Measurements of Higgs bosons decaying to bottom quarks from vector boson fusion production with the ATLAS experiment at $\sqrt{s} = 13$ TeV

ATLAS Collaboration

DOI
10.1140/epjc/s10052-021-09192-8

Publication date
2021

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):
Measurements of Higgs bosons decaying to bottom quarks from vector boson fusion production with the ATLAS experiment at $\sqrt{s} = 13$ TeV

ATLAS Collaboration
CERN, 1211 Geneva 23, Switzerland

Abstract The paper presents a measurement of the Standard Model Higgs Boson decaying to $b$-quark pairs in the vector boson fusion (VBF) production mode. A sample corresponding to 126 fb$^{-1}$ of $\sqrt{s} = 13$ TeV proton–proton collision data, collected with the ATLAS experiment at the Large Hadron Collider, is analyzed utilizing an adversarial neural network for event classification. The signal strength, defined as the ratio of the measured signal yield to that predicted by the Standard Model for VBF Higgs production, is measured to be 0.95$^{+0.38}_{-0.36}$, corresponding to an observed (expected) significance of 2.6 (2.8) standard deviations from the background-only hypothesis. The results are additionally combined with an analysis of Higgs bosons decaying to $b$-quarks, produced via VBF in association with a photon.

1 Introduction

Approximately 4.5 million Higgs bosons [1–8] have decayed into bottom quarks in the nearly 140 fb$^{-1}$ of 13 TeV LHC collisions collected for analysis by the ATLAS experiment during Run 2. While this number is three times larger than that from any other decay mode, this decay channel remains the most poorly measured major Higgs boson decay channel. In the dominant production mode, where the Higgs boson is produced in a virtual top-quark loop connecting two interacting gluons (ggF), the signature of this decay is overwhelmed by the strong production of quarks and gluons, except in cases where the Higgs boson is highly boosted [9]. The ATLAS and CMS experiments have observed Higgs boson decays into $b$-quarks, $H \rightarrow b\bar{b}$, with most of the sensitivity coming from cases where the Higgs boson is produced in association with a vector boson ($VH$, $V=W,Z$), which provides sufficient discrimination against QCD background processes despite its small cross-section. The measurement by the ATLAS experiment of the signal yield relative to the Standard Model expectation, $\mu_{H\rightarrow b\bar{b}}$, is 1.02$\pm$0.12(stat.$)\pm$0.14(syst.$)$, corresponding to a significance of 6.7$s$ [10] relative to the background-only hypothesis. The measurement by the CMS experiment is $\mu_{H\rightarrow b\bar{b}} = 1.04 \pm 0.14$(stat.$) \pm 0.14$(syst.$)$, corresponding to a significance of 5.6$s$ [11].

Vector-boson fusion (VBF), wherein the Higgs boson is produced when quarks from each proton radiate weak vector bosons that fuse to form the Higgs boson, is the second most frequent Higgs boson production mechanism. Its signature, shown in Fig. 1a, is characterized by the presence of jets from each of the quarks with a large rapidity gap between them. Because there is no coloured connection between the two protons, radiative hadronic activity between the two forward jets is suppressed. VBF Higgs boson production has been measured by the ATLAS experiment in several decay channels, and the combined result for the signal strength is $\mu_{VBF} = 1.21 \pm 0.18$(stat.$) \pm 0.15$(syst.$)$ [12]. The most recent CMS measurement yields $\mu_{VBF} = 0.73 \pm 0.23$(stat.$) \pm 0.16$(syst.$)$ also through the combination of several decay channels [13].

Studying Higgs boson decays into $b$-quarks in the vector-boson fusion channel provides an avenue to pursue this challenging signature. Previous measurements of this process by the ATLAS experiment using 31 fb$^{-1}$ of $\sqrt{s} = 13$ TeV data yielded combined results for VBF production with and without a photon in the final state [14]. The resulting signal strength of $H \rightarrow b\bar{b}$ was $\mu_{H\rightarrow b\bar{b}} = 2.5^{+1.4}_{-1.3}$, corresponding to an observed (expected) significance of 1.9$s$ (0.8$s$). The observed signal strength for VBF production of $H \rightarrow b\bar{b}$ when the other Higgs boson production modes are fixed to their Standard Model values, $\mu_{VBF,H\rightarrow b\bar{b}}$, was 3.0$^{+1.7}_{-1.6}$, corresponding to an observed (expected) significance of 1.9$s$ (0.7$s$) relative to the background-only hypothesis, which includes non-VBF Higgs boson production modes.

This paper presents an update in the analysis of VBF production without a photon, using 126 fb$^{-1}$ of proton–proton ($pp$) collision data collected by the ATLAS experiment at the Large Hadron Collider from 2016 to 2018. The analysis has significant improvements in sensitivity with respect...
to the previous version [14]. A brief outline of this analysis is described in the following. The final state is characterized by two $b$-jets from the decay of the Higgs boson, as well as two light-quark jets with a large rapidity gap, coming from the outgoing quarks. In this analysis, two channels corresponding to the available triggers during the data-taking periods and targeting events with and without a high $p_T$ forward jet are used. After preselection, kinematic properties of the events are used as input to an adversarial neural network (ANN) [15], trained on signal Monte Carlo (MC) events and data sidebands. The ANN is constructed such that the classifier output score is independent of the di-jet invariant mass. Each channel is then divided into five regions of varying sensitivity based on the classifier score. Backgrounds consist of two primary components: non-resonant QCD multi-jet events, an example shown in Fig. 1b, and events containing $Z$ boson decays into $b$-jets. The $Z \rightarrow b\bar{b}$ contribution is constrained directly from data using an embedding process in which $Z \rightarrow \mu\mu$ events are selected in data, the muons from simulation, and the analysis selections are applied to determine the number of selected $Z \rightarrow b\bar{b}$ events. The non-resonant background shape is determined from the data and the same shape is used in all regions for each channel. It is primarily determined from the lowest-score regions. Potential bias arising from the assumption that the non-resonant background shape is the same in all regions for each channel is determined through a set of shape-bias control regions. The VBF Higgs boson signal is extracted through a simultaneous fit of the signal and background contributions to the di-$b$-jet invariant mass spectrum in each channel and region. Lastly, the results are combined with the 132 $fb^{-1}$ measurement by the ATLAS experiment in which VBF production of a Higgs boson decaying into $b$-quarks is accompanied by a photon [16], henceforth referred to as the photon analysis, for a joint measurement of $\mu_\nu$ and $H\rightarrow b\bar{b}$.

2 Detector and data samples

The ATLAS detector [17] at the LHC covers nearly the entire solid angle around the collision point.\footnote{ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the $z$-axis points along the beam pipe. The $x$-axis points from the IP to the centre of the LHC ring, and the $y$-axis points upwards. Cylindrical coordinates ($r$, $\phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$.} It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroidal magnets.

The inner-detector system is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$. A high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, the first hit normally being in the insertable B-layer installed before Run 2 of the LHC [18,19]. It is followed by a silicon microstrip tracker, which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$. The TRT also provides electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimized for electromagnetic and hadronic measurements respectively.

The muon spectrometer comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by the superconducting air-core toroids. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. A set of precision chambers covers the region $|\eta| < 2.7$ with three layers of monitored drift tubes, complemented by cathode-strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

Fig. 1 Feynman diagrams of a the VBF signal process and b gluon splitting as an example background process.
Interesting events are selected to be recorded by the first-level trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger [20]. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the high-level trigger reduces in order to record events to disk at about 1 kHz.

The data used in this analysis were collected in 2016, 2017 and 2018 [21]. After applying event-cleaning requirements which ensure that the data are of good quality for triggering and reconstructing jets containing $b$-quarks, these datasets correspond to integrated luminosities of 24.6 fb$^{-1}$, 43.6 fb$^{-1}$, and 57.7 fb$^{-1}$, respectively, for a total integrated luminosity of 125.9 fb$^{-1}$. During the 2016 data-taking, roughly one-fifth of the data taken by ATLAS which was good for physics was affected by an inefficiency in the vertex reconstruction in the high-level trigger, which reduced the efficiency of the algorithms used to identify jets originating from $b$-hadron decays; those events were not retained for further analysis and are not included in the 24.6 fb$^{-1}$.

