Search for Higgs boson decays into a pair of pseudoscalar particles in the $bb\mu\mu$ final state with the ATLAS detector in $pp$ collisions at $\sqrt{s}=13$ TeV

G. Aad et al.*
(ATLAS Collaboration)

(Received 4 October 2021; accepted 22 November 2021; published 11 January 2022)

This paper presents a search for decays of the Higgs boson with a mass of 125 GeV into a pair of new pseudoscalar particles, $H \rightarrow aa$, where one $a$-boson decays into a $b$-quark pair and the other into a muon pair. The search uses 139 fb$^{-1}$ of proton-proton collision data at a center-of-mass energy of $\sqrt{s} = 13$ TeV recorded between 2015 and 2018 by the ATLAS experiment at the LHC. A narrow dimuon resonance is searched for in the invariant mass spectrum between 16 GeV and 62 GeV. The largest excess of events above the Standard Model backgrounds is observed at a dimuon invariant mass of 52 GeV and corresponds to a local (global) significance of 3.3$\sigma$ (1.7$\sigma$). Upper limits at 95% confidence level are placed on the branching ratio of the Higgs boson into invisible and BSM states is determined to be less than 47% at 95% confidence level (CL) [27]. Combined measurements of Higgs boson production cross sections and branching ratios constrains the branching ratios into invisible and undetected states to be $B(H \rightarrow \text{inv}) < 30\%$ and $B(H \rightarrow \text{undetected}) < 21\%$, respectively, whereas the overall branching fraction of the Higgs boson into beyond-the-SM (BSM) states is determined to be less than 47% at 95% confidence level (CL) [27]. Combined measurements of Higgs boson couplings performed by the CMS Collaboration set upper limits of $B(H \rightarrow \text{inv}) < 22\%$ and $B(H \rightarrow \text{undetected}) < 38\%$ at 95% CL [28]. This motivates searches for light states in the Higgs boson decays that probe this potentially large $B(H \rightarrow \text{BSM})$.

I. INTRODUCTION

Light (pseudo) scalars that couple to the 125 GeV Higgs boson [1,2] appear in many well-motivated extensions of the Standard Model (SM) [3–8]. These include models addressing the baryon asymmetry of the universe [9,10], offering a solution to the naturalness problem [11,12], or providing insights into the nature of dark matter [13–19]. Light bosons produced in Higgs boson decays could also be mediators to dark sectors that do not couple to the SM otherwise [20–24]. Furthermore, pseudoscalar mediators appear in models, such as those described in Ref. [25], that were proposed to explain the anomalous muon magnetic moment [26]. A combination of ATLAS measurements of the Higgs boson production cross sections and branching ratios constrains the branching ratios into invisible and undetected states to be $B(H \rightarrow \text{inv}) < 30\%$ and $B(H \rightarrow \text{undetected}) < 21\%$, respectively, whereas the overall branching fraction of the Higgs boson into beyond-the-SM (BSM) states is determined to be less than 47% at 95% confidence level (CL) [27]. Combined measurements of Higgs boson couplings performed by the CMS Collaboration set upper limits of $B(H \rightarrow \text{inv}) < 22\%$ and $B(H \rightarrow \text{undetected}) < 38\%$ at 95% CL [28]. This motivates searches for light states in the Higgs boson decays that probe this potentially large $B(H \rightarrow \text{BSM})$.

This paper presents a search for decays of the 125 GeV Higgs boson into two pseudoscalars, denoted by $a$, in proton-proton ($pp$) collisions at the LHC [29]. The search is performed in events where one $a$-boson decays into two $b$-quarks and the other into two muons, $H \rightarrow aa \rightarrow bb\mu\mu$. The $a$-bosons are assumed to have a decay width that is narrow compared to the detector resolution. As pseudoscalar couplings are generally proportional to mass, which is for example the case in two-Higgs-doublet models [20,30], the $bb\mu\mu$ final state provides a good balance between a high branching ratio from the $a \rightarrow bb$ decay and a clean, high mass-resolution, dimuon resonance signature that is easy to trigger on from the $a \rightarrow \mu\mu$ decay. In scenarios with enhanced lepton couplings, the $a \rightarrow \mu\mu$ branching ratio can also be relatively large, resulting in $B(H \rightarrow aa \rightarrow bb\mu\mu)/B(H \rightarrow aa)$ of up to 0.16% [31].

Light resonances in Higgs boson decays have been searched for by ATLAS and CMS in many different channels, i.e., in the final states involving $4\mu$ [32,33], $2\mu2\tau$ or $4\tau$ [34–38], $2b2\tau$ [39], $4b$ [40,41], $4\gamma$ [42], and $2\gamma + 2$-jets [43]. A search for a dimuon resonance produced in association with b-jets has been performed by CMS [44] and a light resonance decaying to two muons has been searched for by LHCb [45]. CMS has performed a search for $H \rightarrow aa \rightarrow bb\mu\mu$ in 35.9 fb$^{-1}$ of $pp$ collision data at a center-of-mass energy of $\sqrt{s} = 13$ TeV that sets upper limits on $B(H \rightarrow aa \rightarrow bb\mu\mu)$ of $(1-7) \times 10^{-4}$ for $a$-boson masses ($m_a$) in the range $20 \leq m_a \leq 62.5$ GeV [46]. The ATLAS search based on 36 fb$^{-1}$ of Run 2 data [47] sets...
upper limits on $B(H \rightarrow aa \rightarrow bb\mu\mu)$ between $1.2 \times 10^{-4}$ and $8.4 \times 10^{-4}$ for $a$-boson masses in the range $20 \leq m_a \leq 60$ GeV. In this paper, the full Run 2 dataset corresponding to an integrated luminosity of 139 fb$^{-1}$ is used and the search is extended down to $m_a = 16$ GeV and up to $m_a = 62$ GeV. Additionally, boosted decision tree (BDT) techniques are used to improve the separation of the signal from the SM backgrounds, increasing the analysis sensitivity, especially for higher $m_a$.

II. ATLAS DETECTOR

The ATLAS experiment [48,49] is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and a nearly 4π coverage in solid angle. It consists of an inner detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic (EM), and hadron calorimeters, and a muon spectrometer (MS). The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic energy measurements with high granularity. A steel/scintillator-tile hadron calorimeter covers the central pseudorapidity range $|\eta| < 1.7$. The endcap and forward regions are instrumented with LAr calorimeters for both the EM and hadronic energy measurements up to $|\eta| = 4.9$. The MS surrounds the calorimeters and is based on three large superconducting air-core toroidal magnets with eight coils each. The field integral of the toroids ranges between 2.0 Tm and 6.0 Tm across most of the detector. The MS includes a system of precision tracking chambers and fast detectors for triggering. A two-level trigger system is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to accept events at a rate below 100 kHz [50]. This is followed by a software-based trigger that reduces the accepted event rate to 1 kHz on average. An extensive software suite [51] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

III. DATASET AND SIMULATED EVENTS

The data used in this analysis were collected in Run 2 of the LHC during the 2015–2018 data-taking period with $pp$ collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV. The dataset corresponds to an integrated luminosity of 139 fb$^{-1}$. The lowest-threshold unprescaled single-muon and dimuon triggers are used to select the events [52]. Single-muon triggers require the transverse momentum ($p_T$) of the muon to be above 20 or 26 GeV, depending on the data-taking period, while the dimuon trigger requires both muons to have a $p_T$ above 14 GeV.

