EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Submitted to: JHEP

CERN-EP-2022-187
4th November 2022

A search for heavy Higgs bosons decaying into vector bosons in same-sign two-lepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

A search for heavy Higgs bosons produced in association with a vector boson and decaying into a pair of vector bosons is performed in final states with two leptons (electrons or muons) of the same electric charge, missing transverse momentum and jets. A data sample of proton–proton collisions at a centre-of-mass energy of 13 TeV recorded with the ATLAS detector at the Large Hadron Collider between 2015 and 2018 is used. The data correspond to a total integrated luminosity of 139 fb^{-1}. The observed data are in agreement with Standard Model background expectations. The results are interpreted using higher-dimensional operators in an effective field theory. Upper limits on the production cross-section are calculated at 95% confidence level as a function of the heavy Higgs boson's mass and coupling strengths to vector bosons. Limits are set in the Higgs boson mass range from 300 to 1500 GeV, and depend on the assumed couplings. The highest excluded mass for a heavy Higgs boson with the coupling combinations explored is 900 GeV. Limits on coupling strengths are also provided.

Contents

1 Introduction

The discovery of the Higgs boson was a major success for the Standard Model (SM) and an important breakthrough in understanding electroweak symmetry breaking [1, 2]. It also opened new ways to search for physics beyond the Standard Model (BSM physics). Despite its success the SM is not without problems that may require extensions and new concepts. One natural extension common to many BSM physics models is an extended Higgs sector, which leads to the introduction of additional Higgs bosons. In such models, the lightest Higgs boson (h) often has properties similar to those of the observed SM Higgs boson. Additional Higgs bosons have been introduced to explain a very wide range of BSM phenomena, from the observed baryon asymmetry in the universe, and how to solve the strong CP problem with the help of axions, to the generation of non-zero neutrino masses [3–7]. Some models have incorporated recent potential deviations from the SM seen in muon $g-2$ and W mass measurements [8].

Several searches for additional heavy Higgs bosons (H) have already been carried out with the ATLAS and CMS detectors at the Large Hadron Collider (LHC) [9–15]. These searches mainly relied on the gluon–gluon fusion (ggF) and vector-boson fusion (VBF) production mechanisms, which are the dominant production modes for the SM Higgs boson at the LHC. In ggF production, the gluons couple to the Higgs boson mainly via a top-quark loop, so the non-observation of an additional Higgs boson in this channel could point to a reduced fermionic coupling.

This analysis concentrates on the production of a heavy Higgs boson in association with a vector boson (VH, where $V = W, Z$) and $H \to VV$ decays. By utilising a general effective Lagrangian that includes dimension-six operators in an effective field theory (EFT), it can be shown that, relative to VBF production, the VH production mode benefits from having smaller SM backgrounds, and the production cross-section may be enhanced by higher-order contributions, especially at high Higgs boson mass and vector-boson momenta [16–18]. The observed Higgs boson h, is assumed to have the production and decay modes as in the SM, with production through $H \to Zh$ negligible.

Rather than focusing on any specific model, a generic search is performed for the same-sign dilepton signature (SS2L), targeting the $W^{\pm}H \to W^{\pm}W^{\pm}W^{\mp} \to \ell^{\pm}\nu\ell^{\pm}\nu jj$ decay channel. The corresponding Feynman diagram is shown in Figure 1. In this article, the term 'lepton', unless stated otherwise, refers to either an electron or a muon. Electrons and muons from τ-lepton decays are also considered. The hadronically decaying W boson is reconstructed either as two small-radius jets or as a single large-radius jet for higher-momentum W bosons. The presence of neutrinos prevents the full reconstruction of the heavy Higgs boson's mass and is the main drawback of the $W^{\pm}W^{\pm}W^{\mp}$ channel with leptonic decays of the W bosons. It is ameliorated with the help of the reconstruction methods described in Section 6. Compared with other bosonic VH decay channels, the chosen channel has the highest signal sensitivity thanks to low SM backgrounds and a sizeable branching fraction for $H \to WW$ decay [18, 19]. SM processes produce same-sign lepton pairs at the LHC at a rate that is orders of magnitude below that of opposite-sign lepton-pair production in the SM.

2 Phenomenology

In theories with multiple Higgs fields, the fields in the multi-Higgs potential interact and the mass eigenstates are formed from a mixture of the fields. In the simple case of a lightest (h) and next-to-lightest (H) neutral Higgs doublet, the couplings of the Higgs boson to the SM gauge bosons are scaled relative

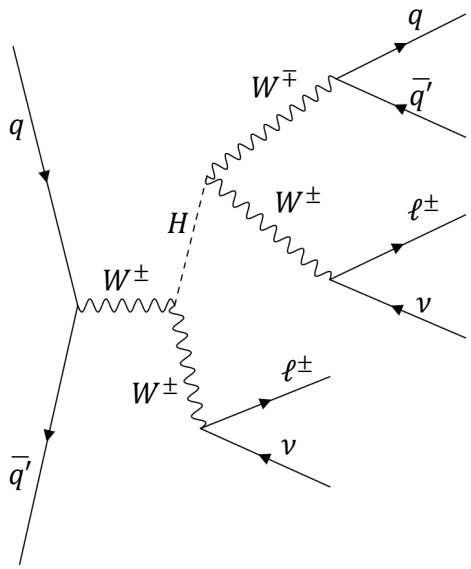

Figure 1: Feynman diagram of the $W^{\pm}H \to W^{\pm}W^{\pm}W^{\mp}$ process.

to SM gauge couplings because of the mixing. In addition to the leading-order dimension-four (dim-4) operators, dimension-six (dim-6) effective operators as described in Refs. [16–18] are also considered. After electroweak symmetry breaking, the effective Lagrangian terms can be written as

$$
\begin{aligned}
\mathcal{L}^{(4)}_{hWW} &= \rho_h g m_W h W^{\mu} W_{\mu}, \\
\mathcal{L}^{(4)}_{hZZ} &= \rho_h \frac{g m_W}{2c_W^2} h Z^{\mu} Z_{\mu}, \\
\mathcal{L}^{(4)}_{HWW} &= \rho_H g m_W H W^{\mu} W_{\mu}, \\
\mathcal{L}^{(4)}_{HZZ} &= \rho_H \frac{g m_W}{2c_W^2} H Z^{\mu} Z_{\mu}, \\
\mathcal{L}^{(6)}_{HWW} &= \rho_H g m_W \frac{f_W}{2\Lambda^2} \left(W^{+}_{\mu\nu} W^{-\mu} \partial^{\nu} H + h.c. \right) - \rho_H g m_W \frac{f_{WW}}{\Lambda^2} W^{+}_{\mu\nu} W^{-\mu\nu} H, \\
\mathcal{L}^{(6)}_{HZZ} &= \rho_H g m_W \frac{c_W^2 f_W + s_W^2 f_B}{2c_W^2 \Lambda^2} Z_{\mu\nu} Z^{\mu} \partial^{\nu} H - \rho_H g m_W \frac{c_W^4 f_{WW} + s_W^4 f_{BB}}{2c_W^2 \Lambda^2} Z_{\mu\nu} Z^{\mu\nu} H,
\end{aligned}
$$

where h, H, W and Z are the fields of the light and heavy Higgs bosons and the W and Z bosons, respectively; m_W is the W boson mass; g is the SM coupling constant of the weak interaction; f_W, f_{WW}, f_B, and f_{BB} are anomalous couplings to W and B fields; ρ_h and ρ_H are scaling factors; $s_W = \sin\theta_W$ and $c_W = \cos\theta_W$, where θ_W is the weak mixing angle; and Λ is the scale below which the effective Lagrangian holds. For a light Higgs boson similar to the one in the SM, ρ_h is close to 1. The simplest two-Higgs-doublet model (2HDM) [20] has $\rho_h = \cos(\beta - \alpha)$ and $\rho_H = \sin(\beta - \alpha)$, where α is the mixing angle between the CP-even Higgs bosons, and β is the rotation angle, with $\tan\beta$ defined as the ratio of vacuum expectation values of the two Higgs doublets. In this analysis, the scaling factor ρ_H is set to 0.05 and the scale Λ is set to 5 TeV, which is much larger than the mass of the heavy Higgs boson in this search. The choice of ρ_H is motivated by the observation that $\rho_h \sim 1$. To further simplify the parameter space, the small terms of $O(s_W^2)$ and $O(s_W^4)$ are neglected, and the anomalous coupling coefficients f_B and f_{BB} are set to zero, following Ref.

[16]. The operator multiplied by the f_W anomalous coupling is proportional to derivatives of the Higgs field and hence production is enhanced with increasing Higgs boson momentum. Results for the heavy Higgs boson's production cross-section are provided as a function of the heavy Higgs boson's mass and the two BSM HVV coupling strengths $\rho_H f_W/\Lambda^2$ and $\rho_H f_{WW}/\Lambda^2$.

3 ATLAS detector and data samples

The ATLAS detector [21] at the LHC covers nearly the entire solid angle around the collision point.[1] It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadron calorimeters, and a muon spectrometer incorporating three large superconducting air-core toroidal magnets.

The inner-detector system (ID) is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, the first hit normally being in the insertable B-layer (IBL) installed before Run 2 [22, 23]. It is followed by the silicon microstrip tracker (SCT), which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$. The TRT also provides electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

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

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

Interesting events are selected by the first-level trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger [24]. The first-level trigger accepts events from the 40 MHz bunch-crossings at a rate below 100 kHz, which the high-level trigger further reduces in order to record events to disk at about 1 kHz.