### 3 Monte Carlo samples

Simulated events with a Higgs boson mass of 125 GeV are used for the signal modelling. The signal models include all major Higgs boson production modes: VBF and ggF, as well as $VH$ and associated production with a pair of top quarks ($t\bar{t}H$). Simulated VBF signal events were generated at next-to-leading order (NLO) in QCD with POWHEG-BOX v2 [22–25], using the NNPDF3.0 [26] parton distribution functions (PDFs). NLO electroweak corrections are calculated using MADGRAPH5_AMC@NLO v3.0.1 [27] and applied as a function of the generated Higgs boson transverse momentum ($p_T$). Simulated ggF samples were generated at next-to-next-to-leading order (NNLO) in QCD with POWHEG-BOX v2 [28–30] using the NNPDF3.0 PDFs. Both the VBF and ggF samples use PYTHIA 8.212 [31] for parton showering and fragmentation with the A3 tuned parameter set and the NNPDF2.3LO PDF set [32]. For the VBF sample, the PYTHIA dipole-recoil scheme [33] is used to improve the modelling of radiation in the central region. POWHEG-BOX v2 interfaced to NNPDF3.0 was used to generate events that were showered with PYTHIA 8.212 to model contributions from $VH$ [34–37] and $t\bar{t}H$ [38]. All samples were interfaced to EVTGEN v1.6.0 [39] for heavy-hadron decays. The $Z$ and non-resonant backgrounds are derived from the data.

Multiple $pp$ collisions were simulated with the soft QCD processes of PYTHIA 8.212 using the A3 tuned parameter set and the NNPDF2.3LO PDFs. These additional interactions are overlaid on the hard-scatter interaction of the signal and background samples according to the luminosity profile of the recorded data to model contributions from $pp$ interactions in both the same bunch crossing and neighbouring bunch crossings (pile-up). The response of the ATLAS detector to the generated events is then modelled with a detailed ATLAS detector simulation software [40] based on GEANT4 [41].

### 4 Event reconstruction and selection

#### 4.1 Event reconstruction

This analysis uses an all-hadronic final state, and therefore the primary objects of interest are jets. Jets are reconstructed by applying the anti-$k_T$ jet clustering algorithm [42,43] with a radius parameter of $R = 0.4$ to particle flow [44] objects which are constructed from calorimeter energy clusters and tracks. In order to remove jets originating from additional $pp$ interactions, the likelihood-based ‘jet vertex tagger’ [45] algorithm is applied to all jets with $p_T < 60$ GeV and $|\eta| < 2.5$. Quality and cleaning cuts are applied to ensure well-measured jets. In addition, $\eta$- and $p_T$-dependent scale factors are applied to correct the jet energy to the hadronic level [46], and a pile-up subtraction algorithm is applied to reduce effects of pile-up contributions to the jet energy.

A multivariate $b$-tagging algorithm, MV2c10 [47], is used to identify jets which contain $b$-hadrons. This algorithm is a boosted decision tree which combines several features of impact parameter distributions of associated tracks with properties of secondary vertices, as well as jet $p_T$ and $\eta$. A similar algorithm is used by the trigger, differing only in the implementation of the jet, track and primary vertex definitions used, and, in 2016 data-taking, by the flavour-tagging discriminant. Two different identification working points of the algorithm are used, corresponding to 77% and 85% efficiency for $b$-jets as measured in a $t\bar{t}$ sample. In the trigger system, the 60%, 70% and 85% working points are used. In order to improve the resolution of the invariant mass of $b$-jet pairs, additional corrections are applied to the energy of $b$-tagged jets. These corrections account for semileptonic decays of the $b$-hadron (the muon-in-jet correction) and energy resolution effects specific to $b$-jets (the PtReco correction). They improve the dijet mass resolution by up to 20% [48].

Photons, electrons and muons are also identified in order to veto events which overlap with other $H \rightarrow b\bar{b}$ analysis channels [16,49]. Muons are additionally used in the embedding procedure described in Sect. 6. Photons are reconstructed from energy clusters in the electromagnetic calorimeter and calibrated to account for energy losses upstream of the calorimeter as well as leakage outside of the clusters [50]. Photons which convert in the detector material are reconstructed separately, utilizing tracks to find the conversion vertex. A tight, cut-based selection is applied, providing good rejection of hadronic jets where a neutral meson carries most of the jet energy. To further suppress the jet back-
grounds, the photons are required to be isolated, meaning that after corrections for pile-up the sum of the transverse energy in a cone of $\Delta R = 0.4$ around the photon candidate must not exceed a threshold which depends on the photon’s $p_T$. Electrons are also reconstructed utilizing energy clusters in the electromagnetic calorimeter [50]. Electron candidates must satisfy a loose likelihood-based electron identification criterion, including a loose isolation requirement. Muons [51] are reconstructed using inner detector, calorimeter and muon spectrometer information. In this analysis they satisfy loose identification requirements, including a loose isolation requirement. Of particular importance for this analysis is a requirement that both electrons and muons are prompt, i.e. their impact parameter significance is required to be small. An overlap removal procedure is applied in order to ensure that jets, electrons, muons and photons are not double counted. Events with identified photons with transverse energy ($E_T > 15$ GeV), or electrons or muons with $p_T > 7$ GeV after overlap removal are vetoed. The objects have identical selection requirements to Refs. [16,49] in order to ensure orthogonality.

4.2 Event selection

There are two channels in this analysis, the Forward and Central channels, which are distinguished by the presence or absence of a high-$p_T$ forward (3.2 < $|\eta|$ < 4.5) jet in the event. The specific selections are noted below and summarized in Table 1. In the following, $b_1$ and $b_2$ refer to the $b$-tagged jets forming the highest-$p_T$ $b$-tagged jet pair in the event, and $j_1$ and $j_2$ are two additional jets referred to as VBF jets. Generally, $j_1$ refers to the highest-$p_T$ and most forward VBF jet, and $j_2$ refers to the second-highest-$p_T$ VBF jet. Additional requirements on the jets are channel specific and detailed below.

The trigger for the Forward channel targets events with at least two $b$-tagged jets, one with trigger level $p_T > 80$ GeV and a second with $p_T > 60$ GeV, and at least one forward jet with $p_T > 45$ GeV. The corresponding event selection requires at least one $b$-tagged jet passing the 77% $b$-tagging working point to have $p_T > 85$ GeV and $|\eta| < 2.5$ ($b_1$), and a second $b$-tagged jet passing the 85% $b$-tagging working point to have $p_T > 65$ GeV and $|\eta| < 2.5$ ($b_2$). One forward jet is required to have $p_T > 60$ GeV ($j_1$) and an additional jet is required with $p_T > 30$ GeV and $|\eta| < 4.5$ ($j_2$). Reconstructed jets in the event are required to be matched to trigger jets.

The Central channel trigger targets events with at least three central jets, two of which are $b$-tagged. The trigger jet thresholds varied over the run period. The corresponding event selection requires at least two $b$-tagged jets passing the 77% $b$-tagging working point to have $p_T > 65$ GeV and $|\eta| < 2.5$ ($b_1, b_2$). Two additional jets are required, one jet having $p_T > 160$ GeV and $|\eta| < 3.1$ ($j_1$) and a second requiring $p_T > 30$ GeV and $|\eta| < 4.5$ ($j_2$). Furthermore, events are rejected if they have any jets with $p_T > 60$ GeV and 3.2 < $|\eta|$ < 4.5 to maintain orthogonality with the Forward channel. As in the Forward channel, reconstructed jets in the event are required to be matched to trigger jets.

In cases where there is ambiguity when identifying the two VBF jets (e.g. when more than two jets not selected as the signal $b$-tagged jets satisfy the $p_T$ and $\eta$ requirements), the pair giving the highest invariant mass of the dijet system, $m_{jj}$, is chosen.

To remove a shaping of the invariant mass of the $b$-tagged dijet system, $m_{bb}$, by the high leading-jet thresholds, a cut on the transverse momentum of the $b$-tagged dijet system, $p_{T,bb}$, is required. Both channels require $p_{T,bb} > 150$ GeV. The Central channel additionally requires $m_{jj}$ to be greater than 800 GeV. The efficiency of the combined selection of both channels for simulated VBF signal events with true Higgs

| Event selection criteria for the two channels used in this analysis. The $p_T$ and $|\eta|$ requirements on the jets are used to match trigger selections and flavour-tagging requirements |
| --- |
| **Forward channel event selection** |
| $b_1$ | $\geq 1$ $b$-tagged jet at 77% efficiency working point with $p_T > 85$ GeV and $|\eta| < 2.5$ |
| $b_2$ | $\geq 1$ $b$-tagged jet at 85% efficiency working point with $p_T > 65$ GeV and $|\eta| < 2.5$ |
| $j_1$ | $\geq 1$ jet with $p_T > 60$ GeV and $3.2 < |\eta| < 4.5$ |
| $j_2$ | $\geq 1$ jet with $p_T > 30$ GeV and $|\eta| < 4.5$ |
| $p_{T,bb} > 150$ GeV |
| **Central channel event selection** |
| $b_1, b_2$ | $\geq 2$ $b$-tagged jets at 77% efficiency working point with $p_T > 65$ GeV and $|\eta| < 2.5$ |
| $j_1$ | $\geq 1$ jet with $p_T > 160$ GeV and $|\eta| < 3.1$ |
| $j_2$ | $\geq 1$ jet with $p_T > 30$ GeV and $|\eta| < 4.5$ |
| No jets with $p_T > 60$ GeV and $3.2 < |\eta| < 4.5$ |
| $p_{T,bb} > 150$ GeV, $m_{jj} > 800$ GeV |
boson rapidity $|y_H| < 2.5$ is 0.8%, with roughly half of the acceptance coming from each channel. With respect to simulated VBF signal events with $|y_H| < 2.5$ and $p_T^{H} > 200$ GeV, this selection is 10% efficient, with roughly 4% and 6% contributed from the Forward and Central channels, respectively.

