Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair at $\sqrt{s} = 13$ TeV with 137 fb$^{-1}$

The CMS Collaboration

Abstract

An inclusive search for the standard model Higgs boson produced with large transverse momentum ($p_T$) and decaying to a bottom quark-antiquark pair is performed using pp collisions data collected by the CMS experiment at the LHC at $\sqrt{s} = 13$ TeV. The data sample corresponds to an integrated luminosity of 137 fb$^{-1}$. Highly Lorentz-boosted Higgs bosons decaying to $b\bar{b}$ are reconstructed as single, large radius jets, and are identified using jet substructure and dedicated $b$ tagging techniques based on a deep neural network. The method is validated with $Z \rightarrow b\bar{b}$ decays. For a Higgs boson mass of 125 GeV, an excess of events above the expected background is observed with a local significance of 2.54 standard deviations, where the expectation is 0.71. The corresponding signal strength is $\mu_H = 3.68 \pm 1.20$ (stat)$^{+0.63}_{-0.66}$ (syst)$^{+0.81}_{-0.46}$ (theo) with respect to the standard model expectation. Additionally, an unfolded differential cross section as a function of Higgs boson $p_T$ is presented. With respect to the previous CMS result, the relative precision of the Higgs boson signal strength measurement improves by approximately a factor of two. The improvement is due to the increased integrated luminosity, improved $b$ tagging, and smaller theoretical uncertainties.
1. Introduction

In the standard model (SM) [1–3], the Brout–Englert–Higgs mechanism [4–8] is responsible for electroweak symmetry breaking and the mass of elementary particles. The Higgs boson (H) has now been observed in several production and decay modes [9–11], including decays to bottom quarks when produced in association with a W or Z boson [12–14]. Recently, there has been considerable interest in the measurement of H bosons produced with high transverse momentum, $p_T$, where the $H \rightarrow b\bar{b}$ decay channel has better sensitivity than traditional channels because of its large branching fraction, $B(H \rightarrow b\bar{b}) = 58.1\%$ [15]. The first searches for high-$p_T$ H($b\bar{b}$) events by the CMS [16] and ATLAS collaborations [17] have demonstrated the experimental power of this channel, with observed significances of 1.5 and 1.6 standard deviations, respectively. Measurements of high-$p_T$ H($b\bar{b}$) events [18] have shown promise in resolving the loop induced and tree-level contributions to the gluon fusion (ggF) process [19], and provide an alternative approach to study the top quark Yukawa coupling in addition to the $t\bar{t}H$ process.

This note reports the results of a search for high-$p_T$ Higgs bosons decaying to b$\bar{b}$ pairs based on a data set of pp collisions at $\sqrt{s} = 13$ TeV collected with the CMS detector at the LHC and corresponding to an integrated luminosity of 137 fb$^{-1}$. The highly Lorentz-boosted H($b\bar{b}$) candidates are reconstructed as single, large-radius jets with jet mass consistent with that of the standard model Higgs boson [20]. The candidate jet is required to have $p_T > 450$ GeV to satisfy restrictive trigger requirements that suppress the large SM background from jets produced through the strong interaction, referred to as quantum chromodynamics (QCD) multijet events. To further distinguish the H boson candidates from the background, the jet is required to have two-prong substructure and b tagging properties consistent with the H($b\bar{b}$) signal. The events are divided into six $p_T$ categories from 450 GeV to 1.2 TeV, with widths ranging from 50 to 400 GeV. The background from QCD multijet production is difficult to model parametrically, and is therefore estimated in data by inverting the b tagging requirement, which is, by design, decorrelated from jet mass and $p_T$. A simultaneous fit to the distributions of the jet mass in all $p_T$ categories is performed in the range 47 to 201 GeV to determine the normalizations and shapes of the jet mass distributions for the backgrounds and to extract the H($b\bar{b}$) production cross sections.

With respect to the previous CMS result, the Higgs boson $p_T$ spectrum from ggF is better modeled with the HJ-MINLO generator [21, 22], which is compatible with the latest theoretical predictions [23, 24] including full finite top mass effect. Another major improvement is the development of a b tagging algorithm based on a deep neural network with better H($b\bar{b}$) signal efficiency.

2. The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T inside its volume. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity ($\eta$) coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid.

Events of interest are selected using a two-tiered trigger system [25]. The first level, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a time interval of less than 4 $\mu$s. The second
level, known as the high-level trigger, consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage.

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [26].

3 Simulated samples

Simulated samples of signal and background events are produced using various Monte Carlo (MC) event generators, with the CMS detector response modeled by GEANT4 [27].

For 2016 running conditions, the QCD multijet and Z+jets processes are modeled at leading order (LO) accuracy using the the MADGRAPH5_aMC@NLO v2.2.2 generator [28]. The W+jets process is modeled at LO accuracy with MADGRAPH5_aMC@NLO v2.3.3. Jets from the matrix element calculation and the parton shower description are matched using the MLM prescription [29]. The tf and single top processes are modeled at next-to-leading order (NLO) using POWHEG 2.0 [30–35]. Diboson processes are modeled with PYTHIA 8.205 [36].

For 2017 and 2018 running conditions, the same configurations are used, but with newer generator versions. The QCD multijet, Z+jets, and W+jets processes are modeled using MADGRAPH5_aMC@NLO v2.4.2, and the diboson processes are modeled with PYTHIA 8.226.