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis coinciding with the axis of the beam pipe. The x-axis points from the IP towards the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r,ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. The distance in (η,ϕ) coordinates, $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, is also used to define cone sizes. Rapidity is defined as $y = (1/2)\ln[(E + p_z)/(E - p_z)]$, where E is the energy and p_z is the z-component of the momentum. Transverse momentum and energy are defined as $p_{\mathrm{T}} = p\sin\theta$ and $E_{\mathrm{T}} = E\sin\theta$, respectively.

The data used in this analysis were collected using single-lepton triggers during the 2015–2018 proton–proton (pp) collision running periods at a centre-of-mass energy of 13 TeV. Events are selected for analysis only if they are of good quality and if all the relevant detector components are known to have been in good operating condition, which corresponds to a total integrated luminosity of 139 fb^{-1} [25, 26]. The recorded events contain an average of 34 inelastic pp collisions per bunch-crossing.

An extensive software suite [27] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

4 Simulation of signal and background processes

Monte Carlo (MC) simulated event samples are used to model heavy Higgs boson signals and to estimate the SM background with two same-sign leptons and/or at least three prompt leptons. Data-driven methods are used to estimate charge-flip, non-prompt and photon-conversion backgrounds, as discussed in Section 7.

The heavy Higgs boson signal process $pp \to VH \to VVV$ was modelled at leading order (LO) in QCD by the MadGraph5_aMC@NLO 2.7.3 generator [28]. The full decays of V bosons were simulated in MadSpin [29]. Parton showers and hadronisation were handled by Pythia 8.244 [30] using the A14 set of tuned parameters [31] and the NNPDF2.3lo [32] parton distribution function (PDF). Events were filtered such that at least one same-sign lepton pair was produced. Each lepton was required to have transverse momentum larger than 18 GeV, and be within $|\eta| < 2.7$. The samples were produced with m_H from 300 GeV to 1.5 TeV, f_W from −2480 to 2510 and f_{WW} from −15000 to 15000. The event samples are normalised to calculations at next-to-leading order (NLO) using a Higgs characterisation model [33]. The NLO K-factor increases the expected event yields by a factor of 1.3, independently of the heavy Higgs boson's mass and BSM HVV coupling strengths.

Samples of diboson final states (VV) were simulated with the Sherpa 2.2.2 [34] generator, including off-shell effects and Higgs boson contributions, where appropriate. Fully leptonic final states were generated using matrix elements at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. Samples for the loop-induced processes $gg \to VV$ were generated using LO-accurate matrix elements for up to one additional parton emission for both the fully leptonic and semileptonic final states. The matrix element calculations were matched and merged with the Sherpa parton shower based on Catani–Seymour dipole factorisation [35, 36] using the MEPS@NLO prescription [37–40]. The virtual QCD corrections are provided by the OpenLoops library [41–43]. The NNPDF3.0nnlo set of PDFs was used [44], along with the dedicated set of tuned parton-shower parameters developed by the Sherpa authors.

The production of triboson (VVV) events was simulated with the Sherpa 2.2.2 generator using factorised gauge-boson decays. Matrix elements, accurate to NLO for the inclusive process and to LO for up to two additional parton emissions, were matched and merged with the Sherpa parton shower based on Catani–Seymour dipole factorisation using the MEPS@NLO prescription. The virtual QCD corrections for matrix elements at NLO accuracy are provided by the OpenLoops library. Samples were generated using the NNPDF3.0nnlo PDF set, along with the dedicated set of tuned parton-shower parameters developed by the Sherpa authors. Contributions with an off-shell W boson through $Wh \to WWW^*$ were generated using Powheg Box v2 [45–48] interfaced to Pythia 8.235 to model the parton shower with the NNPDF2.3lo PDF and the AZNLO set of tuned parameters [49].

The production of $t\bar{t}h$ events was modelled using the POWHEG BOX v2 generator at NLO with the NNPDF3.0NLO PDF set. The events were interfaced to PYTHIA 8.230 with parameters set according to the A14 tune and using the NNPDF2.3LO set of PDFs. The decays of bottom and charm hadrons were performed by EVTGEN 1.6.0 [50]. The production of $t\bar{t}V$, tWZ and tZq events was modelled using the MADGRAPH5_AMC@NLO 2.3.3 [28] generator at NLO with the NNPDF3.0NLO PDF. The events were interfaced to PYTHIA 8.210, which used the A14 tune and the NNPDF2.3LO PDF set. The decays of bottom and charm hadrons were simulated using the EVTGEN 1.2.0 program.

Events from $t\bar{t}$, V+jets and $V\gamma$ processes contribute to the $\ell^{\pm}\ell^{\pm}$ signal region when a lepton charge is mismeasured, or leptons are produced from non-prompt decays or photon conversions. These backgrounds are estimated with data-driven methods, as detailed in Section 7, with MC simulation used for validation and to estimate systematic uncertainties. These MC simulations are briefly introduced in the following paragraphs.

The production of $t\bar{t}$ events was modelled by the POWHEG BOX v2 generator at NLO with the NNPDF3.0NLO PDF set and the h_{damp} parameter[2] set to 1.5 times the top-quark mass, m_{top} [51]. The events were interfaced to PYTHIA 8.230 to model the parton showers, hadronisation, and underlying event, with parameters set according to the A14 tune and using the NNPDF2.3LO set of PDFs. The decays of bottom and charm hadrons were performed by EVTGEN 1.6.0 [50].

The production of V+jets events was simulated with the SHERPA 2.2.1 generator using NLO matrix elements for up to two partons, and LO matrix elements for up to four partons calculated with the Comix and OPENLOOPS libraries. They were matched with the SHERPA parton shower using the MEPS@NLO prescription and the set of tuned parameters developed by the SHERPA authors. The NNPDF3.0NNLO set of PDFs was used and the samples are normalised to a next-to-next-to-leading-order (NNLO) prediction [52].

The production of $V\gamma$ final states was simulated with the SHERPA 2.2.2 generator. Matrix elements at NLO accuracy in QCD for up to one additional parton and LO accuracy for up to three additional parton emissions were matched and merged with the SHERPA parton shower as described for the diboson processes.

The generated events were passed through a simulation of the ATLAS detector and its response [53] based on GEANT4 [54], and reconstructed using the same software framework as for data [27]. The effect of multiple interactions in the same and neighbouring bunch-crossings (pile-up) was modelled by overlaying each simulated hard-scattering event with inelastic pp events generated with PYTHIA 8.186 [55] using the NNPDF2.3LO set of PDFs and the A3 set of tuned parameters [56].

5 Object reconstruction and identification

Proton–proton interaction vertices are reconstructed from charged-particle tracks with transverse momenta $p_{\text{T}} > 500$ MeV [57, 58]. The vertex with the highest sum of squared transverse momenta of associated tracks is selected as the primary vertex of the hard interaction.

Electrons are reconstructed from topological clusters of energy deposits in the electromagnetic calorimeter which are matched to a track in the inner detector [59]. Electrons are required to have $p_{\text{T}} > 20$ GeV and to be reconstructed within $|\eta| < 2.47$, excluding electrons in the transition region between the barrel and endcap calorimeters ($1.37 < |\eta| < 1.52$). The electron identification is based on a multivariate

[2] The h_{damp} parameter is a resummation damping factor and one of the parameters that controls the matching of POWHEG matrix elements to the parton shower and thus effectively regulates the high-p_{T} radiation against which the $t\bar{t}$ system recoils.

likelihood-based discriminant that uses the shower shapes in the electromagnetic calorimeter and the associated ID track properties. Electrons are required to satisfy the *Tight* identification criterion for better rejection of non-prompt electrons [59]. Muon candidates are identified by matching ID tracks to full tracks or track segments reconstructed in the muon spectrometer or by using only information from the muon spectrometer outside of the ID acceptance [60]. Muons are required to have $p_{\mathrm{T}} > 20$ GeV, to be reconstructed within $|\eta| < 2.5$, and to satisfy the *Medium* cut-based identification criterion as defined in Ref. [61]. To have an origin compatible with the primary vertex, electrons (muons) must satisfy $|d_0/\sigma_{d_0}| < 5$ (3) and $|z_0 \sin(\theta)| < 0.5$ mm. Here d_0 is the transverse impact parameter, σ_{d_0} is its uncertainty, z_0 is the distance along the z-axis from the primary vertex to the point where d_0 is measured, and θ is the polar angle of the track. In order to reject leptons likely to have originated from non-prompt hadronic decays, leptons are required to satisfy a criterion based on ID and calorimeter isolation variables and the output of a boosted-decision tree (BDT) in a prompt-lepton-veto tagger algorithm [62]. Electrons must also pass a *charge misidentification suppression BDT* which rejects electrons likely to have a wrongly measured electric charge [59]. Furthermore, in order to reduce the number of electrons likely to have originated from photon conversion, additional requirements are applied to the electron candidate (referred to as 'photon-conversion electron suppression requirements') [59, 63]. It must not be associated with a reconstructed photon-conversion vertex in the detector material nor have a reconstructed displaced vertex with radius $r > 20$ mm whose reconstruction uses the track associated with the electron. Finally the electron candidate's track and the closest track in ΔR at the primary vertex or a conversion vertex must not have an invariant mass below 100 MeV.

