly more stringent than previous results in the $bb\tau\tau$ channel and competitive with limits obtained in other channels.

In summary, a search for resonant and nonresonant Higgs boson pair production in the $bb\tau\tau$ final state is conducted with 36.1 fb$^{-1}$ of $pp$ collision data delivered by the LHC at $\sqrt{s} = 13 \text{ TeV}$ and recorded by the ATLAS detector. The analysis of nonresonant Higgs pair production excludes an enhancement of the SM expectation by more than a factor of 12.7 at 95% C.L. This is the most stringent limit on $HH$ production to date. Upper limits are set on resonant Higgs boson pair production for a narrow-width scalar $X$ and a spin-2 Kaluza-Klein graviton $G_{KK}$ in the bulk RS model.

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; 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, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, 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,

### TABLE III

Observed and expected upper limits on the production cross-section times the $HH \rightarrow bb\tau\tau$ branching ratio for NR $HH$ at 95% C.L., and their ratios to the SM prediction. The ±1σ variations about the expected limit are also shown.

<table>
<thead>
<tr>
<th>$\tau_{lep}\tau_{had}$</th>
<th>$\sigma(HH \rightarrow bb\tau\tau)$ [fb]</th>
<th>Observed</th>
<th>$-1\sigma$</th>
<th>Expected</th>
<th>$+1\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma/\sigma_{SM}$</td>
<td>57</td>
<td>49.9</td>
<td>69</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23.5</td>
<td>20.5</td>
<td>28.4</td>
<td>39.5</td>
</tr>
<tr>
<td>$\tau_{had}\tau_{had}$</td>
<td>$\sigma(HH \rightarrow bb\tau\tau)$ [fb]</td>
<td>40.0</td>
<td>30.6</td>
<td>42.4</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>$\sigma/\sigma_{SM}$</td>
<td>16.4</td>
<td>12.5</td>
<td>17.4</td>
<td>24.2</td>
</tr>
<tr>
<td>Combination</td>
<td>$\sigma(HH \rightarrow bb\tau\tau)$ [fb]</td>
<td>30.9</td>
<td>26.0</td>
<td>36.1</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>$\sigma/\sigma_{SM}$</td>
<td>12.7</td>
<td>10.7</td>
<td>14.8</td>
<td>20.6</td>
</tr>
</tbody>
</table>

![Graph](image-url)
ERDF, 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; CERCA Programme 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. [104].


[37] Equivalently, $Z$ bosons produced in association with at least one light-flavor quark ($u, d,$ or $s$) are referred to as $Z +$ light-flavor.

[38] The notations $\tau\tau$ and $b\bar{b}$ are used throughout this Letter, in place of $\tau^+\tau^-$ and $b\bar{b}$, as charge conjugation is implied.


[71] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the $z$ axis along the beam pipe. The $x$ axis points to the center of the LHC ring and the $y$ axis points upward. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ is the azimuthal angle around the $z$ axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. The distance between two objects in $\eta-\phi$ space is $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. Transverse momentum is defined by $p_T = p \sin \theta$.
[89] The fake-$r_{had}$ contribution of the $t\bar{t}$ background is estimated using the fake-factor method in the analysis, however, for the training of the BDTs the $t\bar{t}$ background is taken from simulation.


Also at The City College of New York, New York, New York, USA.

Also at Departamento de Física Teorica y del Cosmos, Universidad de Granada, Granada (Spain), Spain.

Also at Department of Physics, California State University, Sacramento, California, USA.

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

Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève, Switzerland.

Also at Department of Physics and Astronomy, University of Sheffield, Sheffield, United Kingdom.

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

Also at Department of Applied Physics and Astronomy, University of Sharjah, Sharjah, United Arab Emirates.

Also at Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany.

Also at National Research Nuclear University MEPhI, Moscow, Russia.

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

Also at Giresun University, Faculty of Engineering, Giresun, Turkey.

Also at Institute of Physics, Academia Sinica, Taipei, Taiwan.