Higgs boson production through gluon fusion (ggH) is simulated using the POWHEG+HJ-MINLO [21, 22, 31, 37] event generator with $m_H = 125$ GeV, considering finite top mass effects. The HJ-MINLO sample is chosen to be the nominal gluon fusion signal sample, following the recommendation in Ref. [22]. Additionally, a sample of ggH events generated with POWHEG and corrected for finite top mass effect using the same procedure as described in Ref. [16] is considered. The POWHEG generator is used to model Higgs boson production through vector boson fusion (VBF), VH associated production, and t\(t\)H channels [37–39]. The $p_T$ spectrum of the Higgs boson for the VBF production mode is re-weighted to account for N\(^3\)LO corrections to the cross section [40, 41]. These corrections have a negligible effect on the yield for this process for events with Higgs boson $p_T > 450$ GeV.

For parton showering and hadronization, the POWHEG and MADGRAPH5_aMC@NLO samples are interfaced with PYTHIA 8.205 (8.230) for 2016 (2017 and 2018) running conditions. The PYTHIA parameters for the underlying event description are set to the CUETP8M1 tune [42] (CP5 tune [43]). For 2016 samples, the parton distribution function set NNPDF3.0 [44] is used, with accuracy (LO or NLO) corresponding to that of the generator used, while for 2017 and 2018 samples, NNPDF3.1 [45] at next-to-next-to-leading order accuracy is used for all processes. Differences in jet properties between data and simulation are corrected for year by year.

The total cross sections for the diboson samples are corrected to next-to-next-to-leading order accuracy with the MCFM 7.0 program [46]. The cross sections for W+jets and Z+jets samples include higher-order QCD and electroweak (EW) corrections, which improve the modeling of high-$p_T$ W and Z bosons events [47–50].

4 Event reconstruction and selection

Event reconstruction is based on a particle-flow algorithm [51], which aims to reconstruct and identify each individual particle with an optimized combination of information from the var-
4. Event reconstruction and selection

ious elements of the CMS detector. The algorithm identifies each reconstructed particle as an electron, a muon, a photon, or a charged or neutral hadron. The missing transverse momentum vector is defined as the negative vector sum of the transverse momenta of all the particles identified in the event, and its magnitude is referred to as $p_T^{\text{miss}}$.

Particles are clustered into jets using the anti-$k_T$ algorithm [52] with a distance parameter of 0.8 (AK8 jets) or 0.4 (AK4 jets). The clustering algorithms are implemented by the FASTJET package [53]. To mitigate the effect from the contributions of extraneous pp collisions (pileup), the pileup per particle identification algorithm [54] assigns a weight to each particle prior to jet clustering based on the likelihood of the particle to originate from the hard scattering vertex. Further corrections are applied to the jet energy as a function of jet $\eta$ and $p_T$ to bring the measured response of jets to that of particle level jets on average [55]. Additional selection criteria are applied to each jet to remove jets potentially dominated by anomalous contributions from various subdetector components or reconstruction failures. Specifically, jets are required to have neutral hadron and photon energy fractions less than 90% and a nonzero charged hadron energy fraction.

A combination of several event selection criteria is used to trigger on events, all of which impose minimum thresholds on either the AK8 jet $p_T$ or the event $H_T$, defined as the scalar sum of the $p_T$ of all jets in the event with $|\eta| < 3.0$. For AK8 jets used in the trigger selection, a minimum threshold is also imposed on the trimmed jet mass, where remnants of soft radiation are removed before computing the mass [56], which allows the $H_T$ or $p_T$ thresholds to be reduced while maintaining signal acceptance. The trigger selection efficiency is greater than 95% for events with at least one AK8 jet with $|\eta| < 2.5$, mass greater than 47 GeV and $p_T > 450$ (525, 500) GeV for 2016 (2017, 2018) data.

To reduce backgrounds from SM EW processes, events are vetoed if they contain isolated electrons [57], isolated muons [58], or hadronically decaying $\tau$ leptons with $p_T > 10, 10, or 18$ GeV and $|\eta| < 2.5, 2.4, or 2.3$, respectively. For electrons and muons, the isolation criteria require that the pileup-corrected sum of the $p_T$ of charged hadrons and neutral particles surrounding the lepton divided by the lepton $p_T$ be less than approximately 15 or 25%, respectively, depending on $\eta$ [57, 58]. An AK8 jet is also rejected if a photon with $p_T > 175$ GeV is reconstructed within the jet. For each event, the leading jet in $p_T$ is selected to be the H($b\bar{b}$) candidate, which is around 60% efficient for the ggH($b\bar{b}$) signal process. To reduce the top quark background contamination, events are vetoed if they have $p_T^{\text{miss}} > 140$ GeV, or if they contain a b-tagged [59] AK4 jet with $p_T > 30$ GeV located in the opposite hemisphere from the leading AK8 jet ($\Delta\phi$(AK4, AK8) $> \pi/2$).

The soft-drop algorithm [60] with angular exponent $\beta = 0$ and soft radiation fraction $z = 0.1$ is applied to the Higgs boson jet candidate to remove soft and wide-angle radiation. The parameter $\beta$ controls the grooming profile as a function of subjet separation; for $\beta = 0$, the algorithm is independent of subjet separation, and is equivalent to the modified mass-drop tagger [61]. The resulting jet mass, $m_{SD}$, tends to be lowered for background QCD multijet events, where large jet masses arise from soft gluon radiation. Conversely, the algorithm preserves the mass of jets from heavy boson decays. Corrections to the $m_{SD}$ values from simulation are derived from a comparison of simulated and measured samples in a region enriched with merged $W(q\bar{q})$ decays from $t\bar{t}$ events [62]. The $m_{SD}$ corrections remove a residual dependence on the jet $p_T$, and match the simulated jet mass scale and resolution to those observed in data.

