ATLAS CONF Note

ATLAS-CONF-2021-010

March 26, 2021

# Study of Higgs-boson production with large transverse momentum using the $H \to b\bar{b}$ decay with the ATLAS detector

The ATLAS Collaboration

This note reports studies of Higgs boson production with sizeable transverse momentum decaying to a $b\bar{b}$ quark pair. The analyzed data were recorded with the ATLAS detector in proton–proton collisions with a center-of-mass energy of $\sqrt{s}$ = 13 TeV at the Large Hadron Collider between 2015 and 2018, corresponding to an integrated luminosity of 136 fb$^{-1}$. Higgs bosons decaying to $b\bar{b}$ are reconstructed as single large-radius jets and identified by the experimental signature of two $b$-hadron decays. The experimental techniques are validated in the same kinematic regime using the $Z \to b\bar{b}$ process. For Higgs boson production at transverse momenta above 450 GeV, the production cross section is found to be 13 ± 57(stat.) ± 22(syst.) ± 3(theo.) fb. The differential cross section 95% confidence level upper limits as a function of Higgs boson transverse momentum are $\sigma_H(300 < p_{\mathrm{T}}^H < 450 \text{ GeV}) < 2.8$ pb, $\sigma_H(450 < p_{\mathrm{T}}^H < 650 \text{ GeV}) < 91$ fb, $\sigma_H(p_{\mathrm{T}}^H > 650 \text{ GeV}) < 40.5$ fb, and $\sigma_H(p_{\mathrm{T}}^H > 1 \text{ TeV}) < 10.3$ fb. All results are consistent with the Standard Model predictions.

*Update April 14, 2021: Tables 2 and 15 were updated to correct the fiducial volume $p_T^H$ range in the differential SRL.*

# 1 Introduction

The characterization of the Higgs sector has steadily improved since the Higgs boson ($H$) discovery [1, 2] using proton–proton ($pp$) collision data produced by the Large Hadron Collider (LHC) at CERN. Five production modes: gluon–gluon fusion (ggF), vector–boson fusion (VBF), associated production with a weak vector boson ($WH$ or $ZH$, collectively $VH$), and associated production with a top–antitop pair ($t\bar{t}H$) and five decay modes: $H \to \gamma\gamma$, $ZZ^*$, $WW^*$, $\tau\tau$, $b\bar{b}$ have been observed [3, 4]. Inclusive cross-section measurements have evolved to differential measurements and measurements in kinematic regions defined within the simplified template cross-section framework (STXS) [5–7]. All results agree with the Standard Model (SM) predictions, but sizeable regions of the Higgs sector remain unexplored. In one such region, where the Higgs boson transverse momentum, $p_{\mathrm{T}}^{H}$, reaches the TeV scale, the cross-section hierarchy is very different from the inclusive cross section where ggF is nearly 90% of the total. At the TeV scale, ggF, $VH$, and the sum of the VBF and $t\bar{t}H$ production cross sections are all roughly equal.

The leading effects of many beyond the SM (BSM) scenarios, such as those encapsulated in effective field theories (EFTs), can be parametrized through effective operators suppressed by a new physics scale $\Lambda$. The most energetic parton-level interactions at the LHC allow access to regions where some operators' effects are enhanced by powers of $E/\Lambda$. Here, measurements with an extended energy reach outperform precise low-energy measurements in new physics sensitivity as the signal-to-background ratio increases with $E$. The differential cross section as a function of the Higgs boson transverse momentum exemplifies this potential as it is sensitive to BSM contributions that mainly impact high-$p_{\mathrm{T}}$ Higgs production [8–12]. For example, the ggF process is sensitive to the structure of quasi point-like couplings within the loop nature of the effective $ggH$ coupling. Using the $H \to b\bar{b}$ decay, which exhibits the largest branching fraction, mitigates the impact of the reduced absolute cross section in the high-$p_{\mathrm{T}}$ regime.

This note reports the first ATLAS studies of high-$p_{\mathrm{T}}$ Higgs bosons decaying to a $b\bar{b}$ pair where no restrictions are applied to select a particular Higgs boson production mode and the production yield is determined in different $p_{\mathrm{T}}^{H}$ regions. The fiducial cross section is measured for $p_{\mathrm{T}}^{H} > 450$ GeV and $p_{\mathrm{T}}^{H} > 1$ TeV, allowing for a straightforward comparison to theoretical calculations, such as those reported in Ref. [13]. A differential cross-section measurement within three $p_{\mathrm{T}}^{H}$ bin intervals, 250–450 GeV, 450–650 GeV, and above 650 GeV, affords enhanced sensitivity to BSM physics effects. The dataset used corresponds to 136 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 13$ TeV collected with the ATLAS detector [14] using jet-based trigger requirements during Run 2 (2015–2018) of the LHC. In a similar analysis, the CMS Collaboration measured a signal strength of $\mu_H = 3.7^{+1.6}_{-1.5}$ in the $H \to b\bar{b}$ decay mode, and presented ggF differential cross sections while considering other Higgs boson production modes as a background using events within the high-$p_{\mathrm{T}}^{H}$ regime [15].

Higgs boson candidates with large Lorentz boosts are reconstructed as single large-radius jets having a mass compatible with 125 GeV [16]. To populate the signal region, events must have at least two jets, where one exceeds $p_{\mathrm{T}} > 450$ GeV, to ensure a fully efficient trigger response, and at least one contains evidence of two $b$-hadron decays. Both the leading and subleading jet in the event are considered in the identification of Higgs boson candidates: including the subleading jet increases the sensitivity for $p_{\mathrm{T}}^{H} > 450$ GeV by 11% and permits a fiducial cross-section measurement down to $p_{\mathrm{T}}^{H} = 300$ GeV, overlapping with measurements in other decay channels [17–23]. For a Higgs boson mass ($m_H$) of 125 GeV, the SM predicts ggF-produced events to be the largest contribution reconstructed near $m_H$ in all but one fiducial volume considered. $t\bar{t}H$-produced events are nearly 40% of the total in the lowest $p_{\mathrm{T}}$ region. $VH$-produced events account for around 25% of all Higgs boson events in the subleading jet category. VBF-produced events consistently contribute over 15% of the total yield.

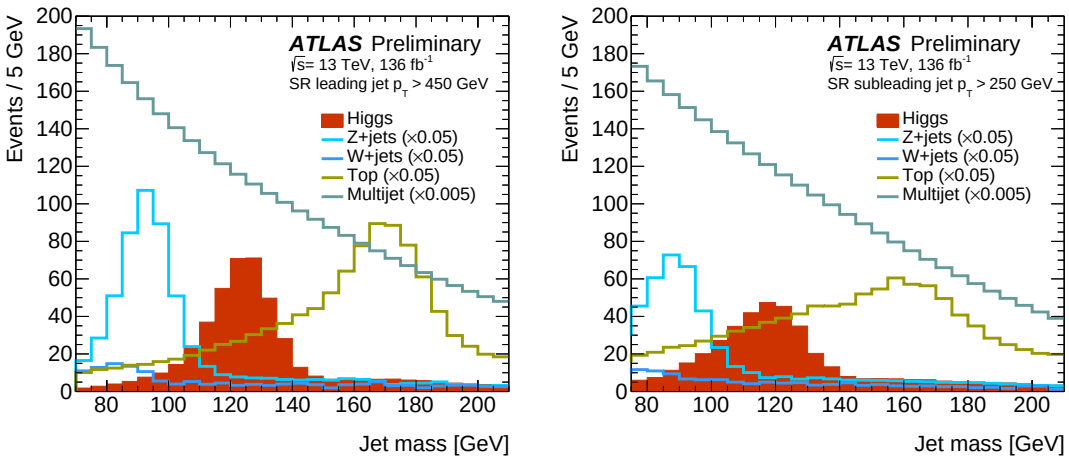



Figure 1: The jet mass distribution for the $H$, $Z$, $W$, and top-quark contributions from the SM prediction as well as the multijet jet mass distribution extracted from data for the signal region (SR) defined by the leading (left) and subleading (right) jets.

A few measurements can be motivated by the minimal definition of the recoil system. For example, dimension-6 EFT operators in the SILH basis [24], such as $c_g$ (modifies the interactions between the Higgs boson and gluons) and $c_{tg}$ (modifies the interactions between gluons and the top quark), motivate measuring $t\bar{t}H$ + ggF production. The corresponding high-$p_{\mathrm{T}}^H$ enhancements from a combination of just these coefficients can be within the sensitivity of this analysis without creating significant deviations from the SM prediction at low $p_{\mathrm{T}}^H$. However, the effect is not uniform in both production modes, with $t\bar{t}H$ being less sensitive. Recent results on VBF production with Higgs boson decays to photons and leptons also included high-$p_{\mathrm{T}}^H$ event categories, but have limited reach [19, 22, 25–28]. The analysis of $VH$ production with leptonic $V$ decays has achieved considerable sensitivity in the high $p_{\mathrm{T}}^H$ regime for a specific production mode [29–31]. The analysis discussed here is designed to be sensitive to ggF production while including the other main production modes. This approach enhances the sensitivity to possible BSM effects and minimizes the dependence on theoretical assumptions.

The dominant background process is multijet production which exhibits a monotonically decreasing jet mass distribution. Hadronically decaying vector bosons, produced in association with jets ($V$+jets) and events with top-quarks ($t\bar{t}$, $tW$ jointly referred to as *Top*) populate the mass regions below and above $m_H$, respectively, as shown in Figure 1. The $Z$ and $H$ resonance structures are distinct from the smoothly falling multijet background while the top-quark resonance spreads over a large portion of the high-mass region. Therefore, the signal extraction uses the reconstructed jet mass distribution as a discriminant and an analytic function is used to model the multijet background. The acceptance of Higgs boson, $V$+jets, and Top events is estimated from simulation. A dedicated control region is used to determine the Top yield while verifying the mass scale and resolution of top-quark jets in a broad range of jet $p_{\mathrm{T}}$. The same is done independently for the mass scale of jets from $V$ decays while all multijet model parameters, the $Z$ normalization, and the Higgs boson normalization are entirely determined, simultaneously, directly from the signal region data. The binned maximum-likelihood fit, referred to as the global likelihood fit, used for signal extraction, is employed in a validation region to extract the $V$+jets yield to ensure the analysis strategy is robust.

## 2 ATLAS detector

The ATLAS experiment [14] at the LHC is a multi-purpose 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. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A hadron (steel/scintillator-tile) calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The forward regions are instrumented with LAr calorimeters for both 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 tracking chambers and fast detectors for triggering. A two-level trigger system is used to select events [32]. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most 100 kHz. A software-based trigger further reduces the accepted event rate to 1 kHz on average.

## 3 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. Events must satisfy a set of triggers requiring a reconstructed anti-$k_t$ jet with radius parameter $R = 1.0$ [33] or a muon with $p_{\mathrm{T}} > 50$ GeV. To adapt to different instantaneous luminosity profiles and the inclusion of pile-up suppressing techniques within the data acquisition system [34], the jet $p_{\mathrm{T}}$ and mass thresholds differ for each year of data taking. The trigger jet $p_{\mathrm{T}}$ threshold varies from 360 GeV to 420 GeV, and the trigger jet mass threshold is either not applied, 30 GeV, or 35 GeV. The total integrated luminosity is 136 fb$^{-1}$ and 139 fb$^{-1}$ for the jet- and muon-triggered data, respectively, with an uncertainty of 1.7% [35, 36].

Monte Carlo (MC) simulated events are used to model the resonant backgrounds ($W$+jets, $Z$+jets, and top-quark production) as well as four Higgs production processes: ggF, VBF, $VH$, and $t\bar{t}H$.

Higgs boson ggF production is simulated at next-to-leading-order (NLO) accuracy in QCD with finite mass effects by using the Hj-MiNLO [37–39] prescription with the Powheg program [40–42] as discussed in Ref. [43]. NLO accuracy in QCD is achieved for VBF and $t\bar{t}H$ production and LO accuracy for $gg \to VH$ production using the Powheg-Box v2 [40–42, 44, 45] program. Using the Powheg-Box v2 program, the improved MiNLO [46] calculation, and the GoSam [47] program, $qq \to VH$ production is also simulated at NLO accuracy in QCD. Corrections for NLO electroweak (EW) effects are applied as a function of the generated Higgs boson transverse momentum for VBF, $VH$, and $t\bar{t}H$ production. The production cross sections used are compatible with those presented in Ref. [13] except for $t\bar{t}H$ production where an overall scale factor is applied to recover the difference. The Higgs boson branching fractions are calculated with HDECAY [48–50] and PROPHECY4F [51–53].

