Search for heavy resonances decaying into a photon and a hadronically decaying Higgs boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

Aad, G.; ATLAS Collaboration

DOI
10.1103/PhysRevLett.125.251802

Publication date
2020

Document Version
Final published version

Published in
Physical Review Letters

License
CC BY

Citation for published version (APA):
https://doi.org/10.1103/PhysRevLett.125.251802
Search for Heavy Resonances Decaying into a Photon and a Hadronically Decaying Higgs Boson in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector

G. Aad et al. * (ATLAS Collaboration)

(Received 16 August 2020; accepted 5 November 2020; published 18 December 2020)

This Letter presents a search for the production of new heavy resonances decaying into a Higgs boson and a photon using proton-proton collision data at $\sqrt{s} = 13$ TeV collected by the ATLAS detector at the LHC. The data correspond to an integrated luminosity of 139 fb$^{-1}$. The analysis is performed by reconstructing hadronically decaying Higgs boson ($H \rightarrow b\bar{b}$) candidates as single large-radius jets. A novel algorithm using information about the jet constituents in the center-of-mass frame of the jet is implemented to identify the two $b$ quarks in the single jet. No significant excess of events is observed above the expected background. Upper limits are set on the production cross-section times branching fraction for narrow spin-1 resonances decaying into a Higgs boson and a photon in the resonance mass range from 0.7 to 4 TeV, cross-section times branching fractions are excluded between 11.6 fb and 0.11 fb at a 95% confidence level.

DOI: 10.1103/PhysRevLett.125.251802

Many extensions to the standard model, such as technicolor [1], little Higgs [2], or a more complex Higgs sector [3], predict new massive bosons. Some of these bosons may decay into a Higgs boson and a photon at the one-loop level [4]. Searches for such particles have been carried out by both the ATLAS [5] and CMS [6] Collaborations at the Large Hadron Collider (LHC).

This Letter reports on a generic search for a narrow, neutral, spin-1 boson ($Z'$) that decays into a photon and a Higgs boson. The Higgs boson subsequently decays hadronically as $H \rightarrow b\bar{b}$, where the hadronic products from both $b$ quarks are reconstructed as a single large-radius jet. The analysis uses data from $\sqrt{s} = 13$ TeV proton-proton ($pp$) collisions that were recorded by the ATLAS detector from 2015 to 2018 with a single-photon trigger [7], corresponding to an integrated luminosity of 139 fb$^{-1}$. The single-photon trigger uses loose photon identification requirements based on calorimetric shower-shape variables [8] and imposes a transverse momentum threshold of 140 GeV. It is fully efficient for events passing the offline analysis selection. The search identifies the two $b$ quarks in the single jet by using a novel methodology based on information about the jet constituents calculated in the center-of-mass frame of the jet. This technique significantly improves the search sensitivity compared to the previous ATLAS [5] and CMS [6] analyses, in addition to the gains from the larger data sample.

The ATLAS detector [9,10] is a general-purpose particle detector with a cylindrical geometry [11]. It consists of an inner detector surrounded by a superconducting solenoid that produces a 2 T magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer with a toroidal magnetic field. The inner detector provides precision tracking of charged particles with pseudorapidity $|\eta| < 2.5$. The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. It comprises sampling calorimeters with either liquid argon or scintillator tiles as the active medium. A two-level trigger system accepts events from the 40 MHz bunch crossings at a rate of 1 kHz for off-line analysis.