The anti-k_t algorithm [64, 65] with a radius parameter of 0.4 is used to reconstruct small-radius (small-R) jets up to $|\eta| = 4.9$. It uses particle-flow objects, which combine tracking and calorimetric information, as input [66]. Energy- and η-dependent correction factors derived from MC simulations are applied in order to correct jets back to the particle level [67]. Jets are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. To suppress jets from pile-up, a jet vertex tagger [68] applied to jets with $p_{\mathrm{T}} < 120$ GeV and $|\eta| < 2.5$ uses information about tracks associated with the primary vertex and pile-up vertices. Jets containing b-flavoured hadrons are identified in the region $|\eta| < 2.5$ with a b-tagging algorithm based on a recurrent neural network [69]. It makes use of the impact parameters of tracks associated with the jet, the position of reconstructed secondary vertices and their compatibility with the decay chains of such hadrons. At the chosen working point, the b-tagging algorithm provides light-flavour (u,d,s-quark and gluon) and c-jet rejection factors of 33 and 3, respectively, for an average 85% b-jet tagging efficiency, as estimated from simulated $t\bar{t}$ events [69].

Hadronically decaying τ-leptons are reconstructed [70, 71] as jets by applying the anti-k_t algorithm with a radius parameter of 0.4 to noise-suppressed energy clusters. They are required to have exactly one or three tracks in the ID within a cone of size $\Delta R = 0.2$ around the jet axis, to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$, and to be outside the transition region between the barrel and endcap calorimeters ($1.37 < |\eta| < 1.52$). To prevent jets from being reconstructed and misidentified as τ-leptons, a multivariate approach using boosted decision trees, based on information from the calorimeters and tracking detectors, is employed. The 'medium' quality criteria described in Ref. [71] are applied. Hadronically decaying τ-leptons are only used in the analysis in the overlap-removal procedure described at the end of this section.

Large-radius (large-R) jets are reconstructed from noise-suppressed topological clusters (topoclusters) of calorimeter energy depositions [72], using the anti-k_t algorithm with a radius parameter of 1.0, with the topoclusters calibrated at the local hadronic scale [72]. Large-R jets are groomed using trimming [73, 74] to improve the jet mass resolution and its stability with respect to pile-up by discarding the softer components of jets that originate from initial-state radiation, pile-up interactions, or the underlying event.

Large-R jets are required to have $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.0$. Only large-R jets with a jet mass m_J between 50 GeV and 200 GeV are considered in the analysis.

The missing transverse momentum ($\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$) is defined as the negative vector sum of the transverse momenta of electrons, muons, hadronically decaying τ-leptons and small-R jets in the event, plus a 'soft-term' built from additional tracks associated with the primary vertex [75, 76]. The magnitude of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$.

An overlap-removal procedure is applied to the selected leptons and jets. Any hadronically decaying τ-lepton reconstructed closer than $\Delta R = 0.2$ to an electron or muon is removed. Electrons that fall within $\Delta R = 0.2$ of a selected muon are also discarded. For electrons and nearby small-R jets, the jet is removed if the separation between the electron and jet satisfies $\Delta R < 0.2$; the electron is removed if the separation satisfies $0.2 < \Delta R < 0.4$. For muons and nearby small-R jets, the jet is removed if the separation between the muon and jet satisfies $\Delta R < 0.2$ and the jet has less than three tracks or the energy and momentum differences between the muon and the jet are small; otherwise the muon is removed if the separation satisfies $\Delta R < 0.4$. Small-R jets that are reconstructed within a cone of size $\Delta R = 0.2$ around the axis of a hadronically decaying τ-lepton are removed. To prevent double-counting of energy from an electron inside the large-R jet, the large-R jet is removed if the separation between the electron and the large-R jet satisfies $\Delta R < 1.0$.

6 Event classification and selection

The experimental signature of the $\ell^{\pm}\nu\ell^{\pm}\nu qq$ signal process requires the presence of two same-sign leptons, $E_{\mathrm{T}}^{\mathrm{miss}}$, and depending on the reconstruction of the qq final state, two small-R jets or one large-R jet with an invariant mass close to 80 GeV. The selection requirements used to define the signal regions are optimised to maximise the sensitivity to the $\ell^{\pm}\nu\ell^{\pm}\nu qq$ signal process while reducing contributions from SM background processes.

Events are required to satisfy a logical OR of single-electron [77] and single-muon [78] triggers with p_{T} thresholds ranging from 20 GeV to 26 GeV and increasing from 2015 to 2018. All events must contain at least one lepton with $p_{\mathrm{T}} > 27$ GeV that triggered the event, which ensures that the trigger efficiency reached its plateau. Events are required to have exactly two nominal leptons meeting the object criteria described in Section 5 and they must have the same electric charge. The invariant mass of the dilepton system is required to be larger than 100 GeV to reduce the contributions coming from the Z + jets process. To reduce the SM background contributions from processes that have more than two leptons, a 'veto lepton' definition is introduced. Compared with the nominal lepton selection criteria, the veto electron (muon) p_{T} threshold is lowered to 7 (4.5) GeV, and the isolation, charge misidentification suppression BDT, and photon-conversion electron suppression requirements are removed. For veto electrons, the *Loose* likelihood-based identification definition is used [59]. For veto muons, the *Loose* cut-based identification definition is used [60]. Events with additional veto leptons are removed.

For the hadronically decaying W boson, the energy deposits of the two resulting jets are either well separated or can largely overlap in the detector, depending on the momentum of the parent boson. Thus the $W \to qq$ decay can either be reconstructed from two resolved small-R jets ($W \to jj$) for low-momentum bosons or identified as one merged large-R jet ($W \to J$) for higher momentum, boosted bosons. An event is assigned to the boosted category if it contains at least one large-R jet satisfying the object criteria described in

Section 5; otherwise the event is assigned to the resolved category. In turn, two signal regions are defined: the boosted signal region (boosted SR) and the resolved signal region (resolved SR).

In the boosted SR, the large-R jet with the highest p_{T} is selected as the candidate for the hadronically decaying W boson and must satisfy $p_{\mathrm{T}} > 200$ GeV. A boson tagger is subsequently applied to distinguish between jets from hadronically decaying W bosons (which decay to two partons), and jets originating from a single quark or gluon [79]. In this analysis, the boson tagger is configured to have 80% identification efficiency for the hadronically decaying W boson. An $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV selection is applied in this region.

In the resolved SR, at least two small-R jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ are required. The invariant mass of the dijet system, formed by the two small-R jets with the largest p_{T}, is required to be consistent with the W boson mass: 50 GeV $< m_{jj} <$ 110 GeV. In both the boosted and resolved SRs, the events are required to have no b-tagged small-R jet present to reduce the background from top-quark production. An $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV selection is applied in this region.

One of the dominant sources of SM background in the SRs is the WZ + jets process. Control regions enriched in this process (WZ CRs) are defined for both the boosted and resolved categories, and are used in the global likelihood fit as detailed in Section 9, in order to constrain the normalisation of this background. The WZ CRs require events with three leptons, of which two form a same-flavour opposite-sign (SFOS) pair. Similarly to the SRs, the WZ CRs veto any events with b-jets to reduce backgrounds coming from top-quark processes and they veto events containing a fourth veto lepton to reduce background coming from the ZZ process. Requirements of $E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV and trilepton invariant mass $m_{\ell\ell\ell} > 100$ GeV are applied in order to further reduce background coming from Drell–Yan and ZZ processes. In the WZ CRs, more than 90% of the events are expected to be from WZ + jets production.

The same-sign WW + jets process is another important SM background in the SRs. A control region enriched with this background is defined (ssWW CR) in the resolved category by requiring a dijet invariant mass $m_{jj} > 200$ GeV, and is also used in the global likelihood fit as detailed in Section 9 to constrain the normalisation of this background. The m_{jj} requirement enhances the fraction of same-sign WW + jets events from electroweak VBF production, as the two forward jets tend to have a large invariant mass [80]. In order to reduce the statistical uncertainties, the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is loosened to $E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV. In the ssWW CR, about 40% of the events are expected to originate from same-sign WW + jets production.

The event selection criteria for the signal and control regions are summarised in Table 1. The average product of acceptance times efficiency for signal events in the combined SR is roughly 0.2%–0.5%, with little variation over the probed mass range for a heavy Higgs boson.

It is not possible to reconstruct the heavy Higgs boson's mass because of the two undetected neutrinos in the final states. The 'effective mass' m_{eff} is found to be a powerful discriminant between the signal and most SM backgrounds since a high mass scale is expected for the signal, and it is thus used as the main observable to extract the signal in the statistical analysis described in Section 9. The effective mass is defined to be the scalar sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse momenta of the leptons and either the leading large-R jet or the leading two small-R jets for the boosted category and resolved category, respectively:

$$\text{boosted category } m_{\mathrm{eff}} = \sum_i p_{\mathrm{T}}^i(\text{lepton}) + p_{\mathrm{T}}(\text{leading J}) + E_{\mathrm{T}}^{\mathrm{miss}},$$

$$\text{resolved category } m_{\mathrm{eff}} = \sum_i p_{\mathrm{T}}^i(\text{lepton}) + p_{\mathrm{T}}(\text{leading j}) + p_{\mathrm{T}}(\text{sub-leading j}) + E_{\mathrm{T}}^{\mathrm{miss}}.$$

Table 1: Overview of the event selection criteria for the signal and control regions.