5 Adversarial neural network for event classification

The sensitivity of the analysis is boosted by using an ANN to divide events into regions of varying signal to background composition. This type of multivariate classifier can be constructed such that the network is penalized for learning a feature of the dataset [52,53]. Because the Higgs boson signal is extracted from a fit to $m_{bb}$, the ANN used in this analysis is penalized for learning the $m_{bb}$ distribution. Using this construction, the non-resonant background shape is independent of the classifier score and therefore the same for each region, which allows for extra constraints in the shape of the non-resonant background.

Twelve variables are used as inputs to the ANN:

- $m_{jj}$: the invariant mass of the VBF jet pair. This is typically larger in the signal than in the background.
- $p_{T,jj}$: the transverse momentum of the VBF jet pair. This is typically larger in the signal than in the background, due to both the harder $p_{T,bb}$ distribution and the lower jet multiplicity in the signal.
- $p_{T}^{balance}$: the ratio of the vectorial and scalar sums of the transverse momenta of $b_1, b_2, j_1$ and $j_2$. Generally, in the signal the four jets tend to be balanced, so the quantity is smaller in the signal than in the background.
- $(p_{T,bb}^{j1} - p_{T,bb}^{j2})/(p_{T,bb}^{j1} + p_{T,bb}^{j2})$: the asymmetry in the VBF jet transverse momenta, which tends to be smaller in the signal than in the background.

- $\Delta \eta(bb,jj)$: the separation in $\eta$ between the $b$-tagged jet pair and the VBF jet pair. Due to the dominant $g(\to bb)\gamma (\to jj)$ background, where the gluons can have a large separation in $\eta$, this quantity tends to be larger in the background than in the signal.
- $\Delta \phi(bb,jj)$: the separation in $\phi$ between the $b$-tagged jet pair and the VBF jet pair. In the signal, the $bb$ and $jj$ systems tend to have a larger separation than in the background.
- $\tan(\Delta \phi(bb/2)) / \tanh(\Delta \eta(bb/2))$: the measure of the relative angle of $\Delta \eta$ and $\Delta \phi$ between the two $b$-tagged jets. This variable takes on higher values in the signal than in the background.
- $n_{jets}$: the number of jets with $p_T > 20$ GeV and $|\eta| < 4.5$. This quantity is, on average, larger in QCD-produced events, such as the background, than those arising from electroweak production, such as the signal.
- $\min \Delta R(j_{1(2)})$: the minimum separation in $\Delta R$ between the (sub)leading VBF jet pair and any jet in the event which is not a part of the $b$-tagged jet pair or VBF jet pair. If there are none of these jets, this quantity takes the default value of 0.
- $N^{trk}_{trk}$: the number of tracks matched to the (sub)leading VBF jet. Signal VBF jets are typically quark-initiated, while background events have a higher gluon composition, which produces higher track-multiplicity jets. This variable is only valid for jets with $p_T > 50$ GeV and $|\eta| < 2.1$. Jets outside this kinematic region take specific values depending on their $p_T$ and $\eta$. This information is useful as jet quark versus gluon composition varies as a function of $p_T$ and $\eta$.

The ANN consists of a classifier and an adversary. The classifier’s role is to determine if the event is signal- or background-like. The adversary’s role is to determine the value of $m_{bb}$ in terms of a binned $m_{bb}$ distribution. Then the two are combined such that the overall network discriminates between signal and background but is penalized if the $m_{bb}$ value is learned, i.e. if it can accurately determine the $m_{bb}$ bin. To achieve this, a three-step training procedure is used [52]. First the classifier is pre-trained with binary cross-entropy loss, while keeping the adversary weights frozen. Next, the adversary is pre-trained with categorical cross-entropy loss, keeping the classifier weights fixed. Third, the classifier and adversary are trained together with a combined loss function, $L$. The combined training proceeds in two sub-steps for each epoch. First the classifier is trained with $L = L_{\text{classifier}} - \lambda L_{\text{adversary}}$, keeping the adversary weights frozen, and then the adversary is trained with loss function $L = L_{\text{adversary}}$. The configurable parameter $\lambda$ controls to what extent the adversary impacts the overall loss function, i.e. how much the network is penalized for learning $m_{bb}$. As a post-training step, the classifier scores are scaled to quantiles of the signal MC distribution with the output values ranging from 0 to 1.

One ANN for each channel is trained in Theano through Keras [54,55]. The Adam optimizer is used with standard parameters from Ref. [53]. Signal events for training data are taken from the VBF MC sample described in Sect. 3. The background events are taken from data sidebands in the range 70 GeV < $m_{bb}$ < 100 GeV (low mass sideband) and 140 GeV < $m_{bb}$ < 200 GeV (high mass sideband). This range is slightly larger than the mass range used in the signal extraction fit (see Sect. 7) to avoid edge effects in the $m_{bb}$ spectrum. Multi-jet MC events are used for determining the network performance and architecture, but data sidebands are used in the final discriminant training due to the fact that the MC sample does not fully capture the data complexity.
Fig. 2 Distribution of ANN output scores for the VBF signal MC events (green crosses), ggF MC events (blue diamonds) and low and high mass data sidebands (black squares, red dots, respectively) for the

(a) Forward channel and (b) Central channel. The bottom panel shows the ratio of the low mass to high mass sidebands. The errors are statistical only.

For events in the low (high) mass sideband, the adversary $m_{bb}$ bins have a width of 5 (10) GeV, leading to six bins per sideband and a roughly equal number of events per bin in each sideband.

The network architecture is optimized for both discrimination of signal from background and decorrelation of the ANN output score with $m_{bb}$. Because a reduction in the correlation between $m_{bb}$ and the classifier score is particularly important in the most sensitive signal regions, both the bulk decorrelation and the decorrelation for events with scores in the top 1% of ANN outputs are used to select the network architecture. The network is constructed with $\lambda = 10$ and two epochs of pre-training for both the adversary and classifier, and the adversary training utilizes only the background events. Four-fold cross validation, in which the data is divided into four random subsamples, or folds, and each fold is tested on an ANN trained with the other three folds, is used to verify that there is no overtraining. This procedure ensures that the data are always categorized with an independently trained ANN. Additional uncertainties due to possible bias arising from using data sidebands in the training in which the Higgs boson mass window cannot be directly checked are assessed in Sect. 8.3. Figure 2 shows the distributions of the ANN scores for the VBF signal, ggF events and the low and high mass data sidebands. It can be seen that good separation is achieved between the VBF signal and data sidebands. Additionally it is noted that the ggF events have a similar score distribution to the data sidebands and that the low and high mass sidebands are in good agreement with each other, as shown in the ratio panel. The variables which contribute most to the ANN’s discrimination power are $m_{jj}$, $n_{jets}$, $p_{T}^{\text{balance}}$, $\Delta \phi (bb, jj)$, and $\min \Delta R (j_{1}(2))$.