The resulting $m_{SD}$ distributions are binned from 47 to 201 GeV with a bin width of 7 GeV. The lower bound is sufficiently above the trigger threshold to be resilient to threshold effects, and the bin width corresponds to the $m_{SD}$ resolution near the W and Z resonances. The dimension-
less mass scale variable for QCD multijet jets, $\rho = 2 \ln(m_{SD}/p_T)$ [61, 63], is used to characterize the correlation between the jet b tagging discriminator, jet mass, and jet $p_T$. Its distribution is roughly invariant in different ranges of jet $p_T$. For each $p_T$ category, only those $m_{SD}$ bins whose barycenter satisfies $-6.0 < \rho < -2.1$ are considered. The upper bound is imposed to avoid instabilities at the edges of the distribution due to finite cone limitations from the jet clustering, while the lower bound avoids the nonperturbative regime of the $m_{SD}$ calculation. This requirement is about 98% efficient for the $H(b\bar{b})$ signal.

The $N_{1,2}$ variable [64] is used to determine how consistent a jet is with having a two-prong substructure. It is based on a ratio of 2-point ($1^e_2$) and 3-point ($2^e_3$) generalized energy correlation functions [65]:

$$1^e_2 = \sum_{1 \leq i < j \leq n} z_i z_j \Delta R_{ij}, \quad (1)$$

$$2^e_3 = \sum_{1 \leq i < j < k \leq n} z_i z_j z_k \min \{ \Delta R_{ij} \Delta R_{ik}, \Delta R_{ij} \Delta R_{jk}, \Delta R_{ik} \Delta R_{jk} \}, \quad (2)$$

where $z_i$ represents the energy fraction of the constituent $i$ in the jet and $\Delta R_{ij}$ is the angular separation between constituents $i$ and $j$. These generalized energy correlation functions are sensitive to correlations of pairwise angles among $n$-jet constituents [64]. For a two-prong structure, signal jets have a stronger 2-point correlation than a 3-point correlation. The discriminant variable $N_{1,2}$ is defined as

$$N_{1,2} = \frac{2^e_3}{(1^e_2)^2}. \quad (3)$$

The calculation of $N_{1,2}$ is based on the jet constituents after application of the soft-drop grooming algorithm to the jet. It provides excellent discrimination between two-prong signal jets and QCD background jets. However, imposing requirements on $N_{1,2}$, or other similar variables, distorts the jet mass distributions differently depending on the $p_T$ of the jet [66]. To minimize this distortion, a transformation is applied to $N_{1,2}$ following the designed decorrelated tagger (DDT) technique [63], reducing its correlation with $\rho$ and $p_T$ in multijet events. The transformed variable is defined as $N_{1,2}^{\text{DDT}} \equiv N_{1,2} - X_{(26\%)}$, where $X_{(26\%)}$ is the 26th percentile of the $N_{1,2}$ distribution in simulated QCD events as a function of $\rho$ and $p_T$. The transformation is derived in bins of $\rho$ and $p_T$. This ensures that the selection $N_{1,2}^{\text{DDT}} < 0$ yields a constant QCD background efficiency across the $\rho$ and $p_T$ range considered in this search. The chosen background efficiency of 26% maximizes the signal sensitivity.

Jets likely to originate from two b quarks are selected using an algorithm based on a deep neural network, referred to here as the deep double-b tagging algorithm [59, 67]. The algorithm takes as inputs several high-level observables that characterize the distinct properties of b hadrons and their flight directions in relation to the two subjet candidate axes, as well as low-level track and vertex observables. Events where the selected AK8 jet is double-b tagged constitute the “passing,” or signal, region, while events failing the double-b tagger form the “failing” region, which is used to estimate the QCD multijet background in the signal region. Specifically, an AK8 jet is considered double-b tagged if its double-b tagger discriminator value exceeds a threshold corresponding approximately to a 1% misidentification probability for QCD jets and a 60% efficiency for reconstructed $H(b\bar{b})$ candidates with $450 < p_T < 1200$ GeV. Compared to the previous double-b tagger algorithm [59] used in a prior CMS result [16], which was based
on a boosted decision tree, the deep double-b tagger improves the H(bb) tagging efficiency by a factor of about 1.5 for the same detector conditions and QCD misidentification probability.

After all selections are applied, the Higgs boson candidate jet is categorized into double-b-tag passing and failing regions, each with 22 \( m_{SD} \) bins evenly dividing the range 47–201 GeV, and split further into six jet \( p_T \) categories with bin boundaries 450, 500, 550, 600, 675, 800, and 1200 GeV. The \( p_T \) binning was optimized for best signal significance, where the upper bound is due to the finite jet \( \rho \) acceptance imposed. Of these bins, any bin where the barycenter in \((p_T, m_{SD})\) space satisfies \( \rho < -6 \) or \( \rho > -2.1 \) are removed, resulting in a total of 124 bins.