Selections	Boosted SR	Resolved SR	*ssWW* CR	Boosted *WZ* CR	Resolved *WZ* CR
Trigger	Single lepton				
Leptons	two same-sign leptons with $p_{\mathrm{T}} > 27, 20$ GeV			three leptons with $p_{\mathrm{T}} > 27, 20, 20$ GeV at least one SFOS lepton pair	
	zero additional veto leptons				
$m_{\ell\ell}$	> 100 GeV			-	
$m_{\ell\ell\ell}$	-			> 100 GeV	
b-jets	zero b-tagged small-R jets				
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 80 GeV	> 60 GeV	> 40 GeV		
Large-R jets	at least one large-R jet with $p_{\mathrm{T}} > 200$ GeV, $\lvert\eta\rvert < 2.0$ 50 GeV $< m_J <$ 200 GeV and pass 80% W-tagger WP	zero large-R jets with $p_{\mathrm{T}} > 200$ GeV, $\lvert\eta\rvert < 2.0$ 50 GeV $< m_J <$ 200 GeV		at least one large-R jet with $p_{\mathrm{T}} > 200$ GeV, $\lvert\eta\rvert < 2.0$ 50 GeV $< m_J <$ 200 GeV	zero large-R jets with $p_{\mathrm{T}} > 200$ GeV, $\lvert\eta\rvert < 2.0$ 50 GeV $< m_J <$ 200 GeV
Small-R jets	-	at least two small-R jets with $p_{\mathrm{T}} > 20$ GeV and $\lvert\eta\rvert < 2.5$		-	at least two small-R jets with $p_{\mathrm{T}} > 20$ GeV and $\lvert\eta\rvert < 2.5$
m_{jj}	-	50 GeV $< m_{jj} <$ 110 GeV	> 200 GeV	-	-

7 Background estimation

The SM processes that mimic the $\ell^{\pm}\nu\ell^{\pm}\nu qq$ signal signature can be mainly grouped into four categories:

- Processes that produce at least three prompt leptons or two prompt leptons with the same electric charge. The main contributions come from WZ + jets (referred to as 'WZ background'), same-sign WW + jets (referred to as 'ssWW background'), WWW (referred to as 'WWW background'), with other small contributions from ttV, tZq, tth, WWZ, WZZ, ZZZ (referred to as 'Other background'). These backgrounds are estimated with MC simulations, except for the backgrounds from WZ and ssWW production, for which the normalisations are corrected using data in dedicated CRs as defined in Section 6.
- Processes that produce two or three prompt charged leptons, but the charge of one lepton is misidentified (referred to as 'charge-flip background'). A data-driven method is used to estimate this background and details are provided in Section 7.1.
- Processes that have one or two non-prompt leptons originating either from misidentified jets or from semileptonic decays of heavy-flavour hadrons (referred to as 'non-prompt background'). A data-driven method is used to estimate this background and details are provided in Section 7.2
- The $W\gamma$ + jets or $Z\gamma$ + jets processes where the photon converts to an electron–positron pair (referred to as 'photon conversion background'). A data-driven method is used to estimate this background and details are provided in Section 7.3

7.1 Electron charge-flip background

The charge-flip background originates from processes that produce oppositely charged prompt leptons, where one lepton's charge is misidentified and results in final states reconstructed as having two same-sign

leptons. The charge-flip background is only significant for electrons and is mainly due to interactions of the electron with material in the ID. The dominant contributions for this background come from $t\bar{t}$, W^+W^-+jets, and Z+jets processes, and are strongly suppressed by the charge misidentification suppression BDT and the kinematic requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{\ell\ell}$.

The electron charge misidentification rate is measured in a data sample enriched in $Z \to e^+e^-$ events (referred to as the Zee CR) selected by requiring two nominal electrons with an invariant mass between 75 GeV and 105 GeV. Non-Zee backgrounds are estimated from the total number of events in two sideband regions, defined by 60 GeV $< m_{ee} <$ 75 GeV and 105 GeV $< m_{ee} <$ 120 GeV. The non-Zee backgrounds are then subtracted from the Z mass region. The sample contains mostly opposite-charge di-electron events, with a small fraction of same-sign di-electron events. The fraction of same-sign di-electron events is used to extract the charge misidentification rate as a function of the electron p_{T} and $|\eta|$ using a likelihood fit method described in Ref. [59], taking into account that either electron in the same-sign pair could be the misidentified one. This rate is found to range between 0.01% and 4%, where higher values are obtained at large rapidities because of the larger amount of material traversed by the electrons, and at high p_{T} because of the larger probability of an incorrect determination of the track curvature. The charge-flip background is estimated in a given region by applying the misidentification rates to data events satisfying all selection criteria except that the two electrons must be oppositely charged.

The statistical uncertainty of this estimate varies between 1% and 10%. Additional systematic uncertainties are considered by comparing the estimated nominal rate with the rate derived by: i) varying the sidebands by 4 GeV, ii) using the Z+jets MC simulation directly, and iii) using MC simulation for background subtraction in the Zee CR. The impact of systematic uncertainties on the charge-flip background yield is approximately 10%, and is dominated by the uncertainty from using the Z+jets MC simulation directly.

7.2 Non-prompt background

The estimation of the non-prompt background assumes that these contributions can be extrapolated from a fake-lepton CR, enriched in non-prompt leptons, with a so-called fake-factor. Events that pass the kinematic requirements of the signal regions but contain one nominal lepton and one 'jet-like' lepton are selected in the fake-lepton CR. Jet-like electrons have to satisfy the likelihood-based *Medium* identification, while the isolation requirement is removed. Jet-like muons have the impact parameter requirement loosened to $|d_0/\sigma_{d_0}| < 10$, and the isolation requirement removed. Jet-like leptons are also required to fail at least one of the nominal lepton selections to ensure that the definitions of nominal and jet-like leptons are mutually exclusive. Simulation shows that the dominant contribution to this background stems from real muons or electrons from heavy-flavour hadrons that undergo semileptonic decays, and is heavily suppressed by the isolation and zero b-tagged small-R jet requirements, as well as the kinematic requirements on $m_{\ell\ell}$.

Events in the fake-lepton CR are scaled by fake-factors to predict the non-prompt-lepton background in the SR. The fake-factors are calculated in control regions with selections designed to enhance the contribution from backgrounds with non-prompt leptons. The control region selections require two same-sign leptons and exactly one b-tagged small-R jet. One of the same-sign leptons must fulfil either the nominal criteria or those of a jet-like lepton, while the other lepton must satisfy the nominal lepton criteria. The fake-factor is defined as the ratio of the number of events in the control region with all selected leptons fulfilling the nominal lepton criteria, to the number of events in the same region with one of the selected leptons satisfying the requirements of a jet-like lepton. The fake-factors are calculated separately for electrons and muons as a function of the lepton p_{T} and $|\eta|$. The SM processes with prompt leptons and the charge-flip

contributions are subtracted in the CR. For the electron and muon fake-factor measurements, the lepton with the second-highest p_{T} is assumed to be the non-prompt one. This assumption is true for more than 90% of events, based on generator-level information in the MC event record, and the potential bias can be covered by the systematic uncertainties as discussed below. The fake-factor dependency on the electron $|\eta|$ is found to be negligible. The electron fake-factors are then measured in three different electron p_{T} bins separated by boundaries at 30 GeV and 40 GeV. The statistical uncertainty is found to be approximately 20% in each bin. A strong $|\eta|$ dependency is found for the muon fake-factors, and their values are estimated in three $|\eta|$ bins: $0 < |\eta| < 0.5$, $0.5 < |\eta| < 1.5$ and $|\eta| > 1.5$. The statistical uncertainty is approximately 20% in the first two bins, and approximately 30% in the last bin. The fake-factor dependency on the muon p_{T} was also checked and found to be negligible.

Apart from the statistical uncertainty, a set of systematic uncertainties is also considered for the estimation of the fake-factors as follows: i) estimating the fake-factors in the inclusive p_{T} and $|\eta|$ region; ii) varying the normalisation of the SM processes with prompt leptons and electron charge-flip background when doing the subtraction; iii) varying the b-tagging algorithm working points used for the $t\bar{t}$-enriched CR definition; iv) estimating the fake-factors with MC simulation directly in both the SRs and $t\bar{t}$-enriched regions, and treating the difference as a systematic uncertainty to take into account any potential fake-factor difference between SRs and $t\bar{t}$-enriched regions. The overall systematic uncertainty amounts to approximately 13% (10%) for the electron (muon) fake-factors, with the dominant contribution coming from fake-factor estimation in the inclusive p_{T} and $|\eta|$ region.

7.3 Photon conversion background

The photon conversion background can contribute in the SRs if the photon is misreconstructed as an electron. This background originates primarily from the $W\gamma$ process and is evaluated using a data-driven method similar to the non-prompt-lepton background estimation by introducing 'photon-like' electrons. A photon-like electron is an object reconstructed like a nominal electron except that the track has no hit in the innermost layer of the pixel detector and the photon-conversion electron suppression requirements are not applied. In order to determine the photon fake-factors, a $Z\gamma$-enriched region is selected by requiring two nominal muons, no b-tagged small-R jets, and one nominal or photon-like electron. The trilepton invariant mass must satisfy 80 GeV $< m_{\mu\mu e} <$ 100 GeV. The photon fake-factor is defined as the ratio of the number of events in the $Z\gamma$-enriched region with the selected electron required to fulfil nominal electron criteria, to the number of events in the same region with the selected electron satisfying photon-like electron requirements. The SM processes with prompt leptons are subtracted in the $Z\gamma$-enriched region.