Simulated events are used in the estimation of the SM backgrounds. SHERPA 2.2.1 [53,54] was used as the baseline generator for the Drell–Yan (DY) + jets, $W(\rightarrow l\nu) +$ jets, diboson and triboson backgrounds. It is a multiparton matrix element and parton shower (PS) generator including hadronization [55–59], with the NNPDF3.0 parton distribution function (PDF) set at next-to-next-to-leading-order (NNLO) accuracy [60]. The DY + jets and multiboson samples were generated with a minimum dilepton mass of 10 and 4 GeV, respectively. The $t\bar{t}$ and single-top-quark samples were generated with Powheg-Box v2 [61–65] using the NNPDF3.0hPDF PDF in matrix element interfaced to PYTHIA 8.230 [66] for the PS. For the underlying-event description a set of tuned parameters called the A14 tune [67] was used, along with the NNPDF2.3LO PDF [68]. The $t\bar{t}$ + vector-boson processes ($t\bar{t} + V$) were generated with MadGraph5_aMC@NLO 2.3.3 [69] interfaced to PYTHIA 8.210 for the PS. The underlying-event tune was the same as for the $t\bar{t}$ sample. EvtGen [70] was used for the properties of the bottom and charm hadron decays in all simulated samples, except those simulated with SHERPA.

Higgs boson production through gluon–gluon fusion (ggF) was generated using the NNLOPS program [71,72] with Powheg-Box v2 [61,63,73,74]. The vector-boson fusion (VBF) processes were generated with Powheg-Box v2 at NLO accuracy [75]. The Higgs boson mass was set to 125 GeV. For both the ggF and VBF production processes, Powheg-Box was interfaced with PYTHIA 8.212 using the AZNLO tune [76] for the simulation of the $H \rightarrow aa \rightarrow bb\mu\mu$ decays, where the $a$-boson is a pseudoscalar, as well as for parton showering, hadronization and the underlying event. The ggF Higgs boson production rate is normalized to the total cross section predicted at next-to-next-to-next-to-leading-order accuracy in QCD with NLO electroweak corrections applied [77–81] and amounts to 48.58 pb. The VBF production rate is normalized to an approximate NNLO cross section with the NLO electroweak corrections applied [82–85], which amounts to 3.8 pb. The contribution from the associated production of a Higgs boson and a vector boson (VH) is calculated to be 3.5% of the total ggF + VBF cross section and is accounted for by scaling the simulated ggF and VBF samples. The contribution from Higgs boson production in association with a pair of top quarks is found to be negligible (below the percent level) and is neglected in the analysis. Thirteen mass points were simulated for the ggF and VBF production modes, with the
$\alpha$-boson mass in the range $m_\alpha = 16$–62 GeV.\(^3\) Below $m_\alpha = 16$ GeV the $b$-quarks coming from the decays of the $\alpha$-boson tend to be so collimated due to its boost that they cannot be reconstructed as two separate $b$-jets (with a radius parameter of $R = 0.4$). Another effect is that in the highly asymmetric decays of low-mass $\alpha$-bosons, the subleading $b$-jet falls below the jet reconstruction threshold of 20 GeV [86]. As a result, the signal acceptance falls below 0.2% and the analysis loses sensitivity.

The effects of additional interactions in the same and neighboring beam-bunch crossings (pileup) were modeled for all simulated events by overlaying additional $pp$ collisions generated with PYTHIA 8.186 using the NNPDF2.3LO PDF set and the A3 tune [87]. Simulated event samples are weighted to reproduce the distribution of the number of pileup interactions observed in the data. All the generated background and signal samples are processed through the ATLAS detector simulation [88] based on GEANT4 [89] and reconstructed using the same software as for the data.

IV. EVENT RECONSTRUCTION AND SELECTION

Muons are reconstructed by combining track information from the MS with tracks found in the ID [90]. They also have to satisfy $p_T > 5$ GeV and $|\eta| < 2.7$ (for $|\eta| > 2.5$, only tracking information from the MS is used), and pass the LowPt working point identification requirement defined in Ref. [90]. Muon tracks must have a longitudinal impact parameter $z_0$ satisfying $|z_0 \sin \theta| < 0.5$ mm and a transverse impact parameter significance $|d_0|/\sigma_{d_0} < 3$ relative to the primary interaction vertex, chosen as the reconstructed vertex with the highest sum of the $p_T^2$ of its associated tracks. Furthermore, muons are required to be isolated from the MS with tracks found in the ID [90]. They also have to satisfy $p_T > \text{LowPt}$ the sum of the transverse momentum of additional inner detector tracks and the vertex with the highest sum of the $p_T^2$ of its associated tracks.

Jets are reconstructed using the anti-$k_T$ algorithm [91] implemented in the FastJet package [92] with a radius parameter of $R = 0.4$. The inputs to the jet clustering are built by combining the information from both the calorimeters and the ID using a particle-flow algorithm [86,93]. Jets with $p_T < 60$ GeV originating from pileup are suppressed with the jet-vertex-tagger (JVT) [94], a multivariate algorithm combining track-based variables. Selected jets are required to have $p_T > 20$ GeV and $|\eta| < 2.5$. An algorithm (MV2c10) relying on multivariate techniques, taking as input the properties of displaced tracks and vertices reconstructed within a jet, is employed to identify (tag) jets containing $b$-hadrons [95]. The MV2c10 tagger is used at 77% $b$-jet identification efficiency, with an approximate misidentification probability of 25% for jets arising from charm quarks, 6.3% for hadronically decaying $\tau$-leptons, and 0.8% for light-flavor jets as measured in simulated $t\bar{t}$ events.

The missing transverse momentum ($E_T^{\text{miss}}$) is calculated as the magnitude of the negative vector sum of the transverse momenta of all the reconstructed and calibrated objects in the event, including a soft term that accounts for charged particles that are associated with the primary vertex, but not with any reconstructed object [96,97].

The events selected for the analysis are required to have two muons of opposite charge, either with the leading and subleading muons satisfying $p_T^{\text{leading}} > 27$ GeV and $p_T^{\text{subleading}} > 5$ GeV, and the event being triggered by a single-muon trigger, or with both muons having $p_T > 15$ GeV, and the event being triggered by a dimuon trigger. The dimuon invariant mass, $m_{\mu\mu}$, is required to be between 15 and 65 GeV. Furthermore, the events must contain exactly two $b$-tagged jets with $p_T$ above 20 GeV.

A kinematic likelihood (KL) [98] fit exploiting the equal invariant masses of the $bb$ and $\mu\mu$ systems in $H \to aa$ decays is performed to improve the four-body invariant mass ($m_{bb\mu\mu}$) resolution and reduce the SM backgrounds. The same fit approach as considered in the previous ATLAS publication [47] is used. The dimuon invariant mass, $m_{\mu\mu}$, is used to constrain the di-$b$-jet mass, as the former has a resolution approximately ten times better than the latter. The $m_{\mu\mu}$ resolution ranges between 0.4 GeV at $m_\alpha = 16$ GeV and 1.3 GeV at $m_\alpha = 62$ GeV. The fit maximizes the likelihood by shifting the $b$-jet energies within the resolution in order to satisfy the constraint $m_{\mu\mu} \approx m_{bb}$. The output of the fit is the logarithm of the maximum likelihood value, $\ln(L_{\text{max}})$, which quantifies how well the event matches the $m_{\mu\mu} = m_{bb}$ hypothesis, characteristic of signal events. The four-body invariant mass, recomputed after the KL fit, is denoted by $m_{\mu\mu}^{\text{KL}}$ and is used for further event categorization.