Monte Carlo (MC) simulated events are used to optimize the event selection and to help validate the analysis. The signal samples, with decays of $Z' \rightarrow H\gamma$ at the one-loop level [4], were generated for eight different mass points in a range from 700 to 4000 GeV via quark-antiquark annihilation, $q\bar{q} \rightarrow Z' \rightarrow H\gamma$, using the MADGRAPH leading-order (LO) v2.6.2 generator [12] interfaced to PYTHIA8.235 [13] with the NNPDF23LO parton distribution functions (PDFs) [14] for both generators and the A14 set of tuned parameters [15] for the underlying event. The total decay widths of the $Z'$ resonances were set to 4.2 MeV, which is much smaller than the experimental mass resolution, which varies from around 35 GeV at the 700 GeV signal mass point to 150 GeV at the 4000 GeV signal mass point. The dominant SM background arises from events with prompt photons produced in association with jets ($\gamma + \text{jets}$). Less dominant SM backgrounds include a prompt photon produced in association with a $W/Z$ boson ($W/Z + \gamma$) or a top-antitop quark pair.
quark pair ($t\bar{t} + \gamma$). The MC sample of $\gamma +$ jets events was simulated using the SHERPA2.2.2 generator [16] with up to two additional parton emissions at next-to-leading-order (NLO) accuracy and up to four additional partons at LO accuracy using Comix [17] and OpenLoops [18]. The events were then merged with the SHERPA parton shower [19] using the ME + PS@NLO prescription [20]. Samples are generated using the NNPDF3.0nlo PDF set [21], along with the dedicated set of tuned parton-shower parameters developed by the SHERPA authors. The W/Z + $\gamma$ events were modeled with SHERPA2.1.1 at LO with the CT10 PDFs [23], then interfaced to PYTHIA8.186 with the A14 parameter tune and the NNPDF23LO PDFs. In the signal samples and $t\bar{t} + \gamma$ background sample, EVTGEN [24] was used to model charm and $b$-hadron decays. The effect of multiple $p p$ interactions in the same and neighboring bunch crossings (pileup) is included by overlaying minimum-bias events simulated with PYTHIA8.186 on each event of interest in all samples. The generated samples were processed through a GEANT4-based detector simulation [25,26] and the same ATLAS reconstruction software as the data.

An event is selected if it contains a $H \rightarrow b\bar{b}$ candidate and at least one isolated photon that satisfies the “tight” identification criteria [27,28]. A selected photon must have transverse momentum ($p_T$) greater than 200 GeV and be within the calorimeter barrel region $|\eta| < 1.37$. Each $H \rightarrow b\bar{b}$ candidate is reconstructed as a single jet using the anti-$k_t$ algorithm [29,30] with a large radius parameter ($R = 1.0$), hereafter referred to as a large-$R$ jet (J). The large-$R$ jets are formed from topological energy clusters (topoclusters) [31] in the calorimeter and are trimmed [32] to mitigate the effects of pileup and soft radiation. The large-$R$ jet constituents are reclustered into subjets using the $k_t$ algorithm [33] with $R = 0.2$, and the subjets that carry less than 5% of the $p_T$ of the original large-$R$ jet are removed. To overcome the limited angular resolution of the calorimeter, the mass of a large-$R$ jet ($m_j$) is computed using a combination of calorimeter and tracking information [34]. Large-$R$ jets are required to have $p_T > 200$ GeV, $|\eta| < 2.0$, 50 GeV < $m_j$ < 200 GeV, and an angular separation of $\Delta R > 1.0$ from photon candidates. For the baseline event selection, at least one large-$R$ jet and one photon are required to pass the selection described above.

The photon and large-$R$ jet with the highest $p_T$ in an event are combined to form a resonance candidate. The invariant mass of the resonance candidate ($m_{J}\gamma$) is used to distinguish signal from background. In addition, the large-$R$ jet mass must be consistent with the Higgs boson mass ($m_H = 125.80$ GeV), $m_T - \Delta_{m,D} < m_J < m_T + \Delta_{m,U}$. The parameters $\Delta_{m,D}$ and $\Delta_{m,U}$ are determined by maximizing the search sensitivity $\mathcal{B}/(\sqrt{B} + 3/2)$ [35], where $\mathcal{B}$ is the resonance signal selection efficiency and $B$ is the number of background events, as estimated from MC samples, within the resonance mass window, $|m_{J}\gamma - m_{Z}^0| < 2\sigma_{m_{J}\gamma}$. Here $m_{Z}^0$ and $\sigma_{m_{J}\gamma}$ are the peak position and the core resolution of the reconstructed $m_{J}\gamma$ distribution of the $Z' \rightarrow H\gamma$ signal MC events, respectively. The above procedure is performed separately for each $m_{J}\gamma$ hypothesis. The optimized mass window of the large-$R$ jets varies from around $[100,130]$ GeV at $p_T = 0.5$ TeV to $[90,160]$ GeV at $p_T = 2$ TeV.