The photon fake-factors are measured in two electron p_{T} bins separated by a boundary at 25 GeV. The statistical uncertainty is found to be approximately 10% in each bin. The fake-factor dependency on the electron $|\eta|$ was also checked and found to be negligible. A photon-conversion electron CR is then filled with events passing the signal region kinematic requirements, but containing one nominal lepton and one photon-like electron. Events in this CR are scaled by the photon fake-factor to predict the photon conversion background in the SR.

In a similar way to the non-prompt background, the photon fake-factor derived from the inclusive p_{T} region is considered as one of the systematic uncertainties, together with the uncertainties from background subtraction. In addition, possible differences between $W\gamma$ and $Z\gamma$ photon fake-factors are checked with MC simulation and found to be negligible. The overall systematic uncertainty is found to be approximately 8%.

7.4 Validation of background estimates

Two validation regions (VR) are used to test the general background predictions in the boosted and resolved categories. They are defined to be close to the signal regions, with the large-R jet W-tagging requirement inverted in the boosted category and the m_{jj} requirement inverted in the resolved category. Events with $m_{jj} > 200$ GeV in the resolved category are removed in order to avoid overlap with the ssWW CR. Kinematic distributions are checked and good agreement between the data and the prediction is observed in the boosted and resolved VRs, as shown in Figure 2. Data-driven methods detailed in this section are used to estimate the charge-flip, non-prompt and photon-conversion backgrounds in the VRs. Predictions from simulation are scaled to the integrated luminosity of the data using the theoretical cross-section of each sample. The WZ and ssWW backgrounds are also scaled by normalisation factors from the global likelihood fit as detailed in Section 9. Only the statistical uncertainty is shown in Figure 2.

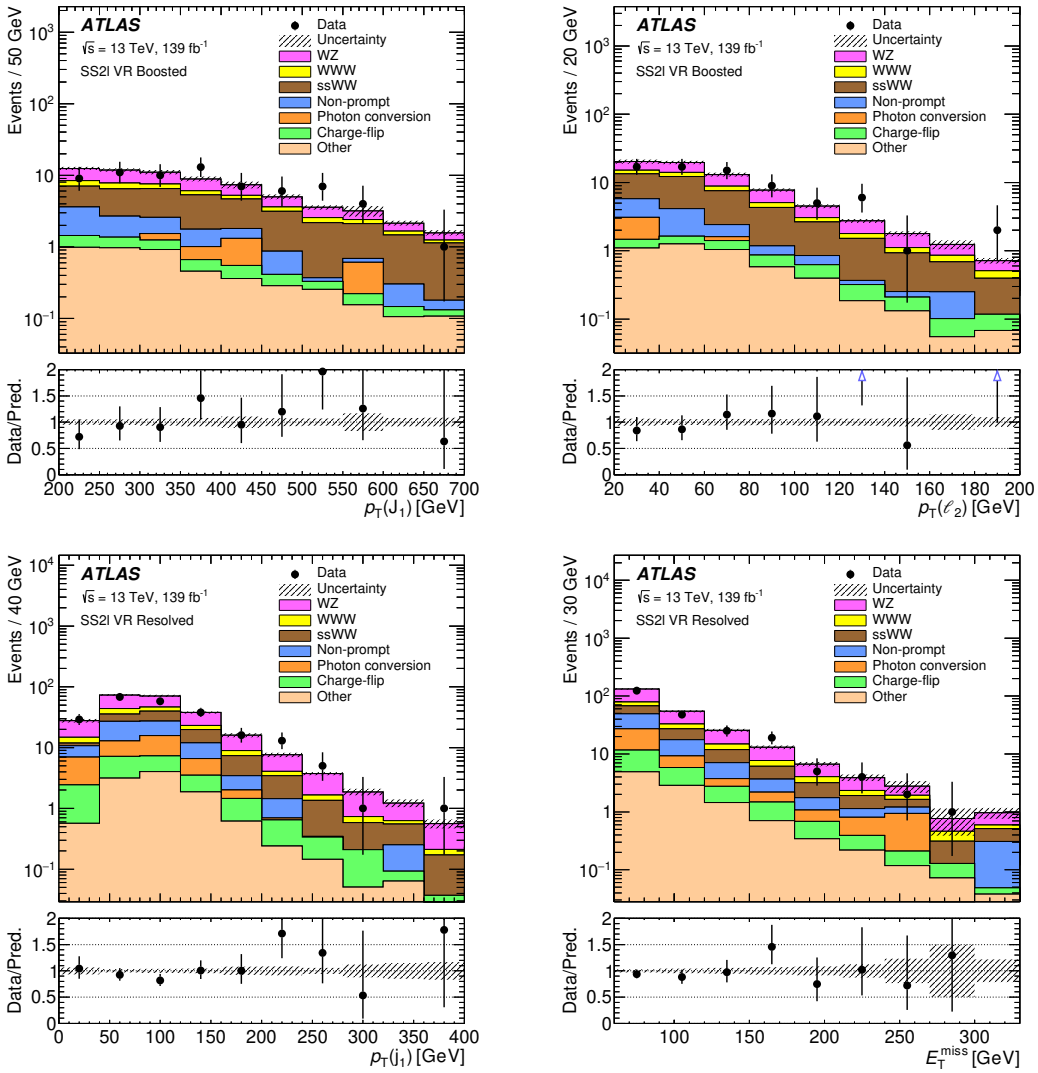

Figure 2: Comparison between data and SM predictions for the large-R jet p_{T} (top left) and the sub-leading lepton p_{T} (top right) in the boosted VR, and for the leading small-R jet p_{T} (bottom left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (bottom right) in the resolved VR. Predictions from simulation are scaled to the integrated luminosity of the data using the theoretical cross-section of each sample. WZ and ssWW backgrounds are also scaled by the normalisation factors from the global likelihood fit. The background predictions are shown as filled histograms. The size of the statistical uncertainty for the sum of the backgrounds is indicated by the hatched band. The lower panel displays the ratio of data to the total prediction. The blue triangles indicate bins where the ratio is non-zero and outside the vertical range of the plot.

8 Systematic uncertainties

The sources of systematic uncertainty can be broadly divided into three groups: those of experimental nature and related to the detector and reconstruction performance, those of theoretical origin and associated with modelling of the simulated background and signal processes, and those related to the data-driven background estimation.

8.1 Experimental uncertainties

Experimental uncertainties arise from the measurement of the luminosity, the modelling of pile-up in the simulation, the trigger selection, the reconstruction and identification of electrons, muons and jets, and the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

The uncertainty in the combined 2015–2018 integrated luminosity is 1.7%. It is derived following a methodology described in Ref. [25], and using the LUCID-detector for the baseline luminosity measurements [81]. An uncertainty associated with the modelling of pile-up in the simulation is included to cover the difference between the predicted and measured inelastic pp collision cross-sections [82].

Uncertainties in the reconstruction, identification, isolation and trigger efficiencies of electrons [59] and muons [60] are considered, along with the uncertainty in their energy scale and resolution. These are found to have only a small impact on the result. The lepton and jet identification efficiencies are well modelled in the simulation, and remaining differences are corrected to values measured in data. The uncertainties in the energy scale and resolution of the jets and leptons are propagated to the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$, which also has additional uncertainties from the modelling of the underlying event and momentum scale, momentum resolution and reconstruction efficiency of the tracks used to compute the soft-term [75, 76].

The uncertainties in the small-R jet energy scale and resolution have contributions from in situ calibration studies, from the dependency on pile-up activity and on the flavour composition of the jets [67, 83].

Uncertainties in the efficiencies for tagging b-jets and for mis-tagging light-flavour jets are determined from $t\bar{t}$ and jet control samples, respectively [69, 84, 85]. For large-R jets, the uncertainties in the energy and mass scales rely on a comparison of the ratio of calorimeter-based to track-based measurements in dijet data and simulation, as described in Ref. [86]. The efficiency of the W boson tagging is estimated using data control samples, following the technique described in Ref. [87]. The efficiency for large-R jet selection from W boson decays is estimated using $t\bar{t}$ control samples for $p_{\mathrm{T}} < 600$ GeV. The measurement is extrapolated to the higher p_{T} region with additional uncertainties estimated from simulations [79]. The efficiency for background large-R jets from gluons or light quarks is estimated using dijet and γ+jets samples.

8.2 Theoretical uncertainties

Theoretical uncertainties affect the normalisations and shapes of m_{eff} distributions of signal and background processes. They arise from sources such as our choices of event generators, PDFs, parton shower models, and underlying-event tunes. The effects of scale and PDF uncertainties are estimated by varying the renormalisation/factorisation scales and PDF sets, respectively. The parton shower uncertainty is evaluated at generator level by comparisons of different parton showers or corresponding scales.

The normalisations of the WZ background, separated into the resolved and boosted categories, as well as the ssWW background normalisation, are free to float in the global likelihood fit, as detailed in Section 9. The theoretical uncertainties of these two backgrounds are not applied to the corresponding CRs, since only normalisation information is used for these CRs in the global fit. Apart from their impact on the shape of the m_{eff} distribution, theoretical uncertainties in the WZ and ssWW backgrounds also impact the SR normalisations, and can be treated as uncertainties in extrapolating from high-purity CRs to the SRs.

The combined effect of the scale and PDF uncertainties, as well as the parton shower uncertainties of the WZ background, is calculated by adding in quadrature the differences between the nominal Sherpa 2.2.2 sample and its associated systematic variations, including variations of i) the renormalisation scale by factors of 0.5 and 2, ii) the factorisation scale by factors of 0.5 and 2, iii) the CKKW merging scale from 30 GeV to 15 GeV, and iv) the parton-shower/resummation scale by factors of 0.5 and 2. The total theoretical uncertainty in the WZ background yield in the boosted SR (resolved SR) is found to be 29% (20%), and is dominated by the scale uncertainty.