The ANN score is used to divide the data into signal regions for the signal extraction fit. Both the number of regions which are used in the fit and the boundaries of those regions are optimized in order to maximize the overall signal sensitivity. The optimization metric is

$$\sum_{n=2}^{N} \frac{s_n}{\sqrt{b_n + \delta(b_n)}} \quad (1)$$

where $N$ is the number of regions, $s_n$ is the number of signal events in the Higgs boson mass window (100 GeV < $m_{bb}$ < 140 GeV) in region $n$, $b_n$ is the number of background events in the Higgs boson mass window in region $n$, and $\delta(b_n)$ is the uncertainty on the number of background events. $s_n$ is determined from the signal MC and $b_n$ is estimated from a multi-jet MC normalized to the data sidebands. The uncertainty on the background, $\delta(b_n)$, is taken to be 10% of $b_n$. The significance scan starts from $n = 2$ because the lowest classifier score region ($n = 1$) is used primarily to constrain the background shape in the high score regions. This optimization results in five signal regions for each channel, with region boundaries, signal yields, $Z$ boson yields, and data sideband yields for each channel shown in Table 2. The four highest-score regions are named SR1 to SR4; the lowest-score region is named RegLow.
6 Resonant background determination

Candidate $Z \to b\bar{b}$ events contribute significantly to the selected event sample in the range $m_{bb} < 120$ GeV. Because of the kinematic requirements of the event selection, simulated $Z$ samples do not accurately reflect the data in the phase space probed by this analysis. Additionally, the $Z \to b\bar{b}$ ANN score distribution is similar to the data sidebands, and consequently there is little constraining power of the resonant contribution in the fit, which is described in Sect. 7. Therefore, the yield of the resonant component is constrained with a data-driven estimate from $Z \to \mu^+\mu^-$ events. The $Z \to \mu^+\mu^-$ candidate data events are selected by requiring a dimuon pair in the $Z$ mass window (81 GeV $< m_{\mu\mu} <$ 101 GeV) recorded by single-muon and double-muon triggers which do not have an isolation requirement. The reconstructed muon pair in the event is replaced with particle-flow objects obtained by showering, hadronizing and reconstructing an MC-generated $b$-$\bar{b}$ quark pair with four-vectors matching the muon pair four-vectors. The $b$-$\bar{b}$ quark pairs are showered and hadronized with PYTHIA 8.226 and run through the ATLAS fast simulation [40], digitization and reconstruction chain. Jet finding is run on the modified events and the event selection cuts are applied. The events are first weighted to correct for the muon trigger and offline reconstruction efficiencies. Then the Forward and Central channel trigger requirements are emulated by applying data-driven trigger efficiency maps as a function of the VBF jet $p_T$ and $\eta$. The resulting sample is called the embedded $Z$ sample. The events are processed through the full event selection and region categorization to yield the number of resonant background events in each region.

The process is validated by applying the embedding method in MC events and then comparing embedded and non-embedded events utilizing the analysis selections. Two samples are used for the validation. Embedded and non-embedded $Z \to b\bar{b}$ MC events are compared; however, the samples are rather small after analysis selections are applied. Therefore, the process is also validated using embedded and non-embedded VBF $H \to bb$ events. The embedded and non-embedded MC samples agree to within 20% for the variables used in the ANN and for the $m_{bb}$ and $p_T, bb$ distributions. In the signal extraction fit described in Sect. 7, this 20% residual non-closure is used as an uncertainty in the resonant component.

7 Signal extraction method

The Higgs boson signal strength is determined from an extended binned maximum-likelihood fit to the $m_{bb}$ distribution in data. The two channels are combined in a joint fit of the $m_{bb}$ distribution where all ten signal regions share the same signal-strength parameter, $\mu$. The negative log-likelihood is constructed as shown in Eq. (2), where $Y_{ijk}$ denotes the yield of $k$th bin of $j$th signal region of $i$th channel. Nuisance parameters, which have external constraints $f(\alpha_i)$, are penalized in the negative log-likelihood.

$$-\log \mathcal{L}(\mu, \{\alpha_i\}) = -\log \prod_{i=1}^{\text{Channels}} \prod_{j=1}^{\text{Regions}} \prod_{k=1}^{\text{Bins}} \frac{e^{-Y_{ijk}(\mu, \{\alpha_i\})}}{N_{ijk}^{\text{Obs}}} \times N_{ijk}^{\text{Obs}} - \log \prod_{i} f(\alpha)$$

(2)

where

$$Y_{ijk} = (\mu_{VBF, H \rightarrow b\bar{b}}) \gamma_{VBF} H_{ijk} + N_{\text{ggF}+VH+nH} H_{ijk} + N_{B} B_{ijk} + \mu Z Z_{ijk}$$

(3)

The normalizations of the non-resonant background in each signal region, denoted with $N_{B}$, are free parameters,
whereas the signal normalizations $N^\text{VBF}_H$ are set to the Standard Model expectation from simulation. $\mu_{\text{VBF}, H \rightarrow b\bar{b}}$ represents the observed signal strength and is an extracted parameter. The contributions from ggF, ttH and VH production are fixed to their Standard Model predictions, $N_{\text{ggF/VH+stH}}$. The measurement is primarily sensitive to differences in the signal contribution relative to the background among the regions, and therefore it is only weakly sensitive to the ggF contribution, which has approximately the same ANN distribution as the backgrounds. The ttH and VH contributions are negligible. The normalizations of the Z events, $N_Z$, are taken from the embedding process. The shapes of Higgs boson, Z boson, and non-resonant background are $H$, $Z$ and $B$, respectively. The Higgs boson signal shape is determined from a binned fit of a Bukin function [56] to MC signal events. The $Z$ boson shape is also taken from a binned fit of a Bukin function to a sample of $Z \rightarrow bb$ events obtained by the embedding process. The normalization of the $Z$ contribution in each region is taken from the number of embedded events passing that region’s selection cuts. For each channel, the non-resonant background $B$ is fit to a binned distribution of arbitrary shape which is subject to the constraint that it is the same in all regions, hence the absence of a $j$ subscript in Eq. (3). The shape is primarily constrained by each channel’s RegLow region, which has approximately 50 times more events than the other regions in each channel. The normalization is allowed to float in each region. The fit range is 80 GeV $< m_{bb} < 200$ GeV. The bin width is 4 GeV. Data in the Higgs boson mass window are kept blinded until all the elements of the analysis are finalized.

Systematic uncertainties affecting the signal and simulation-based background components, described in Sect. 8, are included as nuisance parameters and most are constrained in the likelihood using Gaussian or log-normal probability density functions. All experimental uncertainties, described in Sect. 8.1, are treated as fully correlated across all regions. Uncertainties related to different $b$-tagging working points are taken as uncorrelated with each other; however, it was checked that the impact on the sensitivity of treating them as fully correlated or fully uncorrelated is negligible. Theoretical uncertainties, described in Sect. 8.2, are treated as correlated across regions. The uncertainty in the resonant background normalization and width related to the embedding process is taken as uncorrelated between regions. Uncertainties covering possible biases in the non-resonant background shape are included as an additional uncertainty in the signal normalization per region and are uncorrelated between regions. This bias uncertainty is determined using dedicated control regions described in Sect. 8.3.

The parameter of interest in the fit is the signal strength of the vector-boson fusion production channel, $\mu_{\text{VBF}}$. Tests of Asimov data with signal injected at strengths of $\mu_{\text{VBF}, H \rightarrow b\bar{b}} = 0, 1, 2$ confirmed the linearity of the fit with no bias in $\mu_{\text{VBF}, H \rightarrow b\bar{b}}$. As a cross-check verifying that the analysis is only weakly sensitive to the overall ggF contribution, a 100% correlated uncertainty was added to the ggF uncertainties. The expected $\mu_{\text{VBF}, H \rightarrow b\bar{b}}$ uncertainty changed by only 0.01.

The fit is also performed with the inclusive Higgs boson signal strength $\mu_{H \rightarrow b\bar{b}}$ as the parameter of interest. For this parameter of interest, the ggF, VBF, ttH and VH contributions are fixed to their relative contributions according to the Standard Model values and a single signal strength parameter is applied to all contributions. Signal injection tests also confirm the linearity of the fit with no bias in $\mu_{H \rightarrow b\bar{b}}$.

8 Systematic uncertainties

Systematic uncertainties are divided into three categories: experimental, theoretical and data-driven background uncertainties. The first two categories impact the signal and simulated component of the resonant background estimation. All of these uncertainties except the luminosity uncertainty are propagated to the kinematic event variables prior to signal region classification, such that the uncertainty is estimated both for the ANN classification and for the signal and resonant background shape modelling. The impact of the uncertainties on both the signal region classification and the $m_{bb}$ shape are fully correlated, as are the uncertainties affecting both the signal and the simulated component of the resonant background estimation.

8.1 Experimental systematic uncertainties

The most prominent experimental uncertainty is due to uncertainties in the $b$-jet trigger scale factors which account for differences between data and MC events in the efficiency of the triggers. The per-jet online $b$-tagging efficiency calculated with respect to the offline $b$-tagging efficiency is measured in $t\bar{t}$ events in both data and MC events. A scale factor is then applied to the leading simulated $b$-tagged jet as a function of the leading jet’s $p_T$. For the 2016 dataset, an additional scale factor binned in the leading jet’s $η$ was applied. The total uncertainty in the $b$-jet trigger scale factor is typically 2–5%. These uncertainties are applied to signal events and the embedded Z sample which uses simulated $b$-jets. Additional jet trigger scale-factor uncertainties, which are determined by taking the difference between the leading jet trigger efficiencies as determined in data and MC events using the same procedure, are approximately 1%.

The second leading experimental uncertainty impacts the per-jet track multiplicity distribution used for quark–gluon separation [57] in the ANN. This quantity is affected by the modelling of the charged-particle multiplicity in jet fragmentation models as well as the track reconstruction efficiency in jets. The modelling uncertainties are determined through
a measurement of the per-jet charged-particle multiplicity in
data which is then compared with various models. The experi-
mental uncertainties are evaluated through standard track-
reconstruction efficiency techniques, and their estimation is
fully described in Ref. [57].