[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 $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

Table 1: The generators used for the simulation of the signal and background processes. Matrix element, parton shower and underlying event are abbreviated as ME, PS, and UE respectively. (∗) Powheg was configured to output events with Born $k_T$ above 200 GeV using the *bornktmin* setting. (•) Corrections for NLO EW effects computed with Hawk [57, 58] are applied as a function of the generated Higgs boson transverse momentum. (○) Corrections for NLO EW effects computed with Sherpa+OpenLoops 1 [59–61] are applied as a function of the generated Higgs boson transverse momentum and were provided by Ref. [13]. (†) Sherpa provides 1 additional parton at NLO accuracy and up to 4 additional partons at LO in QCD and custom NNLO QCD corrections were provided by the NNLOJET group.

| Process | ME generator | ME PDF | PS and Hadronisation | UE model tune | Cross-section order |
|---|---|---|---|---|---|
| Higgs Boson | | | | | |
| $gg \to H \to b\bar{b}$ | Powheg-Box v2$^{(*)}$ [40–42]+ MiNLO [37–39] | NNPDF3.0NNLO [62] | Pythia 8.212 [63] | AZNLO [64] | NLO(QCD) + LO(EW) |
| $qq \to H \to q'q'b\bar{b}$ | Powheg-Box v2 [40–42, 44] | NNPDF3.0NLO [62] | Pythia 8.230 | AZNLO | NLO(QCD) + NLO(EW)$^{(\bullet)}$ |
| $qq \to WH$ $\to qq'b\bar{b}$ $\to \ell\nu b\bar{b}$ | Powheg-Box v2 + GoSam [47] + MiNLO [46] | NNPDF3.0NLO | Pythia 8.240 Pythia 8.212 | AZNLO | NNLO(QCD) + NLO(EW)$^{(\bullet)}$ |
| $qq \to ZH$ $\to q\bar{q}b\bar{b}$ $\to \nu\nu b\bar{b}$ $\to \ell\ell b\bar{b}$ | Powheg-Box v2 + GoSam+ MiNLO | NNPDF3.0NLO | Pythia 8.240 Pythia 8.212 | AZNLO | NNLO(QCD) + NLO(EW)$^{(\bullet)}$ |
| $gg \to ZH$ $\to q\bar{q}b\bar{b}$ $\to \nu\nu b\bar{b}$ $\to \ell\ell b\bar{b}$ | Powheg-Box v2 | NNPDF3.0NLO | Pythia 8.240 Pythia 8.212 | AZNLO | LO + NLL(QCD) |
| $gg \to ttH$ $tt \to$ all $H \to$ all | Powheg-Box v2 [45] | NNPDF3.0NLO | Pythia 8.230 | AZNLO | NLO(QCD) + NLO(EW)$^{(\circ)}$ |
| Vector boson + jets | | | | | |
| $W \to qq$ $Z \to qq$ | Sherpa 2.2.8 [60, 65, 66] | NNPDF3.0NNLO | Sherpa 2.2.8 [67, 68] | Default | NNLO(QCD) $^{(\dagger)}$ [55, 56, 69] approx NLO(EW) [70, 71] |
| Top quark, mass set to 172.5 GeV | | | | | |
| $t\bar{t} \to$ all | Powheg-Box v2 [40–42, 72] | NNPDF3.0NLO | Pythia 8.230 | A14 [73] | NNLO+NNLL [74] |
| $tW$ | Powheg-Box v2 [40–42, 75] | NNPDF3.0NLO | Pythia 8.230 | A14 | NLO |
| $t$ t-channel | Powheg-Box v2 [40–42, 76] | NNPDF3.0NLO | Pythia 8.230 | A14 | NLO |
| $t$ s-channel | Powheg-Box v2 [40–42, 77] | NNPDF3.0NLO | Pythia 8.230 | A14 | NLO |
| Multijet | | | | | |
| Dijets | Pythia 8.230 | NNPDF2.3LO [78] | Pythia 8.230 | A14 | LO |

$V$+jets production with hadronic boson decays is simulated with Sherpa to NLO QCD accuracy for 1 additional parton and LO QCD accuracy for up to 4 additional partons. Approximate NLO EW corrections are applied as a function of the generated vector boson momentum $p_{\mathrm{T}}^{V}$. They have a sizeable impact on the differential production cross section decreasing the predicted yield by ~10% at $p_{\mathrm{T}}^{V}$ of 500 GeV and ~20% above 1 TeV. Calculations of next-to-next-to-leading-order (NNLO) QCD corrections to $V$+jets productions are available [54]. The NNLOJET group performed the calculation for $\sqrt{s}$ = 8 TeV [55, 56] and has provided custom corrections for the analysis fiducial region for $\sqrt{s}$ = 13 TeV as a function $p_{\mathrm{T}}^{V}$. They vary from 1.013 to 1.081 and are applied as a multiplicative factor on top of the NLO EW corrections.

The production of top-quark pairs, associated production of top quarks with $W$ bosons ($tW$), and single-top $t$- and $s$-channel production are modeled using the Powheg-Box v2 [40–42, 72, 75–77] generator at NLO in QCD. The diagram removal scheme [79] is used in $tW$ events to account for interference and overlap with $t\bar{t}$ production.

The jet mass distribution of non-resonant multijet events is modeled with an analytic function. Simulated events used to study the multijet model are generated using Pythia 8.230 [63] with leading-order matrix elements for dijet production and interfaced to a $p_{\mathrm{T}}$ ordered parton shower.

All simulated collision particles are processed with the ATLAS detector simulation [80] based on GEANT 4 [81]. Pile-up, multiple interactions in the same and neighboring bunch crossings, is modeled by overlaying simulated inelastic $pp$ events generated with Pythia 8.186 [63] using the NNPDF2.3LO set of parton distribution functions (PDFs) [78] and the A3 tune [82] over the original hard-scattering event. For Higgs boson and top-quark production, the EvtGen v1.2.0 program [83] models the decays of bottom and charm hadrons.

For each sample, Table 1 summarizes the MC generators, parton distribution functions, and underlying event tunes used as well as the order of perturbative QCD computations and EW corrections obtained for the cross section. For additional information, see Ref. [84] for $V$+jets events, Refs. [85–87] for $t\bar{t}$ events, and Ref. [88] for multijet events. Systematic uncertainties for process modeling are described in Section 7.

# 4 Object selection

For large Higgs boson Lorentz boost, the event topology of $pp \to H(\to b\bar{b}) + j$ is characterized by two jets, one of which contains the decay products of the two $b$-hadrons.

## 4.1 Object Reconstruction

Charged particles reconstructed as tracks [89] in the inner detector form interaction vertices [90]. The primary vertex of the hard interaction is defined as the vertex with the highest sum of squared transverse momenta of associated tracks.

The anti-$k_t$ algorithm implemented in FastJet [91] is used to cluster large-radius ($R = 1.0$) jets from noise-suppressed topological energy depositions calibrated to the local hadronic scale [92]. Jet cleaning criteria are used to identify jets arising from non-collision backgrounds or noise in the calorimeters [93] and events containing such jets are removed. Trimming jets reduces pile-up dependence and improves mass resolution [94]; subjets ($R = 0.2$) with $p_{\mathrm{T}}^{i}/p_{\mathrm{T}}^{\mathrm{jet}} < f_{\mathrm{cut}}$ are removed, where $p_{\mathrm{T}}^{i}$ is the transverse momentum of the $i^{\mathrm{th}}$ subjet, and $f_{\mathrm{cut}} = 0.05$. A weighted combination of the jet mass obtained from the calorimeter measurements with that from the charged component within the inner detector defines the jet mass $m_J$ for a trimmed jet [95]. For jets within $|\eta| < 2$, a series of simulation-based corrections are applied to calibrate $p_{\mathrm{T}}$ and $m_J$, while $p_{\mathrm{T}}$ is further corrected based on in situ techniques [95].

The variable radius (VR) jet algorithm [96] forms track-jets from tracks compatible with the primary vertex [97]. They are associated with large-$R$ jets before trimming using ghost association [98, 99]. Simulated track-jets are labeled as $b$-, $c$- or light-flavor according to which hadrons with $p_{\mathrm{T}} > 5$ GeV are found within $\Delta R = 0.3$ of the jet axis [100].

A multivariate discriminant (MV2) [100] is used to tag track-jets containing a $b$-hadron decay ($b$-tagged). The selection is tuned to produce an average efficiency of 77% for $b$-jets in simulated $t\bar{t}$ events, which corresponds to a light-flavor jet ($u$-, $d$-, $s$-quark, and gluon) and $c$-jet misidentification efficiency of

0.9% and 25%, respectively. Both in data and simulated events, jets with overlapping track-jets[2] are not considered for $b$-tagging as flavor-labeling is unreliable [101].

Muons are required to have $|\eta| < 2.4$, $p_{\mathrm{T}}^{\mu} > 10$ GeV, and small impact parameters, as well as satisfy the 'medium' quality criterion [102]. Isolated muons additionally must satisfy loose track- and calorimeter-based isolation conditions.

### 4.2 Analysis Object Definitions

Reconstructed jets possessing properties compatible with an $H \to b\bar{b}$ decay are labeled *candidate jets*. The reconstructed jet containing the Higgs boson decay products, or *H-jet*, is not always the highest $p_{\mathrm{T}}$ jet in the event. Undetected neutrinos from semi-leptonic $b$-hadron decays motivate considering the subleading jet. In around 50% and 47% of simulated ggF events, the $H$-jet is the leading and subleading $p_{\mathrm{T}}$ jet, respectively. Therefore, candidate jets are either of the two leading $p_{\mathrm{T}}$ jets with $|\eta| < 2$, $p_{\mathrm{T}} > 250$ GeV, $m_J > 60$ GeV, and $2m_J/p_{\mathrm{T}} < 1$. Furthermore, they must contain at least two track-jets.

A candidate jet is *double-tagged* if its two leading track-jets are $b$-tagged and *anti-tagged* if neither are $b$-tagged.

A 'muon-in-jet' correction is applied to candidate jets to account for the presence of semi-leptonic $b$-hadron decays. It utilizes the leading $p_{\mathrm{T}}$ muon found within $\Delta R < \min(0.4, 0.04 + 10/p_{\mathrm{T}}^{\mu})$ of each $b$-tagged track-jet. The scheme adds the muon four-momentum to the trimmed jet and removes the energy deposited by the muon in the calorimeter. After correcting 13% of leading and 33% of subleading $H$-jets in simulated ggF events, a 5% and 12% reduction in the $m_J$ width, respectively, is observed. Henceforth, $p_{\mathrm{T}}$ and $m_J$ refer to the corrected jet transverse momentum and mass, respectively, and $p_{\mathrm{T}}^{\mathrm{u}}$ and $m_J^{\mathrm{u}}$ represent the uncorrected versions.

### 4.3 Experimental Systematic Uncertainties

The major experimental uncertainties originate from the jet mass resolution (JMR) modeling and jet mass-scale (JMS) calibration. Uncertainties on $b$-tagging efficiency scaling factors and the jet energy scale play a minor role. The remaining uncertainties, including those arising from muon trigger, reconstruction, identification, and isolation rate modeling [102], are negligible.

A ratio of calorimeter-based to track-based measurements in dijet data and simulation defines the uncertainties in the jet energy and mass scales [95]. Jet energy scale and mass scale uncertainties are divided into 23 and 6 separate components, respectively, to account for different sources of uncertainty. The JMS agreement between data and simulation, while within the systematic uncertainties, displays a process and jet $p_{\mathrm{T}}$ dependence. Therefore, JMS uncertainties for $t\bar{t}$ events remain separated from those for $V$+jets and $H$ events within the global likelihood discussed in Section 8. The dominant component in terms of reconstructed mass scale is further separated to act independently on all processes ($t\bar{t}$, $V$+jets, and $H$) and in all analysis $p_{\mathrm{T}}$ bins. Jet observables in the simulation are smeared to assess the impact of energy scale and mass resolution uncertainties. Consistent with previous studies for trimmed jets [103, 104], the energy resolution has an absolute 2% uncertainty, while the mass resolution has a relative 20% uncertainty. JMR uncertainties act independently on each process ($t\bar{t}$, $V$+jets, and $H$) and in each analysis

[2] Track-jets "overlap" if the $\Delta R$ between them is less than the smaller of the two variable radii.

jet $p_{\mathrm{T}}$ bin to account for generator, process, and $p_{\mathrm{T}}$ dependence. The $V$+jets JMR uncertainty is reduced using independent measurements as described in Section 7.2.

The impact of uncertainties on $b$-tagging rates for $b$-, $c$-, and light-flavor jets are determined separately in various kinematic regions [100, 105, 106]. 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 correction factor to jets with $p_{\mathrm{T}}$ beyond the calibration dataset kinematic reach [107]. The thresholds are 250 GeV, 140 GeV, and 300 GeV for $b$-, $c$- and light-flavor track-jets, respectively.

# 5 Event selection and categorization

Events are classified into three orthogonal regions, a signal region (SR), a $t\bar{t}$ control region ($CR_{t\bar{t}}$), or a validation region (VR) used to study the multijet and $V$+jets background models. In all instances, data are divided into kinematical regions depending on jet $p_{\mathrm{T}}$.

## 5.1 Signal and Validation Regions

A uniform requirement for both the VR and SR in all data taking years of at least one jet with $p_{\mathrm{T}}^{\mathrm{u}} > 450$ GeV and $m_J^{\mathrm{u}} > 60$ GeV removes the kinematic regime biased by the trigger requirements. A second jet with $p_{\mathrm{T}}^{\mathrm{u}} > 200$ GeV is required to ensure a dijet topology. At least one of the two leading jets must satisfy the candidate jet criteria. To categorize events, first, the leading jet is considered. If it is a double-tagged candidate jet, the event populates the leading-jet signal region (SRL). Failed events occupy the subleading-jet signal region (SRS) if the subleading jet is a double-tagged candidate jet. Approximately 40% of the $H$ events surviving the kinematic cuts pass the $b$-tagging requirement. In simulated multijet events satisfying the SR requirements, roughly 70% of the multijet background reconstructed near $m_H$ contains two $b$-hadrons and less than 5% of candidate jets are void of heavy-flavor hadrons.

The SRS has a sensitivity approximately 50% lower than that of the SRL. Jets in which a neutrino from a $b$-hadron decay carries a larger fraction of $p_{\mathrm{T}}$ enter this category. In those cases the mass resolution is degraded since the neutrino is not measured.

The validation region always considers both leading jets. The leading-jet validation region (VRL) includes events where the leading jet is an anti-tagged 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).

The analysis is performed with an inclusive signal region containing jets with $p_{\mathrm{T}} > 250$ GeV, two fiducial signal regions and three differential signal regions. Lower $p_{\mathrm{T}}$ requirements on the selected candidate jet at 450 GeV and 1 TeV define two fiducial signal regions. Three differential signal regions are defined by requiring the candidate jet $p_{\mathrm{T}}$ to be between 250–450 GeV, 450–650 GeV, and 650–1000 GeV. Only the subleading-jet SR populates the lowest $p_{\mathrm{T}}$ region and only the leading-jet SR is used for the highest $p_{\mathrm{T}}$ region, $p_{\mathrm{T}} > 1$ TeV. The VRs follow the same definition. Events with candidate jet $p_{\mathrm{T}} > 3$ TeV are discarded. This cleaning cut has no impact on the reported results.