The same approach as used for the WZ background is also used to estimate the effect of scale and PDF uncertainties for the ssWW background. The effect of the parton shower uncertainty on the ssWW background is estimated by comparing the nominal MC sample with a sample generated with MadGraph5_aMC@NLO + Herwig 7. The total theoretical uncertainty in the ssWW background yield is 31% in the boosted SR and 25% in the resolved SR.

The estimation of the theoretical uncertainties in the on-shell WWW background, and of the effect of PDF and scale uncertainties on the off-shell WWW ($Wh \rightarrow WWW^*$) background, uses the same approach as for the WZ backgrounds. The effect of the parton shower uncertainty on the off-shell WWW background is estimated by comparing the nominal sample generated by Powheg Box + Pythia 8 with a sample generated by Powheg Box + Herwig 7. The total theoretical uncertainty in the WWW background yield is 16% in the boosted SR and 8% in the resolved SR.

The theoretical uncertainties of the WZ, ssWW and WWW backgrounds are decorrelated between the resolved and boosted regions in the global likelihood fit to allow for possible differences between the two regions. Given the small contributions from the processes included in the 'Other' background category, only overall normalisation uncertainties are assigned. The uncertainties vary from 10% to 20% based on the latest measurements of these processes [88–91].

For the signal samples, the effects of scale and PDF uncertainties are estimated by varying the renormalisation/factorisation scales, as well as the PDF set and parameter values used for the nominal MC samples. Parton shower uncertainties are estimated by comparing the nominal samples (MadGraph5_aMC@NLO + Pythia 8) with alternative samples using a different parton-shower generator (MadGraph5_aMC@NLO + Herwig 7). The total theoretical uncertainty in the yields from different signal samples varies between 10% and 40% in the SRs.

8.3 Data-driven background estimation uncertainties

Uncertainties in data-driven background evaluations come mainly from statistical and systematic uncertainties in the charge misidentification rate, lepton fake-factor, and photon-like electron fake-factor. More details can be found in Section 7.

9 Results

9.1 Statistical analysis

The statistical analysis is based on the HistFitter framework [92]. A binned likelihood function is constructed as the product of Poisson probability terms over the bins of the input distributions involving the numbers of data events and the expected signal and background yields, taking into account the effects of the floating background normalisations and the systematic uncertainties. A profile-likelihood-ratio test statistic is used to determine whether the background-only hypothesis is compatible with the observed data. The signal-plus-background hypothesis for the production of a heavy Higgs boson is tested, parameterised with the signal-strength parameter, μ, defined as the ratio of the extracted cross-section to the injected hypothesised signal cross-section. Maximum-likelihood fits to the observed binned distributions of the m_{eff} discriminants in the two SRs and to the numbers of observed events in CRs are performed simultaneously. The m_{eff} distributions are divided into 3 (5) bins for the boosted (resolved) SRs. The bins are of variable size to optimise the fit performance, while keeping the statistical uncertainty of the background contributions in each bin no larger than 10%. The normalisations of the WZ background, separated into boosted and resolved regions, as well as the normalisations of the ssWW background, are free parameters in these fits and are constrained by the data in both the high-purity CRs and the SRs. The effect of systematic uncertainties in the signal and background predictions is described by nuisance parameters, which are constrained by Gaussian or log-normal probability density functions. For each nuisance parameter, the constraint is added as a penalty term to the likelihood, which decreases as soon as the nuisance parameter is shifted away from its nominal value. The statistical uncertainties of background predictions from simulation are included through one nuisance parameter per bin, using the Beeston–Barlow technique [93].

9.2 Data and background comparisons

To test the compatibility of the data and the background expectations, the data are first fit to the background-only hypothesis. Good agreement between the data and the post-fit background contributions is found for the m_{eff} distributions in the SRs and event yields in CRs. The post-fit normalisation factors of the unconstrained WZ background in the boosted and resolved regions are 0.93 ± 0.07 and 0.83 ± 0.03, respectively. For the ssWW background, the extracted normalisation factor is 1.54 ± 0.18. The errors represent the combined statistical and systematic uncertainties, but do not include theoretical uncertainties related to normalisation in the respective CRs. The extracted normalisation factors are consistent with the results from dedicated WZ + jets [94] and ssWW + jets [80] measurements. Table 2 shows the post-fit background event yields from different sources in all SRs and CRs, compared with the numbers of events in data. The post-fit m_{eff} distributions in the SRs are shown in Figure 3, where good agreement between the data and the post-fit background contributions is observed. Figure 4 shows a few selected post-fit kinematic distributions in the SRs. No significant discrepancies are observed.

Good overall normalisation agreement between data and post-fit background contributions in CRs is seen in Table 2, and a few selected post-fit kinematic distributions in the CRs are shown in Figure 5.

Table 3 summarises the systematic uncertainties in the background expectation for each SR. The individual sources of systematic uncertainty detailed in Section 8 are combined into categories. In the resolved SR, the largest uncertainty comes from the data-driven background estimation, followed by theoretical uncertainties in background modelling, the uncertainty due to the limited size of the simulated samples, and

Table 2: Background predictions and data yields for each signal region and control region. The background predictions are obtained through a background-only simultaneous fit. All systematic uncertainties are included. The individual uncertainties can be correlated, and do not necessarily add in quadrature to equal the total background uncertainty. An entry of '–' indicates that a specific background component is negligible in a certain region, or that no simulated events are left after the analysis selections.

Yields	Boosted SR	Resolved SR	Boosted WZ CR	Resolved WZ CR	$ssWW$ CR
Observed events	24	191	236	2094	567
Fitted bkg events	26.8 ± 2.7	189.0 ± 7.8	235 ± 15	2095 ± 46	566 ± 24
WWW	5.8 ± 1.0	30.4 ± 2.9	1.30 ± 0.31	11.2 ± 2.1	28.5 ± 5.5
ssWW	7.5 ± 2.3	16.5 ± 1.9	–	–	254 ± 27
WZ	6.71 ± 0.76	68.7 ± 5.0	221 ± 15	1956 ± 50	150.6 ± 5.7
Non-prompt	3.20 ± 0.36	39.6 ± 6.3	–	–	48.6 ± 8.8
Charge-flip	0.43 ± 0.03	8.61 ± 0.57	–	–	22.8 ± 1.3
Photon conversion	0.73 ± 0.07	17.2 ± 1.7	–	–	46.7 ± 4.7
Other	2.50 ± 0.45	9.0 ± 1.5	12.3 ± 1.6	130 ± 20	14.3 ± 2.0

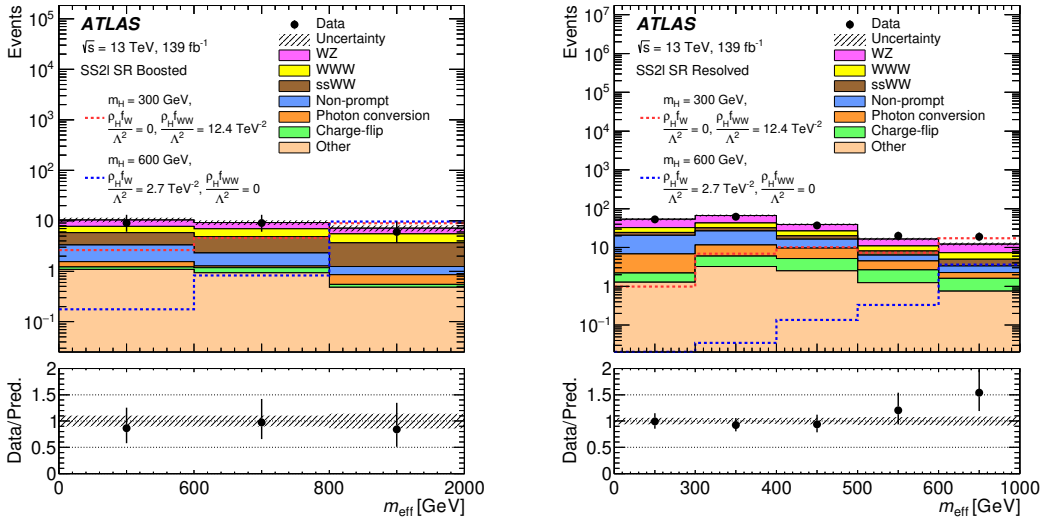

Figure 3: Comparison between data and SM predictions for the m_{eff} distributions in the boosted SR (left) and the resolved SR (right). The background predictions are obtained through a background-only simultaneous fit and are shown as filled histograms. The last bin includes overflow entries. The size of the combined statistical and systematic uncertainty for the sum of the fitted backgrounds is indicated by the hatched band. The ratio of the data to the sum of the fitted backgrounds is shown in the lower panel. Two benchmark signal samples, as indicated in the legend, are also shown as unstacked unfilled histograms normalised to the integrated luminosity of the data using the theoretical cross-sections.