Jet energy scale (JES) and jet energy resolution (JER)
uncertainties comprise the next largest group of experimen-
tal uncertainties. The JES uncertainties are primarily deter-
determined using data-based Z-boson-jet, photon-jet and multi-jet
$pt$-balancing techniques [46]. Additional uncertainties are
applied for the energy scale of jets containing $b$-quarks. The
impact of the uncertainty in the JES is estimated by scaling
the jet energies within their uncertainties. JER uncertainties
are also determined from $in situ$ measurements of $Z$-boson-
jet, photon-jet and dijet $pt$ balancing [46]. The systematic
uncertainty due to the JER is calculated by increasing the res-
olution within its uncertainties, smearing the jet energy by
the resulting change in resolution, and comparing the result
with the nominal shape and normalization in simulation.

Offline reconstructed $b$-tagging scale factors and their
associated uncertainties are determined on a per-jet basis
through comparisons of data and MC $t\bar{t}$, $W+c$, $D^*$, and
multi-jet events [47,58,59]. For $b$-jets the uncertainty in the
scale factors is approximately 2%, for $c$-jets it is 10%, and
for light-quark and gluon jets it is 30%.

The uncertainty in the integrated luminosity is 1.7% [60],
obtained using the LUCID-2 detector [61] for the primary
luminosity measurements. This uncertainty is applied to the
signal process. Other uncertainties include the uncertainty from
the jet vertex tagging requirement, which is measured on a per-jet basis in $Z \rightarrow \ell^+\ell^-$ events with one jet, and the
MC pile-up reweighting uncertainty.

8.2 Theoretical systematic uncertainties

Theoretical uncertainties primarily affect the signal mod-
eling and are divided into several categories. The value of the
$H \rightarrow bb$ branching ratio and its uncertainty are calcul-
dated by the HDECAY program [62] using the LHC Higgs
Cross Section Working Group recommendations [63] with
$m_H = 125$ GeV. To estimate the effect of missing higher-
order terms on the QCD calculations used to predict the VBF
and ggF cross-section and acceptance, the chosen renormal-
ization and factorization scales are independently varied by
factors of 0.5 and 2.0. Additional acceptance uncertainties are
applied to ggF events in which additional radiation pro-
duces a VBF-like topology by having an additional jet pair
with an invariant mass greater than 400 GeV [63]. There is a
20% uncertainty for all events in this category, and additional
uncertainties for events with a fifth jet. The impact of missing
electroweak corrections on the signal cross-section and
acceptance is calculated using the recommendations from
the LHC Higgs Cross Section Working Group, namely that
the uncertainty is calculated from the ratio of the cross-
section calculated with NLO electroweak corrections using
HAWK 2.0 to the leading-order cross-section plus the contri-
butions from photon-induced processes [64]. The uncertainty
is applied as a function of the Higgs boson $pt$. Uncertain-

ties in the cross-section and acceptance due to the choice
of PDF and $\alpha_s$ are evaluated using the error eigenvectors
of the nominal PDF. Additionally, the overall PDF + $\alpha_s$
cross-section uncertainty is applied as a uniform uncertainty. In
order to estimate the uncertainty due to the choice of parton
shower and underlying-event model, the difference between
generator-level signal samples showered with PYTHIA 8.212
and HERWIG 7.0 [65,66] is determined as a ratio in the vari-
ables $m_{bb}$ and $N_{jets}$ which is applied to the nominal signal
weights. These variables are chosen as they show the largest
differences at the truth level. When calculating the signifi-
cance and the signal strength, all uncertainties are used. When
calculating the cross-sections, the cross-section uncertainties
are removed and only acceptance uncertainties are consid-
ered. Contributions from $t\bar{t}H$ and $VH$ are small and a con-
servative 100% uncertainty is assigned to their cross-section
and acceptance.

8.3 Data-driven background systematic uncertainties

Additional uncertainties arise from potential biases in the
non-resonant background shape in the Higgs boson mass
window due to the fact that the mass window is not used in the
ANN training. Studies of multi-jet MC confirm that training
on the sidebands does not induce any bias in the Higgs boson
mass window, however these samples do not have sufficient
statistics to derive an uncertainty. Therefore these uncertain-
ties are assessed with shape-bias control regions (CRs). The
CRs are constructed by inverting one or more event selec-
tion cuts to produce an orthogonal event sample which has
similar kinematics to the SR. Then the CR events undergo a
translation and reweighting, described below, to correct the
CR kinematics back to the SR kinematics. Lastly, the ANN
is applied, generating one control region for each region in
each channel.

The translation is done by smearing the jet $\eta$ of the CR
VBF jets with a Gaussian width of 200 GeV/$p_{T,jet}$. The new
jet $\eta$ is called $\eta'$. The smearing of the original jet $\eta$ value
is repeated until the modified event meets the signal region
selections. Table 3 presents a summary of the event selection
in the CRs. The Forward channel CR uses a trigger requiring
four central jets, at least two of which are $b$-tagged, and the
event selection requires four jets with $|\eta| < 2.5$, at least two
of which are $b$-tagged. In the Central channel CR, the trigger
and event selection is the same as in the SR except that the
CR requires $\Delta\phi_{jj} > \pi/4$ and the $m_{jj}$ cut is inverted, i.e.
$m_{jj} < 800$ GeV. While the $m_{jj}$ cut is sufficient to ensure
orthogonality, the $\Delta\phi_{jj}$ cut improves agreement between CR
and SR kinematic distributions. In order to improve agreement between the SR and CR in the ANN output score distributions, the \textit{Forward} channel CR events are sequentially reweighted to match SR kinematics in $n_{\text{jets}}, m_{jj}$, and $p_{T,jj}$, and the \textit{Central} channel CR events are sequentially reweighted in $\Delta \eta(bb,jj)$, $n_{\text{jets}}$, and $p_{T}^{\text{balance}}$.

To determine the bias uncertainty, the CR $m_{bb}$ distribution is reweighted to match the SR $m_{bb}$ using a reweighting factor derived from RegLow in each channel. This reweighting is performed by first doing a simultaneous fit of each CR and its RegLow to the $Z$ and non-resonant background components. As in the fit described in Sect. 7, the non-resonant background has arbitrary shape and the $Z$ normalization and shape is determined using constraints from the embedding. Next, the $Z$ component is subtracted from the CRs such that only the non-resonant background remains. Then the embedded $Z$ component is subtracted from the SR RegLow, leaving only the SR RegLow non-resonant background, as the contamination from the signal is negligible. The ratio of the $Z$-subtracted SR and CR RegLow $m_{bb}$ shapes is determined. This ratio varies smoothly from 0.95 to 1.1 (1.05) in the \textit{Forward (Central)} channel. A third-order polynomial fits it well in the range $80 \text{ GeV} < m_{bb} < 180 \text{ GeV}$, with a $\chi^2$ probability of 0.92 (0.19) in the \textit{Forward (Central)} region. The ratio is applied to all of the $Z$-subtracted CRs for each channel. Figure 3 shows the distribution of ANN scores for the SR events, the CR events prior to reweighting, the CR events after kinematic reweighting, and the CR events after kinematic and $m_{bb}$ reweighting. It can be seen that after kinematic reweighting the SR and CR ANN distributions agree well. There is good agreement between the reweighted CR and SR $m_{bb}$ sidebands for all regions, with $\chi^2$ probabilities ranging from 0.20 to 0.91. Lastly, the $Z$ component is added back to each CR, and each CR is fit for a bias signal simultaneously with RegLow for that channel’s control region. The shape of the bias signal is the same as the Higgs boson signal shape. The mean of the bias signal is scanned from 100 to 140 GeV and the largest bias signal found in this range is taken as an additional uncertainty in the number of Higgs boson events in that region for the Higgs boson signal-strength fit. True signal contamination in the control regions is negligible. The maximum observed bias ranges from 60% of the expected signal in the least sensitive signal regions to 14% of the expected signal in the most sensitive signal region. The bias uncertainties are uncorrelated between channels and regions because they are dominated by statistical effects. Assuming the bias uncertainties to be correlated leads to a modest increase in the overall uncertainty and no change in the central value of the fit. Several modifications to the CR definitions, smearing procedure and reweighting schemes, including using no reweighting, were tested and no significant change in the bias uncertainty was found.