## 5 Background estimation

The W and Z + jets backgrounds are modeled using simulation. Their overall contribution is less than 1% of the total SM background in the 110 < \( m_{SD} \) < 131 GeV range of the double-b passing region. The normalizations and shapes of the simulated W/Z + jets backgrounds are corrected for NLO QCD and EW effects. Other EW processes, including diboson, triboson, and t\(t\) + W/Z, are estimated from simulation and found to be negligible.

The contribution of t\(t\) production to the total SM background is obtained from simulation, where the normalization and double-b tagging efficiency are corrected with scale factors derived from a t\(t\)-enriched control sample. The control sample targets semileptonic t\(t\) production, consisting of events with an energetic muon with \( p_T > 55 \) GeV and \( |\eta| < 2.1 \), a leading AK8 jet with \( p_T > 400 \) GeV, and an additional b-tagged AK4 jet that is separated from the leading AK8 jet by \( \Delta R > 0.8 \). The AK8 jet with the highest \( p_T \) is taken to be the candidate jet. Using the same candidate jet requirements that define the signal selection, double-b-tag pass and fail regions are constructed in both data and simulation. Due to the relatively low statistics in the control sample, the inclusive event counts for 47 < \( m_{SD} \) < 201 GeV and \( p_T > 400 \) GeV are used. Both the absolute normalization and the deep double-b-tag efficiency of the t\(t\) contribution are allowed to vary without prior constraint from the simulation expectation, but are constrained to vary identically in the t\(t\) control region and the signal region in the simultaneous fit, constraining \textit{in situ} the background expectation and deep double-b mistag efficiency of this process. The net contribution is about 3% of the total SM background in the 110 < \( m_{SD} \) < 131 GeV range of the deep double-b passing region.

The main background in the passing region, QCD multijet production, has a nontrivial jet mass shape that is difficult to model parametrically and depends on jet \( p_T \). Therefore, we constrain it using the background-enriched failing region, i.e., events failing the deep double-b tagger selection, together with a “pass-fail ratio”, \( R_{p/f} \), representing the different mass distributions in the two regions. The pass-fail ratio is factorized into two components. First, the deep double-b tagger discriminator is designed to be uncorrelated from jet mass: the training procedure incorporates a penalty term to the loss function for differences in the jet mass distribution between the passing and failing events. Nonetheless, the tagger exhibits some anticorrelation at high tagger discriminator values and low jet mass, i.e., the mass distributions are different in the passing and failing regions. To account for this, the expected pass-fail ratio is taken from simulated QCD multijet events, specifically by fitting a two-dimensional Bernstein polynomial in \( \rho \) and \( p_T \), \( e^{QCD}(\rho, p_T) \), to the distributions in simulation. Second, residual differences arise from discrepancies in tagger performance between data and simulation, which we parametrize using a Bernstein polynomial in \( \rho \) and \( p_T \). The complete pass-fail ratio in data is given by the
product of these two factors,

\[ R_{p/f}(\rho, p_T) = \sum_{k=0}^{n_\rho} \sum_{\ell=0}^{n_{p_T}} a_{k,\ell} b_{k,n_\rho}(\rho) b_{\ell,n_{p_T}}(p_T) e^{QCD}(\rho, p_T), \tag{4} \]

where \( n_\rho \) is the degree of the polynomial in \( \rho \), \( n_{p_T} \) is the degree of the polynomial in \( p_T \), \( a_{k,\ell} \) is a Bernstein coefficient, and

\[ b_{v,n}(x) = \binom{n}{v} x^v (1-x)^{n-v} \tag{5} \]

is a Bernstein basis polynomial of degree \( n \).

The coefficients \( a_{k,\ell} \) have no external constraints, but are determined from a simultaneous binned fit to data in passing and failing regions across the whole jet mass and \( p_T \) range. The \( p_T \) bin widths, which vary from 50 to 400 GeV, are chosen to provide enough data points to constrain the shape of \( R_{p/f} \). To determine the minimum degree of polynomial necessary to fit the data, a Fisher \( F \)-test \cite{68} is performed. As the magnitude of data-to-simulation discrepancies can vary among the data samples and their corresponding simulation samples, an \( F \)-test is performed independently for each of the three data taking years. For the 2016 data sample, it is found that a polynomial of order \((n_\rho, n_{p_T}) = (2, 1)\) provides a sufficient goodness of fit with respect to increased orders \((p > 0.05)\), while for 2017 and 2018 data, the residual polynomial order is found to be \((n_\rho, n_{p_T}) = (1, 1)\).

In order to validate the background estimation method and associated systematic uncertainties, bias studies are performed on simulated samples and on the background-only fits. Pseudo-experiment data sets are generated, with and without the injection of signal events, and then fit with the signal plus background model. No significant bias in the fitted signal strength is observed; specifically, the means of the differences between the fitted and injected signal strengths divided by the fitted uncertainty are found to be less than 15%.

6 Systematic uncertainties

The systematic uncertainties associated with the jet mass scale, the jet mass resolution, and the \( N_1,\text{DDT}_2 \) selection efficiency are correlated among the \( W, Z, \) and \( H(b\bar{b}) \) processes. These uncertainties are estimated in data using an independent sample of merged \( W \) jets in semileptonic \( t\bar{t} \) events, where the hadronically decaying \( W \) boson is reconstructed as a single AK8 jet.

For this sample, data events are required to have an energetic muon with \( p_T > 100 \text{ GeV} \), \( p_T^{\text{miss}} > 80 \text{ GeV} \), a high-\( p_T \) AK8 jet with \( p_T > 200 \text{ GeV} \), and an additional b-tagged AK4 jet separated from the AK8 jet by \( \Delta R > 0.8 \) with \( p_T > 30 \text{ GeV} \). Using the same \( N_2^{1,\text{DDT}} \) requirement applied in the signal regions, we define samples with events that pass and fail the \( N_2^{1,\text{DDT}} \) selection for merged \( W \) boson jets in data and simulation. A simultaneous fit to the two samples is performed in order to extract the selection efficiency of a merged \( W \) jet in simulation and in data. The data-to-simulation scale factors for the \( N_2^{1,\text{DDT}} \) selection efficiency are measured separately for the three data taking periods, as listed in Table 1.

The jet mass scale and jet mass resolution data-to-simulation scale factors are extracted from the same fit, and are also shown in Table 1. As the semileptonic \( t\bar{t} \) sample does not contain a large population of jets with very high \( p_T \), an additional systematic uncertainty is included to account for the extrapolation to very high \( p_T \) jets. This additional uncertainty is estimated to
be 0.5% per 100 GeV, based on a study of fitting the $m_{SD}$ distributions of $p_T$-binned samples
of merged top quark jets with $p_T > 350$ GeV. Thus, the jet mass scale uncertainty is allowed
to vary in the signal extraction differently depending on the jet $p_T$, and ranges from 1.2% at
450 GeV to 2.1% at 800 GeV.

The efficiency of the deep double-b tagger and the corresponding uncertainty are estimated
in data and simulation in a sample enriched in b$b$ pairs from gluon splitting [59]. The deep
double-b tagger efficiency scale factors, shown in Table 1, are measured in the gluon splitting-
enriched sample, where the uncertainty accounts for various systematic effects including the
calibration of the jet probability tagger algorithm used in the method, the modeling of the
track reconstruction efficiency, the modeling of $b$ quark fragmentation, and others. Given that
differences are to be expected between the behavior of gluon splitting $b$b jets and $b$b jets arising
from Z or H boson decay, the double-b-tag data-to-simulation scale factor is nominally taken as
unity with the difference between the measured tagger efficiency scale factor and unity taken
as a prior uncertainty. The scale factor is further constrained in situ via the observed Z boson
yields in the passing and failing regions.

Table 1: Summary of measured data-to-simulation scale factors for jet mass scale, jet mass
resolution, $N_2^{LDDT}$ selection, and deep double-b selection for different data taking periods.

<table>
<thead>
<tr>
<th>Data period</th>
<th>Jet mass scale</th>
<th>Jet mass resolution</th>
<th>$N_2^{LDDT}$ selection</th>
<th>Deep double-b selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>1.000 ± 0.012</td>
<td>1.084 ± 0.091</td>
<td>0.993 ± 0.043</td>
<td>0.77 ± 0.07</td>
</tr>
<tr>
<td>2017</td>
<td>0.987 ± 0.012</td>
<td>0.905 ± 0.048</td>
<td>0.924 ± 0.018</td>
<td>0.68 ± 0.06</td>
</tr>
<tr>
<td>2018</td>
<td>0.970 ± 0.012</td>
<td>0.908 ± 0.014</td>
<td>0.953 ± 0.016</td>
<td>0.70 ± 0.07</td>
</tr>
</tbody>
</table>

The scale factors described above determine the initial distributions of the jet mass for the
$W(q\bar{q})$, $Z(q\bar{q})$, and $H(b\bar{b})$ processes. In the fit to data, the jet mass scales and resolutions are
further constrained by the presence of the W and Z resonances in the jet mass distribution.

The uncertainty associated with the modeling of the ggF Higgs boson $p_T$ spectrum is propa-
gated to the overall normalization of the ggF Higgs boson signal as a 30% uncertainty for the
POWHEG sample with $p_T$ reweighting [16] and as a 20% uncertainty for the HJ-MNLO sam-
ple. In addition, for the POWHEG sample, the shape of the ggF Higgs boson $p_T$ distribution is
allowed to vary depending on the Higgs boson $p_T$ by up to 30% at 1.2 TeV, without changing
the overall normalization. To account for potential $p_T$-dependent deviations due to missing
higher-order corrections, uncertainties are applied to the $W(q\bar{q})$ and $Z(q\bar{q})$ yields that are $p_T$
dependent and correlated per $p_T$ bin [47, 48, 69–73]. An additional systematic uncertainty is
included to account for potential differences between the W and Z higher-order corrections
(EW W/Z decorrelation) [69].

Finally, additional systematic uncertainties are applied to the $W(q\bar{q})$, $Z(q\bar{q})$, $t\bar{t}$, and $H(b\bar{b})$
yields to account for the uncertainties due to the jet energy scale and resolution [74], variations
in the amount of pileup, the integrated luminosity determination [75], modeling of the
trigger acceptance, and the limited simulation sample sizes. Table 2 lists the major sources of
uncertainty and their observed impact on $\mu_H$ from the combined fit.

7 Results

Figure 1 shows the $m_{SD}$ distributions in the combined data set for the passing and failing re-
gions with the fitted SM background. The bottom panels of Fig. 1 show the difference between
the data and the prediction from the SM background, divided by the statistical uncertainty in
Table 2: Major sources of uncertainty in the measurement of the signal strength $\mu_H$ based on the HJ-MiNLO prediction, and their observed impact ($\Delta \mu_H$) from a fit to the combined data set, are listed. The total uncertainty is separated into three components: statistical, systematic, and theory. Detailed decompositions of the statistical, systematic, and theory components are specified. The impact of each uncertainty is evaluated considering only that source. The sum in quadrature for each source does not in general equal the total uncertainty of each component because of correlations in the combined fit between nuisance parameters in different sources.