Signal-like events are chosen by requiring that $110 < m_{bb\mu\mu}^{\text{KL}} < 140$ GeV, and that $\ln(L_{\text{max}}) > -8$, which ensures that $m_{bb}$ is compatible with $m_{\mu\mu}$. Finally, $E_T^{\text{miss}}$ is required to be less than 60 GeV to reduce the background from $t\bar{t}$ events, which is one of the two major backgrounds and can contain large $E_T^{\text{miss}}$ from neutrinos in top-quark decays. This selection defines the “inclusive” signal region (SRinc) and is summarized in Table I, along with the selection requirements for other analysis regions described later in the text.

A BDT classifier implemented using the TMVA framework [99] is employed to further reduce the SM backgrounds. Its training is done in partially overlapping 8-GeV-wide $m_{\mu\mu}$ windows centered at the $m_\alpha$ values of

---

\(^3\)More specifically, the simulated mass points are at $m_\alpha = 16, 18, 20, 25, 30, 35, 40, 45, 50, 52, 55, 60, \text{and } 62$ GeV.
TABLE I. Summary of the selection requirements for the control (TCR and DYCR), validation (VR1 and VR2), and inclusive signal (SRincl) regions in the analysis, as well as the final SR bins. The control and validation regions are defined in Sec. V.

<table>
<thead>
<tr>
<th>$m_{bb}$ (GeV)</th>
<th>TCR</th>
<th>DYCR</th>
<th>SRincl</th>
<th>VR1</th>
<th>VR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{bb}^{KL}$ (GeV)</td>
<td>[110, 140]</td>
<td>[80, 110] or [140, 170]</td>
<td>[110, 140]</td>
<td>[170, 300]</td>
<td>[110, 140]</td>
</tr>
<tr>
<td>$E_{T}^{\text{miss}}$ (GeV)</td>
<td>&gt; 60</td>
<td>&lt; 60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(L_{\text{max}})$</td>
<td>&gt; - 8</td>
<td></td>
<td></td>
<td>[-11, -8]</td>
<td></td>
</tr>
<tr>
<td>SR bins</td>
<td></td>
<td></td>
<td>SRincl &amp; BDT$<em>{m</em>{a}}$ &gt; 0.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2-GeV-wide (3-GeV-wide) $m_{bb}$ bins for $m_{a}$ ≤ 45 GeV ($m_{a}$ > 45 GeV)

each of the 12 generated signals, in order to fully exploit their kinematic differences. The background sample consists of $t\bar{t}$ and DY + jets events, the two dominant backgrounds, combined in the proportions extracted from the background validation fit described in Sec. VII. The signal samples used for the training include ggF and VBF Higgs boson production samples combined according to their cross sections. The seven kinematic variables included in the training are:

(i) $m_{bb}$,
(ii) $\ln(L_{\text{max}})$,
(iii) $\Delta R_{b_{1}b_{2}}$ (the angular distance between the two $b$-jets),
(iv) $\text{diff} \Delta R_{b_{1}b_{2}} = \Delta R_{b_{1}b_{2}} - \Delta R_{\mu_{1}\mu_{2}}$ (the difference between the angular separations between the two $b$-jets and the two muons),
(v) $\Delta R_{bb\mu}$ (the angular distance between the $bb$ and $\mu\mu$ systems),
(vi) $\overline{\Delta R}_{b_{1}\mu} = |\Delta R_{b_{1}p_{1}} + \Delta R_{b_{1}p_{2}} + \Delta R_{p_{1}p_{2}} + \Delta R_{p_{2}p_{2}}|/4$ (the average angular distance of all four combinations of a $b$-jet and a muon),
(vii) $\overline{m}_{p_{12}} = (m_{b_{1}p_{1}} + m_{b_{1}p_{2}} + m_{b_{2}p_{1}} + m_{b_{2}p_{2}})/4$ (the average mass of all four combinations of a $b$-jet and a muon).

The distributions of these variables for the background and three representative signal masses are shown in Fig. 1.

The $m_{bb}$ variable helps separate the low-mass signal from the backgrounds, as $m_{bb}$ peaks around 60 GeV for the $t\bar{t}$ and DY processes. The $\ln(L_{\text{max}})$ peaks at higher values as the signal mass becomes smaller.

Due to a higher boost of a lighter $a$-boson, its decay products are collimated, resulting in $\Delta R_{b_{1}b_{2}}$ and $\Delta R_{\mu_{1}\mu_{2}}$ being much smaller than for a signal from a heavier $a$-boson or for background processes. As a consequence, $\text{diff} \Delta R_{b_{1}b_{2}}$ shows a narrow distribution centered around zero, while the background and a higher-mass signal exhibit a much broader $\text{diff} \Delta R_{b_{1}b_{2}}$ distribution.

The $\Delta R_{bb\mu}$ variable helps enhance the sensitivity to higher signal masses. Heavier $a$-bosons are produced approximately at rest, resulting in the $\Delta R_{bb\mu}$ distribution being relatively flat with a small peak at low values. As the signal mass decreases, the $\Delta R_{bb\mu}$ distribution transitions into a “back-to-back” topology, characteristic of both a low-mass signal and the background events.

Finally, the $\overline{\Delta R}_{b_{1}\mu}$ and $\overline{m}_{b_{1}\mu}$ variables provide another measure of how close the two $a$-bosons are in $\Delta R$. In the back-to-back topology for lower signal masses, the muons are, on average, further away from the $b$-jets, while for heavier $a$-bosons produced approximately at rest, the average distance between the muons and the $b$-jets is smaller. Consequently, $\overline{\Delta R}_{b_{1}\mu}$ and $\overline{m}_{b_{1}\mu}$ peak at high (low) values for low (high) signal masses, while the backgrounds peak somewhere between the two extreme signal topologies.

The output score of the BDT trained for a signal with mass $m_{a}$ is denoted by BDT$_{m_{a}}$. The BDT$_{m_{a}}$ distributions for $m_{a}$ = 20, 40, and 60 GeV are shown in Fig. 2.

The final signal region (SR) bin for each signal mass is defined by imposing two requirements in addition to the SRincl selection: $m_{a} - X < m_{\mu\mu} < m_{a} + X$ and BDT$_{m_{a}}$ > 0.2, where $X = 1$ GeV ($X = 1.5$ GeV) for $m_{a}$ ≤ 45 GeV ($m_{a}$ > 45 GeV). The widths of the SR bins and the BDT$_{m_{a}}$ cut value are optimized to maximize the significance of signal over background events. For masses at which no signal sample was generated, and, consequently, no BDT was trained, the BDT trained for the $m_{a}$ closest to the one being tested is used. For example, when testing the $m_{a}$ = 32 GeV hypothesis, the requirement BDT30 > 0.2 is applied to select the events for the SR bin. Signal yields for mass points where no signal sample was generated ($m_{a}$ = 32 GeV in this example) are obtained by selecting events with BDT scores above 0.2 for the same BDT$_{m_{a}}$ (BDT30 in this case) in all simulated mass points and interpolating using third-order splines. To assess the uncertainty, the yields of the neighboring simulated mass points ($m_{a}$ = 30 GeV and $m_{a}$ = 35 GeV in this case) are interpolated using a linear function. The difference between the yields obtained using the splines and a linear function for the interpolation is assigned as a systematic uncertainty on the interpolated signal yield.