The uncertainty in the resonant background normalization related to the embedding process is determined from the 20% residual non-closure of the method seen in the variables used in the ANN. Studies show that the uncertainty in the Higgs boson production signal-strength parameter is not sensitive to variations in the resonant background width uncertainty for a

---

Table 3 Event selection criteria for the \textit{Forward} and \textit{Central} channel CRs. The $p_T$ and $|\eta|$ requirements on the offline jets are used to match trigger selections and flavour-tagging requirements. The variables $\eta^{'}$ and $m^{'}_{jj}$ refer to the $\eta$ and $m_{jj}$ of the VBF jets after the smearing procedure described in the text.

<table>
<thead>
<tr>
<th>Channel CR Event Selection</th>
<th>\textit{Forward}</th>
<th>\textit{Central}</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$</td>
<td>$\geq 1$ b-tagged jet at 77% efficiency working point with $p_T &gt; 85 \text{ GeV}$ and $</td>
<td>\eta</td>
</tr>
<tr>
<td>$b_2$</td>
<td>$\geq 1$ b-tagged jet at 85% efficiency working point with $p_T &gt; 65 \text{ GeV}$ and $</td>
<td>\eta</td>
</tr>
<tr>
<td>$j_1$</td>
<td>$\geq 1$ jet with $p_T &gt; 60 \text{ GeV}$, $</td>
<td>\eta^{'}</td>
</tr>
<tr>
<td>$j_2$</td>
<td>$p_{T,bb} &gt; 150 \text{ GeV}$</td>
<td>No jets with $p_T &gt; 60 \text{ GeV}$, $3.2 &lt;</td>
</tr>
<tr>
<td></td>
<td>No jet pair (not including $b_1, b_2$) with $m_{jj} &gt; 800 \text{ GeV}$</td>
<td>no jet pair (not including $j_1, j_2$) with $m_{jj} &gt; 800 \text{ GeV}$</td>
</tr>
<tr>
<td></td>
<td>$m_{jj} &gt; 800 \text{ GeV}$</td>
<td>$m_{jj} &gt; 800 \text{ GeV}$</td>
</tr>
</tbody>
</table>
large range of non-zero values, and therefore a conservative 20% is taken as the width uncertainty. These uncertainties are in addition to the experimental uncertainties related to jet reconstruction, triggering and $b$-tagging described above.

9 Results

The fits to the full mass region are shown in Figs. 4 and 5 for all signal regions in the Forward and Central channels, respectively. A summary plot, with all signal regions summed, weighted by $\ln(1 + s/b)$, is shown in Fig. 6. The values of $s$ and $b$ are the integrated numbers of signal and background events in a region centred on the Higgs boson mass and containing 68% of the signal events. This weighting approximates the relative contribution each region has in the overall significance of the measurement. The quality of the fit was checked by performing an eight-parameter-of-interest fit to separately determine the signal strength in each of the four signal regions of each channel. The probability of compatibility of the signal strengths with a single value is 88% consistent, well within one standard deviation of each other. For this measurement, the contribution from events with $p_T > 200$ GeV is fixed to its Standard Model expectation. The observed signal strength of inclusive Higgs boson production with a true $p_T > 200$ GeV is $0.93^{+0.38}_{-0.38}$ (stat.)$^{+0.24}_{-0.20}$ (syst.), for a significance of $2.2\sigma$, compared with $2.3\sigma$ expected. For this measurement, the contribution from events with true $p_T < 200$ GeV is fixed to its Standard Model expectation. Results of two-dimensional likelihood scans of the signal strength for Higgs boson production with true boson $p_T > 200$ GeV versus the signal strength for Higgs boson production with true boson $p_T < 200$ GeV are shown in the Appendix, where the relative constraining power can be seen. Results of two-dimensional likelihood scans of the signal

Fig. 3 Comparison of the ANN score distributions for the signal and control regions: a shows the Forward channel, b shows the Central channel. The SR selected data are shown in black squares, the CR selected data are shown in red circles prior to reweighting, in green crosses after kinematic reweighting and in blue diamonds after kinematic and $m_{bb}$ reweighting. The bottom panels show the ratio of the CR distributions to the SR distribution

Table 5 summarizes the expected and observed significances and signal-strength parameters for both $\mu_{VBF, H \rightarrow b\bar{b}}$ and $\mu_{H \rightarrow bb}$. Table 5 summarizes the uncertainties in the $\mu_{VBF, H \rightarrow b\bar{b}}$ measurement. The measurement is statistically limited, with the largest uncertainties coming from the data statistics and the non-resonant background bias uncertainty. The experimental systematic uncertainties, which are about a third as large as the statistical uncertainties, are dominated by the trigger and jet energy scale and resolution uncertainties. Taken together, the theory uncertainties are comparable to the largest individual experimental uncertainties, as is the uncertainty due to the embedded Z sample constraint. Due to the requirement on $p_T, bb$, the measurement is primarily sensitive to higher-$p_T$ Higgs boson production. The Higgs boson cross-section at high $p_T$ is sensitive to many BSM models, motivating the definition of dedicated high-$p_T$ (true Higgs boson $p_T > 200$ GeV) regions in the simplified template cross-section framework [67]. The observed signal strength of inclusive Higgs boson production with a true $p_T > 200$ GeV is $0.93^{+0.38}_{-0.38}$ (stat.)$^{+0.24}_{-0.20}$ (syst.), for a significance of $2.2\sigma$, compared with $2.3\sigma$ expected.
Fig. 4 Fit results for the Forward channel. Top left is SR1, top right is SR2, middle left is SR3, middle right is SR4, bottom is RegLow. The data are the black points, the blue line is the non-resonant background shape (continuum), grey is the resonant background ($Z \to b\bar{b}$), red is the fitted VBF Higgs boson signal ($VBF \, H \to b\bar{b}$), and maroon are events attributed to the other Higgs boson production mechanisms (ggF + tH + VH). The bottom panel of each plot shows the difference between the data and the non-resonant background. The hashed band shows the fitted background and signal uncertainties. For RegLow, the dominant background uncertainty is the statistical uncertainty of the shape of the background template. This uncertainty is mainly driven by the statistical uncertainty of data in RegLow, such that the background uncertainty is almost fully correlated with the data uncertainty in RegLow.
Fig. 5 Fit results for the **Central** channel. Top left is SR1, top right is SR2, middle left is SR3, middle right is SR4, bottom is RegLow. The data are the black points, the blue line is the non-resonant background shape (continuum), grey is the resonant background ($Z \rightarrow bb$), red is the fitted VBF Higgs boson signal (VBF $H \rightarrow bb$), and maroon are events attributed to the other Higgs boson production mechanisms ($ggF + tH + VH$). The bottom panel of each plot shows the difference between the data and the non-resonant background. The hashed band shows the fitted background and signal uncertainties. For RegLow, the dominant background uncertainty is the statistical uncertainty of the shape of the background template. This uncertainty is mainly driven by the statistical uncertainty of data in RegLow, such that the background uncertainty is almost fully correlated with the data uncertainty in RegLow.
Fig. 6 The $m_{bb}$ distribution after subtraction of non-resonant background, with signal regions 1 through 4 for both channels weighted by $\ln(1 + s/b)$ within the range $112 < m_{bb} < 136$ GeV and summed. This range corresponds to roughly 68% of the total Higgs boson signal template width. The value of $s$ is calculated from the observed Higgs boson signal (including all $H \rightarrow bb$ production processes) within this range for each region, and $b$ is calculated from the post-fit background yields in each region. The data are the black points, the $Z \rightarrow b\bar{b}$ contribution in grey, the fitted VBF Higgs boson signal is in red, and maroon are events attributed to the other Higgs boson production mechanisms (ggF $+t\bar{t}H + VH$). The hashed band shows the fitted background uncertainties.

Table 4 Expected and observed significances and signal strengths for Higgs boson production, for VBF production, inclusive production, and production with a Higgs boson true $p_T$ of greater than 200 GeV, relative to the Standard Model prediction.

<table>
<thead>
<tr>
<th>Results</th>
<th>VBF Production</th>
<th>Inclusive Production</th>
<th>$p_T &gt; 200$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected significance</td>
<td>2.8$\sigma$</td>
<td>2.9$\sigma$</td>
<td>2.3$\sigma$</td>
</tr>
<tr>
<td>Observed significance</td>
<td>2.6$\sigma$</td>
<td>2.7$\sigma$</td>
<td>2.2$\sigma$</td>
</tr>
<tr>
<td>Expected signal strength</td>
<td>1.00$^{+0.38}_{-0.37}$</td>
<td>1.00$^{+0.37}_{-0.36}$</td>
<td>1.00$^{+0.45}_{-0.43}$</td>
</tr>
<tr>
<td>Observed signal strength</td>
<td>0.95$^{+0.38}_{-0.36}$</td>
<td>0.95$^{+0.37}_{-0.35}$</td>
<td>0.93$^{+0.45}_{-0.43}$</td>
</tr>
</tbody>
</table>

Table 5 Breakdown of the observed uncertainties on the inclusive production signal strength measurement ($\mu_{VBF, H \rightarrow bb}$).