<table>
<thead>
<tr>
<th>Uncertainty source</th>
<th>$\Delta \mu_H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical</td>
<td>+1.20 −1.20</td>
</tr>
<tr>
<td>QCD pass-fail ratio (data correction)</td>
<td>+0.71 −0.70</td>
</tr>
<tr>
<td>$t\bar{t}$ normalisation and misidentification</td>
<td>+0.36 −0.36</td>
</tr>
<tr>
<td>Systematic</td>
<td>+0.63 −0.66</td>
</tr>
<tr>
<td>QCD pass-fail ratio (simulation)</td>
<td>+0.48 −0.52</td>
</tr>
<tr>
<td>Deep double-b tag efficiency</td>
<td>+0.32 −0.41</td>
</tr>
<tr>
<td>Jet mass scale and resolution</td>
<td>+0.11 −0.11</td>
</tr>
<tr>
<td>Jet energy scale and resolution</td>
<td>+0.32 −0.29</td>
</tr>
<tr>
<td>Simulation sample size</td>
<td>+0.14 −0.15</td>
</tr>
<tr>
<td>Other experimental uncertainties</td>
<td>+0.07 −0.09</td>
</tr>
<tr>
<td>Theory</td>
<td>+0.81 −0.46</td>
</tr>
<tr>
<td>V+jets modeling</td>
<td>+0.62 −0.38</td>
</tr>
<tr>
<td>H modeling</td>
<td>+0.54 −0.27</td>
</tr>
<tr>
<td>Total</td>
<td>+1.58 −1.45</td>
</tr>
</tbody>
</table>

The data. These highlight the agreement between the data and the contributions from W and Z boson production, which are clearly visible in the failing and passing regions. The W boson contribution in the passing region is due to the misidentification of W(qq) decays by the deep double-b tagger. In Fig. 2, the $m_{SD}$ distributions are reported for each $p_T$ category.

To validate the substructure and b tagging techniques employed in this search, a maximum likelihood fit is performed using a model where the Z and H signal strengths are left unconstrained. The measured Z boson signal strength is $\mu_Z = 1.01 \pm 0.05$ (stat) $+0.20^{+0.13}_{-0.15}$ (syst) $+0.13^{-0.09}_{-0.07}$ (theo). This demonstrates that the Z boson is clearly separable from the background. For the remainder of results, the Z boson signal strength is fixed to its expectation, with corresponding uncertainties as described in Section 6. To extract the Higgs boson signal, three maximum likelihood fits are performed to the data, each with a different degree of reliance on the modeling of Higgs boson $p_T$ spectrum: an inclusive fit using one signal strength parameter for all jet $p_T$ categories, an alternative fit using one signal strength parameter for each $p_T$ category, and a fit which unfolds detector effects to present results at generator level.

In the inclusive fit and using the HJ-MiNLO sample as the gluon fusion signal model, the measured Higgs boson signal strength is $\mu_H = 3.68 \pm 1.20$ (stat) $+0.63^{-0.46}_{-0.06}$ (syst) $+0.81^{-0.46}_{-0.06}$ (theo). The corresponding observed and expected upper limits on the Higgs boson signal strength at a 95% confidence level are 3.70 and 2.90, respectively, while the observed and expected significances with respect to the background-only hypothesis are 2.54$\sigma$ and 0.71$\sigma$. The measurement exhibits an excess over the SM expectation ($\mu_H = 1$), with a significance of 1.85$\sigma$. Table 3 summarizes the measured signal strengths and significances for the Higgs and Z boson processes. Results for both the ggF Higgs boson $p_T$ spectrum from Ref. [16] and HJ-MiNLO [21, 22] are shown. We note that the prediction used for the Higgs boson $p_T$ spectrum in Ref. [16] is different from that of HJ-MiNLO in both shape and total cross section. In particular, the number of ggF Higgs
Figure 1: The observed and fitted background $m_{SD}$ distributions for the failing (left) and passing (right) regions, combining all the $p_T$ categories, and three data collection years. The fit is performed under the signal-plus-background hypothesis with one inclusive H(\text{b}\text{b}) signal strength parameter floating in all the $p_T$ categories. Because of the finite $\rho$ acceptance, some $m_{SD}$ bins within a given $p_T$ category may be removed, giving rise to the features at 166 and 180 GeV. The shaded blue band shows the systematic uncertainty in the total background prediction. The bottom panel shows the difference between the data and the total background prediction, divided by the statistical uncertainty in the data.

boson events predicted in the fiducial region of the analysis with HJ-MINL0 is approximately a factor of two smaller than that of Ref. [16], which is reflected in the observed difference of the fitted Higgs boson signal strength values based on the two predictions.

To assess the compatibility between the observed signal strengths in the different jet $p_T$ categories, an alternative fit to the data is performed. In this fit, an independent H signal strength is assigned to each of the six reconstructed jet $p_T$ bins. These signal strengths are unconstrained in the fit and are varied simultaneously. All other parameters are profiled, as in the nominal fit. Figure 3 illustrates the compatibility in the best fit signal strengths between the different $p_T$ categories.