Within each region, the discriminating variable is $m_J$. The combination of the $m_J^{\mathrm{u}} > 60$ GeV selection and the muon and neutrino from semi-leptonic $b$-hadron decays causes a reduced acceptance for $m_J$ values

Table 2: A summary of the signal regions and the corresponding measurement fiducial volumes. Left of the break, the jet $p_{\mathrm{T}}$ requirements are given for the four analysis configurations. Each SR has an associated $CR_{t\bar{t}}$ and VR in the same kinematic region. Right of the break, the $p_{\mathrm{T}}^{H}$ and $\eta_H$ requirements used to determine which signal events are counted in the measurement within the fiducial and differential signal regions are given.

| Region | Jet $p_{\mathrm{T}}$ [GeV] | | $p_{\mathrm{T}}^{H}$ [GeV] | | $\lvert\eta_H\rvert$ |
|---|---|---|---|---|---|
| | SRL | SRS | SRL | SRS | SRL/SRS |
| Inclusive | >450 | >250 | – | – | – |
| Fiducial | >450 | >450 | >450 | >450 | < 2 |
| | >1000 | – | >1000 | – | < 2 |
| Differential | 450–650, 650–1000 | 250–450, 450–650, 650–1000 | 450–650, >650 | 300–450, 450–650, >650 | < 2 |

near the selection threshold. Since the $b$-tagging efficiency diminishes with decreasing angular distance to another hadronic object [108], the low $m_J$ region in the signal region is further sculpted. The prevalence of each effect determines the lowest $m_J$ values studied. The SR mass range examined in each region is 70 GeV or 75 GeV depending on the jet being leading or subleading and the $p_{\mathrm{T}}$ range, to 210 GeV. The VRs are studied using the same $p_{\mathrm{T}}$ bins and $m_J$ range as the SRs.

For the measurements within the fiducial and differential regions, the signal events within the fiducial volume(s) defined by requirements on the generator 'truth' record, the Higgs boson pseudorapidity ($\eta_H$) and transverse momentum ($p_{\mathrm{T}}^{H}$), are considered. Further details are given in Section 9.3. Table 2 summarizes the analysis signal regions and fiducial volumes.

## 5.2 $t\bar{t}$ 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 system of a top and an anti-top quark, one decaying leptonically and the other hadronically, in opposite detector hemispheres. An isolated muon, with $p_{\mathrm{T}}^{\mu} > 52.5$ 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 $p_{\mathrm{T}} > 250$ 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 resemble the kinematics of top-quarks reconstructed in the SR, the $CR_{t\bar{t}}$ only contains events with 140 GeV $< m_{J_t} <$ 200 GeV, removing those where $J_t$ does not contain all the top-quark decay products. Residual differences in the $CR_{t\bar{t}}$ $J_t$ and in the SRS candidate jet $p_{\mathrm{T}}$ spectrum below 450 GeV originate from the different trigger requirements. The ratio of the jet $p_{\mathrm{T}}$ spectrum in simulated $t\bar{t}$ events in $CR_{t\bar{t}}$ over the SRS reproduces the difference. It is used to reweight the simulated $t\bar{t}$ event weights and remove data events in the $CR_{t\bar{t}}$ with $p_{\mathrm{T}}^{J_t} <$ 450 GeV.

Table 3: A summary of the $CR_{t\bar{t}}$ selection criteria.

| Jet | $N$ track-jets | $N$ $b$-tags | Angular Selection | Jet Mass [GeV] |
|---|---|---|---|---|
| $J_b$ | $\geq 1$ | 1 | $0.04 + 10/p_{\mathrm{T}}^{\mu} < \Delta R(\mu, J^b) < 1.5$ | – |
| $J_t$ | $\geq 3$ | 1 | $\Delta\phi(J^b, J^t) > \frac{2\pi}{3}$ | 140–200 |

The same $p_{\mathrm{T}}$ boundaries used in the SR are also applied to $p_{\mathrm{T}}^{J_t}$ to define the inclusive and differential $CR_{t\bar{t}}$. The selection criteria, summarized in Table 3, achieve over 95% purity in $t\bar{t}$ events for each $p_{\mathrm{T}}$-binned $CR_{t\bar{t}}$.

# 6 Higgs boson modeling

The limited number of event selection criteria pertaining to properties of the recoil system or other activity in the event results in an inclusive analysis in terms of the Higgs boson production modes. Table 4 shows the relative contribution of the four main production modes as a function of Higgs boson candidate $p_{\mathrm{T}}$, according to SM predictions. Only events within the Higgs boson window, defined by $105 < m_J < 140$ GeV, are considered. In both the SRL and SRS, ggF production contributes most for $p_{\mathrm{T}}^{\mathrm{jet}} > 450$ GeV. For $p_{\mathrm{T}}^{\mathrm{jet}} < 450$ GeV, $t\bar{t}H$ comprises around 40% of the Higgs boson events. A hadronically decaying top-quark can satisfy the jet trigger requirements without a high $p_{\mathrm{T}}^{H}$ value, thus resulting in a significant contribution of $t\bar{t}H$ events with relatively low Higgs boson $p_{\mathrm{T}}$. Almost 90% of $t\bar{t}H$ events in the Higgs boson window arise from $H \to b\bar{b}$ decays. $H \to W^{\pm}W^{\mp}$ provides a majority of the remainder and climbs to almost 15% for larger $m_J$ values.

The acceptance uncertainty on ggF-produced events is 20%. It includes variations of the factorization and

Table 4: The fractional contribution of each production mode to a given analysis region around the Higgs boson peak, defined by $105 < m_J < 140$ GeV. The fraction is given with respect to the total yield in the analysis region in question.

| Process | Jet $p_{\mathrm{T}}$ Range [GeV] | | | |
|---|---|---|---|---|
| | 250–450 | 450–650 | 650–1000 | > 1000 |
| | | SRL | | |
| ggF | – | 0.56 | 0.50 | 0.39 |
| VBF | – | 0.17 | 0.16 | 0.17 |
| $VH$ | – | 0.14 | 0.18 | 0.25 |
| $t\bar{t}H$ | – | 0.13 | 0.16 | 0.19 |
| | | SRS | | |
| ggF | 0.28 | 0.46 | 0.43 | – |
| VBF | 0.07 | 0.19 | 0.21 | – |
| $VH$ | 0.26 | 0.24 | 0.26 | – |
| $t\bar{t}H$ | 0.39 | 0.11 | 0.10 | – |

renormalization scales, PDF, and parton shower model. Ref. [109] demonstrates that the NLO correction is nearly equivalent in the infinite top-mass approximation and full SM calculation; no additional systematic uncertainty is assigned. Acceptance uncertainties on the VBF, $VH$, and $t\bar{t}H$ processes are 0.5%, 5%, and 13%, respectively [13]. Systematic uncertainties on the EW corrections (expressed as $1 + \delta_{\text{EW}}$) are taken as $\delta^2_{\text{EW}}$ following the recommendations in Ref. [5].

# 7 Background process modeling

Multijet production is the dominant background proces. $V$+jets and top-quark resonance peaks flank the Higgs boson signal in low- and high-mass side-bands, respectively, but also leak into the Higgs boson signal window. Within 105 GeV $< m_J <$ 140 GeV, $V$+jets is 1.5% of the total background, top-quarks account for 2%, and multijets provide the rest. The expected signal contribution equates to 0.2% of the background and 20-80% of the data statistical precision.

An accurate and precise determination of the background is paramount and is achieved starting with the determination of the $V$+jets and Top backgrounds.

## 7.1 Top-quark pair production

The $CR_{t\bar{t}}$ design ensures the same underlying physics processes responsible for the SRs' events also populate the control region. The simulated $J_t$ mass distribution in the CR and SR are similar in shape and peak near the top-quark mass because both regions probe a similar top-quark momentum range. Therefore, any adjustment of the simulated top-quark events made to improve the 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 Section 8. The inclusive $CR_{t\bar{t}}$ has a $t\bar{t}$ purity of 97% with similar levels found in the fiducial and differential regions. With such high purity, the $t\bar{t}$ normalization is determined directly from data with better than 10% precision in most regions.

Comparisons between nominal and alternative simulated samples provide modeling systematic uncertainty estimates for the parton shower model (Herwig 7 replaces Pythia 8) and the matrix element calculation (MadGraph5_aMC@NLO replaces Powheg-Box 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 Section 4.3 are utilized. Uncertainties on b-tagging efficiency of $b$-jets and JMS have the largest impact on the $t\bar{t}$ normalization.

Figure 2 shows the $J_t$ mass distribution for each analysis $p_{\text{T}}$ bin after the global likelihood fit in the differential configuration. The simulation agrees well with the data.

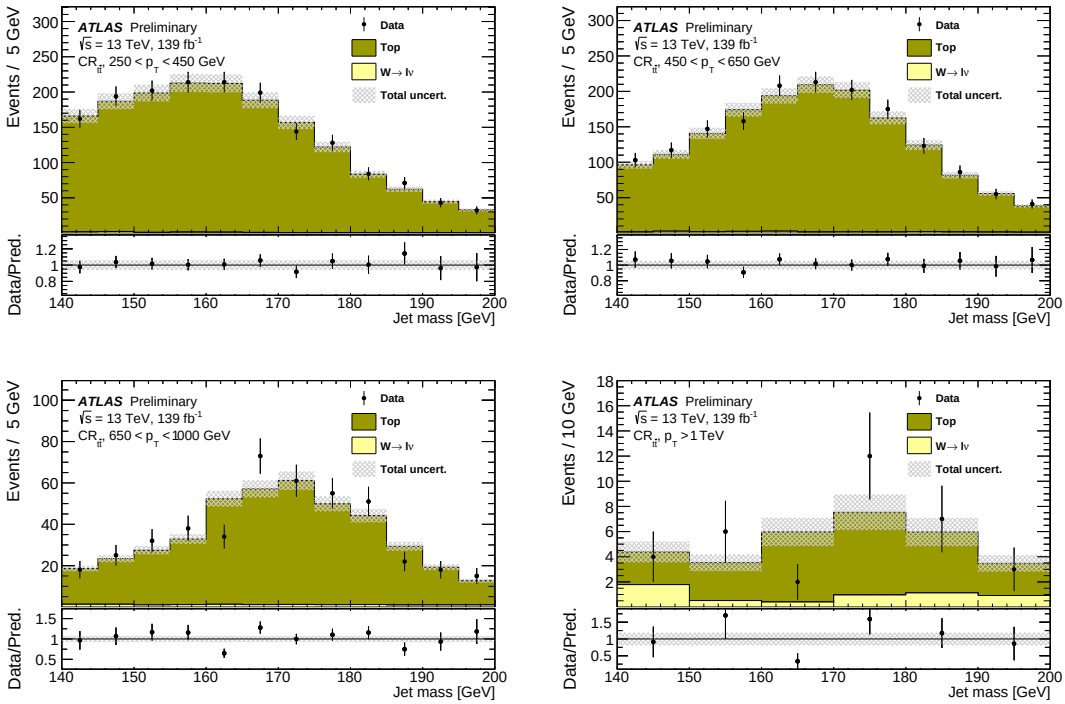



Figure 2: The post-fit $CR_{t\bar{t}}$ $J_t$ mass distribution in the different $p_{\mathrm{T}}$ regions.

## 7.2 $V$+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 experimental performance. $Z$ events outnumber $H$ events by over a factor of 20. Experimental effects 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.

In the VR, $W$ events outnumber $Z$ events nearly three to one due to the larger cross section and comparable acceptance. In the SR, the $Z$ events outnumber $W$ events over three to one due to the sizeable $Z \to b\bar{b}$ branching fraction and flavor tagging requirements. The decay products of the vector boson are reconstructed within the selected candidate jet in approximately 90% and 40% of the $Z$ and $W$ events in the SR. Within the VR, over 60% of candidate jets contain the decay products of a vector boson. In the remainder, the jet is created by the recoil activity resulting in a non-resonant mass similar to the multijet background's shape, which enhances the high mass tail.

As the data directly determines the $Z$+jets normalization, the impact of modeling systematic uncertainties is limited to changes in acceptance. The $W$+jets cross section carries a 10% uncertainty [110]. The maximum of seven independent pairs of renormalization and factorization scale variations by factors of 0.5 and 2 corresponds to a 3–20% error on the expected acceptance. An alternate PDF set (`MMHT2014nlo`), $\alpha_s$ variations within the nominal PDF set, and changing the cluster fragmentation model to the Lund string model [111] did not lead to a significant differences in the acceptance estimate with respect to the nominal model.

All experimental uncertainties described in Section 4.3 are applied. Uncertainties on JMR and JMS have the largest impact on the $V$+jets normalization. Using the multijet model described in Section 7.3 and the likelihood described in Section 8, the leading jet invariant mass distribution in the validation region is described to the level of agreement between simulation and data shown in Figure 3.

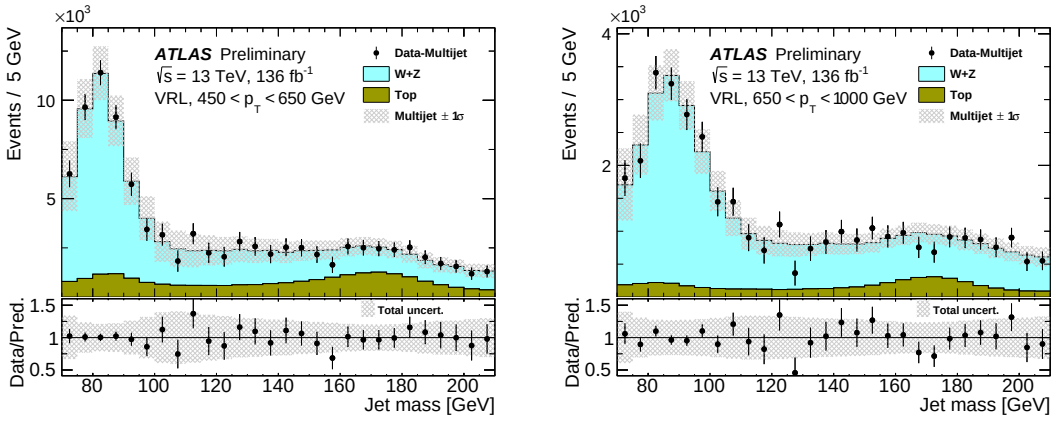



Figure 3: Post-fit leading jet invariant mass distributions after the multijet background subtraction in the validation region for data (points with error bars) and the $V$+jets ($W$+$Z$) and Top components (histograms) for $450 < p_{\mathrm{T}} < 650$ GeV (left) and $650 < p_{\mathrm{T}} < 1000$ GeV (right). The $t\bar{t}$ normalization and its uncertainty are set to the corresponding values from the $CR_{t\bar{t}}$ with the appropriate uncertainty.

### Jet Mass Resolution Uncertainty

The fitted $Z$ normalization in the SR shows a significant correlation with the reconstructed $m_J$ resolution uncertainty because the interaction of the $Z$+jets component in the mass spectrum and the multijet model flexibility can open a local minimum. Tests using subsets of the hybrid validation region, constructed to have a known amount of each process and discussed in the next section, highlight this feature. In some instances, the JMR parameter broadens the $Z$+jets peak. It corresponds to a $Z$+jets normalization increase and a multijet contribution decrease from the expected values.

To stabilize the fit response, the $Z$ and $W$ resonance jet mass widths are measured directly on two data samples and added as an additional constraint in the global likelihood. The two data samples are an alternate $t\bar{t}$ CR ($WCR_{t\bar{t}}$) and the VRL. The $WCR_{t\bar{t}}$ consists of selected semi-leptonic $t\bar{t}$ events having a resolved $Wb$ pair from a hadronically decaying top-quark providing a high-purity reconstructed $W$ peak with jet $p_{\mathrm{T}}$ from 200 GeV up to about 600 GeV. The VRL provides a clear $W$ peak and covers the entire jet $p_{\mathrm{T}}$ range of this analysis but has considerably larger multijet contribution.

The measured jet mass width of the $W$ and $Z$ resonances shows a smooth evolution from low $p_{\mathrm{T}}$ in the $WCR_{t\bar{t}}$ to high $p_{\mathrm{T}}$ of the VRL (see Figure 4). These results differ from the nominal simulated $m_J$ resolution by less than 2.5% and have a precision that is 14–22% of the original JMR uncertainty after systematic uncertainties are incorporated to transfer the result to the $Z \to b\bar{b}$-dominated $V$+jets sample in the SR. When included in the global likelihood, the $Z$+jets normalization and the JMR parameter correlation is reduced. For example, in the inclusive signal region it decreases from ~90% to ~30%.

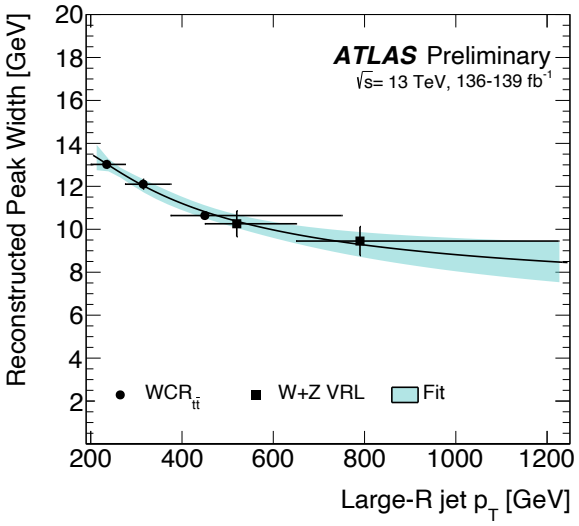



Figure 4: A summary of the $Z$ and $W$ resonance peak reconstructed width measurements as a function of the jet $p_{\mathrm{T}}$ using the resolved $W$ boson in top-quark decays in the $WCR_{t\bar{t}}$ region and the combined $W$ and $Z$ boson mass in the validation region. The horizontal bars cover the jet $p_{\mathrm{T}}$ range used for the corresponding point and the points are centered at the average jet $p_{\mathrm{T}}$ in the same range. The continuous black curve is a fit to the measurements with resultant errors shown as a cyan band.

### 7.3 Multijet production

An exponential of polynomial functions is used to model the multijet contribution

$$f_N\left(x\left|\vec{\theta}\right.\right) = \theta_0 \exp\left(\sum_{i=1}^{N} \theta_i x^i\right), \tag{1}$$

where $x = (m_J - 140\,\mathrm{GeV})/70\,\mathrm{GeV}$ and $\theta_i$ are the parameters of the fit. Parameter values are independently determined in each region simultaneously with the signal extraction. The optimal degree of the polynomial function, $N$ in Equation 1, depends on the mass shape and number of events analyzed. The VRS (VRL) contains 51 (58) times the amount of SRS (SRL) data. Therefore, 'ensemble tests' using modified VR subsets with roughly the same number of events as the corresponding SR, referred to as hybrid VR *slices* ($\mathrm{VR_{hyb}}$), are used to determine $N$.

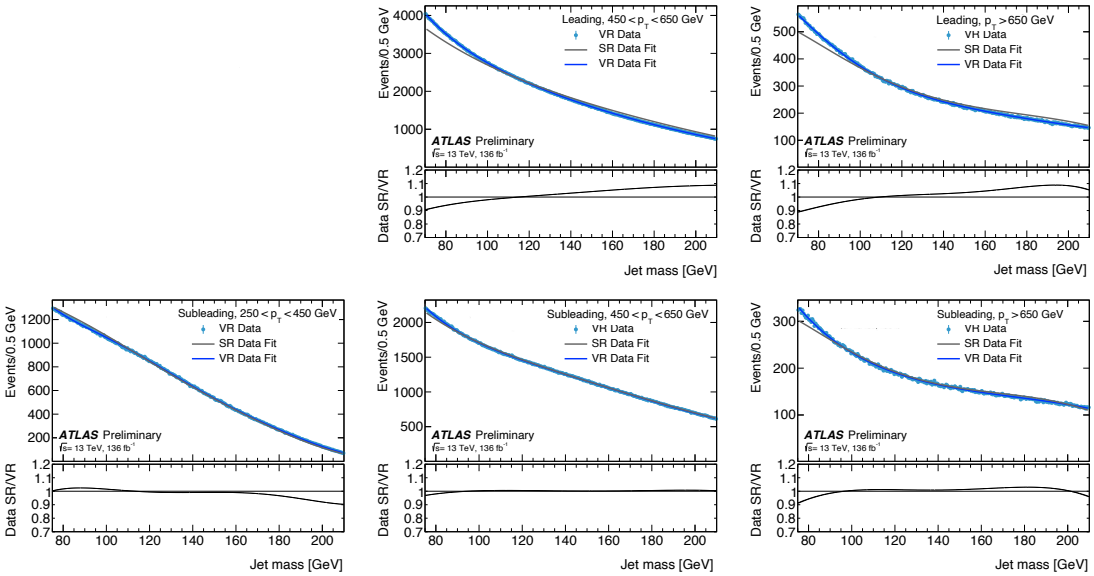
The hybrid VR is the best available proxy for the SR. It is defined by replacing the VR resonance peaks with the SM prediction from the SR and correcting the underlying multijet shape for SR acceptance effects. The following procedure for determining the resonance and multijet shape in the VR and SR data is used. The VR multijet estimate ($\mathrm{MJ_{VR}}$) is created from the average of the multijet model parameter values obtained from likelihood fits to ten[3] random, orthogonal subsets of the VR including all experimental and modeling systematic uncertainties. The input $t\bar{t}$ normalization and associated uncertainty are set to representative values from the $CR_{t\bar{t}}$. The VR $V$+jets and Top estimates ($V_{\mathrm{VR}}$ and $\mathrm{Top_{VR}}$) are created from the average post-fit $V$+jets and Top contributions from the same ten fits to the VR. The SR multijet estimate ($\mathrm{MJ_{SR}}$) is obtained from a global likelihood fit to the SR. The ratio of $\mathrm{MJ_{SR}}$ over $\mathrm{MJ_{VR}}$ defines the acceptance

[3] Ten subsets balance the statistical precision with the need to use greater values of $N$ to model a larger dataset.

differences between the two regions as is shown in Figure 5. Each $\mathrm{VR}^i_{\mathrm{hyb}}$ slice is:

$$\mathrm{VR}^i_{\mathrm{hyb}} = (\mathrm{VR}^i - V_{\mathrm{VR}} - \mathrm{Top}_{\mathrm{VR}}) \times \left(\frac{\mathrm{MJ}_{\mathrm{SR}}}{\mathrm{MJ}_{\mathrm{VR}}}\right) + V_{\mathrm{SR}} + \mathrm{Top}_{\mathrm{SR}} + H_{\mathrm{SR}},$$

where $\mathrm{VR}^i$ is the full distribution in a VR slice and $V_{\mathrm{SR}}$, $\mathrm{Top}_{\mathrm{SR}}$, and $H_{\mathrm{SR}}$ are the nominal MC predictions for $V$+jets, Top, and $H$ production, respectively.




(a) $250 < p_{\mathrm{T}} < 450$ GeV (b) $450 < p_{\mathrm{T}} < 650$ GeV (c) $p_{\mathrm{T}} > 650$ GeV

Figure 5: Comparison of the multijet shapes from the different $p_{\mathrm{T}}$-binned analysis regions. The solid lines show the multijet function shape after a fit to the SR (gray) and VR (blue). The solid points are the data from VR slices with the same number of events as the SR. The $250 < p_{\mathrm{T}} < 450$ GeV region is only populated in the subleading jet SR.

A procedure to choose the optimal $N$ for each region utilizes three metrics evaluated with the $\mathrm{VR}_{\mathrm{hyb}}$ collection. First, a log-likelihood ratio (LLR) test compares the result of an $N$-parameter fit (null hypothesis) to an $N+1$-parameter fit (alternative hypothesis) in each $\mathrm{VR}_{\mathrm{hyb}}$ slice without any injected SM resonances. By Wilks' theorem, the likelihood ratio follows an asymptotic $\chi^2$ distribution 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 of a free normalization parameter and its associated error (generalized as $\mu^{\mathrm{VR}} \pm \sigma^{\mathrm{VR}}_{\mathrm{stat}}$) on either the $Z$+jets process or the Higgs boson process. Both utilize $\mathrm{VR}_{\mathrm{hyb}}$ slices with all SR resonances injected at the SM rates. $\mathrm{F}_{2\sigma}$, the fraction of results where $|\mu^{\mathrm{VR}} - 1|$ is beyond twice its error $\sigma^{\mathrm{VR}}_{\mathrm{stat}}$, estimates the probability that the multijet model allows a substantial artificial excess or deficit. A $2\sigma$ threshold ensures some results from the full set of VR slices cross the boundary. $\overline{\mu/\sigma}$ is the average value of $(\mu^{\mathrm{VR}} - 1)/\sigma^{\mathrm{VR}}_{\mathrm{stat}}$ and indicates a bias in determining the signal strength. $N$ chosen by the LLR test is incremented until $\mathrm{F}_{2\sigma}$ is compatible with 0.05, and $\overline{\mu/\sigma}$ stabilizes for both $Z$+jets and $H$ production. The resultant $N$ agrees when including systematic uncertainties and when not injecting the $Z$+jets and $H$ processes into the $\mathrm{VR}_{\mathrm{hyb}}$ slices, with

appropriate changes to the definition of $F_{2\sigma}$ and $\overline{\mu/\sigma}$. A fifth degree polynomial function is used in both inclusive regions, while either a fourth or fifth degree polynomial is used in the analysis $p_T$ bins.

$\overline{\mu/\sigma}$ for the $N$ chosen in each region indicates a bias in the background model. It defines the spurious signal systematic uncertainty and ranges from 0.01–0.33 for $H$ and 0.15–0.65 for $Z$ + jets production. In both cases, it has an insignificant impact on the sensitivity.

An alternative function, the Laurent series, 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.

# 8 Statistical analysis

Signal yields are extracted by minimizing the negative logarithm of a likelihood function $\mathcal{L}(\mu, \vec{\theta})$ with the RooStats framework [112]. The likelihood function is defined as the product of Poisson probability terms, one for each bin of the $m_J$ distribution. Bin widths are set to 5 GeV, necessitating technical advancements to fit an analytic function to a wide-binned dataset [113, 114]. Systematic uncertainties enter the likelihood as nuisance parameters, $\vec{\theta}$, constrained with Gaussian or log-normal probability density function (PDF) priors. The JMR constraints obtained from the $WCR_{t\bar{t}}$ and VRL regions are included as Gaussian PDF priors. Unconstrained, or *free*, parameters control the normalization of the MC templates within each jet $p_T$ region or within a given fiducial volume and are common to the SRL, SRS, and $CR_{t\bar{t}}$. For the multijet model, the function normalization and its polynomial coefficients are free parameters and independent for each region. 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 production cross section are derived using the $CL_s$ method [115, 116].

Table 5 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 technique [117]. Systematic variations yielding large statistical fluctuations are smoothed using custom algorithms which also remove variations resulting in effects below 2% within each region.

Table 5: 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, "$p_{\mathrm{T}}$ bins" indicates a decorrelation between the analysis $p_{\mathrm{T}}$ bins, and "LS" means it is decorrelated between the SRL and SRS. For the inclusive analysis, "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 $V$+jets events. ($\bullet$) The spurious signal uncertainty is only applied to $Z$+jets when the procedure to extract signal strengths in fiducial volumes is tested using $Z$+jets events in the SR.

| Description | Processes | Category | Effect |
|---|---|---|---|
| Experimental Systematic Uncertainties | | | |
| JMR | $t\bar{t}$, $V$+jets, $H$ | $p_{\mathrm{T}}$ bins | N+S |
| JMS (dominant) | $t\bar{t}$, $V$+jets, $H$ | $p_{\mathrm{T}}$ bins | N+S |
| JMS (rest) | $t\bar{t}$, $V$+jets $+H$ | all | N+S |
| Jet energy scale | all$^{(*)}$ | all | N+S |
| Jet energy resolution | all | all | N+S |
| $b$-tag efficiency $b$-jets | all | all | N+S |
| $b$-tag efficiency $c$-jets | all | all | N+S |
| $b$-tag efficiency $l$-jets | all | all | N+S |
| Modeling Systematic Uncertainties | | | |
| Cross-section and acceptance | $W$+jets | all | N |
| Renormalization and factorization scale | $V$+jets | all | N+S |
| Parton shower model | $t\bar{t}$ | all | N+S |
| Matrix element calculation | $t\bar{t}$ | all | N+S |
| Initial-/Final-state radiation | $t\bar{t}$ | all | N+S |
| Acceptance | $H$ | all | N |
| NLO EW corrections | VBF+$VH$+$t\bar{t}H$ | all | N |
| | $V$+jets | all | N |
| Spurious signal | $H$ | $p_{\mathrm{T}}^{H}$ bins $\times$ LS | N |
| | $Z$+jets $^{(\bullet)}$ | $p_{\mathrm{T}}^{Z}$ bins $\times$ LS | N |