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>$\sigma(\mu_{VBF, H \rightarrow bb})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistics</td>
<td>±0.32</td>
</tr>
<tr>
<td>NR background bias</td>
<td>±0.15</td>
</tr>
<tr>
<td>Embedded Z</td>
<td>±0.05</td>
</tr>
<tr>
<td>Experimental</td>
<td>+0.10/−0.06</td>
</tr>
<tr>
<td>Trigger</td>
<td>+0.07/−0.03</td>
</tr>
<tr>
<td>Jet</td>
<td>+0.07/−0.04</td>
</tr>
<tr>
<td>Flavor tagging</td>
<td>+0.02/−0.01</td>
</tr>
<tr>
<td>Other</td>
<td>+0.03/−0.02</td>
</tr>
<tr>
<td>Signal Theory</td>
<td>+0.06/−0.03</td>
</tr>
</tbody>
</table>

The measured $\mu_{VBF}$ corresponds to a cross-section $\sigma_{VBF} \times B_{H \rightarrow bb} = 2.07^{+0.70}_{-0.70}$ (stat.)$^{+0.46}_{-0.37}$ (syst.) pb. Assuming a branching fraction $B_{H \rightarrow bb} = 0.5809$ [63], this corresponds to an inclusive VBF cross-section $\sigma_{VBF} = 3.56^{+1.21}_{-1.12}$ (stat.)$^{+0.80}_{-0.64}$ (syst.) pb. The predicted inclusive VBF cross-section calculated by the LHC Higgs Working Group at NNLO in QCD with proVBFH [68], including corrections at NLO from electroweak and photon processes determined with HAWK 2.0, is 3.77 ± 0.08 pb. A fiducial cross-section is determined by requiring $|y_H| < 2.5$ and is measured to be $\sigma_{VBF, fid} = 3.31^{+1.12}_{-1.12}$ (stat.)$^{+0.74}_{-0.70}$ (syst.) fb.

These results are combined with the photon analysis [16] described at the end of Sect. 1. The photon analysis yields a signal strength of 1.3 ± 1.0 with a significance of 1.3$\sigma$, compared with an expectation of 1.0$\sigma$ for both inclusive and VBF production. The two analyses are orthogonal, as the measurement described in this paper vetoes events with photons. They are combined in a joint likelihood fit, sharing a common parameter of interest. All experimental systematic uncertainties are correlated with the exception of the jet energy resolution and offline flavour-tagging uncertainties, as the two analyses use different jet definitions. The jet energy scale is partially correlated because both jet definitions use similar datasets to derive the uncertainties. Theoretical uncertainties are correlated except for the parton shower uncertainties, which are determined using a variation of the HERWIG HardScale parameter in the photon analysis, and a comparison between PYTHIA and HERWIG in this analysis. The combined measured signal strength for VBF production is $0.99^{+0.30}_{-0.30}$ (stat.)$^{+0.18}_{-0.18}$ (syst.), corresponding to a significance of 2.9$\sigma$ (2.9$\sigma$ expected). This corresponds to an observed $\sigma_{VBF} \times B_{H \rightarrow bb} = 2.16^{+0.07}_{-0.06}$ (stat.)$^{+0.40}_{-0.35}$ (syst.) pb. When fitting for inclusive Higgs boson production the observed signal strength is $0.99^{+0.30}_{-0.30}$ (stat.)$^{+0.19}_{-0.15}$ (syst.), corresponding to an observed (expected) significance of 3.0$\sigma$ (3.0$\sigma$). The combination significances and signal strengths are summarized in Table 6.
Table 6 Expected and observed results for the Higgs boson production rate, for both inclusive production and VBF production only, relative to the Standard Model prediction for the combined analyses

<table>
<thead>
<tr>
<th>Results</th>
<th>VBF production</th>
<th>Inclusive production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected significance</td>
<td>2.9σ</td>
<td>3.0σ</td>
</tr>
<tr>
<td>Observed significance</td>
<td>2.9σ</td>
<td>3.0σ</td>
</tr>
<tr>
<td>Expected signal strength</td>
<td>1.00±0.36</td>
<td>1.00±0.35</td>
</tr>
<tr>
<td>Observed signal strength</td>
<td>0.99±0.36</td>
<td>0.99±0.35</td>
</tr>
</tbody>
</table>

10 Conclusion

This paper presents a measurement of Standard Model VBF Higgs boson production in the b̄b decay mode using 126 fb⁻¹ of 13 TeV proton–proton data collected with the ATLAS detector at the LHC. Significant improvements relative to the previous analysis using 31 fb⁻¹ lead to a measurement of VBF Higgs boson production with a significance of 2.6σ relative to the background-only hypothesis, with an observed signal strength of 0.95±0.32(stat.)±0.20(syst.) compared to an expected value of 1.00±0.32(stat.)±0.21(syst.). The improvements include the usage of an ANN to decorrelate m_hb from the event classifier, boosting the statistical power of the fit by allowing the same background shape to be used in all signal regions. Additionally, the usage of embedding to estimate the Z boson contribution directly from data results in significantly less uncertainty on the Z boson contribution to the fit. The event selection and classifier provide sensitivity primarily to VBF production at high Higgs boson p_T. The measured cross-section is σ_{VBF} × B_{H→b̄b} = 2.07±0.76(stat.)±0.31(syst.) fb, and the inclusive VBF cross-section σ_{VBF} = 3.56±1.21(stat.)±0.80(syst.) fb. Requiring |y_H| < 2.5, this corresponds to a fiducial cross-section, σ_{VBF-fid} = 3.31±1.12(stat.)±0.74(syst.) fb. These results are combined with a complementary measurement of VBF H → b̄b production in association with a photon, leading to a 2.9σ measurement of VBF H → b̄b production.

Acknowledgements We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SNU, Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. Funded by SCOAP3.
Fig. 7 Likelihood scan as a function of the signal strength for Higgs boson production with true boson $p_T > 200$ GeV versus the signal strength for Higgs boson production with true boson $p_T < 200$ GeV. Shown is the best-fit value (blue plus sign), and the 68% (solid blue) and 95% (dashed blue) confidence level (CL) contours.

Appendix

This analysis is particularly sensitive to high Higgs boson $p_T$ where true Higgs boson $p_T, p_{T,H}$, is greater than 200 GeV, but is unable to separately constrain the cross-sections at high and low $p_T$. In the forward channel, 54% of events have both $p_{T,bb}$ and $p_{T,H} > 200$ GeV, 34% have both $p_{T,bb}$ and $p_{T,H} < 200$ GeV, and 8 (4)% have $p_{T,bb} < (>) 200$ GeV and $p_{T,H} > (>) 200$ GeV. In the central channel, 79% of events have both $p_{T,bb}$ and $p_{T,H} > 200$ GeV, 13% have both $p_{T,bb}$ and $p_{T,H} < 200$ GeV, and 7 (2)% have $p_{T,bb} < (>) 200$ GeV and $p_{T,H} > (>) 200$ GeV. The two-dimensional likelihood scan of high- and low-$p_T$ Higgs boson production is shown in Fig. 7, where the relative size of the ranges for each $p_T$ bin gives a measure of the relative sensitivity to high and low Higgs boson $p_T$. Because the total VBF contribution is rather well-constrained, the production at high and low $p_T$ is anti-correlated.

This analysis does not strongly constrain the ggF Higgs boson production cross-section, as the ggF kinematics are similar to the kinematics of the SM background, resulting in a similar ANN score distribution as the background. The VBF production cross-section cannot be measured precisely if the ggF cross-section is unconstrained, particularly due to large uncertainties for ggF events with at least two extra jets and high $m_{jj}$. The two-dimensional likelihood scan of VBF and ggF production signal strengths is shown in Fig. 8. A large range in the ggF signal strength is observable in the figure, and the impact of the lack of ggF constraint on the VBF production signal strength is only found to be large when the ggF production signal strength is far from the SM expectation.