To further reduce the background, a novel algorithm [36–38] is applied to the large-$R$ jet to identify the two $b$-quarks that originated from the Higgs boson. It uses the kinematics of the jet constituents in the center-of-mass (c.m.) frame of the large-$R$ jet (jet rest frame), where the final products of a two-body $H \rightarrow b\bar{b}$ decay can be easily separated into a back-to-back topology. In this approach, the topoclusters of the large-$R$ jet and the tracks associated with the jet are boosted to the large-$R$ jet's rest frame. In the jet rest frame, the topoclusters of the large-$R$ jet are reclustered using the EEKT jet algorithm [39] to form exactly two c.m. subjets, assumed to originate from the Higgs boson decay. A track is considered to be associated with a c.m. subjet if the opening angle $\Delta \theta$ between the track and the c.m. subjet, calculated in the jet rest frame, satisfies the requirement that $2 \times (1 - \cos \Delta \theta) < 0.8$. The c.m. subjets and their associated tracks are then boosted back to the laboratory frame and the standard ATLAS $b$-tagging algorithm based on a multivariate technique, MV2c10 [40,41], is applied to each c.m. subjet to identify those containing a $b$ hadron (called c.m. $b$ subjets). For this analysis, the working point of the MV2c10 tagger output is chosen to have an overall efficiency of 77%. This was determined using simulated Randall-Sundrum graviton [42] ($G \rightarrow HH$, $H \rightarrow b\bar{b}$) events, in which the $p_T$ distribution of the large-$R$ jets that contain a Higgs boson is reweighted to match the inclusive jet $p_T$ distribution observed in data [43]. Compared to the previous method used to identify $H \rightarrow b\bar{b}$ reconstructed as large-$R$ jets, MC studies [43] show that $b$-tagging based on c.m. subjets can reject more background than the $b$ tagging based on the other subjet algorithm at a given signal identification efficiency: by 20%–50% for large-$R$ jets with $p_T \lesssim 1.5$ TeV and up to a factor of 10 or more for large-$R$ jets with $p_T > 1.5$ TeV. Among several tagging techniques [43] developed to improve the identification of $H \rightarrow b\bar{b}$ with $p_T > 1$ TeV, the c.m. algorithm typically rejects 20% more background at a given signal efficiency.

Studies using MC simulated events show that the correlation between the $b$-tagging efficiencies of two c.m. $b$ subjets is negligible, and thus the $b$-tagging efficiency of each c.m. $b$ subjet in a large-$R$ jet can be calibrated using boosted hadronic top-quark decays $t \rightarrow Wb$ from $t\bar{t} \rightarrow WbWb\bar{b}$ events where one $W$ boson...
decays hadronically and the other decays leptonically. The hadronic products of the boosted $t \rightarrow Wb$ decay are reconstructed as a single large-$R$ jet, in which exactly two c.m. subjets are reconstructed in the jet rest frame: one corresponding to the $b$ quark, and the other corresponding to the $W$ boson. MC studies show that the $b$-tagging performance is almost identical for c.m. $b$ subjets in the boosted hadronic top-quark decay events and $H \rightarrow bb$ events. A standard combinatorial likelihood approach [44] is applied to extract the c.m. $b$-subjett tagging efficiency in order to calculate an MC-to-data scale factor, defined as the ratio of the c.m. $b$-subjett tagging efficiencies measured in data and simulated $t\bar{t}$ events [45]. The scale factor is found to be consistent with unity within its uncertainty and has no significant dependence on the kinematics of the c.m. subjet and the large-$R$ jet. The uncertainty of the scale factor is about 5%, dominated by the systematic uncertainties such as the dependence of the calibration scale factor on the choice of the $t\bar{t}$ MC generators, and the dependence of the MV2c10 [40,41] $b$-tagging scale factors on the jet flavor.