4One BDT was trained for each generated signal MC sample, except for $m_{a}$ = 52 GeV, as this sample was produced only at a later analysis stage.
Using a BDT at a mass for which the training was not performed results in a negligible loss of significance relative to a BDT that was optimized for that mass point.

The signal acceptance × efficiency varies between 0.3% and 2.5% for ggF Higgs boson production and between 0.2% and 3.0% for VBF production, where the lowest
acceptance × efficiency is obtained for the lowest $m_{\tau\tau}$, and grows as $m_{\tau\tau}$ increases. The largest loss of acceptance occurs when requiring that there are two $b$-jets in the event, as one of the signal jets tends to fall below the reconstruction threshold of 20 GeV. The fraction of signal events passing the two-$b$-jet requirement is less than 20% for all mass points.

V. BACKGROUND ESTIMATION

The dominant backgrounds in the analysis arise from the DY dimuon process in association with $b$-quarks and pair production of top quarks ($t\bar{t}$) where each $W$ boson decays into a muon and a neutrino. These two backgrounds account for more than 96% of background events in all analysis regions.

Two control regions are designed to constrain the $t\bar{t}$ and DY backgrounds. They are chosen so that they have negligible signal contamination, are kinematically as close as possible to SRincl, and maximize the contribution of one of the respective background processes. A top-quark control region (TCR) is defined by inverting the $E_T^{miss}$ selection criterion in SRincl to $E_T^{miss} > 60$ GeV. This results in an event sample approximately 93% pure in $t\bar{t}$ events. The DY control region (DYCR) is defined in the 30 GeV-wide $m_{bb\mu\mu}^{KL}$ sidebands of SRincl, i.e., by requiring $80 < m_{bb\mu\mu}^{KL} < 110$ GeV or $140 < m_{bb\mu\mu}^{KL} < 170$ GeV. Approximately 50% of the events in DYCR originate from the DY process, whereas the rest mostly come from $t\bar{t}$ production. Two validation regions (VR1 and VR2) are used to validate the normalizations of the backgrounds. VR1 is defined in the $170 < m_{bb\mu\mu}^{KL} < 300$ GeV range, while VR2 is obtained by inverting the $\ln(L_{max})$ selection criterion of SRincl to $-11 < \ln(L_{max}) < -8$. All the analysis regions are summarized in Table I and illustrated in Fig. 3.

The shapes of the $t\bar{t}$ kinematic variable distributions are obtained from simulation, while the overall normalization

![Figure 2](https://example.com/figure2.png)

**FIG. 2.** Three BDT $m_{\tau\tau}$ distributions, BDT20, BDT40, and BDT60, plotted in the $m_{\mu\mu}$ windows of SRincl, as indicated in the figures. The distributions are normalized to unit area. The background histogram is the sum of the $t\bar{t}$ and DY event templates, combined in the proportions extracted from the background validation fit described in Sec. VII.

![Figure 3](https://example.com/figure3.png)

**FIG. 3.** Illustration of the signal, control, and validation regions used in the analysis. VR2 (not shown) is defined by the same selection as SRincl, except that the requirement on $\ln(L_{max})$ is inverted to $-11 < \ln(L_{max}) < -8$. 

G. AAD et al. PHYS. REV. D 105, 012006 (2022)
is extracted from the fits described in Sec. VII. The distributions for the DY background are taken from data templates because the limited sizes of the simulated event samples do not allow a reliable estimate. The template regions are defined in the same way as the analysis regions in Table I, except that the two-b-tag requirement is replaced by a zero-b-tag requirement. The template regions are >95% pure in DY events. Contributions from other processes, namely $t\bar{t}$, $W +$ jets, diboson and single-top, are subtracted using simulation. Following the subtraction, the DY templates are corrected to account for kinematic differences between event samples dominated by jets originating from light quarks or gluons (template regions) and event samples dominated by $b$-jets (analysis regions). The correction is applied as a per-event weight, where the reweighting is derived from a comparison between two-b-tag and zero-b-tag kinematic distributions in simulated DY events. Two sets of event weights are derived and applied sequentially. First, the jet multiplicity of the zero-b-tag MC sample is reweighted to the one in the two-b-tag sample. It is the distribution with the largest difference between the zero- and two-b-tag samples and was hence corrected first. Second, a BDT-based reweighting is employed to further correct the zero-b-tag template kinematics. A BDT is trained on the zero-b-tag versus the two-b-tag simulated DY samples. The BDT input consists of kinematic properties and angular distributions of the $b$-jets, muons and the two corresponding $a$-boson candidates, as well as $E_T^{\text{miss}}$ and $m_{bbpp}^{\text{KL}}$. The ratio of the BDT score distributions obtained for the two-b-tag and zero-b-tag simulated events is then applied as a weight to every event from the zero-b-tag DY template, as a function of its BDT score. Following the BDT-based reweighting, the $m_{bbpp}^{\text{KL}}$ and $E_T^{\text{miss}}$ distributions are corrected by up to 20%. The DY templates are normalized to data in the fits described in Sec. VII.

Minor backgrounds include diboson and single-top-quark production, production of a $t\bar{t}$ pair in association with a vector boson, and $W$ boson production in association with $b$-jets. The estimation of these minor backgrounds relies purely on simulation normalized to the best available theoretical prediction. The events where a jet is misidentified as a muon are taken into account as follows: nonprompt/missidentified muons in $W +$ jets and $t\bar{t}$ events are included in the analysis on the basis of simulation, any contribution of nonprompt/missidentified muons in the DY + jets component is accounted for by the data template, and the potential contribution from multijet events is found to be negligible.

**VI. SYSTEMATIC UNCERTAINTIES**

Systematic uncertainties in the analysis are divided into three categories: experimental uncertainties affecting the simulated background and signal processes, uncertainties in the modeling of the DY template, and theoretical uncertainties of the simulated background and signal samples. Table II shows a summary of the dominant systematic uncertainties in the total background and signal yields in the signal region bins, as resulting from the fits described in Sec. VII and hereafter denoted by “postfit”.

 Among the experimental uncertainties, the leading effects come from those associated with the calibration and resolution of jet energies [100], and with the measurement of the $b$-tagging efficiency [95]. The impact of these uncertainties on the total background (signal) yields in the SR bins is as large as 3% (10%). The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [101], obtained using the LUCID-2 detector [102] for the primary luminosity measurement. Other uncertainties, such as those arising from the muon identification efficiency, momentum scale and resolution [90,103], and pileup are found to have a negligible impact on the final yields.

The uncertainty arising from limited MC sample sizes ranges from 8% to as large as 40% in the low $m_{\mu\mu}$ mass bins due to there being few $t\bar{t}$ events in this region.

Five sources of uncertainty in the data-driven DY template are considered. The uncertainty in subtracting non-DY events from the non-reweighted template in the signal region is expressed as a percentage of the total background and signal yields in the signal region bins, as resulting from the fits described in Sec. VII and hereafter denoted by “postfit”.