References


55. F. Chollet et al., Keras, https://keras.io (2015)


22 (a) Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá, Colombia; (b) Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia
23 (a) Dipartimento di Fisica, INFN Bologna and Universita’ di Bologna, Bologna, Italy; (b) INFN Sezione di Bologna, Bologna, Italy
24 Physikalisches Institut, Universität Bonn, Bonn, Germany
25 Department of Physics, Boston University, Boston, MA, USA
26 Department of Physics, Brandeis University, Waltham, MA, USA
27 (a) Transilvania University of Brasov, Brasov, Romania; (b) Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania; (c) Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi, Romania; (d) Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj-Napoca, Romania; (e) University Politehnica Bucharest, Bucharest, Romania; (f) West University in Timisoara, Timisoara, Romania
28 (a) Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia; (b) Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
29 Physics Department, Brookhaven National Laboratory, Upton, NY, USA
30 Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina
31 California State University, Long Beach, CA, USA
32 Cavendish Laboratory, University of Cambridge, Cambridge, UK
33 (a) Department of Physics, University of Cape Town, Cape Town, South Africa; (b) iThemba Labs, Cape Town, Western Cape, South Africa; (c) Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg, South Africa; (d) National Institute of Physics, University of the Philippines Diliman, Quezon City, Philippines; (e) Department of Physics, University of South Africa, Pretoria, South Africa; (f) School of Physics, University of the Witwatersrand, Johannesburg, South Africa
34 Department of Physics, Carleton University, Ottawa, ON, Canada
35 (a) Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies-Université Hassan II, Casablanca, Morocco; (b) Faculté des Sciences, Université Ibn-Tofail, Kenitra, Morocco; (c) Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA-Marrakech, Marrakech, Morocco; (d) Moroccan Foundation for Advanced Science Innovation and Research (MAscIR), Rabat, Morocco; (e) LPMR, Faculté des Sciences, Université Mohamed Premier, Oujda, Morocco; (f) Faculté des sciences, Université Mohammed V, Rabat, Morocco
36 CERN, Geneva, Switzerland
37 Enrico Fermi Institute, University of Chicago, Chicago, IL, USA
38 LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand, France
39 Nevis Laboratory, Columbia University, Irvington, NY, USA
40 Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
41 (a) Dipartimento di Fisica, Università della Calabria, Rende, Italy; (b) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Frascati, Italy
42 Physics Department, Southern Methodist University, Dallas, TX, USA
43 Physics Department, University of Texas at Dallas, Richardson, TX, USA
44 National Centre for Scientific Research “Demokritos”, Agia Paraskevi, Greece
45 (a) Department of Physics, Stockholm University, Stockholm, Sweden; (b) Oskar Klein Centre, Stockholm, Sweden
46 Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen, Germany
47 Lehrstuhl für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany
48 Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany
49 Department of Physics, Duke University, Durham, NC, USA
50 SUPA-School of Physics and Astronomy, University of Edinburgh, Edinburgh, UK
51 INFN e Laboratori Nazionali di Frascati, Frascati, Italy
52 Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
53 II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
54 Département de Physique Nucléaire et Corpusculaire, Université de Genève, Geneva, Switzerland
55 (a) Dipartimento di Fisica, Università di Genova, Genoa, Italy; (b) INFN Sezione di Genova, Genoa, Italy
56 II. Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany
57 SUPA-School of Physics and Astronomy, University of Glasgow, Glasgow, UK
58 LPSC, Université Grenoble Alpes, CNRS/IN2P3, Grenoble INP, Grenoble, France
Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge, MA, USA

Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei, China; (b) Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao, China; (c) School of Physics and Astronomy, Shanghai Jiao Tong University, Key Laboratory for Particle Astrophysics and Cosmology (MOE), SKLPPC, Shanghai, China; (d) Tsung-Dao Lee Institute, Shanghai, China

Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany; (b) Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany

Department of Physics, Chinese University of Hong Kong, Shatin N.T., Hong Kong, China; (b) Department of Physics, University of Hong Kong, Hong Kong, Hong Kong; (c) Department of Physics and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

Department of Physics, National Tsing Hua University, Hsinchu, Taiwan

ICLlab, Université Paris-Saclay, CNRS/IN2P3, 91405, Orsay, France

Department of Physics, Indiana University, Bloomington, IN, USA

INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine, Italy; (b) ICTP, Trieste, Italy; (c) Dipartimento Politecnico di Ingegneria e Architettura, Università di Udine, Udine, Italy

INFN Sezione di Lecce, Lecce, Italy; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce, Italy

INFN Sezione di Milano, Milano, Italy; (b) Dipartimento di Fisica, Università di Milano, Milano, Italy

INFN Sezione di Napoli, Naples, Italy; (b) Dipartimento di Fisica, Università di Napoli, Naples, Italy

INFN Sezione di Pavia, Pavia, Italy; (b) Dipartimento di Fisica, Università di Pavia, Pavia, Italy

INFN Sezione di Pisa, Pisa, Italy; (b) Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa, Italy

INFN Sezione di Roma, Rome, Italy; (b) Dipartimento di Fisica, Sapienza Università di Roma, Rome, Italy

INFN Sezione di Roma Tor Vergata, Rome, Italy; (b) Dipartimento di Fisica, Università di Roma Tor Vergata, Rome, Italy

INFN Sezione di Roma Tre, Rome, Italy; (b) Dipartimento di Matematica e Fisica, Università Roma Tre, Rome, Italy

INFN-TIFPA, Trento, Italy; (b) Università degli Studi di Trento, Trento, Italy

Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck, Austria

University of Iowa, Iowa City, IA, USA

Department of Physics and Astronomy, Iowa State University, Ames, IA, USA

Joint Institute for Nuclear Research, Dubna, Russia

Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora, Brazil; (b) Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil; (c) Instituto de Física, Universidade de São Paulo, São Paulo, Brazil

KEK, High Energy Accelerator Research Organization, Tsukuba, Japan

Graduate School of Science, Kobe University, Kobe, Japan

AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland; (b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Kraków, Poland

Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland

Faculty of Science, Kyoto University, Kyoto, Japan

Kyoto University of Education, Kyoto, Japan

Research Center for Advanced Particle Physics and Department of Physics, Kyushu University, Fukuoka, Japan

Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina

Physics Department, Lancaster University, Lancaster, UK

Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK

Department of Experimental Particle Physics, Jožef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana, Slovenia

School of Physics and Astronomy, Queen Mary University of London, London, UK

Department of Physics, Royal Holloway University of London, Egham, UK

Department of Physics and Astronomy, University College London, London, UK

Louisiana Tech University, Ruston, LA, USA

Fysiska institutionen, Lunds universitet, Lund, Sweden

Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France
Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz, CA, USA
(a) Departamento de Física, Pontificia Universidad Católica de Chile, Santiago, Chile; (b) Universidad Andres Bello, Department of Physics, Santiago, Chile; (c) Instituto de Alta Investigación, Universidad de Tarapacá, Arica, Chile; (d) Departamento de Física, Universidad Técnica Federico Santa María, Valparaiso, Chile
Universidade Federal de São João del Rei (UFSJ), São João del Rei, Brazil
Department of Physics, University of Washington, Seattle, WA, USA
Department of Physics and Astronomy, University of Sheffield, Sheffield, UK
Department of Physics, Shinshu University, Nagano, Japan
Department Physik, Universität Siegen, Siegen, Germany
Department of Physics, Simon Fraser University, Burnaby, BC, Canada
SLAC National Accelerator Laboratory, Stanford, CA, USA
Physics Department, Royal Institute of Technology, Stockholm, Sweden
Departments of Physics and Astronomy, Stony Brook University, Stony Brook, NY, USA
Department of Physics and Astronomy, University of Sussex, Brighton, UK
School of Physics, University of Sydney, Sydney, Australia
Institute of Physics, Academia Sinica, Taipei, Taiwan
(a) E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi, Georgia; (b) High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia
Department of Physics, Technion, Israel Institute of Technology, Haifa, Israel
Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece
International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo, Japan
Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
Tomsk State University, Tomsk, Russia
Department of Physics, University of Toronto, Toronto, ON, Canada
(a) TRIUMF, Vancouver, BC, Canada; (b) Department of Physics and Astronomy, York University, Toronto, ON, Canada
Division of Physics and Tonomaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba, Japan
Department of Physics and Astronomy, Tufts University, Medford, MA, USA
Department of Physics and Astronomy, University of California Irvine, Irvine, CA, USA
Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
Department of Physics, University of Illinois, Urbana, IL, USA
Instituto de Física Corpuscular (IFIC), Centro Mixto Universidad de Valencia - CSIC, Valencia, Spain
Department of Physics, University of British Columbia, Vancouver, BC, Canada
Department of Physics and Astronomy, University of Victoria, Victoria, BC, Canada
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg, Würzburg, Germany
Department of Physics, University of Warwick, Coventry, UK
Waseda University, Tokyo, Japan
Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot, Israel
Department of Physics, University of Wisconsin, Madison, WI, USA
Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal, Germany
Department of Physics, Yale University, New Haven, CT, USA
a Also at Borough of Manhattan Community College, City University of New York, New York, NY, USA
b Also at Center for High Energy Physics, Peking University, China
c Also at Centro Studi e Ricerche Enrico Fermi, Rome, Italy
d Also at CERN, Geneva, Switzerland
e Also at CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
f Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Geneva, Switzerland
g Also at Departament de Física de la Universitat Autònoma de Barcelona, Barcelona, Spain
h Also at Department of Financial and Management Engineering, University of the Aegean, Chios, Greece
i Also at Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA