Constraints on Higgs boson production with large transverse momentum using $H \rightarrow b\bar{b}$ decays in the ATLAS detector

G. Aad et al.*
(ATLAS Collaboration)

(Received 17 November 2021; accepted 2 March 2022; published 11 May 2022)

This paper reports constraints on Higgs boson production with transverse momentum above 1 TeV. The analyzed data from proton–proton ($pp$) collisions produced by the Large Hadron Collider (LHC) at CERN. Four production modes have been observed, gluon–gluon fusion (ggF), vector-boson fusion (VBF), associated production with a weak vector boson ($VH$), and associated production with a top quark–antiquark pair ($ttH$), along with five decay modes $H \rightarrow \gamma\gamma$, $ZZ^\ast$, $WW^\ast$, $\tau\tau$, $b\bar{b}$ [3,4]. The initial measurements of inclusive cross sections have evolved to include differential cross section measurements, and measurements in the simplified template cross section framework [5–7]. All results agree with the Standard Model (SM) predictions within the current precision, but sizable regions of the Higgs sector remain unexplored. In one such region, where the Higgs boson transverse momentum $p_T^H$ reaches the TeV scale, the cross section hierarchy is very different from that in the inclusive cross section, where ggF is nearly 90% of the total. At the TeV scale, the SM predicts the cross sections for the ggF and $VH$ production processes to be roughly equal, while the VBF and $ttH$ production cross sections are around 60% and 30% of the ggF process, respectively.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP3.

I. INTRODUCTION

The characterization of the Higgs sector has steadily improved since the Higgs boson ($H$) discovery [1,2] using data from proton–proton ($pp$) collisions produced by the Large Hadron Collider (LHC) at CERN. Four production modes have been observed, gluon–gluon fusion (ggF), vector-boson fusion (VBF), associated production with a weak vector boson ($VH$), and associated production with a top quark–antiquark pair ($ttH$), along with five decay modes $H \rightarrow \gamma\gamma$, $ZZ^\ast$, $WW^\ast$, $\tau\tau$, $b\bar{b}$ [3,4]. The initial measurements of inclusive cross sections have evolved to include differential cross section measurements, and measurements in the simplified template cross section framework [5–7]. All results agree with the Standard Model (SM) predictions within the current precision, but sizable regions of the Higgs sector remain unexplored. In one such region, where the Higgs boson transverse momentum $p_T^H$ reaches the TeV scale, the cross section hierarchy is very different from that in the inclusive cross section, where ggF is nearly 90% of the total. At the TeV scale, the SM predicts the cross sections for the ggF and $VH$ production processes to be roughly equal, while the VBF and $ttH$ production cross sections are around 60% and 30% of the ggF process, respectively.

*Full author list given at the end of the article.

The leading effects of many beyond-the-SM (BSM) scenarios can be parametrized through effective field theories (EFTs), whose operators are suppressed by a new physics scale $\Lambda$ [8]. Measured observables at the LHC would only be affected through effective interactions among SM particles. For example, the ggF production mode is sensitive to the structure of quasi-point-like couplings within the loop nature of the effective $ggH$ coupling. Studies of Higgs bosons produced with large transverse momentum access regions where some potential BSM effects are enhanced by powers of $p_T^H/\Lambda$ [9–13]. Differential cross section measurements with an extended reach may be more sensitive than higher precision, low energy measurements also because the signal-to-background ratio increases with $p_T^H$.

In the high-$p_T^H$ regime, the CMS Collaboration measured a signal yield relative to the SM prediction, or signal strength, of $\mu_{ggF} = 3.7_{-1.5}^{+1.6}$. In the $H \rightarrow b\bar{b}$ decay mode for events containing a large-radius jet with $p_T > 450$ GeV, and presented ggF differential cross sections while considering other Higgs boson production modes as a background [14]. In the $H \rightarrow \gamma\gamma$ decay mode, ggF production with $p_T^H > 200$ GeV was measured to a precision of less than 50% relative to the SM prediction [15]. The analysis of $VH$ production with leptonic $V$ decays has achieved considerable sensitivity in the high-$p_T^H$ regime [16–18]. The ATLAS Collaboration measured a signal strength of $\mu_{VH} = 0.72_{-0.36}^{+0.39}$ in the $H \rightarrow b\bar{b}$ decay mode targeting events with $p_T^H > 250$ GeV, and presented differential cross sections in two exclusive vector-boson transverse momentum regions, 250–400 GeV and above...
400 GeV [17]. However, it is sensitive to different EFT operators than those for ggF and \(t\bar{t}H\) production. Recent results on VBF production with Higgs boson decays to photons and leptons also included high-\(p_T\) event categories, but have limited reach [19–23]. Similarly, measurements of \(t\bar{t}H\) production have yet to reach the high-\(p_T\) regime [24,25].

This paper reports the first ATLAS studies of Higgs bosons produced with transverse momentum above 1 TeV. The yield of Higgs bosons decaying into \(b\bar{b}\) pairs is determined in several \(p_T^H\) regions. No restrictions are applied to select a particular Higgs boson production mode aside from requiring an energetic hadronic recoil system. The cross section is measured for \(p_T^H > 450\) GeV, allowing a straightforward comparison with theoretical calculations, such as those reported in Ref. [26]. In addition a differential analysis is performed to extract the cross section in four Higgs boson \(p_T\) intervals, 300–450 GeV, 450–650 GeV, 650–1000 GeV, and above 1 TeV.

Data used correspond to 136 fb\(^{-1}\) of \(pp\) collisions at \(\sqrt{s} = 13\) TeV and were collected with the ATLAS detector [27] using jet-based trigger requirements during Run 2 (2015–2018) of the LHC. Higgs bosons with a large Lorentz boost are reconstructed as single large-radius jets having a mass compatible with 125 GeV [28]. To populate the signal region, events must have at least two large-radius jets. At least one jet must have \(p_T > 450\) GeV to ensure a fully efficient trigger response. Either of the two leading-\(p_T\) jets in the event must contain evidence of two \(b\)-hadron decays. Including the subleading jet as a possible Higgs boson candidate increases the sensitivity for \(p_T^H > 450\) GeV by 11\% and permits a cross section measurement down to \(p_T^H = 300\) GeV, overlapping with measurements in other decay channels [15,20,22,24,29–31].

For a Higgs boson mass \((m_H)\) of 125 GeV, the SM predicts the ggF production mode contributes nearly half of the Higgs boson events reconstructed near \(m_H\) when they are summed over all signal regions. VBF-, \(VH\)-, and \(t\bar{t}H\)-produced events each contribute approximately 15\%–20\%. Since some of the same dimension-6 EFT operators modify the \(p_T^H\) spectrum in the ggF and \(t\bar{t}H\) processes, it could be advantageous to consider \(t\bar{t}H\) and ggF production together. These operators can produce enhancements at high \(p_T^H\) that are within the sensitivity of the present analysis without inducing significant deviations from the SM prediction at low \(p_T^H\). However, the expected yield enhancement differs for the two production modes, with ggF production having a larger increase. While this analysis is primarily sensitive to ggF production, all the main production modes are considered as the signal. This approach enhances the sensitivity to possible BSM effects and minimizes the dependence on theoretical assumptions. Using the \(H \rightarrow b\bar{b}\) decay, which has the largest branching fraction, mitigates the impact of the smaller absolute cross section in the high-\(p_T\) regime.

The dominant background process is multijet production, which exhibits a monotonically decreasing jet mass distribution. Hadronically decaying vector bosons, produced in association with jets (\(V + \text{jets}\)) and events with one or two top quarks (jointly referred to as \(\text{Top}\)) populate the jet mass regions below and above \(m_H\), respectively, as shown in Fig. 1. The \(Z\) and \(H\) resonance structures are distinct from the smoothly falling multijet background, while the top quark resonance is spread over a large portion of the high-mass region. Therefore, the signal is extracted from the reconstructed jet mass distribution. An analytic function is used to model the multijet background. The jet mass spectra and acceptance for Higgs boson, \(V + \text{jets}\), and \(\text{Top}\) events are estimated from simulation.

A binned maximum-likelihood fit, referred to as the global likelihood fit, is used to measure the signal strength. Unconstrained, or free, parameters for the multijet model’s yield and shape, the \(Z + \text{jets}\) yield, and the Higgs boson yield are determined simultaneously from the signal region.

FIG. 1. Jet mass distributions for the Higgs boson, \(Z + \text{jets}\), \(W + \text{jets}\), and \(\text{Top}\) contributions from the SM prediction as well as the multijet jet mass distribution extracted from data in the signal region (SR) defined by the leading (left) and subleading (right) jets.

092003-2
data. Control region data, fit simultaneously with the signal region, determine the unconstrained yield for top quark pair-production ($t\bar{t}$) events. Concurrently, the modeling of the jet mass scale and resolution from $V$ and top quark decays are separately validated in a broad range of jet $p_T$. Validation region data are used to study $V + $ jets events, the parametrization of the multijet model, and the robustness of the global likelihood fit.

II. ATLAS DETECTOR

The ATLAS experiment [27] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4$\pi$ coverage in solid angle.\(^1\) It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer. The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. The Insertable B-Layer, the innermost pixel layer at a mean radius of 3.3 cm, was installed before Run 2 of the LHC [32,33]. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A steel/scintillator-tile hadron calorimeter covers the central pseudorapidity range $|\eta| < 1.7$. The forward regions are instrumented with LAr calorimeters for EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large air-core toroidal superconducting magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. The muon spectrometer includes a system of precision chambers for tracking and fast detectors for triggering. A two-level trigger system is used to select events [34]. The first-level trigger is implemented in hardware and uses a subset of the detector information to accept events at a rate below 100 kHz. A software-based trigger reduces the accepted event rate to 1 kHz on average. An extensive software suite [35] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

III. DATA AND SIMULATED SAMPLES

The data were collected with the ATLAS detector in $pp$ collisions with a center-of-mass energy of 13 TeV during

Run 2 (2015–2018) of the LHC [36]. Events must satisfy a set of triggers requiring a jet reconstructed with the anti-$k_T$ algorithm with radius parameter $R = 1.0$ [37]. To adapt to different instantaneous luminosity profiles and the inclusion of pileup suppression techniques within the data acquisition system [38], the jet-$p_T$ and mass thresholds differ for each year of data taking. The trigger jet-$p_T$ threshold varies from 360 GeV to 420 GeV, and the trigger jet mass threshold is either not applied, 30 GeV, or 35 GeV. Events which pass a trigger requiring a muon with $p_T > 50$ GeV [39] populate a control region for top quark production. The total integrated luminosities are 136 fb$^{-1}$ and 139 fb$^{-1}$ for the jet- and muon-triggered data, respectively, with an uncertainty of 1.7% [40,41].

Monte Carlo (MC) simulated events are used to model the resonant backgrounds ($W + $ jets, $Z + $ jets, and Top production) as well as four Higgs production processes: ggF, VBF, $VH$, and $ttH$. Higgs boson ggF production was simulated at next-to-leading-order (NLO) accuracy in QCD with finite mass effects by using the Hj-MiNLO [42–44] prescription with the POWHEG program [45–47] as discussed in Ref. [48]. NLO accuracy in QCD for VBF and $ttH$ production and leading order (LO) accuracy for $gg \rightarrow VH$ production was achieved using the POWHEGBOX v2 [45–47,49,50] program. Using the POWHEGBOX v2 program, the improved MINL0 [51] calculation, and the GOSAM [52] program, $qq \rightarrow VH$ production was also simulated at NLO accuracy in QCD. Corrections for NLO electroweak (EW) effects were applied as a function of the generated Higgs boson transverse momentum for VBF, $VH$, and $ttH$ production. The production cross sections used are compatible with those presented in Ref. [26], except for $ttH$ production where a scale factor is applied to match the reported value [26]. The Higgs boson branching fractions were calculated with HDECAY [53–55] and PROPHECY4F [56–58].

Production of $V + $ jets events with hadronic boson decays was simulated with SHERPA to NLO QCD accuracy for one additional parton and LO QCD accuracy for up to four additional partons. Approximate NLO EW corrections [59] were applied as a function of the generated vector-boson momentum $p_T^ V$. They have a sizable impact on the differential production cross section, reducing the predicted yield by $\sim 10\%$ at a $p_T^ V$ of 500 GeV and $\sim 20\%$ above 1 TeV. Calculations of next-to-next-to-leading-order (NNLO) QCD corrections to $V + $ jets production are available [60]. The NNLOJET group performed the calculation for $\sqrt{s} = 8$ TeV [61,62] and has provided custom corrections for the analysis kinematic region for $\sqrt{s} = 13$ TeV as a function of $p_T^ V$. They vary from 1.01 to 1.09 and are applied as a multiplicative factor on top of the NLO EW corrections. Diboson production was found to make a negligible contribution to the present analysis.

The production of top quark pairs, associated production of a top quark with a $W$ boson ($tW$), and single-top $t$- and $s$-channel production were modeled using the

---

\(^1\)ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the center 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 $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. 

---
POWHEGBOX v2 [45–47,63–66] generator at NLO in QCD. The diagram subtraction scheme [67] was used in \( tW \) events to account for interference and overlap with \( \bar{t} \bar{t} \) production.

The jet mass distribution of nonresonant multijet events is modeled with an analytic function. Simulated events used to study the multijet model were generated using PYTHIA 8.230 [68] with leading-order matrix elements for dijet production and interfaced to a \( p_T \)-ordered parton shower.

All simulated particles from collisions were processed with the ATLAS detector simulation [69] based on GEANT4 [70]. Pileup, multiple interactions in the same and neighboring bunch crossings, was modeled by overlaying the hard-scatter event with inelastic \( pp \) events generated with PYTHIA8.186 [68] using the NNPDF2.3LO set of parton distribution functions (PDFs) [71] and a set of tuned parameters called the A3 tune [72]. For Higgs boson, Top, and dijet production, the EVTGEN1.2.0 program [73] was used to model the decays of bottom and charm hadrons.

For each sample, Table I summarizes the MC generator, parton distribution functions, parton shower and hadronization model, and underlying event tune used, as well as the order of perturbative QCD computations and EW corrections obtained for the cross section. For additional information, see Ref. [74] for \( V + \text{jets} \) events, Refs. [75–78] for top quark events, and Ref. [79] for multijet events. Systematic uncertainties for process modeling are described in Sec. VII.

### Table I. The generators used for the simulation of the signal and background processes. The matrix element, parton shower, and underlying event are abbreviated as ME, PS, and UE, respectively.

<table>
<thead>
<tr>
<th>Process</th>
<th>ME generator</th>
<th>ME PDF</th>
<th>PS and hadronization</th>
<th>UE model</th>
<th>Cross section order</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g g \rightarrow H \rightarrow b \bar{b} )</td>
<td>POWHEGBOX v2 (^{(*)} ) [45–47] + NNPDF3.0NNLO [85]</td>
<td>NNPDF3.0NNLO [85]</td>
<td>PYTHIA8.212 [68]</td>
<td>AZNLO [86]</td>
<td>NLO(QCD) + LO(EW)</td>
</tr>
<tr>
<td>( q q \rightarrow H \rightarrow q'q'^b\bar{b} )</td>
<td>POWHEGBOX v2 [45–47,49]</td>
<td>NNPDF3.0NNLO [85]</td>
<td>PYTHIA8.230</td>
<td>AZNLO</td>
<td>NLO(QCD) + NLO(EW)(^{(*)} )</td>
</tr>
<tr>
<td>( q q \rightarrow W + H )</td>
<td>POWHEGBOX v2 + GOSAM [52] + MINLO [51]</td>
<td>NNPDF3.0NNLO</td>
<td>PYTHIA8.240</td>
<td>PYTHIA8.212</td>
<td>AZNLO NLO(QCD) + NLO(EW)(^{(c)} )</td>
</tr>
<tr>
<td>( q q \rightarrow Z+H )</td>
<td>POWHEGBOX v2 + GOSAM + MINLO</td>
<td>NNPDF3.0NNLO</td>
<td>PYTHIA8.240</td>
<td>PYTHIA8.212</td>
<td>AZNLO NLO(QCD) + NLO(EW)(^{(c)} )</td>
</tr>
<tr>
<td>( g g \rightarrow Z+H )</td>
<td>POWHEGBOX v2 [50]</td>
<td>NNPDF3.0NNLO</td>
<td>PYTHIA8.240</td>
<td>PYTHIA8.212</td>
<td>AZNLO LO + NLL(QCD)</td>
</tr>
<tr>
<td>( t \bar{t} \rightarrow all )</td>
<td>POWHEGBOX v2</td>
<td>NNPDF3.0NNLO</td>
<td>PYTHIA8.230</td>
<td>AZNLO</td>
<td>NLO(QCD) + NLO(EW)(^{(c)} )</td>
</tr>
</tbody>
</table>

### Vector boson + jets

<table>
<thead>
<tr>
<th>Process</th>
<th>ME generator</th>
<th>ME PDF</th>
<th>PS and hadronization</th>
<th>UE model</th>
<th>Cross section order</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W \rightarrow q \bar{q} )</td>
<td>SHERPA 2.2.8 [83,87,88]</td>
<td>NNPDF3.0NNLO</td>
<td>SHERPA2.2.8 [89,90]</td>
<td>Default</td>
<td>NNLO(QCD) (^{(f)} ) [61,62,91] approx NLO(EW) [59,92,93]</td>
</tr>
<tr>
<td>( Z \rightarrow q \bar{q} )</td>
<td>SHERPA 2.2.8 [83,87,88]</td>
<td>NNPDF3.0NNLO</td>
<td>SHERPA2.2.8 [89,90]</td>
<td>Default</td>
<td>NNLO(QCD) (^{(f)} ) [61,62,91] approx NLO(EW) [59,92,93]</td>
</tr>
</tbody>
</table>

### Top quark, mass set to 172.5 GeV

<table>
<thead>
<tr>
<th>Process</th>
<th>ME generator</th>
<th>ME PDF</th>
<th>PS and hadronization</th>
<th>UE model</th>
<th>Cross section order</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t \bar{t} \rightarrow all )</td>
<td>POWHEGBOX v2 [45–47,63]</td>
<td>NNPDF3.0NNLO</td>
<td>PYTHIA8.230</td>
<td>A14 [94]</td>
<td>NNLO + NLL [95]</td>
</tr>
<tr>
<td>( tW )</td>
<td>POWHEGBOX v2 [45–47,64]</td>
<td>NNPDF3.0NNLO</td>
<td>PYTHIA8.230</td>
<td>A14</td>
<td>NLO</td>
</tr>
<tr>
<td>( t ) t-channel</td>
<td>POWHEGBOX v2 [45–47,65]</td>
<td>NNPDF3.0NNLO</td>
<td>PYTHIA8.230</td>
<td>A14</td>
<td>NLO</td>
</tr>
<tr>
<td>( t ) s-channel</td>
<td>POWHEGBOX v2 [45–47,66]</td>
<td>NNPDF3.0NNLO</td>
<td>PYTHIA8.230</td>
<td>A14</td>
<td>NLO</td>
</tr>
</tbody>
</table>

### Multijet

<table>
<thead>
<tr>
<th>Process</th>
<th>ME generator</th>
<th>ME PDF</th>
<th>PS and hadronization</th>
<th>UE model</th>
<th>Cross section order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dijets</td>
<td>PYTHIA 8.230</td>
<td>NNPDF2.3LO [71]</td>
<td>PYTHIA8.230</td>
<td>A14</td>
<td>LO</td>
</tr>
</tbody>
</table>

092003-4
IV. OBJECT SELECTION

For Higgs bosons with a large Lorentz boost, the event topology of \( pp \rightarrow H(\rightarrow b\bar{b}) + J \) is characterized by two jets, one of which contains the decay products of the two \( b \)-hadrons.

A. Object reconstruction

Charged-particle tracks [96] are reconstructed in the inner detector and used to form interaction vertices [97]. The primary vertex of the hard interaction is defined as the vertex with the highest sum of squared transverse momenta of associated tracks.

Large-radius \((R = 1.0)\) jets are formed by applying the anti-\( k_T \) algorithm implemented in FASTJET [98] to topological clusters of noise-suppressed calorimeter energy depositions calibrated to the local hadronic scale [99]. Jet cleaning criteria are applied to identify jets arising from noncollision backgrounds or noise in the calorimeters [100], and events containing such jets are removed. A jet trimming procedure reduces pileup dependence and improves the mass resolution [101]. It produces a collection of subjets by reclustering the constituents of each jet using the \( k_T \) algorithm [102] with \( R = 0.2 \). Subjets with \( p_T^{\text{subj}}/p_T^{\text{jet}} < 0.05 \) are removed and the jet four-momentum is recalculated. The trimmed jet’s mass \( m_J \) is computed as a weighted combination of the jet mass obtained from the calorimeter measurements and that from the charged component in the inner detector [103]. For jets within \(|\eta| < 2\), simulation-based corrections are applied to calibrate \( p_T \) and \( m_J \), while \( p_T \) also has corrections based on in situ techniques [103].

The anti-\( k_T \) algorithm with a variable, \( p_T \)-dependent radius parameter is used to form track-jets from tracks compatible with the primary vertex [104,105]. The effective jet radius \( R_{\text{eff}} = \rho / p_T \) where the \( \rho \)-parameter is set to 30 GeV. The lower and upper bounds on the track-jet radius are 0.02 and 0.4, respectively. Track-jets are matched to large-\( R \) jets by ghost association before trimming [106,107]. In simulated events, track-jets are labeled as \( b \)-, \( c \)- or light-flavor according to which hadrons with \( p_T > 5 \) GeV are found within \( \Delta R = 0.3 \) of the jet axis [108].

A multivariate discriminant (denoted MV2 in Ref. [108]) is used to tag track-jets containing a \( b \)-hadron decay (\( b \)-tagging). Track-jets with \( p_T > 10 \) GeV, \(|\eta| < 2.5 \), and at least two tracks are considered. The operating point is tuned to produce an average efficiency of 77% for \( b \)-jets in simulated \( t\bar{t} \) events, which corresponds to light-flavor jet (\( u, d, s \)-quark, and gluon) and \( c \)-jet misidentification efficiencies of 0.9% and 25%, respectively. If the \( \Delta R \) between two track-jets with \( p_T > 5 \) GeV in a large-\( R \) jet is smaller than either of their respective variable radii, the jet is not considered for \( b \)-tagging [109].

Muons are required to have \(|\eta| < 2.5 \), \( p_T^{\mu} > 10 \) GeV, and small impact parameters relative to the primary vertex, as well as to satisfy the “medium” quality criterion [110]. Isolated muons must also satisfy loose track- and calorimeter-based isolation conditions [110].

B. Analysis object definitions

Reconstructed jets possessing properties compatible with an \( H \rightarrow b\bar{b} \) decay are labeled \textit{candidate jets}. The reconstructed jet containing the Higgs boson decay products, or \( H\text{-jet} \), is not always the highest-\( p_T \) jet in the event. The Higgs boson and the hadronic recoil system have equal \( p_T \), but the \( p_T \) ordering of the reconstructed jets is affected by final-state radiation, jet resolution, and activity outside the jet cone. Undetected neutrinos from semileptonic \( b \)-hadron decays may also cause the \( H \)-jet to be reconstructed as the subleading jet. In around 50% (47%) of selected simulated ggF events, the \( H \)-jet is the leading (subleading) jet. Simulated \( VH \) events contain similar proportions of leading and subleading \( H \)-jets. For \( t\bar{t}H \) production, any one of at least three final-state particles can be reconstructed as the leading jet. Therefore, candidate jets are defined as \( \text{either of} \) the two leading-\( p_T \) jets with \( |\eta| < 2 \), \( p_T > 250 \) GeV, \( m_J > 60 \) GeV, and \( 2m_{\ell}/p_T < 1 \). The last requirement selects jets compatible with a boosted decay. Furthermore, they must contain at least two track-jets. A candidate jet is \textit{double-tagged} if its two leading track-jets are \( b \)-tagged and \textit{anti-tagged} if neither are \( b \)-tagged. Each signal region requires a double-tagged candidate jet, as discussed in Sec. VA.

A “\( \mu \)-in-jet” correction is applied to candidate jets to account for the presence of semileptonic \( b \)-hadron decays. It utilizes the leading-\( p_T \) muon passing a minimum set of quality criteria and found within \( \Delta R = \min(0.4, 0.04 + 10/p_T^\mu) \) of a \( b \)-tagged track-jet. The muon four-momentum is added to the trimmed jet while the energy deposited by the muon in the calorimeter is removed. After correcting 13% (33%) of leading (subleading) \( H \)-jets in simulated ggF events, the \( m_J \) width is reduced by 5% (12%). Henceforth, \( p_T \) and \( m_J \) refer to the muon-corrected jet transverse momentum and mass, respectively, and \( p_T^\mu \) and \( m_J^\mu \) represent the corresponding uncorrected versions.

C. Reconstructed object systematic uncertainties

The most important experimental uncertainties originate from the jet mass resolution (JMR) modeling and jet mass scale (JMS) calibration. Uncertainties in \( b \)-tagging efficiency scaling factors and the jet energy scale are found to play a minor role. The remaining uncertainties, including those arising from muon trigger, reconstruction, identification, and isolation rate modeling [111], are negligible.

With appreciable reconstructed \( V \) and top quark resonance peaks, adjustments of the JMS and JMR central values and uncertainties are possible. Considerations about the validity of transferring these improvements between processes, or along jet \( p_T \), inform the correlation scheme described below. It was verified that the chosen configuration is conservative
in terms of the expected sensitivity to the $V + \text{jets}$ and $H$ signal strengths. Jet observables in the simulation are smeared to assess the impact of scale and resolution uncertainties. A ratio of calorimeter-based to track-based measurements in dijet data and simulation defines the uncertainties in the jet energy (relative 1%–2%) and mass scales (relative 2%–10%) [103]. Jet energy scale and mass scale uncertainties are divided into 23 and 6 separate components, respectively, to account for different sources of uncertainty. The level of JMS agreement between data and simulation, while within the systematic uncertainties, displays a process and jet-$p_T$ dependence. Therefore, JMS uncertainties for $t\bar{t}$ events are separated from those for $V+\text{jets}$ and $H$ events within the global likelihood, discussed in Sec. VIII. The dominant component in terms of reconstructed mass scale is further separated so as to act independently on all processes ($t\bar{t}$, $V+\text{jets}$, and $H$) and in all analysis jet-$p_T$ bins. Consistent with previous studies of trimmed jets [112,113], the energy resolution has an absolute 2% uncertainty, while the mass resolution has a relative 20% uncertainty. JMR uncertainties act independently on each process ($t\bar{t}$, $V+\text{jets}$, and $H$) and in each analysis jet-$p_T$ bin to account for generator, process, and $p_T$ dependence. The $V+\text{jets}$ JMR uncertainty is reduced using independent measurements as described in Sec. VII B.

The impact of uncertainties in $b$-tagging rates for $b$-, $c$-, and light-flavor jets is determined separately in various kinematic regions [108,114,115]. Each flavor category uncertainty is decomposed into independent components. A specific component for each jet flavor, based on the impact of experimental and theoretical uncertainties, accounts for an extrapolation of the scaling factors to jets with $p_T$ beyond the calibration dataset’s kinematic reach [116]. The thresholds are 250 GeV, 140 GeV, and 300 GeV for $b$-, $c$-, and light-flavor track-jets, respectively.

V. EVENT SELECTION AND CATEGORIZATION

Events are classified into three orthogonal regions: a signal region (SR) used to extract the signal strength, a control region used to study top quark events (CR$_T$), and a validation region (VR) used to study the multijet and $V+\text{jets}$ background models. Each region is further configured depending on jet $p_T$ to determine the various signal strengths as described below.

A. Signal and validation regions

A uniform requirement for both the VR and SR in all data-taking years of at least one jet with $p_T^j > 450$ GeV and $m_H^j > 60$ GeV removes the kinematic regime biased by the trigger requirements. A second jet with $p_T^j > 200$ GeV is required. At least one of the two leading-$p_T$ jets must satisfy the candidate jet criteria.

The event categorization first considers the leading jet. If it is a double-tagged candidate jet, the event populates the leading-jet signal region (SRL). Only if the leading jet is not a double-tagged candidate jet, and the subleading jet is, the event is categorized into the subleading-jet signal region (SRS). Approximately 40% of the $H$ events surviving the kinematic cuts pass the $b$-tagging requirement [117]. In simulated multijet events satisfying the SR requirements, roughly 70% of the candidate jets with mass close to $m_H$ contain two $b$-hadrons and less than 5% of candidate jets do not contain any heavy-flavor hadrons. The SRS has a sensitivity approximately 50% lower than that of the SRL, due in part to inferior mass resolution.

Both the leading and subleading jets are always considered when creating the validation regions. The leading-jet validation region (VRL) includes events where the leading jet is an antitagged candidate jet, and the subleading jet either has the same distinction or is not a candidate jet. An analogous definition defines the subleading-jet validation region (VRS). Figure 2 summarizes the event categorization.

The signal strength is extracted in an inclusive signal region containing candidate jets with $p_T > 250$ GeV. For the cross section measurements, the signal region is further configured into a fiducial signal region containing candidate jets with $p_T > 450$ GeV and four differential signal regions defined by requiring the candidate jet $p_T$ to be in the ranges 250–450 GeV, 450–650 GeV, 650–1000 GeV, or above 1 TeV. Only the leading-jet SR is used for the highest-$p_T$ differential SR, $p_T > 1$ TeV. Only the subleading-jet SR is populated for regions with $p_T < 450$ GeV. The VRs follow the same definition. Table II summarizes the analysis signal regions. For measurements within the fiducial and differential regions, the signal yields within *volumes* defined by requirements on the generator’s event “truth” record are extracted as described in Sec. IX C.

![FIG. 2. Diagram showing the event categorization criteria.](Image)
TABLE II. A summary of the candidate jet $p_T$ requirements is given for the three analysis SR configurations. Each SR has an associated CR$_t$ and VR with the same jet $p_T$ requirements.

<table>
<thead>
<tr>
<th>Region</th>
<th>Candidate jet $p_T$ (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SRL</td>
</tr>
<tr>
<td>Inclusive</td>
<td>$&gt;450$</td>
</tr>
<tr>
<td>Fiducial</td>
<td>$&gt;450$</td>
</tr>
</tbody>
</table>

Within each region, the discriminating variable, $m_J$, is studied in 5 GeV bins. As a simple analytic function is used to model the multijet background, $m_J$ is studied in a restricted range where this background decreases monotonically. The combination of the $m_J > 60$ GeV selection and the presence of a muon and neutrino from semileptonic $b$-hadron decays reduces the acceptance for $m_J$ values near the selection threshold. Since the $b$-tagging efficiency diminishes with decreasing angular distance to another hadronic object \cite{118}, the low $m_J$ region in the signal region is further sculpted. The prevalence of each effect determines the minimum $m_J$ requirement. The SR mass range in each region is either 70 GeV to 210 GeV or 75 GeV to 210 GeV, depending on the jet being leading or subleading and the $p_T$ range. The VRs use the same $p_T$ bins and $m_J$ ranges.

**B. $t\bar{t}$ control region**

A dedicated $t\bar{t}$ control region, CR$_{t\bar{t}}$, using muon-trigger events, provides data with a high purity of top quark pair events to determine the $t\bar{t}$ yield in conditions equivalent to those of the SRs. The reconstructed final state is a top quark and a top antiquark with one semileptonic decay and one hadronic decay, with a minimum angular separation in the transverse plane. An isolated muon, with $p_T > 52.5$ GeV, close to a jet $J_b$ containing at least one track-jet defines the former, and a jet $J_t$ with at least three track-jets defines the latter. Both jets are required to have $p_T > 250$ GeV and exactly one of the leading two (three) track-jets in $J_b$ ($J_t$) must be $b$-tagged. Considering multiple track-jets within $J_b$ improves the identification efficiency of a $b$-quark reconstructed as a large-$R$ jet by 7.5%.

To ensure the kinematics of top quarks reconstructed in the CR$_{t\bar{t}}$ resemble those in the SR, only events with $140 < m_J < 200$ GeV are considered, removing those where $J_t$ does not contain all the top quark decay products. Residual differences between the CR$_{t\bar{t}}$ $J_t$ and SRS candidate jet $p_T$ spectra below 450 GeV originate from the different trigger requirements. The ratio of the jet $p_T$ spectrum in simulated $t\bar{t}$ events in the CR$_{t\bar{t}}$ to the jet $p_T$ spectrum in the SRS reproduces the difference. It is used to adjust the simulated $t\bar{t}$ event weights and remove data events in the CR$_{t\bar{t}}$ with $p_T^\text{jet} < 450$ GeV.

The same $p_T$ boundaries used in the SR are also applied to $p_T^\text{jet}$ to define the inclusive and differential CR$_{t\bar{t}}$ regions. The selection criteria, summarized in Table III, achieve over 95% purity in $t\bar{t}$ events for each CR$_{t\bar{t}}$ $p_T$ bin.

**VI. HIGGS BOSON MODELING**

The limited number of event selection criteria pertaining to properties of the recoil system or other activity in the event result in an inclusive analysis in terms of the Higgs boson production modes. Table IV shows the relative contributions of the four main production modes as a function of Higgs boson candidate $p_T$, according to SM predictions and within the Higgs boson window ($105 < m_J < 140$ GeV). In both the SRL and SRS, ggF production contributes the most for $p_T > 450$ GeV. For $p_T < 450$ GeV, $t\bar{t}H$ accounts for around 40% of the selected Higgs boson events. A hadronically decaying top quark can satisfy the jet trigger requirements without a high $p_T^\text{jet}$ value, thus resulting in a significant contribution of $t\bar{t}H$ events with relatively low Higgs boson $p_T$. Almost 90% of $t\bar{t}H$ events in the Higgs boson window arise from $H \rightarrow b\bar{b}$ decays. The majority of the remainder are $H \rightarrow W^\pm W^\mp$ events and these climb to almost 15% for larger $m_J$ values.

The uncertainty on the cross section and acceptance for ggF-produced events is 20%. It includes variations of the factorization and renormalization scales, the PDF, and the parton shower model to account for their respective uncertainties. Reference \cite{119} demonstrates that the NLO correction is nearly the same in the infinite top-mass approximation and full SM calculation, so no additional systematic uncertainty is assigned. Uncertainties on the cross section and acceptance for the VBF, VH, and $t\bar{t}H$ processes are taken to be 0.5%, 5%, and 13%, respectively \cite{26}. Systematic uncertainties in the EW corrections (expressed as $1 + \delta_\text{EW}$) are taken as $\delta_\text{EW}$ following the recommendations in Ref. \cite{5}.

TABLE III. A summary of the CR$_{t\bar{t}}$ selection criteria.

<table>
<thead>
<tr>
<th>Jet</th>
<th>$N$ track-jets</th>
<th>$N$ $b$-tags</th>
<th>Angular selection</th>
<th>Jet mass (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_b$</td>
<td>$\geq 1$</td>
<td>1</td>
<td>$0.04 + 10/p_T^\text{jet} &lt; \Delta R(\mu, J^b) \leq 1.5$</td>
<td>$140-200$</td>
</tr>
<tr>
<td>$J_t$</td>
<td>$\geq 3$</td>
<td>1</td>
<td>$\Delta \phi(J^b, J^t) &gt; \frac{\pi}{3}$</td>
<td></td>
</tr>
</tbody>
</table>
TABLE IV. The fractional contribution of each production mode to a given analysis region around the Higgs boson mass peak, defined as $105 < m_J < 140$ GeV. The fraction is given relative to the total signal yield in the analysis region in question.

<table>
<thead>
<tr>
<th>Process</th>
<th>250–450</th>
<th>450–650</th>
<th>650–1000</th>
<th>&gt;1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ggF</td>
<td>0.56</td>
<td>0.50</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>VBF</td>
<td>0.17</td>
<td>0.16</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>VH</td>
<td>0.14</td>
<td>0.18</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>tH</td>
<td>0.13</td>
<td>0.16</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>SRS</td>
<td>0.28</td>
<td>0.46</td>
<td>0.43</td>
<td>⋯</td>
</tr>
<tr>
<td>VBF</td>
<td>0.07</td>
<td>0.19</td>
<td>0.21</td>
<td>⋯</td>
</tr>
<tr>
<td>VH</td>
<td>0.26</td>
<td>0.24</td>
<td>0.26</td>
<td>⋯</td>
</tr>
<tr>
<td>tH</td>
<td>0.39</td>
<td>0.11</td>
<td>0.10</td>
<td>⋯</td>
</tr>
</tbody>
</table>

VII. BACKGROUND PROCESS MODELING

Multijet production is the dominant background process. The $V+$ jets and top quark resonance peaks flank the Higgs boson signal in low- and high-mass sidebands, respectively, but also leak into the Higgs boson signal window. Within $105 < m_J < 140$ GeV, $V+$ jets account for about 1% of the total background, top quarks for about 3%, and multijets provide the rest. The expected Higgs boson signal contribution corresponds to $0.2\%$ of the background for jet $p_T > 450$ GeV in the mass window and $20\%$–$80\%$ of the data statistical precision in the analysis jet-$p_T$ bins.

Therefore, an accurate and precise determination of the background is paramount and is achieved starting with the determination of the $V+$ jets and top quark backgrounds. These backgrounds are determined by maximizing a binned likelihood function defined as the product of Poisson probability terms for each bin of the $m_J$ distributions in the SRL, SRS, and CR$_{t\bar{t}}$. Unconstrained normalization parameters are common to the signal and control regions within each jet $p_T$ range. Systematic uncertainties are included as constrained nuisance parameters. Details are given in Sec. VIII.

A. Top quark production

The candidate jet in SR top quark events usually contains the remnants of a $b$-quark and the two hadronic decay products of a $W$ boson. As discussed above, the CR$_{t\bar{t}}$ design ensures the same physics processes also populate the control region. The simulated jet mass distributions in the CR and SR are similar in shape and peak near the top quark mass because both regions probe a comparable top quark momentum range. Therefore, any adjustment of the simulated top quark events made to improve their agreement with data in the CR$_{t\bar{t}}$ can be directly applied to the SR. This is achieved by including the CR$_{t\bar{t}}$ in the global likelihood described in Sec. VIII. The inclusive CR$_{t\bar{t}}$ has a $t\bar{t}$ purity of $97\%$ with similar levels found in the fiducial and differential region configurations. With such high purity, the $t\bar{t}$ normalization is determined directly from data with better than $10\%$ precision in most regions.

In the SR, $tW$-produced events where a top quark is matched to a candidate jet contribute $2\%$–$3\%$ relative to the $t\bar{t}$ yield. For $t$-channel production, the ratio to $t\bar{t}$ is $1\%$–$5\%$. Both have a candidate jet mass distribution similar to that in $t\bar{t}$ production. To reduce effects due to limited MC sample size, the mass spectrum’s shape for events with one top quark used in the likelihood described in Sec. VIII is obtained by scaling the $t\bar{t}$ MC sample to the number of events predicted by the dedicated $tW$ and $t$-channel MC samples within each jet $p_T$ bin. The contribution from $s$-channel production is negligible.

For $t\bar{t}$ production, comparisons between nominal and alternative simulated samples provide systematic uncertainty estimates for the parton shower model (HERWIG 7 replaces PYTHIA8) and the matrix element calculation (MADGRAPH5_AMC@NLO replaces POWHEGBOX v2). The comparisons show a $6\%$–$20\%$ and $1\%$–$19\%$ difference in yield in the various analysis regions, respectively. Within the nominal sample, variations of internal weights are used to estimate the systematic uncertainties associated with initial- and final-state radiation ($1\%$–$7\%$), as well as the renormalization and factorization scales (negligible). All experimental uncertainties described in Sec. IV C are utilized. Uncertainties in the $b$-tagging efficiency for $b$-jets and the JMS have the largest impact on the $t\bar{t}$ normalization. A $50\%$ normalization uncertainty is applied to the estimated number of events with a single-top quark, mainly driven by the comparison between the diagram subtraction and diagram removal schemes [67] in $tW$ events.

Figure 3 shows the jet mass distribution for each analysis $p_T$ bin in the CR$_{t\bar{t}}$ after the global likelihood fit described in Sec. VIII in the differential configuration. The simulation, shown with a $68\%$ confidence level (CL) uncertainty band, agrees well with the data.

B. $V+$ jets production

With a decay structure and relative experimental resolution similar to that of the Higgs boson, the vector-boson mass peaks offer a unique opportunity to validate the experimental performance. Events with Z bosons in the signal region outnumber $H$ events by over a factor of 20. Experimental effects that are challenging to discern in a statistically limited $H$ production measurement will be evident in the $Z$ observation. A well-understood $Z$ measurement is therefore a precursor to a robust $H$ measurement. Furthermore, the validation region offers a sample with a topology similar to the SR with which to study $V+$ jets production with a larger event sample but a lower signal-to-background ratio.
FIG. 3. The post-fit CR, $m_J$, mass distribution in the four $p_T$ regions used in the global likelihood of the differential fit (Sec. IX C). The $W(\ell\nu)$ contribution is flat in jet mass and for events with $p_T < 1$ TeV it is estimated to be 1–3% of the total. The $p_T > 1$ TeV region is shown in 10 GeV jet mass bins. The ratio of the data to the background prediction is shown in the lower panel. The shaded areas indicate the 68% CL for all background processes.

In the VR, $W + jets$ events outnumber $Z$ events nearly three to one due to the larger cross section and comparable acceptances. The decay products of the vector boson are reconstructed within the selected candidate jet in over 60% of events. In the remainder, the selected candidate jet is created by the recoil hadronic system, resulting in a nonresonant mass distribution, similar in shape to the multijet background, that enhances the high-mass tail. In the SR, the $Z + jets$ event contribution is dominant and exceeds that of $W + jets$ events by more than a factor of three, because of the sizable $Z \rightarrow b\bar{b}$ branching fraction and the flavor-tagging requirements. About 90% of the candidate jets in $Z + jets$ events contain the decay products of a vector boson. Due to the low misidentification rate for $b$-tagging, only 40% of candidate jets in $W + jets$ events are from the boson decay. The prevalence of candidate jets from the recoil system in $W + jets$ events leads to a broad $m_J$ distribution in the SR.

The $Z + jets$ normalization is determined directly from the data with the global likelihood described in Sec. VIII. Therefore, the impact of the considered systematic uncertainties in the modeling is limited to relative changes in acceptance across regions and the $m_J$ distribution shape. The $W + jets$ cross section is assigned a 10% uncertainty in the signal region [120]. For both processes, the largest effect from seven independent pairs of renormalization and factorization scale variations by factors of 0.5 and 2 corresponds to a 3%–20% error in the expected relative acceptance across regions. An alternative PDF set (MMHT2014NLO), $\alpha_s$ variations within the nominal PDF set, and changing the cluster fragmentation model to the Lund string model [121] does not lead to a significant change in the acceptance estimate relative to the nominal model. In the CR, the minor $W + jets$ contribution, referred to as $W(\ell\nu)$, has a total uncertainty of 30%.

All experimental uncertainties described in Sec. IV C are applied. Uncertainties in the JMR and JMS have the largest impact on the $V + jets$ normalization. Using the multijet model described in Sec. VII C and the likelihood described in Sec. VIII, the jet mass distribution in the leading-jet validation region is described to the level of agreement between simulation and data shown in Fig. 4.

1. Jet mass resolution uncertainty

The fitted $Z + jets$ normalization in the SR had a significant correlation with the reconstructed $m_J$ resolution as the flexibility of the $Z + jets$ template and the multijet
model could open a local minimum in the likelihood fit minimization procedure. Tests using subsets of a hybrid validation region, constructed to have a known amount of each process and discussed in the next section, highlighted this feature. In some instances, the best-fit value of the JMR parameter broadened the $Z^\text{+ jets}$ peak, corresponding to an increase of the $Z^\text{+ jets}$ normalization and a decrease of the multijet contribution compared to the expected values.

To stabilize the fit response, the $Z$ and $W$ resonance jet mass width is measured directly in two data samples as a function of $p_T$ and used to create additional constraints on the $V^\text{+ jets}$ JMR parameter in the global likelihood (see Sec. VIII). The two data samples are an alternative $t\bar{t}$ CR (WCR$_{t\bar{t}}$) and the VRL. The WCR$_{t\bar{t}}$ consists of selected semileptonic $t\bar{t}$ events having a resolved $Wb$ pair from a hadronically decaying top quark, providing a high-purity reconstructed $W$ peak with jet $p_T$ from 200 GeV up to about 600 GeV. The VRL covers the entire jet-$p_T$ range above 450 GeV, providing a clear $V$ peak but with considerably more multijet background.

The measured jet mass width of the $W$ and $Z$ resonances shows a smooth evolution from low $p_T$ in the WCR$_{t\bar{t}}$ to high $p_T$ in the VRL (see Fig. 5). These results differ from the nominal simulated $m_J$ resolution by less than 2.5% and have a precision that is around one fifth of the original JMR uncertainty after systematic uncertainties are incorporated to transfer the result to the $Z \rightarrow b\bar{b}$-dominated $V^\text{+ jets}$ sample in the SR. When included in the global likelihood, the correlation between the $Z^\text{+ jets}$ normalization and the JMR is reduced compared to when the auxiliary mass

![Image of post-fit leading-jet invariant mass distributions](image)

FIG. 4. Post-fit leading-jet invariant mass distributions after the multijet background was subtracted in the validation region for data (points with error bars) and the $V^\text{+ jets}$ ($W + Z$) and Top components (histograms) for (a) $450 < p_T < 650$ GeV, (b) $650 < p_T < 1000$ GeV, and (c) $p_T > 1$ TeV shown in wider 10 GeV jet mass bins. The $V^\text{+ jets}$ contribution is split into five generator “truth” $p_T^V$ volumes labeled $p_T^V-p_T^V$ for $p_T^V < 300$ GeV, 300–450 GeV, 450–650 GeV, 650–1000 GeV, and $> 1000$ GeV, respectively. The $t\bar{t}$ normalization and its uncertainty are set to the corresponding values from the CR$_{t\bar{t}}$. The $m_J$ range has been extended down to 60 GeV for only this fit to show the level of agreement along the rising edge of the $V^\text{+ jets} m_J$ distribution. The ratio of the data to the background prediction is shown in the lower panel. The shaded areas indicate the 68% CL for all background processes.
measurement is not considered. For example, in the inclusive signal region it decreases from \(-90\%\) to \(-30\%\).

C. Multijet production

The monotonically decreasing jet mass spectrum of the multijet background is modeled with an exponential function of a polynomial:

\[
f_N(x|\Theta) = \Theta_0 \exp \left( \sum_{i=1}^{N} \Theta_i x^i \right). \tag{1}
\]

where \(x = (m_j - 140 \text{ GeV})/70 \text{ GeV}\) and \(\Theta_i\) are the parameters of the fit. Parameter values are determined independently in each region simultaneously with the signal extraction. The optimal degree of the polynomial, \(N\) in Eq. (1), depends on the mass spectrum’s shape and the number of events analyzed. Values of \(N\) which are too large increase correlations between the multijet and resonant process models, decreasing their statistical significance. Values of \(N\) which are too small induce biases in the resonant process yields because the function is too rigid. These effects are studied in the VRL (VRS), which contains 58 (51) times the amount of SRL (SRS) data. To determine \(N\), statistical tests are performed on modified VR subsets, referred to as hybrid VR slices, with roughly the same number of events as the corresponding SR.

The hybrid VR is the best available proxy for the SR. The slices are generated from the VR subsets by replacing the VR resonance peaks with the corresponding SM prediction in the SR and correcting the underlying multijet mass spectrum shape to match that of the SR. The multijet shape difference between the SR and the VR is defined as the ratio of the SR multijet model (MJ_{SR}) to the VR multijet model (MJ_{VR}) (see Fig. 6). The MJ_{SR} is obtained from a global likelihood fit to the SRs and CR_{\bar{t}t}, including all experimental and modeling systematic uncertainties. The Higgs boson signal normalization parameter is blinded and common unconstrained normalization parameters are used within each jet \(p_T\) region for the \(V + \) jets and \(\bar{t}t\) templates separately. The MJ_{VR} is created from the average of the multijet model parameter values obtained from likelihood fits to ten\(^2\) random, orthogonal subsets of the VR including the full set of systematic uncertainties. The input \(\bar{t}t\) normalization and associated uncertainty are set to representative values from the CR_{\bar{t}t}. The VR \(V + \) jets and Top estimates (V{\text{VR}} and Top_{\text{VR}}) are created from the average post-fit contributions from the same ten fits to the VR. Each slice of the hybrid VR, VR_{hyb}^{i}, is built as

\[
V R_{hyb}^{i} = \left( V R^{i} - V_{VR}^{i} - Top_{VR}^{i} \right) \times \frac{M J_{SR}}{M J_{VR}}
+ V_{SR}^{i} + \bar{t}t_{SR}^{i} + H_{SR}^{i},
\]

where VR\(^i\) is the jet mass distribution of the data events in the \(\bar{t}t\) VR slice and V_{SR}^{i}, \bar{t}t_{SR}^{i}, and H_{SR}^{i} are the nominal MC predictions for \(V + \) jets, \(\bar{t}t\), and \(H\) production in the SR, respectively. The ratio of MJ_{SR} over MJ_{VR} is stable versus \(N\).

The optimal \(N\) for each region is chosen by considering three metrics evaluated with the VR_{hyb} collection. First, a log-likelihood ratio (LLR) test compares the result of an \(N\)-parameter fit (null hypothesis) with the result of an \((N + 1)\)-parameter fit (alternative hypothesis) in each VR_{hyb} slice without any injected SM resonances. By Wilks’ theorem, the likelihood ratio follows an asymptotic \(\chi^2\) distribution with one degree of freedom when the data corresponds to the null hypothesis. The corresponding distribution of \(p\)-values is flat. The smallest \(N\) that yields a uniform distribution of \(p\)-values is selected.

The LLR test ensures a good description of the data over the full mass range, but resonance measurements are sensitive to local effects. Two rate tests sensitive to local effects rely on the fit result for a free normalization parameter and its associated error (generalized as \(\mu_{VR}^{i} \pm \sigma_{stat}^{VR}\)) for either the \(Z + \) jets process or the Higgs boson process. Both utilize VR_{hyb} slices with all SR resonances injected at the SM rates. The fraction of results \(F_{2\sigma}\), where \(|\mu_{VR}^{i} - 1|\) is beyond twice its error \(\sigma_{stat}^{VR}\), gives an estimate of the probability that the multijet model allows a substantial artificial excess or deficit. A 2\(\sigma\) threshold

\(^{2}\)Ten subsets balance the statistical precision with the need to use greater values of \(N\) to model a larger dataset.
ensures some results from the full set of VR slices cross the boundary. The average value of \((\mu_{VR} - 1)/\sigma_{VR,stat}\) is denoted by \(\mu/\sigma\) and indicates a bias in determining the signal strength. The value of \(N\) chosen by the LLR test is incremented until \(F_{2\sigma}\) is compatible with 0.05, and \(\mu/\sigma\) stabilizes for both \(Z +\) jets and \(H\) production. The chosen values of \(N\) do not change when including systematic uncertainties and when not injecting the \(Z +\) jets and \(H\) processes into the \(VR_{hyb}\) slices (with appropriate changes to the definition of \(F_{2\sigma}\) and \(\mu/\sigma\)). A fifth-degree polynomial is used in both inclusive regions, while either a fourth- or fifth-degree polynomial is used in the analysis of \(p_T\) bins.

A nonzero value of \(\mu/\sigma\) for the \(N\) chosen in each region indicates a bias in the background model. It defines the spurious-signal systematic uncertainty and is in the range 0.01–0.33 for \(H\) and 0.15–0.65 for \(Z +\) jets production. A compatible estimate of the spurious-signal systematic uncertainty is found when not injecting the \(Z +\) jets or \(H\) processes into the \(VR_{hyb}\) slices. In all cases, the spurious-signal systematic uncertainty has an insignificant impact on the sensitivity.

Above 1 TeV, the \(V +\) jets yield extracted from the VR has a sizable statistical uncertainty, necessitating a change in the above procedure (see Fig. 4). Comparisons of fit quality and \(\mu/\sigma\) for both \(Z +\) jets and the Higgs boson in the hybrid VR for an \(N\) of 4, 5, and 6 show no improvement in modeling from the higher-order functions. The \(\mu/\sigma\) is calculated in two ways to estimate the spurious-signal systematic uncertainty for the chosen value of \(N = 4\). The \(V_{\tau\gamma}\) subtracted from each \(VR_i\) slice is taken from the average post-fit contributions to five subsets of the VR or the SM prediction. The spurious-signal systematic uncertainty is found to be 0.5 and 0.3 for \(Z +\) jets and \(H\), respectively, and has an insignificant impact on the sensitivity.

Finally, an alternative function for the multijet background modeling, the Laurent series, was tested. It provides a similar level of agreement with the multijet background in the \(VR_{hyb}\) collection. The differences between the two models are much smaller than the data statistical uncertainty and the expected Higgs boson yield. Therefore, the model choice does not motivate an additional systematic uncertainty.

VIII. STATISTICAL ANALYSIS

Signal yields are extracted by minimizing the negative logarithm of a likelihood function \(\mathcal{L}(\mu, \theta)\) with the
TABLE V. A summary of the systematic uncertainties included within the profile likelihood for the $H$ and $Z$ signal strength extraction. For a given uncertainty, the second column lists each process which has independent nuisance parameters within the likelihood. The third column describes how the systematic uncertainty is correlated across regions: “all” indicates a fully correlated parameter, “$p_T$ bins” indicates a decorrelation between the analysis $p_T$ bins, and “LS” means it is decorrelated between the SRL and SRS. For the inclusive and fiducial analysis configurations, “bins” does not apply, and should be understood to mean the same as “all.” The fourth column describes the change induced by the parameter. “S” means the $m_J$ shape will change while “N” denotes parameters which change the normalization and can result in a migration of events between regions. (•) Two minor components separately apply to $t\bar{t}$ and $V + j$ events. (•*) For cross section limits and measurements, this uncertainty covers relative acceptance across regions instead of the absolute cross section uncertainty. (•) The spurious-signal uncertainty is only applied to $Z + j$ events when the procedure to extract signal strengths in “truth”-based volumes is tested using $Z + j$ events in the SR.

<table>
<thead>
<tr>
<th>Description</th>
<th>Processes</th>
<th>Category</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reconstructed object systematic uncertainties</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JMR</td>
<td>$t\bar{t}$, $V + j$, $H$</td>
<td>$p_T$ bins</td>
<td>N + S</td>
</tr>
<tr>
<td>JMS (dominant)</td>
<td>$t\bar{t}$, $V + j$, $H$</td>
<td>$p_T$ bins</td>
<td>N + S</td>
</tr>
<tr>
<td>JMS (rest)</td>
<td>$t\bar{t}$, $V + j$, $H$</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>All</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>All</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>$b$-tag efficiency for $b$-jets</td>
<td>All</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>$b$-tag efficiency for $c$-jets</td>
<td>All</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>$b$-tag efficiency for light-flavor jets</td>
<td>Process modeling systematic uncertainties</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td><strong>Process modeling systematic uncertainties</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renormalization and factorization scale</td>
<td>$V + j$</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>Cross section</td>
<td>$W + j$</td>
<td>All</td>
<td>N</td>
</tr>
<tr>
<td>Cross section and acceptance</td>
<td>$W(f_W)$</td>
<td>All</td>
<td>N</td>
</tr>
<tr>
<td>Parton shower model</td>
<td>$t\bar{t}$</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>Matrix element calculation</td>
<td>$t\bar{t}$</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>Initial- and final-state radiation</td>
<td>$t\bar{t}$</td>
<td>All</td>
<td>N + S</td>
</tr>
<tr>
<td>Cross section and acceptance</td>
<td>$tH$</td>
<td>All</td>
<td>N</td>
</tr>
<tr>
<td>Cross section and acceptance(*)</td>
<td>$V_{BF} + VH + t\bar{t}H$</td>
<td>All</td>
<td>N</td>
</tr>
<tr>
<td>NLO EW corrections</td>
<td>$H$</td>
<td>All</td>
<td>N</td>
</tr>
<tr>
<td>Spurious signal</td>
<td>$H$</td>
<td>$p_T^H$ bins × LS</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>$Z + j$ (•)</td>
<td>$p_T^Z$ bins × LS</td>
<td>N</td>
</tr>
</tbody>
</table>

The likelihood function is defined as the product of Poisson probability terms, one for each bin of the $m_J$ distribution of the SRL, SRS, and CR$_R$ regions. Bin widths are set to 5 GeV, necessitating technical advancements within RooStats to fit an analytic function to a wide-binned dataset [124]. Systematic uncertainties enter the likelihood as nuisance parameters, $\theta$, constrained with Gaussian or log-normal probability density function priors. The JMR constraints obtained from the WCR$_{t\bar{t}}$ and VRL regions are included as Gaussian probability density function priors. Unconstrained, or free, parameters control the normalization of the MC templates within each jet $p_T$ region or within a given “truth”-based volume and are common to the SRL, SRS, and CR$_R$. For the multijet model, the function normalization and its polynomial coefficients are free parameters and independent for each jet mass distribution. Signal yields are expressed as signal strengths, $\mu$, obtained by normalizing the fitted number of signal events to the corresponding SM predictions. Upper limits on the Higgs boson signal strengths and production cross section are derived using the CL$_{s}$ method [125,126]. The expected limits assume no Higgs boson contribution.

Table V summarizes the systematic uncertainties considered in the likelihood fit. In addition, uncertainties due to the limited number of events in the simulated samples used for the background predictions are parametrized using the Beeston–Barlow “lite” technique [127]. Systematic variations yielding large statistical fluctuations are smoothed using custom algorithms which also remove variations resulting in effects below 2% within a given region.

**IX. RESULTS**

The three analysis configurations designed to probe Higgs boson production with considerable transverse momentum are summarized in Table II. The inclusive region provides a measure of the $H$ signal strength, the fiducial region allows a measurement of the fiducial cross section, and the differential regions are used to measure the cross section in four $p_T$ bins. All $H$ production modes are considered for the signal strength extraction. Signal
TABLE VI. A summary of the fiducial and STXS volumes used to determine which signal events are considered for the signal strength measurement in the fiducial and differential regions, respectively. Signal cross sections outside of these volumes are constrained to their SM prediction.

| Volume | $p_T^H$ (GeV) | $|y_H|$ |
|--------|----------------|---------|
| Fiducial | $>450$ | $<2$ |
| STXS | $300-450$, $450-650$, $650-1000$, $>1000$ | $<2$ |

strengths within the fiducial region consider events within a fiducial volume defined by requirements on the generator’s event “truth” record, the Higgs boson transverse momentum ($p_T^H$), and rapidity ($y_H$). Using the same “truth” information within the differential regions, volumes like the bins used in the simplified template cross section (STXS) framework and simply referred to as STXS volumes, are considered; the $p_T^H$ requirements match those for ggF production in the STXS scheme [5–7], but the $y_H$ requirement is more stringent and all production modes are included. The analysis jet-$p_T$ bins align well with the $p_T^H$-defined volumes. The yield of signal events outside the targeted volume(s) are constrained to their SM prediction within the theoretical and experimental uncertainties. The cross sections are obtained from the fitted signal yields divided by the integrated luminosity, corrected by the product of the estimated selection efficiency and fiducial or STXS volume acceptance. Using the same $p_T$ boundaries, cross section measurements of $V +$ jets production in the VRL and of $Z +$ jets production in the SR validate the method. For these tests, the fiducial and STXS volumes are defined by the generator “truth” vector-boson transverse momentum ($p_T^V$). Table VI summarizes the fiducial and STXS volumes and the SM predicted cross sections are in the HEPData repository [128].

A. Inclusive region

The inclusive region yields a Higgs boson signal strength of $\mu_H = 0.8 \pm 3.2$ for the combination of SRL, SRS and CR_{ij}. The results are summarized in Table VII and the yields in Table VIII. The post-fit SRL and SRS jet mass distributions are shown in Fig. 7. The Higgs boson signal sensitivity is limited by the size of the data sample. The leading sources of systematic uncertainty are the jet mass resolution and mass scale. The $\mu_{ij}$ value is compatible with $t \bar{t}$ event measurements in a similar kinematic phase-space [129]. With nearly 99% purity in top-jet events, the CR_{ij} data reduces the top quark jet JMR uncertainty to a relative 7% (from 20%). For $V +$ jets (Top) events, the JMS uncertainty’s nuisance parameter is pulled by ~80% (20%) of its original width, and its width is reduced by 50% (50%). The $V +$ jets resonance peak position moves by about 2 GeV from the MC prediction. No other nuisance parameters are significantly modified.

B. Fiducial region

In the fiducial region, the Higgs boson yield and cross section are determined within the fiducial volume defined by the Higgs boson transverse momentum $p_T^H > 450$ GeV and rapidity $|y_H| < 2.0$. Compared to the inclusive measurement discussed above, the fiducial region does not include the SRS region below 450 GeV. The acceptance times efficiency values for the different SM Higgs boson production processes are given in Table IX.

Two Higgs boson mass templates are used in each SR. The first describes the jet mass distribution of events within the fiducial volume; the second includes events outside the fiducial volume, i.e. those with $p_T^H$ below 450 GeV. The first component accounts for more than 80% of the Higgs boson signal selected and has a free normalization parameter common to the SRL and SRS. The second component is constrained to the SM value within the theoretical and experimental uncertainties and also tends to have a broader mass spectrum shifted to higher values.

This procedure is first tested with $W \rightarrow qq'$ and $Z \rightarrow qq'$ in the VR and $Z \rightarrow b\bar{b}$ in the SR. For these tests, the $V$ and $Z$ mass templates are structured similarly to those of the Higgs boson described above. The Higgs boson yields are kept fixed to the SM expectations in the fit and the result is insensitive to this choice. The fitted signal strength for $V +$ jets with $p_T^V > 450$ GeV is 1.01 ± 0.09. In the SR, the

TABLE VIII. Event yields and associated uncertainties after the global likelihood fit in the inclusive region. The total background yield can differ from the sum of the individual components due to rounding.

<table>
<thead>
<tr>
<th>Process</th>
<th>SRL</th>
<th>SRS</th>
<th>CR_{ij}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>590 700 ± 4200</td>
<td>529 300 ± 3500</td>
<td>⋯</td>
</tr>
<tr>
<td>$Z +$ jets</td>
<td>16 100 ± 2800</td>
<td>12 000 ± 2100</td>
<td>⋯</td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>3050 ± 720</td>
<td>2510 ± 500</td>
<td>⋯</td>
</tr>
<tr>
<td>Top</td>
<td>16 300 ± 1900</td>
<td>15 900 ± 2000</td>
<td>3737 ± 68</td>
</tr>
<tr>
<td>$W(\ell\nu)$</td>
<td>⋯</td>
<td>⋯</td>
<td>53 ± 16</td>
</tr>
<tr>
<td>$H$</td>
<td>400 ± 1500</td>
<td>300 ± 1300</td>
<td>⋯</td>
</tr>
<tr>
<td>Total</td>
<td>626 530 ± 820</td>
<td>560 090 ± 770</td>
<td>3790 ± 66</td>
</tr>
<tr>
<td>Data</td>
<td>626 532</td>
<td>560 083</td>
<td>3791</td>
</tr>
</tbody>
</table>
signal strength for $Z$ events with $p_T > 450$ GeV is $1.35 \pm 0.25$. These results are in agreement with the SM.

When extracting the Higgs boson signal strength, the likelihood fit result for $p_T > 450$ GeV provides a signal strength of $\mu_H = -0.1 \pm 3.5$. The corresponding observed (expected) 95% CL upper limit on the Higgs boson production cross section is

$$\sigma_H(p_T > 450 \text{ GeV}) < 115(128) \text{ fb}$$

and the SM prediction is 18.4 fb.

The results are summarized in Table X. The $\mu_H$ value is compatible with $t\bar{t}$ event measurements in a similar kinematic phase-space [129]. The Higgs boson signal strength sensitivity is limited by the data sample size, and the impact from the main sources of uncertainty is given in Table XI. The jet uncertainties give the largest contribution of systematic uncertainty, driven by JMS.

### TABLE IX. Signal acceptance times efficiency within the fiducial volume used in the fiducial region.

| Process | $p_T^H > 450 \text{ GeV}$ | $|y_H| < 2$ |
|---------|-----------------|-----------|
| All     | 0.24            |           |
| ggF     | 0.26            |           |
| VBF     | 0.22            |           |
| VH      | 0.27            |           |
| $t\bar{t}H$ | 0.20   |           |

### TABLE X. Expected and observed values of the signal strengths for the $H$, $Z$, and $t\bar{t}$ components in the fiducial fits.

<table>
<thead>
<tr>
<th>Result</th>
<th>$\mu_H$</th>
<th>$\mu_Z$</th>
<th>$\mu_{t\bar{t}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected</td>
<td>1.0 ± 3.4</td>
<td>1.00 ± 0.18</td>
<td>1.00 ± 0.08</td>
</tr>
<tr>
<td>Observed</td>
<td>-0.1 ± 3.5</td>
<td>1.30 ± 0.22</td>
<td>0.75 ± 0.06</td>
</tr>
</tbody>
</table>

### TABLE XI. Contributions to the systematic uncertainties for the measurement of the fiducial volume signal strength, defined as the signal yield relative to the SM prediction. The total uncertainty is also given for comparison.

<table>
<thead>
<tr>
<th>Uncertainty Contribution</th>
<th>$p_T^H &gt; 450 \text{ GeV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>3.5</td>
</tr>
<tr>
<td>Statistical</td>
<td>2.6</td>
</tr>
<tr>
<td>Systematic</td>
<td>2.3</td>
</tr>
<tr>
<td>Jet systematic uncertainties</td>
<td>2.2</td>
</tr>
<tr>
<td>Modeling and theory systems</td>
<td>0.8</td>
</tr>
<tr>
<td>Flavor-tagging systems</td>
<td>0.2</td>
</tr>
</tbody>
</table>

FIG. 7. Post-fit jet mass distributions for the various components in the inclusive SRL (left) and SRS (right) regions. The middle panels show the distributions after subtraction of the multijet distribution. The shaded areas indicate the 68% CL for the multijet background from the fitted parameters and normalizations of the exponentiated polynomials. The lower panels show the distributions after subtraction of all the fitted background processes: multijet, $V + $ jets, and $t\bar{t}$. The shaded areas indicate the 68% CL for all background processes.
FIG. 8. For each of the $p_T^H$ differential volumes (x-axis), the expected signal event yield for all Higgs boson events (left) and the fraction of signal in percent (right) in each reconstructed jet $p_T$ region (y-axis) is shown. The leading jet’s $p_T$ in the SRL is denoted by $p_T^L$ and the subleading jet’s $p_T$ in the SRS is denoted by $p_T^S$.

effects. Half of the JMS contribution comes from the JMS uncertainty for both the $V + \text{jets}$ and $t\bar{t}$ events, while the other half corresponds to the uncertainty in the JMS of reconstructed Higgs bosons. With nearly 99% purity in top-jet events, the CR$_{ij}$ data reduces the top quark jet JMR uncertainty to a relative 5%, one quarter of the original size. In $V + \text{jets}$ events, the dominant JMS uncertainty’s nuisance parameter is pulled by $-80\%$ of its original width, and its width is reduced by 50%. This moves the $V + \text{jets}$ resonance peak position about 2 GeV. For Top events, the corresponding parameter is also constrained to 50% of its original value. No other nuisance parameters are modified significantly.

**C. Differential regions**

Possible deviations from the SM predictions could manifest with an amplitude increasing with $p_T^H$. The differential regions aim to measure the Higgs boson transverse momentum spectrum in four $p_T^H$ volumes $300-450$ GeV, $450-650$ GeV, $650-1000$ GeV, and greater than 1 TeV using four signal regions requiring the candidate jet $p_T$ to be in the range $250-450$ GeV, $650-450$ GeV, $650-1000$ GeV, or above 1 TeV (see Table II). Extending the procedure adopted for the fiducial region measurement, a Higgs boson mass template for each $p_T^H$ volume is used within each jet $p_T$ region in the global likelihood. The $p_T^H > 1$ TeV volume probes a new domain of highly boosted $Z$ and Higgs boson reconstruction. The expected sensitivity in the SRS with jet $p_T$ above 1 TeV is marginal because the muon-in-jet correction and $b$-tagging turn-on effects are more significant than in the SRL. Therefore, only the SRL and CR$_{ij}$ regions are included for measurements above 1 TeV. Figure 8 presents the expected signal yield in each candidate jet $p_T$ region for each STXS volume and the corresponding fraction of signal events. The acceptance times efficiency values for the different Higgs boson production processes are given in Table XII.

Again, the procedure is tested with $W \rightarrow q\bar{q}'$ and $Z \rightarrow q\bar{q}'$ in the VR and $Z \rightarrow b\bar{b}$ in the SR with the $V$ and $Z$ mass templates structured similarly to those of the Higgs boson described above. Due to the larger sample size, the VRL is divided into five slices, the fit is performed independently on each slice, and the results are averaged. In the SR, the $Z$ fit is performed in

---

**Table XII.** Signal acceptance times efficiency for the STXS volumes in the differential measurement. Along with the $p_T^H$ requirements shown, $|y_H| < 2$ is required. For events with $p_T^H < 300$ GeV, the acceptance times efficiency is less than $0.1 \times 10^{-2}$.

<table>
<thead>
<tr>
<th>Process</th>
<th>$300 &lt; p_T^H &lt; 450$ GeV</th>
<th>$450 &lt; p_T^H &lt; 650$ GeV</th>
<th>$650 &lt; p_T^H &lt; 1000$ GeV</th>
<th>$p_T^H &gt; 1$ TeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>$1.3 \times 10^{-2}$</td>
<td>0.23</td>
<td>0.31</td>
<td>0.23</td>
</tr>
<tr>
<td>ggF</td>
<td>$0.7 \times 10^{-2}$</td>
<td>0.25</td>
<td>0.35</td>
<td>0.28</td>
</tr>
<tr>
<td>VBF</td>
<td>$0.4 \times 10^{-2}$</td>
<td>0.21</td>
<td>0.32</td>
<td>0.25</td>
</tr>
<tr>
<td>VH</td>
<td>$1.7 \times 10^{-2}$</td>
<td>0.26</td>
<td>0.30</td>
<td>0.20</td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>$4.7 \times 10^{-2}$</td>
<td>0.19</td>
<td>0.24</td>
<td>0.19</td>
</tr>
</tbody>
</table>
the SRL, SRS, and CR$_t$ regions with the Higgs boson contribution fixed to the SM prediction. Again, the results do not change when using a freely floating Higgs boson normalization. Results of the two differential fits are shown in Fig. 9, where they are compared with the predictions for the EW NLO and QCD NNLO corrections as a function of reconstructed $p_T^V$. Both results agree with SM expectations.

To extract the four Higgs boson signal strengths within the STXS volumes, ten differential SR and CR regions defined in Table II are fitted simultaneously, exploiting the corresponding systematic uncertainty configurations shown in Table V. The results are summarized in Table XIII and Fig. 10. The $t\bar{t}$ normalization corrections determined from the data in each jet $p_T$ region are compatible with $t\bar{t}$ event measurements in a similar kinematic phase-space [129]. The four Higgs boson signal strengths are compatible, with a $p$-value of 0.53. Post-fit jet mass distributions from the SRL are shown in Fig. 11.

The resulting Higgs boson production cross section for $p_T^H > 1$ TeV is

![Diagram](image_url)

**Table XIII.** Expected and observed values of the signal strengths for the $H$, $Z$, and $t\bar{t}$ components in the differential fits. The value for $\mu_H$ refers to a STXS volume, while those for $\mu_Z$ and $\mu_{t\bar{t}}$ pertain to the corresponding jet $p_T$ regions. Absolute cross section uncertainties are not included in the reported $\mu_Z$ and $\mu_{t\bar{t}}$ values.

<table>
<thead>
<tr>
<th>$p_T^H$ (GeV)</th>
<th>$\mu_H$</th>
<th>Expected</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>300–450</td>
<td></td>
<td>$1 \pm 18$</td>
<td>$-6 \pm 18$</td>
</tr>
<tr>
<td>450–650</td>
<td></td>
<td>$1.0 \pm 3.3$</td>
<td>$-3 \pm 5$</td>
</tr>
<tr>
<td>650–1000</td>
<td></td>
<td>$1 \pm 6$</td>
<td>$5 \pm 7$</td>
</tr>
<tr>
<td>$&gt;1000$</td>
<td></td>
<td>$1 \pm 30$</td>
<td>$18 \pm 32$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Jet $p_T$ (GeV)</th>
<th>$\mu_Z$</th>
<th>Expected</th>
<th>Observed</th>
<th>$\mu_{t\bar{t}}$</th>
<th>Expected</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>250–450</td>
<td>$1.0 \pm 1.1$</td>
<td>$1.8 \pm 1.1$</td>
<td>$1.00 \pm 0.07$</td>
<td>$0.85 \pm 0.06$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>450–650</td>
<td>$1.00 \pm 0.17$</td>
<td>$1.28 \pm 0.22$</td>
<td>$1.00 \pm 0.07$</td>
<td>$0.76 \pm 0.06$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>650–1000</td>
<td>$1.00 \pm 0.33$</td>
<td>$1.4 \pm 0.4$</td>
<td>$1.00 \pm 0.09$</td>
<td>$0.74 \pm 0.08$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&gt;1000$</td>
<td>$1.0 \pm 1.6$</td>
<td>$2.4 \pm 1.7$</td>
<td>$1.00 \pm 0.22$</td>
<td>$0.57 \pm 0.18$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

092003-17
The Higgs boson signal strengths correspond to the following observed (expected) 95% CL and the SM prediction is 0.13 fb. The differential results are labeled with the corresponding $p_T$ range as a superscript. The $Z +$ jets signal strengths $\mu_Z$ are labeled with the corresponding jet $p_T$ range as a superscript.

$$\sigma_H(p_T^H > 1 \text{ TeV}) = 2.3 \pm 3.9(\text{stat}) \pm 1.3(\text{syst}) \pm 0.5(\text{theory}) \text{ fb}$$

and the SM prediction is 0.13 fb. The differential results correspond to the following observed (expected) 95% CL upper limits on the Higgs boson differential production cross sections:

$$\sigma_H(300 < p_T^H < 450 \text{ GeV}) < 2.9(3.1) \text{ pb},$$

$$\sigma_H(450 < p_T^H < 650 \text{ GeV}) < 89(102) \text{ fb},$$

$$\sigma_H(650 < p_T^H < 1000 \text{ GeV}) < 39(34) \text{ fb},$$

$$\sigma_H(p_T^H > 1000 \text{ GeV}) < 9.6(7.4) \text{ fb}.$$
The largest source of uncertainty is the data sample size. The contributions from the main sources of uncertainty are summarized in Table XIV. The jet uncertainties tend to give the largest contribution of systematic uncertainty, driven by JMS effects. The dominant JMS uncertainty component in V + jets events is reduced by 30%–40% separately in the three SRs with jet $p_T > 450$ GeV. The same component is reduced by 20%–65% for Top events over the full jet $p_T$ range. The JMS nuisance parameter receives small pulls that differ between V + jets and Top events and with $p_T$. The impact of JMR increases above 1 TeV, where the JMR measurement constraints on V + jets are looser due to the extrapolation uncertainties (see Fig. 5). The JMR uncertainty for jets in Top events is constrained to a relative 6%–9% (from 20%) below 1 TeV but is 18% above 1 TeV. Similarly, the flavor-tagging uncertainties increase above 1 TeV due to the extrapolation from the $p_T$ range of the calibration regions. The $t\bar{t}$ modeling systematic uncertainties are more relevant in the first jet-$p_T$ bin and decrease above 450 GeV, where the top quark decay products become more collimated, thus reducing the contamination around the Higgs boson mass peak. The fit reduces these systematic uncertainties by 20%–30%, driven by the purity of the CR$_{tt}$. Other than the few mentioned, no other nuisance parameter has a posterior probability very different from the prior.

X. CONCLUSIONS

High-$p_T$ Higgs boson production is studied in the $b\bar{b}$ decay channel probing a new highly Lorentz-boosted boson reconstruction domain with transverse momentum above 1 TeV. The results are based on $pp$ collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector during Run 2 of the LHC, corresponding to an integrated luminosity of 136 fb$^{-1}$. The Higgs boson is reconstructed as a single large-$R$ jet and identified with $b$-tagging techniques. The measured signal strengths of $W \to q\bar{q}$ and $Z \to q\bar{q}$ in the validation region and $Z \to b\bar{b}$ in the signal region agree with the SM predictions and validate the experimental techniques.

The observed (expected) 95% CL limit on the Higgs boson production cross section for $p_T^H > 450$ GeV obtained from the fiducial signal region is 115 (128) fb. From the four differential signal regions, the observed (expected) 95% CL limits on the Higgs boson production cross section are:

\[
\sigma_H(300 < p_T^H < 450 \text{ GeV}) < 2.9(3.1) \text{ pb},
\]

\[
\sigma_H(450 < p_T^H < 650 \text{ GeV}) < 89(102) \text{ fb},
\]

\[
\sigma_H(650 < p_T^H < 1000 \text{ GeV}) < 39(34) \text{ fb},
\]

\[
\sigma_H(p_T^H > 1000 \text{ GeV}) < 9.6(7.4) \text{ fb}.
\]

The Higgs boson production cross section for $p_T^H > 1$ TeV is found to be

\[
\sigma_H(p_T^H > 1 \text{ TeV}) = 2.3 \pm 3.9(\text{stat}) \pm 1.3(\text{syst}) \pm 0.5(\text{theory}) \text{ fb}.
\]

All of the Higgs boson results are consistent with the Standard Model predictions.

<table>
<thead>
<tr>
<th>Uncertainty Contribution</th>
<th>300 &lt; $p_T^H$ &lt; 450 GeV</th>
<th>450 &lt; $p_T^H$ &lt; 650 GeV</th>
<th>650 &lt; $p_T^H$ &lt; 1000 GeV</th>
<th>$p_T^H$ &gt; 1 TeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>18</td>
<td>5.0</td>
<td>6.5</td>
<td>32</td>
</tr>
<tr>
<td>Statistical</td>
<td>16</td>
<td>3.0</td>
<td>5.5</td>
<td>30</td>
</tr>
<tr>
<td>Systematic</td>
<td>7</td>
<td>3.9</td>
<td>3.4</td>
<td>10</td>
</tr>
<tr>
<td>Jet systematic uncertainties</td>
<td>6</td>
<td>3.8</td>
<td>3.4</td>
<td>9.5</td>
</tr>
<tr>
<td>Modeling and theory systs.</td>
<td>4</td>
<td>0.7</td>
<td>0.7</td>
<td>2</td>
</tr>
<tr>
<td>Flavor-tagging systs.</td>
<td>0.2</td>
<td>0.4</td>
<td>0.4</td>
<td>2</td>
</tr>
</tbody>
</table>

TABLE XIV. Contributions to the systematic uncertainties for the differential measurements of the signal strength, defined as the signal yield relative to the SM prediction within STXS volumes. The total uncertainty is also given for comparison.
ACKNOWLEDGMENTS

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MEiN, Poland; FCT, Portugal; MNE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; PNRA, Slovenia; DSI/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; Norwegian Financial Mechanism 2014-2021, Norway; NRC and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [130].


ATLAS Collaboration, Higgs boson production cross-section measurements and their EFT interpretation in the $4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector, Eur. Phys. J. C 80, 957 (2020); 81, 29(E) (2021).


G. Lusioni, P. Nason, C. Oleari, and F. Tramontano, $HW^\pm/HZ + 0$ and $1$ jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO, J. High Energy Phys. 10 (2013) 083.
S. Frixione, P. Nason, and G. Ridolfi, A Positive-weight
Decay of supersymmetric particles: The program SUSY-HIT
38, 635 (2007).

A. Breidenstein, A. Denner, S. Dittmaier, and M. M. Weber,
Radiative corrections to the semileptonic and hadronic
Higgs-boson decays \( H \to WW/ZZ \to 4 \) fermions, J. High

A. Breidenstein, A. Denner, S. Dittmaier, and M. M. Weber,
Precise predictions for the Higgs-boson decay \( H \to WW/ZZ \to 4 \) leptons, Phys. Rev. D 74, 013004 (2006).

A. Breidenstein, A. Denner, S. Dittmaier, and M. M. Weber,


U. Blumenschein et al., Pushing the precision frontier at the
LHC with \( V + \) jets, arXiv:1802.02100.

A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover,
A. Huss, and T. A. Morgan, Precise QCD Predictions for
the Production of a Z Boson in Association with a

A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover,
A. Huss, and T. A. Morgan, The NNLO QCD corrections
cern.ch/record/2216168.

ATLAS Collaboration, Simulation of top-quark production
for the ATLAS experiment at \( \sqrt{s} = 13 \) TeV, Report
cern.ch/record/2216168.

ATLAS Collaboration, Studies on top-quark Monte Carlo

ATLAS Collaboration, Simulation of top-quark production for the ATLAS experiment at \( \sqrt{s} = 13 \) TeV, Report
cern.ch/record/2216168.

ATLAS Collaboration, Studies on top-quark Monte Carlo
modelling with Sherpa and MG5_aMC@NLO, Report
ch/record/2261938.

ATLAS Collaboration, Improvements in \( \ell \ell \) modelling
using NLO + PS Monte Carlo generators for Run 2, Report
cds.cern.ch/record/2630327.

ATLAS Collaboration, Multijet simulation for 13 TeV
cds.cern.ch/record/2672252.

A. Denner, S. Dittmaier, S. Kallweit, and A. Muck,
Electroweak corrections to Higgs-strahlung off \( W/Z 
\) bosons at the Tevatron and the LHC with HAWK, J. High

A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, HAWK
2.0: A Monte Carlo program for Higgs production in
vector-boson fusion and Higgs strahlung at hadron

F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S.

F. Cascioli, P. Maierhöfer, and S. Pozzorini, Scattering

A. Denner, S. Dittmaier, and L. Hofer, Collier: A fortran-
based complex one-loop library in extended regularizations,

R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. D.
Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. J. Latorre,
J. Rojo, and M. Ubiali, Parton distributions for the LHC run


CONSTRAINTS ON HIGGS BOSON PRODUCTION WITH LARGE \( Z \) \( \text{BR}_{\text{SM}} \) \( \times \) \( \text{BR}_{\text{BSM}} \) 

PHYS. REV. D 105, 092003 (2022)

(Atlas Collaboration)
25d National Institute for Research and Development of Isotopic and Molecular Technologies, Physics Department, Cluj-Napoca, Romania
25e University Politehnica Bucharest, Bucharest, Romania
25f West University in Timisoara, Timisoara, Romania

26a Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovak Republic
26b Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic

27 Physics Department, Brookhaven National Laboratory, Upton, New York, USA
28 Departamento de Física (FCEN) and IFIBA, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina

29 California State University, Long Beach, California, USA
30 Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
31a Department of Physics, University of Cape Town, Cape Town, South Africa
31b Thembalabs, Western Cape, South Africa
31c Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg, South Africa

31d National Institute of Physics, University of the Philippines Diliman (Philippines), Diliman, Philippines
31e University of South Africa, Department of Physics, Pretoria, South Africa
31f School of Physics, University of the Witwatersrand, Johannesburg, South Africa
32 Department of Physics, Carleton University, Ottawa, Ontario, Canada

33a Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies—Université Hassan II, Casablanca, Morocco
33b Faculté des Sciences, Université Ibn-Toufaily, Kénitra, Morocco
33c Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA-Marrakech, Morocco
33d LPMR, Faculté des Sciences, Université Mohamed Premier, Oujda, Morocco
33e Faculté des sciences, Université Mohammed V, Rabat, Morocco
33f Mohammed VI Polytechnic University, Ben Gurir, Morocco
34 CERN, Geneva, Switzerland

35 Enrico Fermi Institute, University of Chicago, Chicago, Illinois, USA
36 LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand, France
37 Nevis Laboratory, Columbia University, Irvington, New York, USA
38 Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
39a Dipartimento di Fisica, Università della Calabria, Rende, Italy
39b INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Italy
39c Physics Department, Southern Methodist University, Dallas, Texas, USA
40 Physics Department, University of Texas at Dallas, Richardson Texas, USA
41 National Centre for Scientific Research “Demokritos”, Agia Paraskevi, Greece
42 Department of Physics, Stockholm University, Stockholm, Sweden
43a Oskar Klein Centre, Stockholm, Sweden
44 Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen, Germany
45a Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany
46 Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany
47 Department of Physics, Duke University, Durham, North Carolina, USA
48 SUPA—School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
49 INFN e Laboratori Nazionali di Frascati, Frascati, Italy
50 Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
51 II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
52 Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève, Switzerland
53a Dipartimento di Fisica, Università di Genova, Genova, Italy
53b INFN Sezione di Genova, Italy
54 II. Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany
55 SUPA—School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
56 LPSC, Université Grenoble Alpes, CNRS/IN2P3, Grenoble INP, Grenoble, France
57 Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge, Massachusetts, USA
58a Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei, China
58b Institute of Frontier and Interdisciplinary Science and State Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao, China
58c School of Physics and Astronomy, Shanghai Jiao Tong University, Key Laboratory for Particle Astrophysics and Cosmology (MOE), SKLPPC, Shanghai, China