# 9 Results

The analysis regions designed to probe Higgs boson production with considerable transverse momentum are summarized in Table 2. They provide one all-encompassing region to determine the $H$ signal strength, two regions for fiducial cross-section measurements, and three bins for a differential measurement. All $H$ production modes are considered for the signal strength extraction. The analysis jet $p_{\mathrm{T}}$ bins align well with the $p_{\mathrm{T}}^{H}$-defined fiducial volumes and the yield of signal events outside the targeted fiducial volume(s) are constrained to their SM prediction within the theoretical and experimental uncertainties. The cross sections are derived from the fitted signal yields divided by the integrated luminosity corrected by the product of the estimated selection efficiency and fiducial acceptance. Using the same category definitions, differential cross-section measurements of $V$+jets production in the VRL and $Z$+jets production in the SR validate

Table 6: Expected and observed values of the signal strengths for the $H$, $Z$ and $t\bar{t}$ components in the inclusive fit. The values of $\mu_{t\bar{t}}$ obtained are in agreement with unfolded measurements of $t\bar{t}$ events in a similar kinematic phase-space [118].

| Result | $\mu_H$ | $\mu_Z$ | $\mu_{t\bar{t}}$ |
|---|---|---|---|
| Expected | $1.0 \pm 3.0$ | $1.00 \pm 0.17$ | $1.00 \pm 0.07$ |
| Observed | $1.1 \pm 3.6$ | $1.25 \pm 0.22$ | $0.81 \pm 0.06$ |

Table 7: Event yields and associated uncertainties after the global likelihood fit in the inclusive region. The rate of the $W(\ell\nu)$ background is assumed to be fixed. The total background yield can differ from the sum of the individual components due to rounding.

| Process | SRL | | SRS | | $CR_{t\bar{t}}$ | |
|---|---|---|---|---|---|---|
| Multijet | 591 400 | ± 4200 | 530 000 | ± 3500 | – | |
| $Z$ + jets | 15 900 | ± 2700 | 11 800 | ± 2000 | – | |
| $W$+ jets | 3070 | ± 700 | 2520 | ± 500 | – | |
| Top | 15 700 | ± 1900 | 15 300 | ± 2000 | 3740 | ± 70 |
| $W(\ell\nu)$ | – | | – | | 50 | |
| $H$ | 500 | ± 1800 | 400 | ± 1400 | – | |
| Total | 626 530 | ± 870 | 560 090 | ± 810 | 3790 | ± 70 |
| Data | 626 532 | | 560 083 | | 3791 | |

the method.

### 9.1 Inclusive region

The inclusive region yields a Higgs boson signal strength for the combination of SRL, SRS and $CR_{t\bar{t}}$ of $\mu_H = 1.1 \pm 3.6$. The fit $\chi^2$ probability values are 0.19 and 0.77 for the SRL and SRS, respectively. Results are summarized in Table 6 and the yields in Table 7. The Higgs boson signal strength uncertainty is statistically dominated. The leading sources of systematic uncertainties are jet mass resolution and mass scale.

### 9.2 Fiducial $p_{\mathrm{T}}^H > 450$ GeV and $p_{\mathrm{T}}^H > 1$ TeV regions

The two fiducial regions determine the Higgs boson yield and cross section in the phase space defined by the Higgs boson pseudorapidity range $|\eta_H| < 2.0$ and transverse momentum $p_{\mathrm{T}}^H > 450$ GeV and $p_{\mathrm{T}}^H > 1$ TeV. Compared to the inclusive measurement discussed above, the fiducial region for $p_{\mathrm{T}}^H > 450$ GeV does not include the SRS region below 450 GeV. The $p_{\mathrm{T}}^H > 1$ TeV region probes a new domain of highly boosted $Z$ and Higgs boson reconstruction. Since the expected sensitivity in the SRS above 1 TeV is marginal as the muon-in-jet correction and $b$-tagging turn-on effects are more significant compared to the SRL, the

Table 8: Signal acceptance times efficiency for the signal regions in the fiducial measurements.

| Process | $p_{\mathrm{T}}^{H} > 450$ GeV | $p_{\mathrm{T}}^{H} > 1$ TeV |
|---|---|---|
| All | 0.25 | 0.18 |
| $ggF$ | 0.26 | 0.22 |
| $VH$ | 0.27 | 0.19 |
| $VBF$ | 0.22 | 0.15 |
| $ttH$ | 0.20 | 0.16 |

measurement above 1 TeV is based only on the SRL and $CR_{t\bar{t}}$ regions. The acceptance times efficiency values for the different SM Higgs boson production processes are given in Table 8.

Two Higgs boson mass templates are used in the fiducial fits. The first describes the mass distribution of events with signal jet and Higgs boson $p_{\mathrm{T}}$ above the cut of choice; the second those with signal jet above the $p_{\mathrm{T}}$ cut but Higgs boson $p_{\mathrm{T}}$ below it. The first component accounts for more than 80% (70%) of the Higgs boson signal selected by the 450 GeV (1 TeV) jet $p_{\mathrm{T}}$ cut and the yield is left free in the fit, while the second is constrained to the SM value within the theoretical and experimental uncertainties. The latter contribution also tends to have a broader mass spectrum shifted to higher values.

This procedure is first tested with $W \to qq'$ and $Z \to q\bar{q}$ in the VR and $Z \to b\bar{b}$ in the SR. For these tests, the $V$ and $Z$ mass templates are structured similarly to those of the Higgs boson decribed above. The Higgs boson yields are kept fixed to the SM expectations in the fit. In the VR, the fitted signal strengths for $V$+jets with $p_{\mathrm{T}}^{V} > 450$ GeV and 1 TeV are $1.01 \pm 0.09$ and $1.56 \pm 0.50$, respectively. In the SR, the signal strength for $Z$ events with $p_{\mathrm{T}}^{Z} > 450$ GeV and 1 TeV are $1.35 \pm 0.23$ and $1.8 \pm 1.4$, respectively. These results are in agreement with the SM.

When extracting the Higgs boson signal strength, the likelihood fit result for $p_{\mathrm{T}}^{H} > 450$ GeV provides a signal strength of $\mu_H = 0.7 \pm 3.3$ and for $p_{\mathrm{T}}^{H} > 1$ TeV gives $\mu_H = 26 \pm 31$. These yields correspond to Higgs boson production cross-section values in the fiducial region of

$$\begin{aligned} \sigma_H(p_{\mathrm{T}}^{H} > 450\ \mathrm{GeV}) &= 13 \pm 52\ (\mathrm{stat.}) \pm 32\ (\mathrm{syst.}) \pm 3\ (\mathrm{theory})\ \mathrm{fb}, \\ \sigma_H(p_{\mathrm{T}}^{H} > 1\ \mathrm{TeV}) &= 3.4 \pm 3.9\ (\mathrm{stat.}) \pm 1.0\ (\mathrm{syst.}) \pm 0.8\ (\mathrm{theory})\ \mathrm{fb}, \end{aligned}$$

and 95% confidence level (CL) upper limits of

$$\begin{aligned} \sigma_H(p_{\mathrm{T}}^{H} > 450\ \mathrm{GeV}) &< 144\ \mathrm{fb}, \\ \sigma_H(p_{\mathrm{T}}^{H} > 1\ \mathrm{TeV}) &< 10.3\ \mathrm{fb}. \end{aligned}$$

The post-fit SRL and SRS jet mass distributions are shown in Figure 6. Results are summarized in Table 9. The uncertainties are statistically dominated. The contributions of the main categories of systematic uncertainties are given in Table 10. The jet uncertainties give the largest contribution, driven by the JMS effects at lower $p_{\mathrm{T}}$ and by JMR above 1 TeV where the JMR measurement constraints are looser due to the extrapolation uncertainties (see Figure 4). Similarly, the flavor tagging uncertainties increase above 1 TeV due to the extrapolation from the $p_{\mathrm{T}}$ range of the calibration regions.

Table 9: Expected and observed values of the signal strengths for the $H$, $Z$, and $t\bar{t}$ components in the fiducial fits. The value for $\mu_H$ refers to the fiducial volume $|\eta_H| < 2.0$, $p_T^H > 450$ GeV and $p_T^H > 1$ TeV, while those for $\mu_Z$ and $\mu_{t\bar{t}}$ pertain to regions of jet $p_T$ above 450 GeV and 1 TeV. The values of $\mu_{t\bar{t}}$ obtained are in agreement with unfolded measurements of $t\bar{t}$ events in a similar kinematic phase-space [118].

| $p_T^H$/Jet $p_T$ | $\mu_H$ | | $\mu_Z$ | | $\mu_{t\bar{t}}$ | |
|---|---|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. | Exp. | Obs. |
| $> 450$ GeV | $1.0 \pm 3.3$ | $0.7 \pm 3.3$ | $1.00 \pm 0.18$ | $1.27 \pm 0.22$ | $1.00 \pm 0.07$ | $0.81 \pm 0.06$ |
| $> 1$ TeV | $1.0 \pm 29.0$ | $26 \pm 31$ | $1.0 \pm 1.6$ | $2.4 \pm 1.7$ | $1.0 \pm 0.3$ | $0.51 \pm 0.19$ |

Table 10: Contributions to the systematic uncertainties for the fiducial signal strength measurements. The total uncertainty is also given for comparison.

| Uncertainty Contribution | $p_T^H > 450$ GeV | $p_T^H > 1$ TeV |
|---|---|---|
| Total | 3.3 | 31 |
| Statistical | 2.8 | 30 |
| Jet Systematics | 1.2 | 7 |
| Modeling and Theory Systs. | 1.0 | 1 |
| Flavor Tagging Systs. | 0.5 | 3 |
| Total Systematics | 1.7 | 8 |

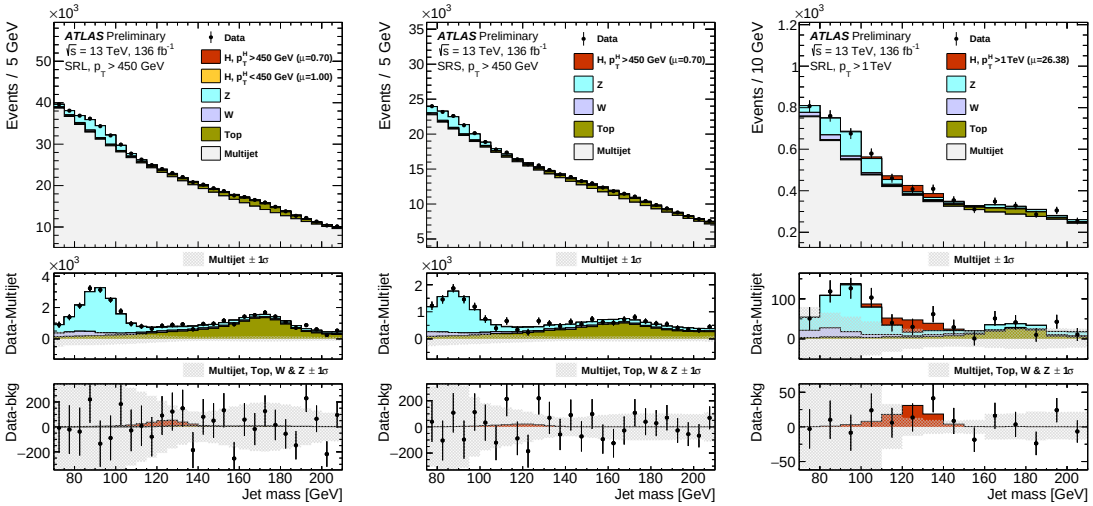



Figure 6: Post-fit signal jet mass distributions with the various component for the fiducial regions with $p_{\mathrm{T}}^{H} > 450$ GeV in the SRL (left) and SRS (middle) as well as with $p_{\mathrm{T}}^{H} > 1$ TeV in the SRL (right) shown in 10 GeV bins. The middle panels show the distributions after subtraction of the multijet distribution. The shaded areas indicate the 68% CL for the multijet background from the fitted parameters and normalizations of the exponential polynomials. The lower panels show the distributions after subtraction of the fitted background processes: multijet, $V$+jets and Top. The shaded areas indicate the 68% CL for all background processes. Contributions below 0.5 per mille of the total yield are not shown.

### 9.3 Differential regions

The differential regions aim to measure the differential spectrum of the Higgs boson transverse momentum, where possible deviations from the SM predictions could manifest with an amplitude increasing with $p_{\mathrm{T}}^{H}$. Extending the procedure adopted for the fiducial measurements, several Higgs boson mass templates corresponding to same jet $p_{\mathrm{T}}$ but different $p_{\mathrm{T}}^{H}$ ranges are used in the fits. Figure 7 presents the expected signal yield in each reconstructed-event category for each fiducial volume and the corresponding fraction of signal events.

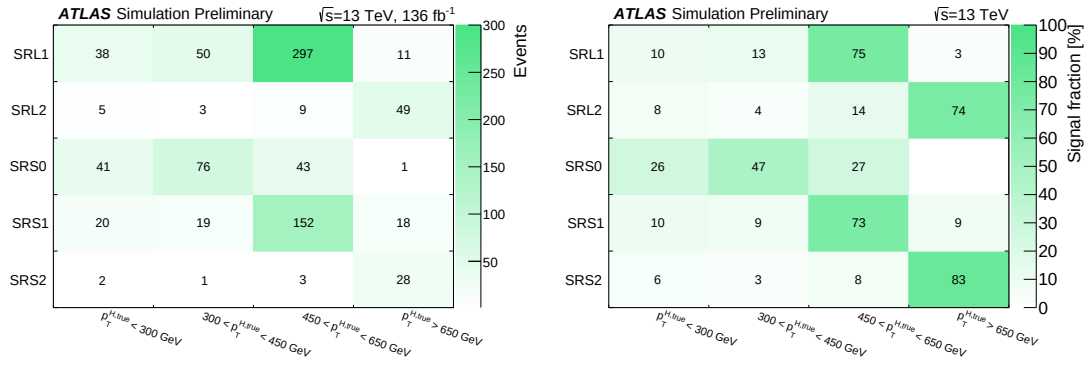



Figure 7: For each of the differential regions, the expected signal event yield for all Higgs boson events (left) and the fraction of signal in percent (right) in each reconstructed region vs. the differential fiducial volumes.

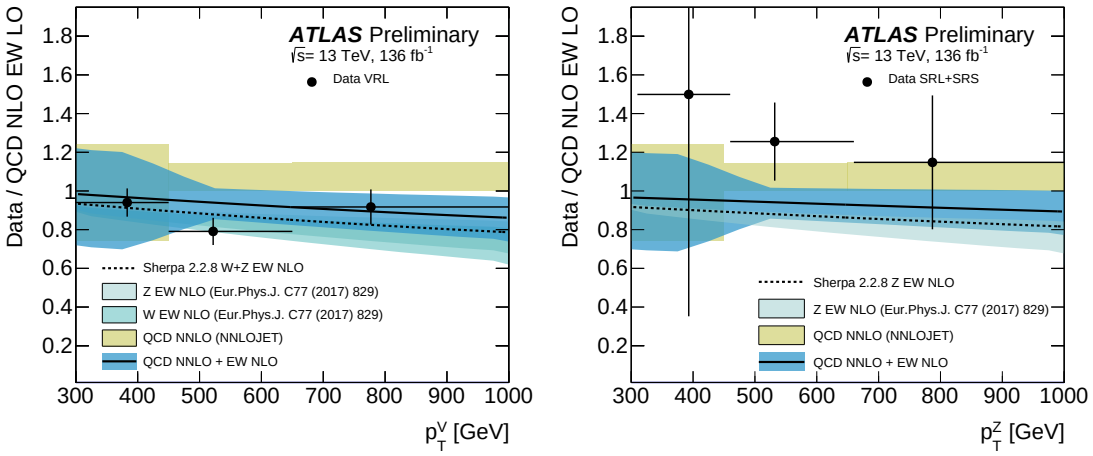
The procedure is tested with $W \to qq'$ and $Z \to q\bar{q}$ in the VR and $Z \to b\bar{b}$ in the SR. The VR offers a sample with larger statistics although lower signal-to-background ratio. For the differential $V$+jets analysis, the VRL is divided in five slices, the fit performed independently on each slice and the results combined. In the SR the differential $Z$ fit is performed to the SRL, SRS, and $CR_{t\bar{t}}$ regions with the Higgs boson contribution fixed to the SM prediction. Results of the two fits are shown in Figure 8 where they are compared to the predictions for the EW NLO and QCD NNLO corrections as a function of reconstructed $p_{\mathrm{T}}^{V}$. Both results agree with SM expectations.

To extract the Higgs boson signal strength, eight differential SR and CR regions defined in Table 2 are simultaneously fitted exploiting the corresponding systematic configurations shown in Table 5. Results are summarized in Tables 11 and 12. The three Higgs boson signal strengths are compatible with a $p$-value of 0.56. The post-fit jet mass distribution from the most sensitive category in each jet $p_{\mathrm{T}}$ bin is shown in Figure 9.

The acceptance times efficiency values for the different Higgs boson production processes are given in Table 13. The resulting Higgs boson production cross section for $p_{\mathrm{T}}^{H} > 650$ GeV is 13 ± 16 (stat.) ± 7 (syst.) ± 3 (theory) fb. The differential results correspond to the following 95% CL upper limits on the Higgs boson differential production cross sections:

$$\begin{aligned} \sigma_H(300 < p_{\mathrm{T}}^{H} < 450\ \mathrm{GeV}) &< 2.8\ \mathrm{pb}, \\ \sigma_H(450 < p_{\mathrm{T}}^{H} < 650\ \mathrm{GeV}) &< 91\ \mathrm{fb}, \\ \sigma_H(p_{\mathrm{T}}^{H} > 650\ \mathrm{GeV}) &< 40.5\ \mathrm{fb}. \end{aligned}$$

The uncertainties are statistically dominated. The contributions of the main categories of systematic uncertainties are summarized in Table 14.




(a) $V$+jets in VRL (b) $Z$+jets in SR

Figure 8: Comparison of differential fit signal strengths for $V$+jets in the VRL (a) and $Z$+jets in the SR (b). The signal strength is calculated with respect to the prediction at NLO QCD and LO EW accuracy. They are compared to the NLO EW correction provided by SHERPA, the NNLO QCD correction provided by the NNLOJET group, and their product. The Higgs boson yields are kept fixed to the SM expectation when extracting the $Z$+jets signal strength within the fiducial volumes.

Table 11: 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 fiducial $p_{\mathrm{T}}^H$ volume, while those for $\mu_Z$ and $\mu_{t\bar{t}}$ pertain to the corresponding jet $p_{\mathrm{T}}$ regions. The values of $\mu_{t\bar{t}}$ obtained are in agreement with unfolded measurements of $t\bar{t}$ events in a similar kinematic phase-space [118].

| $p_{\mathrm{T}}^H$ [GeV] | $\mu_H$ | |
|---|---|---|
| | Exp. | Obs. |
| 300–450 | 1 ± 18 | −7 ± 17 |
| 450–650 | 1.0 ± 3.3 | −2.9 ± 4.7 |
| >650 | 1.0 ± 6.3 | 4.8 ± 6.4 |

| Jet $p_{\mathrm{T}}$ [GeV] | $\mu_Z$ | | $\mu_{t\bar{t}}$ | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| 250–450 | 1.00 ± 1.11 | 1.77 ± 1.13 | 1.00 ± 0.07 | 0.87 ± 0.06 |
| 450–650 | 1.00 ± 0.17 | 1.28 ± 0.22 | 1.00 ± 0.07 | 0.76 ± 0.06 |
| 650–1000 | 1.00 ± 0.33 | 1.34 ± 0.42 | 1.00 ± 0.09 | 0.74 ± 0.07 |

The $t\bar{t}$ modeling systematic uncertainties are more relevant in the first jet $p_{\mathrm{T}}$ 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.

Table 12: Correlations amongst the three Higgs boson signal strengths and between the three Higgs boson and $Z$ + jets signal strengths. The Higgs boson signal strengths $\mu_H$ are labeled with the corresponding $p_{\mathrm{T}}^H$ range as a superscript. The $Z$ + jets signal strengths $\mu_Z$ are labeled with the corresponding region name as a superscript, where SR$i$ refers to SRL$i$ and SRS$i$.

| | $\mu_H^{300-450}$ | $\mu_H^{450-650}$ | $\mu_H^{>650}$ |
|---|---|---|---|
| $\mu_H^{300-450}$ | 1.00 | −0.18 | 0.07 |
| $\mu_H^{450-650}$ | −0.18 | 1.00 | −0.05 |
| $\mu_H^{>650}$ | 0.07 | −0.05 | 1.00 |

| | $\mu_Z^{SRS0}$ | $\mu_Z^{SR1}$ | $\mu_Z^{SR2}$ |
|---|---|---|---|
| $\mu_H^{300-450}$ | −0.07 | −0.02 | −0.03 |
| $\mu_H^{450-650}$ | 0.04 | 0.51 | −0.21 |
| $\mu_H^{>650}$ | 0.03 | −0.05 | 0.54 |

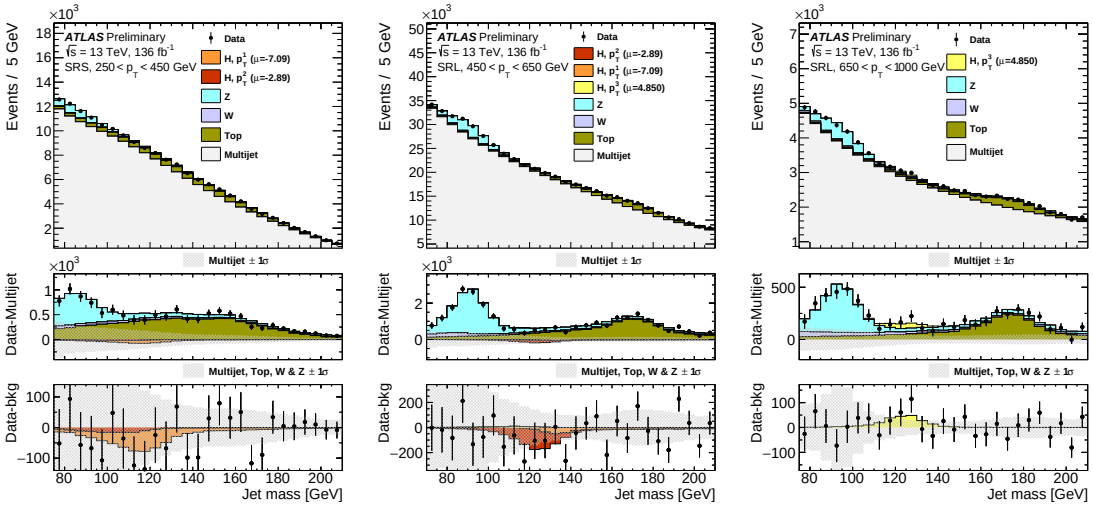



Figure 9: Post-fit signal jet mass distributions for the differential signal region defined by the subleading jet with $250 < p_{\mathrm{T}} < 450$ GeV (left) and leading jet with $450 < p_{\mathrm{T}} < 650$ GeV (middle) and $650 < p_{\mathrm{T}} < 1000$ GeV (right) with the various components. The 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 exponential polynomials. The lower panels show the distributions after subtraction of the fitted background processes: multijet, $V$+jets and Top. The shaded areas indicate the 68% CL for all background processes. The four fiducial volumes are labeled $p_{\mathrm{T}}^0$–$p_{\mathrm{T}}^3$ corresponding to $p_{\mathrm{T}}^H < 300$ GeV, 300 – 450 GeV, 450 – 650 GeV, and > 650 GeV, respectively. The $p_{\mathrm{T}}^0$ event yield is constrained to its SM value within the theoretical and experimental uncertainties and free parameters act independently on the remaining three volumes. Contributions below 0.5 per mille of the total yield are not shown.

Table 13: Signal acceptance times efficiency for the signal regions in the differential measurements. For events with $p_{\mathrm{T}}^{H} < 300$ GeV, the acceptance times efficiency is less than $0.1 \times 10^{-2}$.

| Process | $300 < p_{\mathrm{T}}^{H} < 450$ GeV | $450 < p_{\mathrm{T}}^{H} < 650$ GeV | $p_{\mathrm{T}}^{H} > 650$ GeV |
|---|---|---|---|
| All | $1.4 \times 10^{-2}$ | 0.25 | 0.33 |
| ggF | $0.7 \times 10^{-2}$ | 0.26 | 0.37 |
| $VH$ | $1.8 \times 10^{-2}$ | 0.28 | 0.31 |
| VBF | $0.2 \times 10^{-2}$ | 0.14 | 0.18 |
| $t\bar{t}H$ | $4.9 \times 10^{-2}$ | 0.20 | 0.25 |

Table 14: Contributions to the systematic uncertainties for the differential signal strength measurements. The total uncertainty is also given for comparison.

| Uncertainty Contribution | $300 < p_{\mathrm{T}}^{H} < 450$ GeV | $450 < p_{\mathrm{T}}^{H} < 650$ GeV | $p_{\mathrm{T}}^{H} > 650$ GeV |
|---|---|---|---|
| Total | 17 | 4.7 | 6.4 |
| Statistical | 16 | 4.0 | 5.9 |
| Jet Systematics | 5.4 | 2.6 | 2.3 |
| Modeling and Theory Systs. | 3.8 | $< 0.1$ | 0.1 |
| Flavor Tagging Systs. | $< 0.1$ | $< 0.1$ | $< 0.1$ |
| Total Systematics | 6 | 2.7 | 2.5 |

## 10 Conclusions

Studies of the Higgs boson produced at high $p_{\mathrm{T}}$ and decaying into a $b\bar{b}$ pair are performed. The results are based on the Run 2 dataset of $pp$ collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC, corresponding to an integrated luminosity of 136 fb$^{-1}$. The Higgs boson is reconstructed from a single large-$R$ jet and identified with $b$-tagging techniques. The cross section is extracted in multiple regions of increasing $p_{\mathrm{T}}^{H}$. Within the same kinematic regimes, measurements of the $Z \to b\bar{b}$ process agree with the Standard Model predictions, validating the methods.

The fitted values and upper limits at 95% CL of the Higgs boson signal strengths in three $p_{\mathrm{T}}^{H}$ regions are shown in Figure 10.

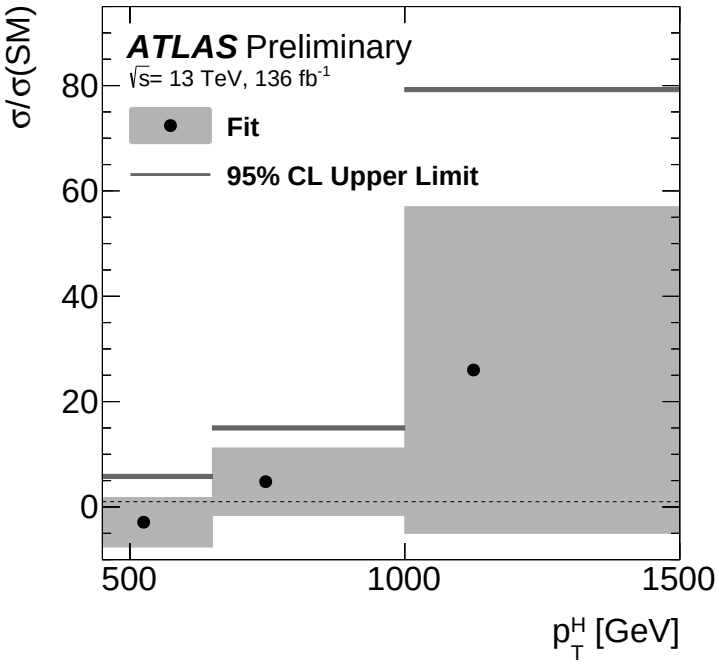



Figure 10: A summary of the signal strengths measured in the various $p_{\mathrm{T}}^{H}$ fiducial volumes.

The Higgs boson production cross section, in three regions, is found to be:

$$\begin{aligned} \sigma_H(p_{\mathrm{T}}^{H} > 450\ \mathrm{GeV}) &= 13 \pm 57\ (\mathrm{stat.}) \pm 22\ (\mathrm{syst.}) \pm 3\ (\mathrm{theory})\ \mathrm{fb}, \\ \sigma_H(p_{\mathrm{T}}^{H} > 650\ \mathrm{GeV}) &= 13 \pm 16\ (\mathrm{stat.}) \pm 7\ (\mathrm{syst.}) \pm 3\ (\mathrm{theory})\ \mathrm{fb}, \\ \sigma_H(p_{\mathrm{T}}^{H} > 1\ \mathrm{TeV}) &= 3.4 \pm 3.9\ (\mathrm{stat.}) \pm 1.0\ (\mathrm{syst.}) \pm 0.8\ (\mathrm{theory})\ \mathrm{fb}. \end{aligned}$$

Higgs boson cross-section production limits at 95% CL obtained in four differential volumes are: $\sigma_H(300 < p_{\mathrm{T}}^{H} < 450\ \mathrm{GeV}) < 2.8$ pb, $\sigma_H(450 < p_{\mathrm{T}}^{H} < 650\ \mathrm{GeV}) < 91$ fb, $\sigma_H(p_{\mathrm{T}}^{H} > 650\ \mathrm{GeV}) < 40.5$ fb, and $\sigma_H(p_{\mathrm{T}}^{H} > 1\ \mathrm{TeV}) < 10.3$ fb. All of the Higgs boson results are consistent with the Standard Model predictions.

# Appendix

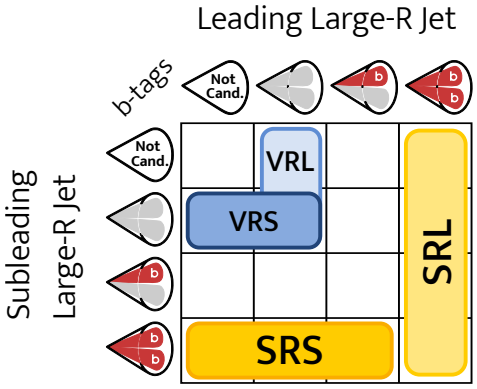



Figure 11: Diagram showing the event categorization criteria. The columns (rows) are divided into 4 categories: when the leading (subleading) jet is not a candidate jet (see Section 4.2), when neither of the first two $p_{\mathrm{T}}$-ordered track-jet is $b$-tagged, when one of the track-jets is $b$-tagged and when both track-jets are $b$-tagged.

Table 15: A summary of selection requirements for the SR and $CR_{t\bar{t}}$ used within the likelihood fit. The fiducial volumes within the SR are defined, where appropriate, and all require $|\eta_H| < 2$. Following the definitions in Section 5.1, each SR has an associated VR in the same kinematic region. (∗) Jets are ordered before the muon-in-jet correction is applied.

| | Inclusive | | Fiducial | | | | Differential | |
|---|---|---|---|---|---|---|---|---|
| Signal Regions | | | | | | | | |
| Jet Order* | Lead | Sublead | Lead | Sublead | Lead | Sublead | Lead | Sublead |
| Jet $p_{\mathrm{T}}$ [GeV] | >450 | >250 | >450 | >450 | >1000 | – | 450–650, 650–1000 | 250–450, 450–650, 650–1000 |
| Fiducial Volumes | | | | | | | | |
| $p_{\mathrm{T}}^{H}$ [GeV] | – | – | >450 | >450 | >1000 | – | 450–650, >650 | 300–450, 450–650, >650 |
| $t\bar{t}$ Control Regions | | | | | | | | |
| $J_t$ $p_{\mathrm{T}}$ [GeV] | >250 | | >450 | | >1000 | | 250–450, 450–650, 650–1000 | |

Table 16: Event yields and associated uncertainties after the global likelihood fit in the $p_{\mathrm{T}}^{H} > 450$ GeV fiducial region. The rate of the $W(\ell\nu)$ background is assumed to be fixed. The total background yield can differ from the sum of the individual components due to rounding.

| Process | SRL | | SRS | | $CR_{t\bar{t}}$ | |
|---|---|---|---|---|---|---|
| Multijet | 591 800 | ± 4000 | 371 500 | ± 2300 | – | |
| $Z$ + jets | 16 000 | ± 2700 | 9100 | ± 1500 | – | |
| $W$+ jets | 3110 | ± 740 | 1700 | ± 350 | – | |
| Top | 15 200 | ± 1900 | 8100 | ± 1100 | 3740 | ± 70 |
| $W(\ell\nu)$ | – | | – | | 50 | |
| $H$ $p_{\mathrm{T}}^{H} < 450$ GeV | 100 | ± 20 | 40 | ± 9 | – | |
| $H$ $p_{\mathrm{T}}^{H} > 450$ GeV | 300 | ± 1200 | 140 | ± 670 | – | |
| Total | 626 530 | ± 820 | 390 670 | ± 650 | 3790 | ± 70 |
| Data | 626 532 | | 390 665 | | 3791 | |

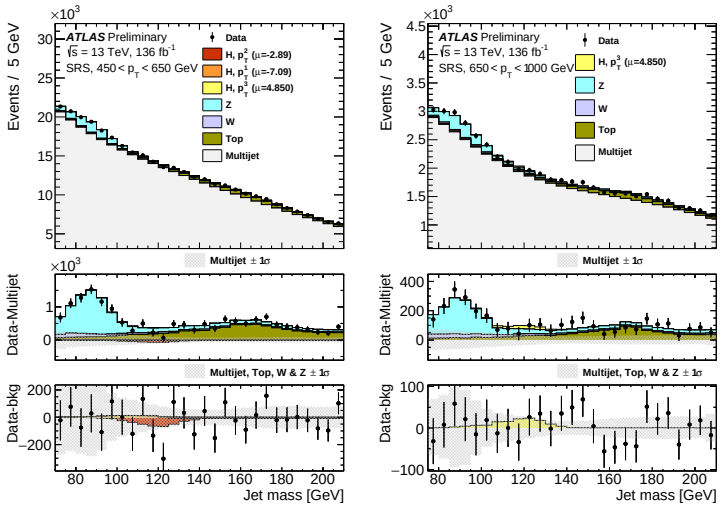



Figure 12: Post-fit signal jet mass distributions for the differential regions defined by the subleading jet $450 < p_{\mathrm{T}} < 650$ GeV (left) and $650 < p_{\mathrm{T}} < 1000$ GeV (right) with the various components. The middle panels show the distributions after subtraction of the multijet distribution. The shaded areas indicate the 68% CL for the multijet background from the fitted parameters and normalizations of the exponential polynomials. The lower panels show the distributions after subtraction of the fitted background processes: multijet, $V$+jets and Top. The shaded areas indicate the 68% CL for all background processes. The four fiducial volumes are labeled $p_{\mathrm{T}}^{0}$–$p_{\mathrm{T}}^{3}$ corresponding to $p_{\mathrm{T}}^{H} < 300$ GeV, 300 – 450 GeV, 450 – 650 GeV, and > 650 GeV, respectively. The $p_{\mathrm{T}}^{0}$ event yield is constrained to its SM value within the theoretical and experimental uncertainties and free parameters act independently on the remaining three volumes. Contributions below 0.5 per mille of the total yield are not shown.

Table 17: Event yields and associated uncertainties after the global likelihood fit in the $p_{\mathrm{T}}^{H} > 1$ TeV fiducial region. The rate of the $W(\ell\nu)$ background is assumed to be fixed. The total background yield can differ from the sum of the individual components due to rounding.

| Process | SRL | $CR_{t\bar{t}}$ |
|---|---|---|
| Multijet | 5570 ± 330 | – |
| $Z$ + jets | 400 ± 240 | – |
| $W$+ jets | 100 ± 30 | – |
| Top | 130 ± 40 | 28 ± 6 |
| $W(\ell\nu)$ | – | 6 |
| $H$ $p_{\mathrm{T}}^{H} < 1$ TeV | 1.0 ± 0.3 | – |
| $H$ $p_{\mathrm{T}}^{H} > 1$ TeV | 90 ± 100 | – |
| Total | 6280 ± 80 | 34 ± 6 |
| Data | 6283 | 34 |

Table 18: Event yields and associated uncertainties after the global likelihood fit of the differential regions. The rate of the $W(\ell\nu)$ background is assumed to be fixed. The total background yield can differ from the sum of the individual components due to rounding. The four fiducial volumes are labeled $p_{\mathrm{T}}^{0}$–$p_{\mathrm{T}}^{3}$ corresponding to $p_{\mathrm{T}}^{H} < 300$ GeV, 300 – 450 GeV, 450 – 650 GeV, and > 650 GeV, respectively. The $p_{\mathrm{T}}^{0}$ event yield is constrained to its SM value within the theoretical and experimental uncertainties and free parameters act independently on the remaining three volumes.

| | SRL | | SRS | | | $CR_{t\bar{t}}$ | | |
|---|---|---|---|---|---|---|---|---|
| Jet $p_{\mathrm{T}}$ [GeV] | 450–650 | 650–1000 | 250–450 | 450–650 | 650–1000 | 250–450 | 450–650 | 650–1000 |
| Multijet | 510 200 ± 3600 | 72 500 ± 1100 | 157 600 ± 2900 | 341 100 ± 2100 | 48 160 ± 710 | – | – | – |
| $Z$ + jets | 13 100 ± 2200 | 2570 ± 770 | 3900 ± 2400 | 8000 ± 1300 | 1610 ± 520 | – | – | – |
| $W$+ jets | 2280 ± 500 | 630 ± 160 | 770 ± 150 | 1380 ± 260 | 365 ± 84 | – | – | – |
| Top | 12 600 ± 1200 | 2290 ± 290 | 7800 ± 1000 | 6870 ± 740 | 1060 ± 160 | 1675 ± 42 | 1607 ± 41 | 427 ± 21 |
| $W(\ell\nu)$ | – | – | – | – | – | 12 | 22 | 14 |
| $H\ p_{\mathrm{T}}^{0}$ | 38 ± 8 | 5 ± 1 | 40 ± 8 | 20 ± 4 | 2.0 ± 0.4 | – | – | – |
| $H\ p_{\mathrm{T}}^{1}$ | −360 ± 870 | −19 ± 45 | −500 ± 1300 | −130 ± 320 | −6 ± 16 | – | – | – |
| $H\ p_{\mathrm{T}}^{2}$ | −900 ± 1400 | −27 ± 44 | −120 ± 200 | −440 ± 710 | −8 ± 13 | – | – | – |
| $H\ p_{\mathrm{T}}^{3}$ | 54 ± 71 | 240 ± 310 | 3 ± 4 | 90 ± 110 | 140 ± 180 | – | – | – |
| Total | 537 040 ± 740 | 78 180 ± 280 | 169 420 ± 420 | 356 910 ± 600 | 51 320 ± 230 | 1690 ± 40 | 1630 ± 40 | 440 ± 20 |
| Data | 537 038 | 78 177 | 169 418 | 356 910 | 51 324 | 1687 | 1628 | 442 |

## References


[1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on p. 2).

[2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on p. 2).

[3] ATLAS Collaboration, *Combined measurements of Higgs boson production and decay using up to 80 $fb^{-1}$ of proton–proton collision data at $\sqrt{s}$ = 13 TeV collected with the ATLAS experiment*, Phys. Rev. D **101** (2020) 012002, arXiv: 1909.02845 [hep-ex] (cit. on p. 2).

[4] CMS Collaboration, *Combined measurements of Higgs boson couplings in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **79** (2019) 421, arXiv: 1809.10733 [hep-ex] (cit. on p. 2).

[5] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph] (cit. on pp. 2, 11).

[6] J. R. Andersen et al., "Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report," *9th Les Houches Workshop on Physics at TeV Colliders (PhysTeV 2015) Les Houches, France, June 1-19, 2015*, 2016, arXiv: 1605.04692 [hep-ph], URL: http://lss.fnal.gov/archive/2016/conf/fermilab-conf-16-175-ppd-t.pdf (cit. on p. 2).

[7] S. Amoroso et al., "Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report," *11th Les Houches Workshop on Physics at TeV Colliders: PhysTeV Les Houches*, 2020, arXiv: 2003.01700 [hep-ph] (cit. on p. 2).

[8] C. Grojean, E. Salvioni, M. Schlaffer, and A. Weiler, *Very boosted Higgs in gluon fusion*, JHEP **05** (2014) 022, arXiv: 1312.3317 [hep-ph] (cit. on p. 2).

[9] A. Biekötter, A. Knochel, M. Krämer, D. Liu, and F. Riva, *Vices and virtues of Higgs effective field theories at large energy*, Phys. Rev. D **91** (2015) 055029, arXiv: 1406.7320 [hep-ph] (cit. on p. 2).

[10] K. Mimasu, V. Sanz, and C. Williams, *Higher Order QCD predictions for Associated Higgs production with anomalous couplings to gauge bosons*, JHEP **08** (2016) 039, arXiv: 1512.02572 [hep-ph] (cit. on p. 2).

[11] M. Grazzini, A. Ilnicka, M. Spira, and M. Wiesemann, *Modeling BSM effects on the Higgs transverse-momentum spectrum in an EFT approach*, JHEP **03** (2017) 115, arXiv: 1612.00283 [hep-ph] (cit. on p. 2).

[12] A. Banfi, A. Bond, A. Martin, and V. Sanz, *Digging for Top Squarks from Higgs data: from signal strengths to differential distributions*, JHEP **11** (2018) 171, arXiv: 1806.05598 [hep-ph] (cit. on p. 2).

[13] K. Becker et al., *Precise predictions for boosted Higgs production*, (2020), arXiv: 2005.07762 [hep-ph] (cit. on pp. 2, 4, 5, 11).

[14] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 2, 4).

[15] CMS Collaboration, *Inclusive search for highly boosted Higgs bosons decaying to bottom quark-antiquark pairs in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **12** (2020) 085, arXiv: 2006.13251 [hep-ex] (cit. on p. 2).

[16] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100** (2008) 242001, arXiv: 0802.2470 [hep-ph] (cit. on p. 2).

[17] ATLAS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with* $36\,fb^{-1}$ *of pp collision data at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **98** (2018) 052005, arXiv: 1802.04146 [hep-ex] (cit. on p. 2).

[18] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173, arXiv: 1806.00425 [hep-ex] (cit. on p. 2).

[19] 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** (2020) 957, arXiv: 2004.03447 [hep-ex] (cit. on pp. 2, 3).

[20] ATLAS Collaboration, *Measurements of the Higgs boson inclusive and differential fiducial cross sections in the* $4\ell$ *decay channel at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **80** (2020) 942, arXiv: 2004.03969 [hep-ex] (cit. on p. 2).

[21] C. Collaboration, *Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at* $\sqrt{s} = 13$ *TeV*, (2021), arXiv: 2103.06956 [hep-ex] (cit. on p. 2).