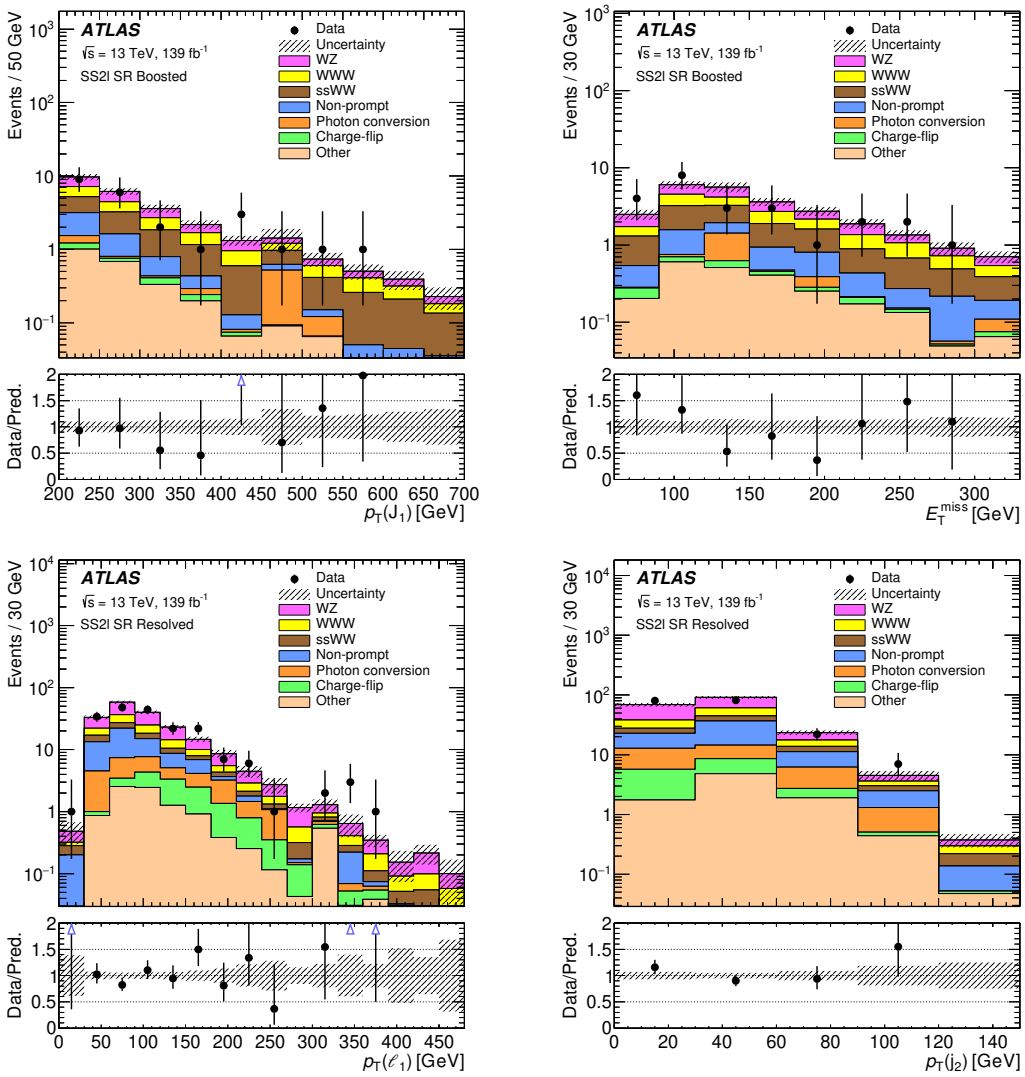

Figure 4: Comparison between data and SM predictions for the large-R jet p_{T} (top left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (top right) in the boosted SR, and for the leading lepton p_{T} (bottom left) and the sub-leading small-R jet p_{T} (bottom right) in the resolved SR. The background predictions are obtained through a background-only simultaneous fit and are shown as filled histograms. The size of the combined statistical and systematic uncertainty for the sum of the fitted backgrounds is indicated by the hatched band. The ratio of the data to the sum of the fitted backgrounds is shown in the lower panel. The blue triangles indicate bins where the ratio is non-zero and outside the vertical range of the plot.

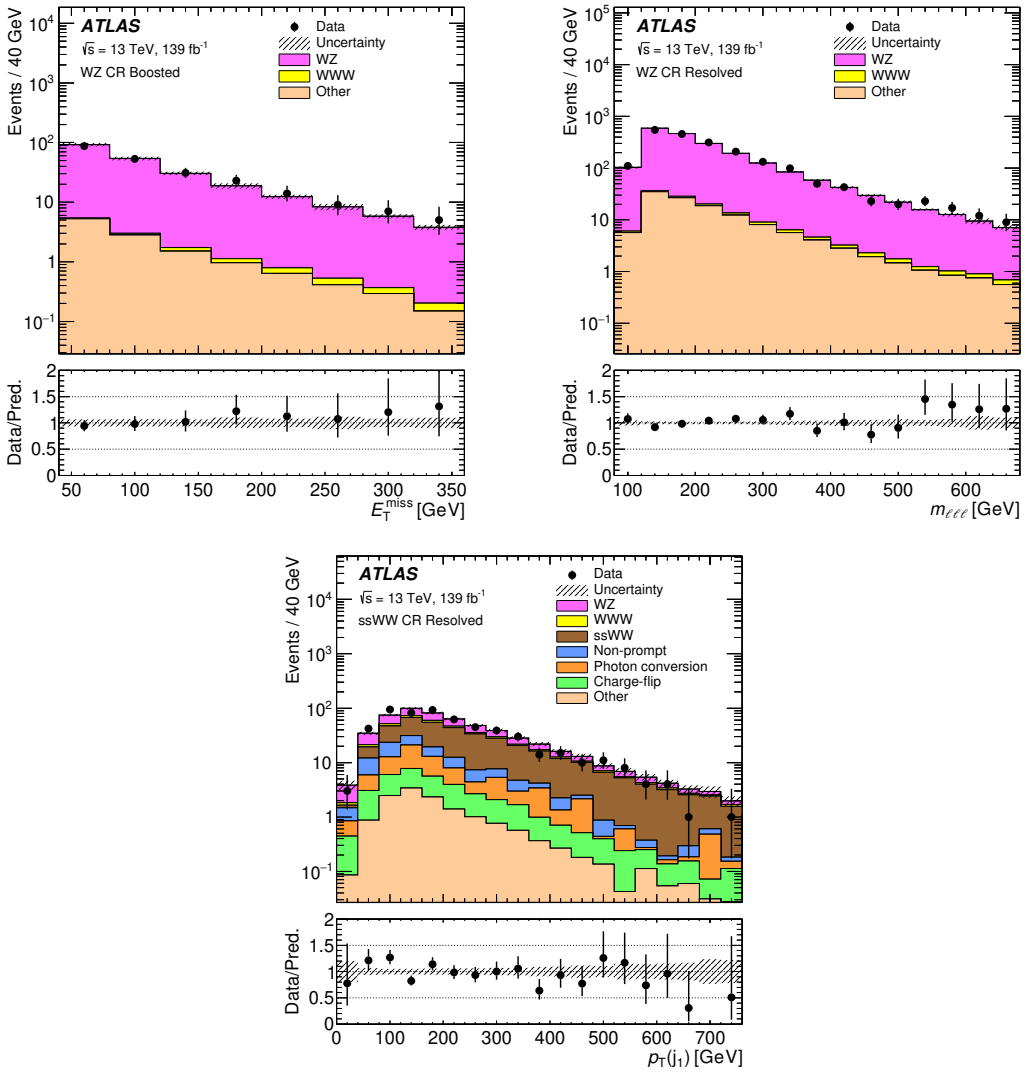

Figure 5: Comparison between data and SM predictions for $E_{\mathrm{T}}^{\mathrm{miss}}$ (top left) in the boosted WZ CR, for the invariant mass of the three-lepton system (top right) in the resolved WZ CR, and for the leading small-R jet p_{T} (bottom) in the ssWW CR. The background predictions are obtained through a background-only simultaneous fit and are shown as filled histograms. The size of the combined statistical and systematic uncertainty for the sum of the fitted backgrounds is indicated by the hatched band. The ratio of the data to the sum of the fitted backgrounds is shown in the lower panel.

Table 3: Breakdown of the dominant systematic uncertainties in background estimates in both the boosted and resolved signal regions. The background predictions are obtained through a background-only simultaneous fit. The individual uncertainties can be correlated, and do not necessarily add in quadrature to equal the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background. An entry of '–' indicates that a specific uncertainty component is not relevant in a certain region.

Uncertainty of channel	Boosted SR	Resolved SR
Total systematic uncertainties	10.0%	4.1%
Data driven non-prompt	1.3%	3.3%
Theoretical uncertainties	8.9%	2.6%
MC statistical uncertainties	3.0%	1.9%
Floating normalisations	3.5%	1.2%
Data driven photon conversion	0.2%	0.9%
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.2%	0.7%
b-tagging	0.8%	0.5%
Data driven charge-flip	0.1%	0.3%
Electron	0.5%	0.2%
Muon	0.6%	0.2%
Pile-up reweighting	0.2%	0.2%
Large-R jet	1.1%	0.2%
W-tagger	3.7%	–
Small-R jet	–	1.1%

the small-R jet uncertainty. In the boosted SR, the systematic uncertainties associated with the theoretical modelling of the background and with the W-tagger play a dominant role, followed by sizeable effects from the limited size of the simulated samples and the large-R jet uncertainty.

9.3 Limits on the production of heavy Higgs bosons

Constraints on the production of heavy Higgs bosons are derived by repeating the fit to the signal-plus-background hypothesis. Upper limits on the production cross-sections of heavy Higgs bosons are calculated with a modified frequentist method [95], known as $\mathrm{CL_s}$, using the $\tilde{q}_\mu$ test statistic in the asymptotic approximation [96].