**TABLE II. Summary of the dominant postfit systematic uncertainties in the background and signal yields. The uncertainties are expressed as a percentage of the total background and signal yields per $m_{\mu\mu}$ bin of the signal region. Only uncertainties exceeding 2% in at least one SR bin are shown.**

<table>
<thead>
<tr>
<th>Category</th>
<th>Source</th>
<th>Total background (%)</th>
<th>Signal (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DY</td>
<td>BDT$m_{\mu\mu}$ selection</td>
<td>7–14</td>
<td>⋯</td>
</tr>
<tr>
<td></td>
<td>Normalization</td>
<td>5–10</td>
<td>⋯</td>
</tr>
<tr>
<td></td>
<td>$m_{\mu\mu}$ shape</td>
<td>1–8</td>
<td>⋯</td>
</tr>
<tr>
<td></td>
<td>Kinematics</td>
<td>0.3–6</td>
<td>⋯</td>
</tr>
<tr>
<td></td>
<td>Background subtraction</td>
<td>0.6–3</td>
<td>⋯</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>Hadronization/PS</td>
<td>0.3–4</td>
<td>⋯</td>
</tr>
<tr>
<td></td>
<td>Hard-scatter generation</td>
<td>0.2–3</td>
<td>⋯</td>
</tr>
<tr>
<td></td>
<td>Normalization</td>
<td>0.2–3</td>
<td>⋯</td>
</tr>
<tr>
<td>Overall MC</td>
<td>Sample statistics</td>
<td>8–40</td>
<td>1–2</td>
</tr>
<tr>
<td>Jets</td>
<td>$b$-tagging</td>
<td>0.03–0.7</td>
<td>9–10</td>
</tr>
<tr>
<td></td>
<td>Jet-energy resolution</td>
<td>1–3</td>
<td>6–7</td>
</tr>
<tr>
<td></td>
<td>Jet-energy scale</td>
<td>1–3</td>
<td>4–5</td>
</tr>
<tr>
<td>Signal</td>
<td>FSR</td>
<td>⋯</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>⋯</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>VH contribution</td>
<td>⋯</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>MPI</td>
<td>⋯</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>QCD scale</td>
<td>⋯</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>ISR</td>
<td>⋯</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>ggF cross section</td>
<td>⋯</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>-missing higher-order QCD</td>
<td>⋯</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-PDF &amp; $a_S$</td>
<td>⋯</td>
<td>3</td>
</tr>
</tbody>
</table>
analysis regions is assessed by comparing the nominal template, for which the simulated non-DY backgrounds had been subtracted before reweighting, with an alternative template for which no subtraction had been performed. The uncertainties in the template kinematics modeling are derived by comparing the DY template with simulation in two key variables: $E_{T}^{miss}$ and $m_{bb\mu\mu}^{KL}$. The ratios of the template to the simulated DY events are fit with linear functions and used in assigning uncertainties to the shapes of the $E_{T}^{miss}$ and $m_{bb\mu\mu}^{KL}$ distributions. Similarly, the uncertainty in the $m_{bb\mu\mu}$ template shape is assessed by comparing the template with the smoothed simulated sample and applying the observed difference as a systematic uncertainty. The uncertainty in the normalization of the DY template is obtained from the fits to data. Finally, the uncertainty in the efficiency of the $BDT_{m_{a}}$ selection criteria is evaluated by taking the difference in the $BDT_{m_{a}}$ cut efficiency, $\frac{N_{DY \text{ events SR}}}{N_{DY \text{ events SR}}}$, between the template and the simulation. All one-sided DY template uncertainties are symmetrized around the nominal value.

To assess the uncertainties in the generation of the hard-scatter $t\bar{t}$ process, the Powheg sample is compared with a sample generated using MadGraph5_aMC@NLO 2.3.3. The hadronization and fragmentation uncertainties in the PS are evaluated by comparing the nominal sample showered by PYTHIA 8.230 with an alternative sample generated by Powheg using the same PDF in matrix element as for the nominal sample, but showered with HERWIG 7.0.4 [104,105]. The initial- and final-state radiation (ISR and FSR) uncertainties of the $t\bar{t}$ sample are assessed by varying the internal PYTHIA 8.230 showering parameters. Finally, the uncertainties due to the PDF choice are evaluated using the internal variations of the nominal PDF4LHC15_NLO_30 set [106].

Uncertainties in the calculation of the ggF and VBF Higgs boson production cross sections are assessed by following the recommendations of the LHCHiggs Working Group given in Refs. [77,82]. As no VH signal sample was generated, a conservative 100% uncertainty is assigned to the estimated VH yield. To evaluate the uncertainties due to the PDF choice, the yields obtained with the baseline NNPDF30_NLO_AS_0118 set are compared with the yields obtained using the internal variations of NNPDF30_NLO_AS_0118 and with the yields obtained with the nominal MMHT2014NLO68CLAS118 [107] and CT14NLO [108] sets. The largest difference is taken as the overall PDF uncertainty for all signal mass points. Furthermore, the effects of uncertainties in the ISR, FSR, multiparton interactions (MPI) in PYTHIA, parton showering, and renormalization and factorization scales are also assessed. Uncertainties from these sources have an impact of 1–6% on the signal yields, with the largest contributions arising from the ggF production cross section and FSR uncertainties.

VII. ANALYSIS AND RESULTS

The final background and signal estimates are obtained in a set of binned likelihood fits [109] using the HistFitter [110] package. The likelihood is a product of Poisson probability functions, describing the observed and predicted numbers of events in each region, and Gaussian distributions that constrain the nuisance parameters associated with the systematic uncertainties. In the background validation fit, the data in TCR and DYCR are used to extract the normalization of the $t\bar{t}$ and DY backgrounds, respectively. As the $t\bar{t}$ sample in TCR is modeled very well, it is implemented as only one bin in the fit, whereas DYCR is divided into five equal-width bins in $m_{bb\mu\mu}$ to provide greater sensitivity to the DY template shape. The purpose of this fit is to validate the modeling of the background in the control and validation regions and in SRincl. The fitted $t\bar{t}$ normalization factor is $\mu_{t\bar{t}} = 1.07^{+0.06}_{-0.07}$, while the value of $\mu_{DY}$ has no physical meaning because it is scaled from a template region and is thus not quoted. Figures 4 and 5 show postfit distributions of $m_{bb\mu\mu}^{KL}$, $E_{T}^{miss}$, $\ln(L_{max}^{\mu})$, and $m_{bb\mu\mu}$ spanning various analysis regions, while Fig. 6 shows BDT20 and BDT50 in SRincl. Good agreement between the estimated backgrounds and the data is observed in the kinematic distributions. In SRincl, 1185 events are observed, which is compatible with the total estimated background of 1155.3 ± 13.6. The yields in several representative SR bins, i.e., $m_{bb\mu\mu}$ windows after applying the BDT selection, as obtained from the background validation fit above, are shown in Table III. When comparing the systematic uncertainty with the statistical uncertainty, it can be seen that the analysis is clearly statistically limited. Figure 7 shows the data and the estimated backgrounds in all final SR bins. Due to the limited statistics of the background samples, the estimates are not perfectly smooth; however, the bin-to-bin fluctuations are much smaller than the statistical uncertainty of the data. Larger jumps, which occur at $m_{bb\mu\mu} = 23$, 28, 33, 38 GeV etc., appear when the BDT discriminant used for the selection changes from the one trained in the lower mass range to the one trained in the higher mass range.

To test for the presence of new phenomena, fits are performed for each of the 47 hypothesized signal masses in the range 16 ≤ $m_{bb\mu\mu}$ ≤ 62 GeV in 1 GeV steps. It was verified that the analysis is also sufficiently sensitive to a signal with $m_{bb\mu\mu}$ centered in between these 1 GeV steps. TCR, DYCR, and the respective SR bin are included in each fit in order to constrain the backgrounds and the signal to the data.

A model-independent fit, i.e., not including any signal sample, is performed to test whether the data are compatible with the background-only hypothesis. The result is a scan of $p_{0}$-values as shown in Fig. 8. The largest discrepancy is found at $m_{bb\mu\mu} = 52$ GeV, corresponding to a local (global) $p_{0}$-value of 0.00054 (0.048) and a local (global)
The hatched bands show the total postfit statistical and systematic uncertainties of the backgrounds. The histogram labeled as “Other” in the legend includes the contributions from the diboson, single-top-quark, $t\bar{t} + V$ and $W + \text{jets}$ backgrounds.

FIG. 5. Postfit $\ln(L_{\text{max}})$ in VR2 and SRincl (left); $m_{\mu\mu}$ in SRincl (right). No selection based on the BDT discriminants is applied in the analysis regions shown in the figures. The signal distributions are normalized to the SM Higgs boson cross section (including ggF, VBF, and VH production) and assume $B(H \to \mu\mu)$ as indicated in the legends of the figures (chosen to ensure good visibility in the plot). The hatched bands show the total postfit statistical and systematic uncertainties of the backgrounds. The histogram labeled as “Other” in the legend includes the contributions from the diboson, single-top-quark, $t\bar{t} + V$ and $W + \text{jets}$ backgrounds.

FIG. 6. Postfit BDT20 (left) and BDT50 (right) distributions in SRincl. The signal distributions are normalized to the SM Higgs boson cross section (including ggF, VBF, and VH production) and assume $B(H \to \mu\mu)$ as indicated in the legends of the figures (chosen to ensure good visibility in the plot). The hatched bands show the total postfit statistical and systematic uncertainties of the backgrounds. The histogram labeled as “Other” in the legend includes the contributions from the diboson, single-top-quark, $t\bar{t} + V$, and $W + \text{jets}$ backgrounds.
significance of $3.3\sigma$ ($1.7\sigma$). The global significance was calculated from the asymptotic formulas in Refs. [109,111].

Upper limits, derived using the CLs technique [112,113], are set on $B(H \to aa \to b\bar{b}\mu\mu)$ in a series of conditional fits, this time also including the signal samples. The limits as a function of $m_a$ are shown in Fig. 9. Uniform sensitivity is achieved for all masses above 18 GeV, while for lower signal masses, $m_a \leq 18$ GeV, the sensitivity of the analysis decreases due to $b$-jets falling below the reconstruction threshold or merging into one reconstructed jet. Figure 10 shows $m_{\mu\mu}$ and BDTr limits distributions after the signal + background fit for two SR bins, $m_{\mu\mu} = 35$ GeV and $m_{\mu\mu} = 52$ GeV, where the two largest deviations from the background-only hypothesis are observed. The signal in the plots is scaled to the best-fit value, corresponding to $B(H \to aa \to b\bar{b}\mu\mu) = 6.4 \times 10^{-5} (1.9 \times 10^{-4})$ for $m_a = 35$ GeV ($m_a = 52$ GeV).

The upper limits at 95% CL on $B(H \to aa \to b\bar{b}\mu\mu)$ range between $0.2 \times 10^{-4}$ and $4.0 \times 10^{-4}$, depending on $m_a$. These limits improve upon the previous ATLAS result based on 36 fb$^{-1}$ of data [47] by a factor of 2–5 over the full $m_{\mu\mu}$ range. A factor of ~2 improvement in sensitivity comes from the larger dataset, and a further factor of ~2 is achieved thanks to the use of multivariate techniques to discriminate between the signal and the SM backgrounds. Due to small number of background events at lower signal masses $m_a$, the BDTr training is less efficient in this region, and the gain from applying the BDTr selection criteria is higher at higher $m_a$. Taking as an example the favorable scenario with $B(H \to aa \to b\bar{b}\mu\mu)/B(H \to aa) = 0.16\%$, the analysis probes the Higgs boson branching fraction into pseudoscalars down to $B(H \to aa) = 1.3\%$, much lower than the limits derived from combinations of the Higgs boson measurements.

So as not to restrict the analysis sensitivity solely to models where the $a$-particle is a pseudoscalar, upper limits obtained without employing the BDTr discriminants are also derived as shown in Fig. 11. In addition to being less sensitive to the particle’s CP properties, the limits in SRincl without the BDTr selection also facilitate reinterpretations of the analysis. These limits are derived in the same way as described above, i.e., by scanning the $m_{\mu\mu}$ windows of

![Image of a graph showing the number of events in SR bins](image-url)

**FIG. 7.** Postbackground-validation-fit number of events in all SR bins (after applying the BDTr selection) that are tested for the presence of signal. The bin widths are 2 GeV (3 GeV) in $m_{\mu\mu}$ for $m_a \leq 45$ GeV ($m_a > 45$ GeV). Neighboring bins partially overlap, hence they are not statistically independent. The bottom panel shows the pull in each bin, defined as $(n_{\text{obs}} - n_{\text{pred}})/\sigma_{\text{tot}}$, where $n_{\text{obs}}$ is the number of events in the data, $n_{\text{pred}}$ is the number of fitted background events and $\sigma_{\text{tot}}$ is the total (systematic and statistical, added in quadrature) uncertainty in the fitted background yield. Discontinuities in the background predictions appear when the BDTr discriminant used for the selection changes from the one trained in the lower mass range to the one trained in the higher mass range. The histogram labeled as “Other” in the legend includes the contributions from the diboson, single-top-quark, $t\bar{t} + V$, and $W +$ jets backgrounds.
FIG. 8. The local $p_0$-values are quantified in standard deviations $\sigma$ and plotted as a function of the signal mass hypothesis. Between the points, the $p_0$-values are interpolated and may not be fully representative of the actual sensitivity.

FIG. 9. Upper limits on $B(H \to aa \to bb\mu\mu)$ at 95% CL, including the BDT selection, as a function of the signal mass hypothesis. Black and red dots show masses for which the hypothesis testing was done. Between these points, the limits are interpolated and may not be fully representative of the actual sensitivity.

FIG. 10. $m_{\mu\mu}$ distributions in the SRincl after the BDT35 > 0.2 selection (top left) and BDT50 > 0.2 selection (bottom left), and BDT35 (top right) and BDT50 (bottom right) distributions in the SRincl in the $m_{\mu\mu}$ window 34–36 GeV and 50.5–53.5 GeV, respectively. The signal is scaled to the best-fit value, $B(H \to aa \to bb\mu\mu) = 6.4 \times 10^{-5}$ for the top plots, and $1.9 \times 10^{-4}$ for the bottom plots, assuming the SM Higgs boson cross section (including ggF, VBF, and VH production). The hatched bands show the total postfit statistical and systematic uncertainties of the backgrounds and the signal. The histogram labeled as “Other” in the legend includes the contributions from the diboson, single-top-quark, $t\bar{t} + V$, and $W + jets$ backgrounds.
A search for light pseudoscalar particles (denoted by $a$) in the decays of the 125 GeV Higgs boson in the final state with two muons and two $b$-tagged jets, $H \rightarrow aa \rightarrow bb\mu\mu$, is presented. The analysis is performed using 139 fb$^{-1}$ of $\sqrt{s} = 13$ TeV $pp$ collision data recorded by the ATLAS detector at the LHC between 2015 and 2018. A narrow resonance is searched for in the dimuon invariant mass spectrum in the range $16 \leq m_{\mu\mu} \leq 62$ GeV. BDT classifiers are trained to distinguish the $H \rightarrow aa$ signal, where $a$ is a pseudoscalar, from the SM backgrounds. Additionally, the result without selection on the BDT discriminants is also provided to ensure sensitivity to models where the $a$-particle is not necessarily a pseudoscalar, as well as to facilitate reinterpretations of the analysis. No significant excess of the data above the SM backgrounds is observed. In the BDT analysis, the lowest local $p_{0}$-value of 0.00054 is observed at $m_{\mu\mu} = 52$ GeV and corresponds to a local significance of $3.3\sigma$. The global significance of that excess is determined to be $1.7\sigma$. Upper limits at 95% CL (excluding) the BDT selection criteria are set on $B(H \rightarrow aa \rightarrow bb\mu\mu)$ and range between $0.2 \times 10^{-4}$ and $4.0 \times 10^{-4}$ ($0.5 \times 10^{-4}$ and $5.0 \times 10^{-4}$), depending on $m_{a}$. The result including the BDT selection criteria improves upon previous ATLAS and CMS limits by about a factor of 2–5 for $m_{a} > 20$ GeV, while both results (with and without the BDT) extend the search down to $m_{a}$ values of 16 GeV.

**FIG. 11.** Upper limits on $B(H \rightarrow aa \rightarrow bb\mu\mu)$ at 95% CL, with no BDT selection applied, as a function of the signal mass hypothesis. The dash-dotted blue line indicates the expected limit set in the analysis with the BDT selection. Black and red dots show masses for which the hypothesis testing was done. Between these points, the limits are interpolated and may not be fully representative of the actual sensitivity.

**FIG. 12.** Postbackground-validation-fit number of events in all SR bins (without applying the BDT selection) that are tested for the presence of signal. The bin widths are 2 GeV (3 GeV) in $m_{\mu\mu}$ for $m_{a} \leq 45$ GeV ($m_{a} > 45$ GeV). Neighboring bins partially overlap, hence they are not statistically independent. The bottom panel shows the pull in each bin, defined as $(n_{obs} - n_{pred})/\sigma_{tot}$, where $n_{obs}$ is the number of events in the data, $n_{pred}$ is the number of fitted background events, and $\sigma_{tot}$ is the total (systematic and statistical, added in quadrature) uncertainty in the fitted background yield. The discontinuity at $m_{a} = 45$ GeV appears where the $m_{\mu\mu}$ window size is changed. The histogram labeled as “Other” in the legend includes the contributions from the diboson, single-top-quark, $tt + V$, and $W + jets$ backgrounds.

**VIII. CONCLUSION**

A search for light pseudoscalar particles (denoted by $a$) in the decays of the 125 GeV Higgs boson in the final state with two muons and two $b$-tagged jets, $H \rightarrow aa \rightarrow bb\mu\mu$, is presented. The analysis is performed using 139 fb$^{-1}$ of $\sqrt{s} = 13$ TeV $pp$ collision data recorded by the ATLAS detector at the LHC between 2015 and 2018. A narrow resonance is searched for in the dimuon invariant mass spectrum in the range $16 \leq m_{\mu\mu} \leq 62$ GeV. BDT classifiers are trained to distinguish the $H \rightarrow aa$ signal, where $a$ is a pseudoscalar, from the SM backgrounds. Additionally, the result without selection on the BDT discriminants is also provided to ensure sensitivity to models where the $a$-particle is not necessarily a pseudoscalar, as well as to facilitate reinterpretations of the analysis. No significant excess of the data above the SM backgrounds is observed. In the BDT analysis, the lowest local $p_{0}$-value of 0.00054 is observed at $m_{\mu\mu} = 52$ GeV and corresponds to a local significance of $3.3\sigma$. The global significance of that excess is determined to be $1.7\sigma$. Upper limits at 95% CL (excluding) the BDT selection criteria are set on $B(H \rightarrow aa \rightarrow bb\mu\mu)$ and range between $0.2 \times 10^{-4}$ and $4.0 \times 10^{-4}$ ($0.5 \times 10^{-4}$ and $5.0 \times 10^{-4}$), depending on $m_{a}$. The result including the BDT selection criteria improves upon previous ATLAS and CMS limits by about a factor of 2–5 for $m_{a} > 20$ GeV, while both results (with and without the BDT) extend the search down to $m_{a}$ values of 16 GeV.
ACKNOWLEDGMENTS

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; 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; Minciencias, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, CR, Czech Republic; DNRF and DNSRC, Denmark; Minciencias, Colombia; MSMT CR, MPO CR and VSC ANID, Chile; CAS, MOST and NSFC, China; FAPERJ, Brazil; NSERC, NRC and CFI, Canada; CERN; AEypi, Greece; RGC and Hong Kong SAR, China; ISTF and Georgia; BMBF, HGF and MPG, Germany; GSRI, IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, CR, Czech Republic; DNRF and DNSRC, Denmark; Minciencias, Colombia; MSMT CR, MPO CR and VSC ANID, Chile; CAS, MOST and NSFC, China; MEXT and JSPS, Greece; RGC and Hong Kong SAR, China; ISF and Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MEiN, Poland; FCT, Portugal; MNE/IFA, Romania; JINR; MES of Russia and NRC KE, Russian Federation; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DSI/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, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; COST, ERC, ERDF, Horizon 2020 and Marie Sklodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleiteos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; Norwegian Financial Mechanism 2014–2021, Norway; NCN and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; G"oran 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), 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. [114].


[26] ATLAS Collaboration, Search for Higgs bosons decaying to aa in the $\mu\mu\tau\tau$ final state in p p collisions at $\sqrt{s} = 8$ TeV with the ATLAS experiment, Phys. Rev. D 92, 052002 (2015).


[76] ATLAS Collaboration, Measurement of the $Z/\gamma^* \to \ell\ell$ transverse momentum distribution in $p\bar{p}$ collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, J. High Energy Phys. 09 (2014) 145.


SEARCH FOR HIGGS BOSON DECAYS INTO A PAIR OF ...

PHYS. REV. D 105, 012006 (2022)
SEARCH FOR HIGGS BOSON DECAYS INTO A PAIR OF ... PHYS. REV. D 105, 012006 (2022)

11\textsuperscript{f} İstinye University, Sariyer, Istanbul, Turkey
12\textsuperscript{i} Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain
13\textsuperscript{a} Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China
13\textsuperscript{b} Physics Department, Tsinghua University, Beijing, China
13\textsuperscript{c} Department of Physics, Nanjing University, Nanjing, China
13\textsuperscript{d} University of Chinese Academy of Science (UCAS), Beijing, China
13\textsuperscript{e} Institute of Physics, University of Belgrade, Belgrade, Serbia
15\textsuperscript{f} Department for Physics and Technology, University of Bergen, Bergen, Norway
16\textsuperscript{g} Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, California, USA
17\textsuperscript{a} Institut für Physik, Humboldt Universität zu Berlin, Berlin, Germany
18\textsuperscript{a} Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
19\textsuperscript{a} School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
20\textsuperscript{a} Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá, Colombia
20\textsuperscript{b} Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia
21\textsuperscript{a} Dipartimento di Fisica e Astronomia A. Righi, Università di Bologna, Bologna, Italy
21\textsuperscript{b} INFN Sezione di Bologna, Bologna, Italy
22\textsuperscript{a} Physikalisches Institut, Universität Bonn, Bonn, Germany
23\textsuperscript{a} Department of Physics, Boston University, Boston, Massachusetts, USA
24\textsuperscript{a} Department of Physics, Brandeis University, Waltham, Massachusetts, USA
25\textsuperscript{a} Transilvania University of Brașov, Brașov, Romania
25\textsuperscript{b} Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania
25\textsuperscript{c} Department of Physics, Alexandru Ioan Cuza University of Iași, Iași, Romania
25\textsuperscript{d} National Institute for Research and Development of Isotopic and Molecular Technologies, Physics Department, Cluj-Napoca, Romania
25\textsuperscript{e} University Politehnica Bucharest, Bucharest, Romania
25\textsuperscript{f} West University in Timișoara, Timișoara, Romania
26\textsuperscript{a} Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovak Republic
26\textsuperscript{b} Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
27\textsuperscript{a} Physics Department, Brookhaven National Laboratory, Upton, New York, USA
28\textsuperscript{a} Departamento de Física (FCEN) and IFIBA, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina
29\textsuperscript{a} California State University, California, USA
30\textsuperscript{a} Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
31\textsuperscript{a} Department of Physics, University of Cape Town, Cape Town, South Africa
31\textsuperscript{b} Thembalab, Western Cape, South Africa
31\textsuperscript{c} Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg, South Africa
31\textsuperscript{d} National Institute of Physics, University of the Philippines Diliman (Philippines), Quezon City, Philippines
31\textsuperscript{e} University of South Africa, Department of Physics, Pretoria, South Africa
31\textsuperscript{f} University of Zululand, KwaDlangezwa, South Africa
31\textsuperscript{g} School of Physics, University of the Witwatersrand, Johannesburg, South Africa
32\textsuperscript{a} Department of Physics, Carleton University, Ottawa, Ontario, Canada
33\textsuperscript{a} Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Énergies—Université Hassan II, Casablanca, Morocco
33\textsuperscript{b} Faculté des Sciences, Université Ibn-Tofail, Kénitra, Morocco
33\textsuperscript{c} Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA-Marrakech, Morocco
33\textsuperscript{d} LPMR, Faculté des Sciences, Université Mohamed Premier, Oujda, Morocco
33\textsuperscript{e} Faculté des sciences, Université Mohammed V, Rabat, Morocco
33\textsuperscript{f} Mohammed VI Polytechnic University, Ben Guerir, Morocco
33\textsuperscript{g} CERN, Geneva, Switzerland
35\textsuperscript{a} Enrico Fermi Institute, University of Chicago, Chicago, Illinois, USA
36\textsuperscript{a} LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand, France
37\textsuperscript{a} Nevis Laboratory, Columbia University, Irvington, New York, USA
38\textsuperscript{a} Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
39\textsuperscript{a} Dipartimento di Fisica, Università della Calabria, Rende, Italy

012006-25
G. AAD et al. PHYS. REV. D 105, 012006 (2022)

123Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman, Oklahoma, USA
124Department of Physics, Oklahoma State University, Stillwater, Oklahoma, USA
125Palacký University, Joint Laboratory of Optics, Olomouc, Czech Republic
126Institute for Fundamental Science, University of Oregon, Eugene, Oregon, USA
127Graduate School of Science, Osaka University, Osaka, Japan
128Department of Physics, University of Oslo, Oslo, Norway
129Department of Physics, Oxford University, Oxford, United Kingdom
130LPNHE, Sorbonne Université, Université de Paris, CNRS/IN2P3, Paris, France
131Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania, USA
132Konstantinov Nuclear Physics Institute of National Research Centre "Kurchatov Institute", PNPI, St. Petersburg, Russia
133Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
134Laboratório de Instrumentação e Física Experimental de Partículas—LIP, Lisboa, Portugal
135Département de Physique, Faculté des Sciences, Tunis, Tunisie
136University of York, York, United Kingdom
137Department of Physics and Astronomy, York University, Toronto, Ontario, Canada
138Department of Physics and Astronomy, University of Sheffield, Sheffield, United Kingdom
139Department of Physics, Shinshu University, Nagano, Japan
140Department of Physics, Royal Institute of Technology, Stockholm, Sweden
141Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York, USA
142Department of Physics and Astronomy, University of Sussex, Brighton, United Kingdom
143School of Physics, University of Sydney, Sydney, Australia
144Department of Physics, Academy Sinica, Taipei, Taiwan
145E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi, Georgia
146High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia
147Department of Physics, Technion, Israel Institute of Technology, Haifa, Israel
148Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
149Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece
150International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo, Japan
151Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
152Department of Physics, Tomsk State University, Tomsk, Russia
153Department of Physics, University of Toronto, Toronto, Ontario, Canada
154TRIUMF, Vancouver, British Columbia, Canada
155Department of Physics and Astronomy, York University, Toronto, Ontario, Canada

012006-28
SEARCH FOR HIGGS BOSON DECAYS INTO A PAIR OF …  PHYS. REV. D 105, 012006 (2022)

Deceased.

Also at Department of Physics, King’s College London, London, United Kingdom.

Also at Istanbul University, Department of Physics, Istanbul, Turkey.

Also at Instituto de Fisica Teorica, IFT-UAM/CSIC, Madrid, Spain.

Also at TRIUMF, Vancouver, British Columbia, Canada.

Also at Department of Physics, University of Fribourg, Fribourg, Switzerland.

Also at Department of Physics and Astronomy, University of Louisville, Louisville, Kentucky, USA.

Also at Department of Physics, Ben Gurion University of the Negev, Beer Sheva, Israel.

Also at Universita di Napoli Parthenope, Napoli, Italy.

Also at Institute of Particle Physics (IPP), Canada.

Also at Bruno Kessler Foundation, Trento, Italy.

Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia.

Also at Borough of Manhattan Community College, City University of New York, New York, New York, USA.

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

Also at Department of Financial and Management Engineering, University of the Aegean, Chios, Greece.

Also at Centro Studi e Ricerche Enrico Fermi, Italy.

Also at Department of Physics, California State University, East Bay, California, USA.

Also at Instituto Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain.

Also at Graduate School of Science, Osaka University, Osaka, Japan.

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

Also at University of Chinese Academy of Sciences (UCAS), Beijing, China.

Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan.

Also at Yeditepe University, Physics Department, Istanbul, Turkey.

Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia.

Also at CERN, Geneva, Switzerland.

Also at Joint Institute for Nuclear Research, Dubna, Russia.

Also at Hellenic Open University, Patras, Greece.

Also at Center for High Energy Physics, Peking University, China.

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

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

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

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

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 Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan, USA.