To facilitate comparisons with theoretical predictions, we isolate and remove the effects of limited detector acceptance and response to the ggF Higgs boson production cross section using a maximum-likelihood unfolding technique. In our treatment, the remaining Higgs boson production modes are assumed to occur at SM rates, as predicted by the HJ-MINL0 simulation. The ggF Higgs boson signal is split into several bins according to the generated Higgs boson $p_T$ ($p_T^H$), and each $p_T^H$ bin is considered as a separate process with a freely floating signal strength parameter in the likelihood model. The respective $p_T^H$ bins are 300–450, 450–650, and $> 600$ GeV. This binning choice follows the simplified template cross section (STXS) recommendation [76]. As the minimum reconstructed jet $p_T$ is 450 GeV, a negligible signal contribution is expected from events with $p_T^H < 300$ GeV. The fiducial cross section in each $p_T^H$ bin is then extracted by scaling the cross section found in simulation, imposing no selection requirements other than those on $p_T^H$, by the corresponding signal strength parameter. The uncertainty in this value is taken from the correspondingly scaled signal strength uncertainty. For the theoretical uncertainties, only those that affect the acceptance of signal events into the reconstructed selection are taken into account. Based on the envelope of acceptance values from varying the
Figure 2: The observed and fitted background $m_{SD}$ distributions in each $p_T$ category in the passing regions. The fit is performed under the signal-plus-background hypothesis with one inclusive $H(b\bar{b})$ signal strength parameter floating in all the $p_T$ categories. The shaded blue band shows the systematic uncertainty in the total background prediction. The bottom panel shows the difference between the data and the total background prediction, divided by the statistical uncertainty in the data.
Table 3: Fitted signal strength, expected and observed significance of the Higgs and Z boson signals. The Higgs boson results are presented with two gluon fusion signal models, one using the nominal HJ-MiNLO sample and the other simulated with the same procedure described in Ref. [16]. The 95% confidence level upper limit (UL) on the Higgs boson signal strength is also listed. In the results for the Higgs boson, the Z boson yield is fixed to the SM prediction value with the corresponding theoretical uncertainties to better constrain in situ the data-to-simulation scale factor of the deep double-b tagger. For the expected and observed signal strength of the Z boson, the Higgs boson signal strength is freely floating.

<table>
<thead>
<tr>
<th></th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected signal strength $\mu_Z$</strong></td>
<td>1.00+0.38−0.28</td>
<td>1.00+0.42−0.29</td>
<td>1.00+0.43−0.29</td>
<td>1.00+0.23−0.19</td>
</tr>
<tr>
<td><strong>Observed signal strength $\mu_Z$</strong></td>
<td>0.86+0.32−0.24</td>
<td>1.11+0.48−0.33</td>
<td>0.91+0.37−0.26</td>
<td>1.01+0.24−0.20</td>
</tr>
<tr>
<td><strong>HJ-MiNLO [21, 22]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Expected signal strength $\mu_H$</strong></td>
<td>1.00+3.34−3.54</td>
<td>1.00+2.46−2.50</td>
<td>1.00+2.30−2.37</td>
<td>1.00+1.42−1.40</td>
</tr>
<tr>
<td><strong>Observed signal strength $\mu_H$</strong></td>
<td>7.92+3.42−3.19</td>
<td>4.76+2.64−2.47</td>
<td>1.70+2.31−2.33</td>
<td>3.68+1.58−1.46</td>
</tr>
<tr>
<td><strong>Expected H significance ($\mu_H = 1$)</strong></td>
<td>0.29σ</td>
<td>0.41σ</td>
<td>0.43σ</td>
<td>0.71σ</td>
</tr>
<tr>
<td><strong>Observed H significance</strong></td>
<td>2.42σ</td>
<td>1.90σ</td>
<td>0.73σ</td>
<td>2.54σ</td>
</tr>
<tr>
<td><strong>Expected UL signal strength ($\mu_H = 0$)</strong></td>
<td>&lt; 6.84</td>
<td>&lt; 5.02</td>
<td>&lt; 4.67</td>
<td>&lt; 2.90</td>
</tr>
<tr>
<td><strong>Observed UL signal strength</strong></td>
<td>&lt; 7.97</td>
<td>&lt; 4.78</td>
<td>&lt; 1.70</td>
<td>&lt; 3.70</td>
</tr>
</tbody>
</table>

Ref. [16] $p_T$ spectrum

<table>
<thead>
<tr>
<th></th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected signal strength $\mu_H$</strong></td>
<td>1.00+1.52−1.50</td>
<td>1.00+1.14−1.06</td>
<td>1.00+1.12−1.05</td>
<td>1.00+0.68−0.58</td>
</tr>
<tr>
<td><strong>Observed signal strength $\mu_H$</strong></td>
<td>3.97+1.89−1.55</td>
<td>2.18+1.37−1.19</td>
<td>1.13+1.12−1.06</td>
<td>1.90+0.86−0.70</td>
</tr>
<tr>
<td><strong>Expected H significance ($\mu_H = 1$)</strong></td>
<td>0.68σ</td>
<td>0.94σ</td>
<td>0.95σ</td>
<td>1.73σ</td>
</tr>
<tr>
<td><strong>Observed H significance</strong></td>
<td>2.62σ</td>
<td>1.84σ</td>
<td>1.06σ</td>
<td>2.85σ</td>
</tr>
<tr>
<td><strong>Expected UL signal strength ($\mu_H = 0$)</strong></td>
<td>&lt; 3.41</td>
<td>&lt; 2.42</td>
<td>&lt; 2.29</td>
<td>&lt; 1.42</td>
</tr>
<tr>
<td><strong>Observed UL signal strength</strong></td>
<td>&lt; 4.00</td>
<td>&lt; 2.18</td>
<td>&lt; 1.13</td>
<td>&lt; 1.90</td>
</tr>
</tbody>
</table>