[22] C. Collaboration, *Measurements of production cross sections of the Higgs boson in the four-lepton final state in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, (2021), arXiv: 2103.04956 [hep-ex] (cit. on pp. 2, 3).

[23] CMS Collaboration, *Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at* $\sqrt{s} = 13$ *TeV*, (2020), arXiv: 2011.03652 [hep-ex] (cit. on p. 2).

[24] M. Grazzini, A. Ilnicka, and M. Spira, *Higgs boson production at large transverse momentum within the SMEFT: analytical results*, Eur. Phys. J. **C78** (2018) 808, arXiv: 1806.08832 [hep-ph] (cit. on p. 3).

[25] ATLAS Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of* $\tau$*-leptons in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072001, arXiv: 1811.08856 [hep-ex] (cit. on p. 3).

[26] ATLAS Collaboration, *Measurements of Higgs Bosons Decaying to Bottom Quarks from Vector Boson Fusion Production with the ATLAS Experiment at* $\sqrt{s} = 13$ *TeV*, (2020), arXiv: 2011.08280 [hep-ex] (cit. on p. 3).

[27] CMS Collaboration, *Observation of the Higgs boson decay to a pair of* $\tau$ *leptons*, Phys. Lett. B **779** (2018) 283, arXiv: 1708.00373 [hep-ex] (cit. on p. 3).

[28] CMS Collaboration, *Search for the standard model Higgs boson produced through vector boson fusion and decaying to* $b\bar{b}$, Phys. Rev. D **92** (2015) 032008, arXiv: 1506.01010 [hep-ex] (cit. on p. 3).

[29] ATLAS Collaboration, *Measurements of $WH$ and $ZH$ production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector*, (2020), arXiv: 2007.02873 [hep-ex] (cit. on p. 3).

[30] ATLAS Collaboration, *Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2020), arXiv: 2008.02508 [hep-ex] (cit. on p. 3).

[31] CMS Collaboration, *Observation of Higgs Boson Decay to Bottom Quarks*, Phys. Rev. Lett. **121** (2018) 121801, arXiv: 1808.08242 [hep-ex] (cit. on p. 3).

[32] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv: 2007.12539 [physics.ins-det] (cit. on p. 4).

[33] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-$k_t$ jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 4).

[34] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex] (cit. on p. 4).

[35] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: https://cds.cern.ch/record/2677054 (cit. on p. 4).

[36] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 4).

[37] K. Hamilton, P. Nason, and G. Zanderighi, *MINLO: Multi-Scale Improved NLO*, JHEP **10** (2012) 155, arXiv: 1206.3572 [hep-ph] (cit. on pp. 4, 5).

[38] J. M. Campbell et al., *NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM*, JHEP **07** (2012) 092, arXiv: 1202.5475 [hep-ph] (cit. on pp. 4, 5).

[39] K. Hamilton, P. Nason, C. Oleari, and G. Zanderighi, *Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching*, JHEP **05** (2013) 082, arXiv: 1212.4504 [hep-ph] (cit. on pp. 4, 5).

[40] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph] (cit. on pp. 4, 5).

[41] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146 (cit. on pp. 4, 5).

[42] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on pp. 4, 5).

[43] K. Hamilton, P. Nason, and G. Zanderighi, *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*, JHEP **05** (2015) 140, arXiv: 1501.04637 [hep-ph] (cit. on p. 4).

[44] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **02** (2010) 037, arXiv: 0911.5299 [hep-ph] (cit. on pp. 4, 5).

[45] H. B. Hartanto, B. Jäger, L. Reina, and D. Wackeroth,
*Higgs boson production in association with top quarks in the POWHEG BOX*,
Phys. Rev. D **91** (2015) 094003, arXiv: 1501.04498 [hep-ph] (cit. on pp. 4, 5).

[46] G. Luisoni, 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*, JHEP **10** (2013) 083,
arXiv: 1306.2542 [hep-ph] (cit. on pp. 4, 5).

[47] G. Cullen et al., *Automated One-Loop Calculations with GoSam*, Eur. Phys. J. **C72** (2012) 1889,
arXiv: 1111.2034 [hep-ph] (cit. on pp. 4, 5).

[48] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56,
arXiv: hep-ph/9704448 (cit. on p. 4).

[49] M. Spira, *QCD Effects in Higgs physics*, Fortsch. Phys. **46** (1998) 203, arXiv: hep-ph/9705337
(cit. on p. 4).

[50] A. Djouadi, M. M. Mühlleitner, and M. Spira,
*Decays of supersymmetric particles: The Program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface)*,
Acta Phys. Polon. B **38** (2007) 635, arXiv: hep-ph/0609292 (cit. on p. 4).

[51] A. Bredenstein, 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*,
JHEP **02** (2007) 080, arXiv: hep-ph/0611234 (cit. on p. 4).

[52] A. Bredenstein, 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** (2006) 013004, arXiv: hep-ph/0604011 [hep-ph] (cit. on p. 4).

[53] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber,
*Precision calculations for the Higgs-boson decay $H \to ZZ/WW \to 4$ leptons*,
Nucl. Phys. Proc. Suppl. **160** (2006) 131, arXiv: hep-ph/0607060 [hep-ph] (cit. on p. 4).

[54] U. Blumenschein et al., "Pushing the precision frontier at the LHC with V+jets," 2018,
arXiv: 1802.02100 [hep-ex] (cit. on p. 5).

[55] 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 hadronic jet*,
Phys. Rev. Lett. **117** (2016) 022001, arXiv: 1507.02850 [hep-ph] (cit. on p. 5).

[56] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan,
*The NNLO QCD corrections to Z boson production at large transverse momentum*,
JHEP **07** (2016) 133, arXiv: 1605.04295 [hep-ph], We thank Nigel Glover and Alexander Huss for providing $\sqrt{s}$ = 13 TeV k-factors for the analysis fiducial volume. (Cit. on p. 5).

[57] 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*, JHEP **03** (2012) 075,
arXiv: 1112.5142 [hep-ph] (cit. on p. 5).

[58] 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 colliders*,
Comput. Phys. Commun. **195** (2015) 161, arXiv: 1412.5390 [hep-ph] (cit. on p. 5).

[59] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019) 866, arXiv: 1907.13071 [hep-ph]
(cit. on p. 5).

[60] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph] (cit. on p. 5).

[61] A. Denner, S. Dittmaier, and L. Hofer, *COLLIER: A fortran-based complex one-loop library in extended regularizations*, Comput. Phys. Commun. **212** (2017) 220, arXiv: 1604.06792 [hep-ph] (cit. on p. 5).

[62] R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 5).

[63] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on pp. 5, 6).

[64] ATLAS Collaboration, *Measurement of the $Z/\gamma^*$ boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014) 145, arXiv: 1406.3660 [hep-ex] (cit. on p. 5).

[65] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 5).

[66] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph] (cit. on p. 5).

[67] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph] (cit. on p. 5).

[68] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph] (cit. on p. 5).

[69] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, *Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO*, Phys. Rev. Lett. **103** (2009) 082001, arXiv: 0903.2120 [hep-ph] (cit. on p. 5).

[70] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr, *NLO electroweak automation and precise predictions for W+multijet production at the LHC*, JHEP **04** (2015) 012, arXiv: 1412.5157 [hep-ph] (cit. on p. 5).

[71] B. Biedermann et al., *Automation of NLO QCD and EW corrections with Sherpa and Recola*, Eur. Phys. J. C **77** (2017) 492, arXiv: 1704.05783 [hep-ph] (cit. on p. 5).

[72] S. Frixione, P. Nason, and G. Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph] (cit. on p. 5).

[73] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419 (cit. on p. 5).

[74] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph] (cit. on p. 5).

[75] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, arXiv: 1009.2450 [hep-ph] (cit. on p. 5).

[76] R. Frederix, E. Re, and P. Torrielli,
*Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*,
JHEP **09** (2012) 130, arXiv: 1207.5391 [hep-ph] (cit. on p. 5).

[77] S. Alioli, P. Nason, C. Oleari, and E. Re,
*NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*,
JHEP **09** (2009) 111, arXiv: 0907.4076 [hep-ph] (cit. on p. 5), Erratum: JHEP **02** (2010) 011.

[78] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244,
arXiv: 1207.1303 [hep-ph] (cit. on pp. 5, 6).

[79] S. Frixione, E. Laenen, P. Motylinski, C. White, and B. R. Webber,
*Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029,
arXiv: 0805.3067 [hep-ph] (cit. on p. 5).

[80] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823,
arXiv: 1005.4568 [physics.ins-det] (cit. on p. 6).

[81] S. Agostinelli et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250
(cit. on p. 6).

[82] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*,
ATL-PHYS-PUB-2016-017, 2016, URL: https://cds.cern.ch/record/2206965
(cit. on p. 6).

[83] D. J. Lange, *The EvtGen particle decay simulation package*,
Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 6).

[84] ATLAS Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*,
ATL-PHYS-PUB-2017-006, 2017, URL: https://cds.cern.ch/record/2261937
(cit. on p. 6).

[85] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*,
ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168
(cit. on p. 6).

[86] ATLAS Collaboration,
*Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO*,
ATL-PHYS-PUB-2017-007, 2017, URL: https://cds.cern.ch/record/2261938
(cit. on p. 6).

[87] ATLAS Collaboration,
*Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2*,
ATL-PHYS-PUB-2018-009, 2018, URL: https://cds.cern.ch/record/2630327
(cit. on p. 6).

[88] ATLAS Collaboration, *Multijet simulation for 13 TeV ATLAS Analyses*,
ATL-PHYS-PUB-2019-017, 2019, URL: https://cds.cern.ch/record/2672252
(cit. on p. 6).

[89] ATLAS Collaboration,
*Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s} = 13$ TeV*,
ATL-PHYS-PUB-2015-051, 2015, URL: https://cds.cern.ch/record/2110140
(cit. on p. 6).

[90] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332, arXiv: 1611.10235 [hep-ex] (cit. on p. 6).

[91] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph] (cit. on p. 6).

[92] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex] (cit. on p. 6).

[93] ATLAS Collaboration, *Selection of jets produced in* 13 *TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015, URL: https://cds.cern.ch/record/2037702 (cit. on p. 6).

[94] D. Krohn, J. Thaler, and L.-T. Wang, *Jet Trimming*, JHEP **02** (2010) 084, arXiv: 0912.1342 [hep-ph] (cit. on p. 6).

[95] ATLAS Collaboration, *In situ calibration of large-radius jet energy and mass in* 13 *TeV proton–proton collisions with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 135, arXiv: 1807.09477 [hep-ex] (cit. on pp. 6, 7).

[96] D. Krohn, J. Thaler, and L.-T. Wang, *Jets with Variable R*, JHEP **06** (2009) 059, arXiv: 0903.0392 [hep-ph] (cit. on p. 6).

[97] ATLAS Collaboration, *Variable Radius, Exclusive-$k_T$, and Center-of-Mass Subjet Reconstruction for Higgs($\to b\bar{b}$) Tagging in ATLAS*, ATL-PHYS-PUB-2017-010, 2017, URL: https://cds.cern.ch/record/2268678 (cit. on p. 6).

[98] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph] (cit. on p. 6).

[99] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. **B659** (2008) 119, arXiv: 0707.1378 [hep-ph] (cit. on p. 6).

[100] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: 1907.05120 [hep-ex] (cit. on pp. 6, 8).

[101] ATLAS Collaboration, *Identification of boosted Higgs bosons decaying into b-quark pairs with the ATLAS detector at* 13 *TeV*, Eur. Phys. J. C **79** (2019) 836, arXiv: 1906.11005 [hep-ex] (cit. on p. 7).

[102] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex] (cit. on p. 7).

[103] ATLAS Collaboration, *Jet mass and substructure of inclusive jets in $\sqrt{s}$ = 7 TeV pp collisions with the ATLAS experiment*, JHEP **05** (2012) 128, arXiv: 1203.4606 [hep-ex] (cit. on p. 7).

[104] ATLAS Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, ATLAS-CONF-2016-035, 2016, URL: https://cds.cern.ch/record/2200211 (cit. on p. 7).

[105] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018, URL: https://cds.cern.ch/record/2306649 (cit. on p. 8).

[106] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s}$ = 13 TeV*, ATLAS-CONF-2018-006, 2018, URL: https://cds.cern.ch/record/2314418 (cit. on p. 8).

[107] *Simulation-based extrapolation of b-tagging calibrations towards high transverse momenta in the ATLAS experiment*, tech. rep. ATL-PHYS-PUB-2021-003, CERN, 2021, URL: https://cds.cern.ch/record/2753444 (cit. on p. 8).

[108] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: https://cds.cern.ch/record/2273281 (cit. on p. 9).

[109] J. M. Lindert, K. Kudashkin, K. Melnikov, and C. Wever, *Higgs bosons with large transverse momentum at the LHC*, Phys. Lett. B **782** (2018) 210, arXiv: 1801.08226 [hep-ph] (cit. on p. 11).

[110] J. Lindert et al., *Precise predictions for V+ jets dark matter backgrounds*, Eur. Phys. J. C **77** (2017) 829, arXiv: 1705.04664 [hep-ph] (cit. on p. 12).

[111] T. Sjöstrand, *Jet fragmentation of multiparton configurations in a string framework*, Nucl. Phys. B **248** (1984) 469 (cit. on p. 12).

[112] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: physics/0306116 [physics.data-an] (cit. on p. 16).

[113] S. Hageboeck, "What the new RooFit can do for your analysis," *40th International Conference on High Energy Physics*, 2020, arXiv: 2012.02746 [physics.data-an] (cit. on p. 16).

[114] *RooBinSamplingPdf*, URL: https://root.cern.ch/doc/master/classRooBinSamplingPdf.html (cit. on p. 16).

[115] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Meth. A **434** (1999) 435, arXiv: hep-ex/9902006 (cit. on p. 16).

[116] A. L. Read, *Presentation of search results: the $CL_S$ technique*, J. Phys. G **28** (2002) 2693 (cit. on p. 16).

[117] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219 (cit. on p. 16).

[118] ATLAS Collaboration, *Measurements of top-quark pair differential and double-differential cross-sections in the $\ell$+jets channel with pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector*, Eur. Phys. J. C **79** (2019) 1028, arXiv: 1908.07305 [hep-ex] (cit. on pp. 18, 20, 23).