The selected resonance candidates are retained for further analysis if one or both of the c.m. subjets in the large-$R$ jet pass the $b$-tagging requirement, and are assigned to the single- or double-$b$-tagged category, respectively. Afterwards, optimizations of the selection requirements on the photon $p_T (p_T^γ)$ and the large-$R$ jet $p_T (p_T^J)$ are carried out in sequence in order to further improve the search sensitivity. The optimizations are performed separately for the selected events in the single- and double-$b$-tagged categories with the same procedure as used for the large-$R$ jet mass-window optimization described above. It yields $p_T^γ > p_T^0 + a \times m_{J^*/T^*/H}$ and $p_T^J > 0.8 \times (p_T^0 + a \times m_{J^*/T^*/H})$, where $p_T^0 = 12.0 (121.8)$ GeV and $a = 0.35 (0.22)$ for the selected events with $m_{J^*/T^*/H} \leq 2000 (1500)$ GeV in the single-$b$-tagged (double-$b$-tagged) category. For events with $m_{J^*/T^*/H} > 2000 (1500)$ GeV, the selection requirements on the photon and the large-$R$ jet $p_T$ are the same as those for events with $m_{J^*/T^*/H} = 2000 (1500)$ GeV. Depending on the resonance mass, the final signal efficiency in the single- and double-$b$-tagged categories varies between 10% and 20%.

The final discrimination between signal and background is achieved by a simultaneous fit to the $m_{J^*/T^*/H}$ distributions of the selected data events in the single- and double-$b$-tagged categories. The signal probability density function (SPDF) is modeled as a sum of a Crystal Ball function [46] and a small Gaussian component that describes the tails produced by poorly reconstructed resonance candidates. The SPDF parameters extracted from MC simulated events are interpolated as polynomial functions of the resonance mass up to the third order. Afterwards, the parameters of the SPDF at a given resonance mass are fixed to the values determined using the parameterization. The background probability density function (BPDF) is modeled as $B(m_{J^*/T^*/H}) = (1-x)^{p_1} x^{p_2} + p_3 \log(x)$ [47], where $x = m_{J^*/T^*/H}/\sqrt{s}$, $\sqrt{s} = 13$ TeV is the center-of-mass energy, and the three dimensionless shape parameters $p_1$, $p_2$, and $p_3$ are allowed to float in the fit. The choice of the BPDF is motivated and validated by using control data samples containing events that satisfy all the signal selection criteria in either the single- or double-$b$-tagged category, except for the $b$-tagging and large-$R$ jet mass requirements. The selected large-$R$ jet candidates in the control data samples are required to have masses lying in sidebands, whose width varies from 10 GeV to 30 GeV, separated from the Higgs boson signal band by 5 GeV, and to have both of the c.m. subjets failing the $b$-tagging requirement at the 85%-efficiency working point. MC simulated events show that the background $m_{J^*/T^*/H}$ distributions in the single- and double-$b$-tagged categories are well described by the events in the corresponding control sample.

The effect of systematic uncertainties from various sources was studied. The uncertainty of the integrated luminosity is 1.7% [48,49]. Uncertainties resulting from detector effects only affect the calculation of the signal selection efficiencies since the background is estimated from the data. Those uncertainties include effects from the energy and mass scales (2%–6.5%) of the large-$R$ jets [50], the large-$R$ jet energy resolution ($< 0.2\%$) and mass resolution ($18\%–30\%$), the trigger efficiencies ($< 0.1\%$), the photon energy scale and resolution ($< 2\%$) [28], the photon reconstruction, identification and isolation efficiencies ($< 0.1\%$) [8], the $b$-tagging efficiency of the c.m. subjet ($3\%–15\%$), and the pileup modeling ($< 0.5\%$) [51]. In principle, the detector modeling may also affect the SPDF. However, such effects are found to be negligible. The signal efficiency and acceptance are also affected by theoretical uncertainties, such as the PDF choice and initial- and final-state radiation modeling. These are also found to be small ($< 5\%$ from the PDF, $< 1\%$ from parton showering, and $< 1\%$ from renormalization-factorization scale).