Figure 3: The best-fit signal strength $\mu_H$ (black squares) and uncertainty (red lines) per $p_T$ category based on the HJ-MiNLO [21, 22] prediction. The dashed black line indicates the SM expectation. The solid blue line and green band represents the combined best-fit signal strength and uncertainty, respectively, of $\mu_H = 3.68^{+1.58}_{-1.46}$ extracted from a simultaneous fit of all channels.
renormalization and factorization scales by factors of two, this theoretical acceptance uncertainty is estimated to be 2%.

The result of this unfolding procedure is shown in Fig. 4 and Table 4, along with the predicted cross sections from Ref. [22] and the predictions of the signal event generators described in Section 3. The correlation coefficients among the three $p_T^H$ bins are shown in Table 5. The measured cross section uncertainty in the first $p_T^H$ bin is larger because of limited acceptance. The first and second $p_T^H$ bins have a mild anti-correlation, primarily due to jet energy scale uncertainties, which inflates the corresponding per-bin uncertainties in the unfolded cross section. The observed cross section in the third $p_T^H$ bin has a smaller relative uncertainty than that in the second bin because of the larger magnitude of the central value in that bin. With respect to the SM, the upward deviation of the cross section in the third $p_T^H$ bin, when profiling the other two, corresponds to a significance of 2.64σ. Instead, when considering all three cross section parameters of interest simultaneously, the total deviation from the SM corresponds to a significance of 1.91σ.

Figure 4: Measured differential fiducial cross section as a function of Higgs boson $p_T$, in comparison to the predictions of Ref. [22], shown in red, and HJ-MINLO [21], shown in blue. The two predictions are nearly identical. In the bottom two panels, the dotted line corresponds to a ratio of one. The measured cross section uncertainty in the first bin is larger because of limited acceptance. The cross section measurements in the first and second bins have a mild anti-correlation, primarily due to jet energy scale uncertainties, which inflates the corresponding uncertainties. The observed cross section in the third bin has a smaller relative uncertainty than that in the second bin because of the larger magnitude of the central value in that bin. The relative uncertainties in the predictions of Ref. [22] and HJ-MINLO are approximately 10% and 20%, respectively.
### Summary

An inclusive search for the standard model Higgs boson decaying to bottom quark-antiquark pairs and reconstructed as a single, large-radius jet with $p_T > 450$ GeV has been presented. The search uses a data sample of proton-proton collisions at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 137 fb$^{-1}$. The $Z$+jets process is used to validate the method and is measured to be consistent with the standard model prediction. The Higgs boson signal strength is measured to be $\mu_H = 3.68 \pm 1.20$ (stat) $^{+0.63}_{-0.66}$ (syst) $^{+0.81}_{-0.46}$ (theo) = 3.68 $^{+1.38}_{-0.41}$ based on the theoretical prediction from the HJ-MiNLO generator. The measured signal strength corresponds to an observed significance of 2.54$\sigma$ with respect to the background-only hypothesis, where the expected significance of the standard model signal is 0.71$\sigma$. The significance of the observed excess with respect to the standard model expectation is 1.85$\sigma$. With respect to the previous CMS result, the relative precision of the Higgs boson signal strength measurement improves by approximately a factor of two because of the increased integrated luminosity, improved $b$ tagging based on a deep neural network, and smaller theoretical uncertainties. Finally, the differential cross section for the Higgs boson transverse momentum in the phase space regions recommended by the LHC simplified template cross section framework has also been presented.

### Table 4: Measured and predicted differential fiducial cross section as a function of Higgs boson $p_T$. All cross sections are in units of fb.

<table>
<thead>
<tr>
<th></th>
<th>$300–450$ GeV</th>
<th>$450–650$ GeV</th>
<th>$&gt; 650$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured</td>
<td>$600^{+800}_{-800}$</td>
<td>$5^{+42}_{-43}$</td>
<td>$29^{+12}_{-11}$</td>
</tr>
<tr>
<td>LHCHXSWG [22]</td>
<td>—</td>
<td>$15.4^{+1.4}_{-1.8}$</td>
<td>$1.9^{+0.2}_{-0.2}$</td>
</tr>
<tr>
<td>HJ-MiNLO [21]</td>
<td>$89^{+20}_{-18}$</td>
<td>$13.5^{+3.0}_{-2.7}$</td>
<td>$1.9^{+0.4}_{-0.4}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$300–450$ GeV</th>
<th>$450–650$ GeV</th>
<th>$&gt; 650$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$300–450$ GeV</td>
<td>1.0</td>
<td>−0.18</td>
<td>0.002</td>
</tr>
<tr>
<td>$450–650$ GeV</td>
<td>−0.18</td>
<td>1.0</td>
<td>0.06</td>
</tr>
<tr>
<td>$&gt; 650$ GeV</td>
<td>−0.002</td>
<td>0.06</td>
<td>1.0</td>
</tr>
</tbody>
</table>
References