The above systematic uncertainties degrade the final limits by 10% at 700 GeV, increasing to around 20% at 2.5 TeV and back to 10% at 4 TeV. Another kind of uncertainty, referred to as the spurious signal, arises from a potential bias in the estimated number of signal events due to the choice of background parameterization. It was estimated by fitting the signal-plus-background model to control data sample $m_{J^*/T^*/H}$ distributions with a control region to signal region background-shape correction factor derived from simulation. The absolute number of fitted signal events at a given $m_{Z'}$ hypothesis value is taken as the number of spurious-signal events, which varies from a few events in the low mass region to less than 0.1 events in the high mass region, and is parameterized as an exponential function of $m_{Z'}$. The signal from a hypothetical $Z'$ resonance is extracted as $\sigma \times B$, defined as its production cross-section times the decay branching fraction $B(Z' \rightarrow H_p)$, by performing an unbinned extended maximum-likelihood fit to the $m_{J^*/T^*/H}$ distributions of the selected events in the
double- and single-\(b\)-tagged categories. The predicted SM value of the \(H \rightarrow b\bar{b}\) decay branching ratio, \(0.582 \pm 0.007\) [52], is used to calculate the upper limit on \(\sigma \times B\) from \(\sigma \times B(Z' \rightarrow H\gamma) \times B(H \rightarrow b\bar{b})\). The fitting range for the double-\(b\)-tagged category is from 0.6 TeV to 4.2 TeV, while for the single-\(b\)-tagged category it is from 1.4 TeV to 4.2 TeV because of poor sensitivity in the low mass region. Systematic uncertainties are taken into account as nuisance parameters with Gaussian sampling distributions [5]. The lowest local (global) \(p\) value is 0.005 (0.412) at 775 GeV, corresponding to a local (global) significance of 2.6\(\sigma\) (0.22\(\sigma\)). No significant signal-like excess is observed and the data are found to be described very well by a background-only fit, as shown in Figs. 1(a) and 1(b). Hypothetical signal distributions for \(m_{Z'} = 2\) TeV and \(m_{Z'} = 3\) TeV with arbitrary normalizations are also plotted in Figs. 1(a) and 1(b) for illustration purposes. Combined upper limits on the signal \(\sigma \times B\) at the 95% confidence level are derived using a modified frequentist method [53,54], with toy MC experiment, taking into account both the statistical and systematic uncertainties. The result as a function of the resonance mass is shown in Fig. 1(c). The better sensitivity and larger integrated luminosity (139 fb\(^{-1}\)) of this search lowers the expected upper limits of this search as compared to that of the previous ATLAS search (139 fb\(^{-1}\)) [5]. The ratio of the current expected upper limits to that of the previous result is about 1/3 (1/15) for resonances with masses below 1.2 TeV (above 2.5 TeV). A similar comparison with that of the previous CMS search (139 fb\(^{-1}\)) [6], where a multivariable approach based on a boosted decision tree was used to identify \(H \rightarrow b\bar{b}\) decays, finds a ratio that varies between 2/5 and 1/3 for masses below 2.5 TeV.

In conclusion, this Letter reports on a search for the production of new heavy resonances decaying into a Higgs boson and a photon, using 139 fb\(^{-1}\) of \(\sqrt{s} = 13\) TeV \(pp\) collision data collected by the ATLAS detector at the LHC. The analysis is performed by reconstructing the hadronic decay of the Higgs boson as a single large-radius jet, targeting the \(H \rightarrow b\bar{b}\) mode. A novel algorithm using information about the jet constituents in the center-of-mass frame of the jet is implemented to identify the two \(b\) quarks in the jet and enhances the sensitivity of the search. No significant excess of events is observed above the expected background. Upper limits are set on the production cross-section times branching fraction for resonance decays into a Higgs boson and a photon in the resonance mass range from 0.7 to 4 TeV, which is significantly wider than in the previous ATLAS and CMS searches.

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; ANID, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DSNRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRT, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MICINN, 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, USA. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada, CRC and IVADO, Canada; Beijing Municipal Science & Technology Commission, China; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CERN 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. [56].


[11] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point 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, φ) are used in the transverse plane, φ being the azimuthal angle around the z axis. The pseudorapidity is defined in terms of the polar angle θ as η = −ln tan(θ/2). The distance between two objects in η − φ space is ΔR = (∆η)2 + (∆φ)2. Transverse momentum is defined by pT = p sin θ.


mm Also at TRIUMF, Vancouver, British Columbia, Canada.
nn Also at Universita di Napoli Parthenope, Napoli, Italy.
o Also at University of Chinese Academy of Sciences (UCAS), Beijing, China.