Figure 6 shows the expected and observed exclusion contours at the 95% confidence level (CL) for signals from heavy Higgs bosons with masses of 300 GeV, 600 GeV and 900 GeV as a function of the coupling strengths $\rho_H f_W/\Lambda^2$ and $\rho_H f_{WW}/\Lambda^2$. The scaling factor ρ_H is set to 0.05 and the scale Λ is set to 5 TeV as mentioned in Section 2. The hypotheses are tested for each mass value in each of the 16 radial directions of the (f_W, f_{WW}) space. The local p_0-value for the observation to be compatible with the background-only hypothesis reaches its smallest value at 300 GeV with $(\rho_H f_W/\Lambda^2, \rho_H f_{WW}/\Lambda^2) = (0, 4.9\ \mathrm{TeV}^{-2})$, corresponding to 1.3 standard deviation. For a heavy Higgs boson with a mass of 300 GeV, $|\rho_H f_W/\Lambda^2| > 2.7\ \mathrm{TeV}^{-2}$ and $|\rho_H f_{WW}/\Lambda^2| > 10\ \mathrm{TeV}^{-2}$ can be excluded at 95% CL. Couplings of $|\rho_H f_W/\Lambda^2| > 2.5\ \mathrm{TeV}^{-2}$ and $|\rho_H f_{WW}/\Lambda^2| > 12\ \mathrm{TeV}^{-2}$ can be excluded for the production of a heavy

Higgs boson with a mass of 600 GeV. Similarly, for a heavy Higgs boson with a mass of 900 GeV, $|\rho_H f_W/\Lambda^2| > 2.9$ TeV^{-2} and $|\rho_H f_{WW}/\Lambda^2| > 15$ TeV^{-2} can be excluded.

The overall excess, at the level of approximately 1σ, observed for a 300 GeV heavy Higgs boson is mostly due to the small excess observed in data in the rightmost bin of the resolved SR's m_{eff} distribution, as shown in Figure 3. This is because the resolved SR dominates the sensitivity to lower-mass heavy Higgs bosons.

From the ellipse of expected limits in Figure 6, two sets of couplings $(\rho_H f_W/\Lambda^2, \rho_H f_{WW}/\Lambda^2)$ with values $(0, 12.4\ \text{TeV}^{-2})$ and $(2.7\ \text{TeV}^{-2}, 0)$ are chosen as benchmark examples with which to explore the dependence of the results on the heavy Higgs boson's mass. Coupling combinations on this ellipse are expected to have similar phenomenology. Although the two points chosen on the ellipse are somewhat arbitrary, they are representative. Figure 7 shows the expected and observed 95% CL upper limits on the heavy Higgs boson's production cross-section as a function of its mass for those two sets of anomalous couplings. The mass hypotheses are tested at 60 GeV steps between 300 and 900 GeV, and three additional mass points at 1000 GeV, 1200 GeV and 1500 GeV. For both of sets of couplings, the local p_0-value is smallest at 300 GeV, corresponding to 1.3 and 0.9 standard deviations, respectively. For the coupling choice $(\rho_H f_W/\Lambda^2, \rho_H f_{WW}/\Lambda^2) = (0, 12.4\ \text{TeV}^{-2})$, a heavy Higgs boson with mass up to 700 GeV can be excluded, while for the choice $(2.7\ \text{TeV}^{-2}, 0)$, the range of excluded masses extends to 900 GeV.

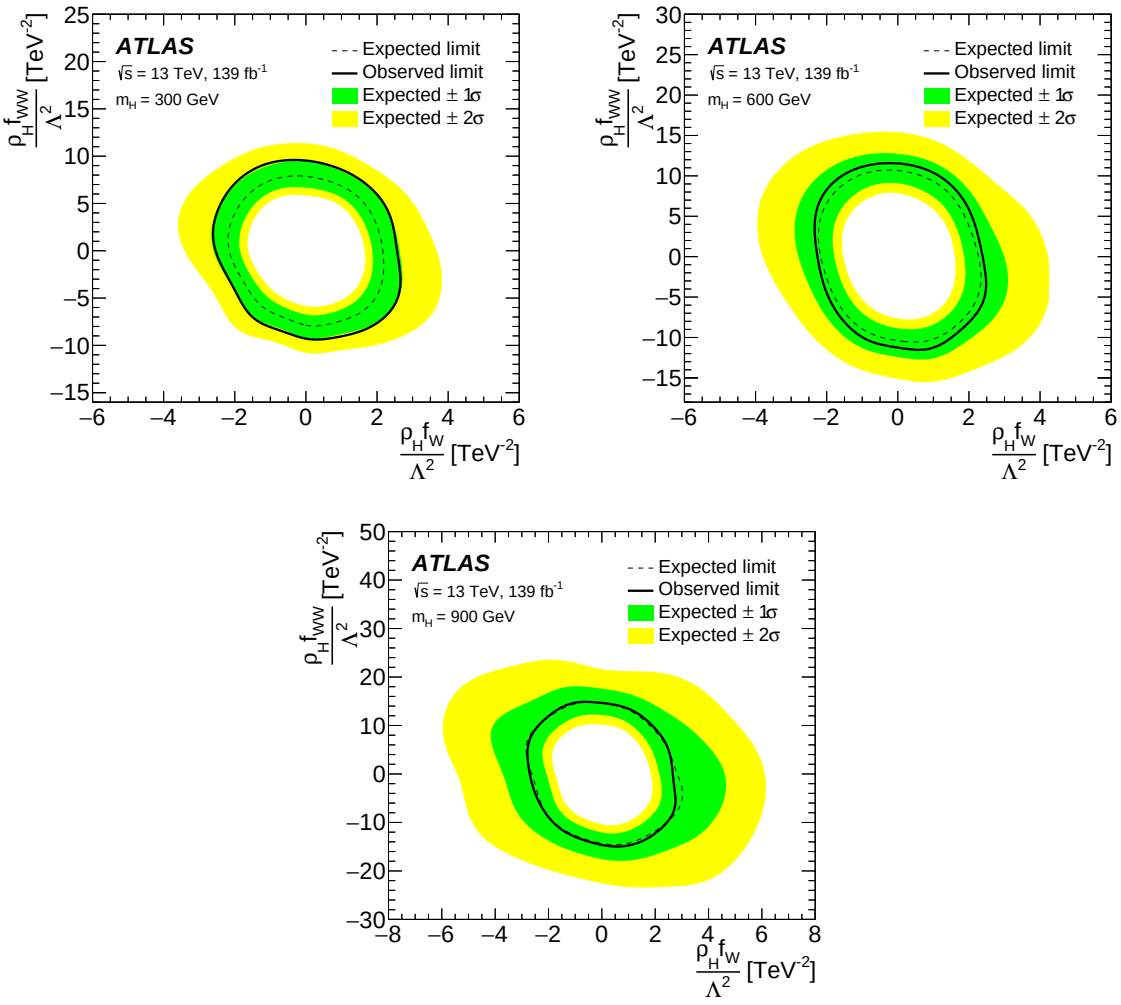

Figure 6: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production of a heavy Higgs boson as a function of $\rho_H f_W/\Lambda^2$ and $\rho_H f_{WW}/\Lambda^2$ for a mass of 300 GeV (top left), 600 GeV (top right) and 900 GeV (bottom). The green (inner) and yellow (outer) bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties of the expected limits.

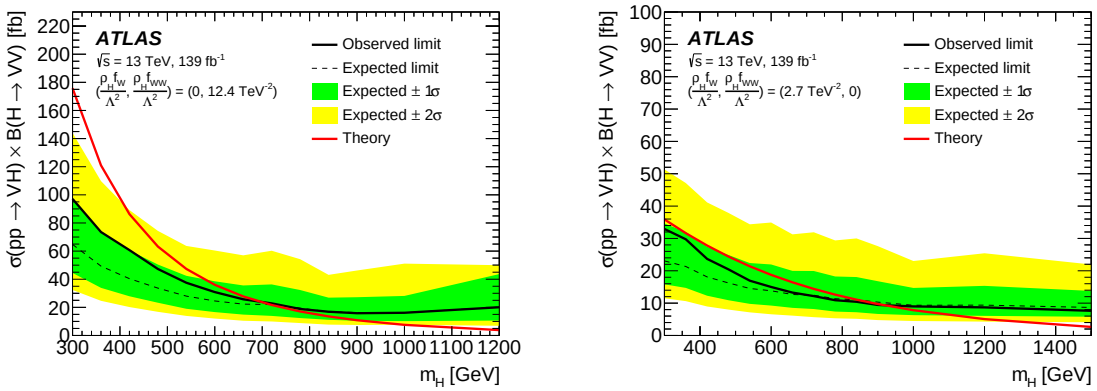

Figure 7: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross-section times decay branching fraction of a heavy Higgs boson as a function of its mass with $(\rho_H f_W/\Lambda^2, \rho_H f_{WW}/\Lambda^2)$ fixed to $(0, 12.4\ \text{TeV}^{-2})$ (left) and $(2.7\ \text{TeV}^{-2}, 0)$ (right). The green (inner) and yellow (outer) bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties of the expected limits. The unevenness in the expected limits reflects the variations in the estimated systematic uncertainties. Theoretical predictions (red solid curve) as a function of the heavy Higgs boson's mass are overlaid.

10 Summary

This paper presents a search for heavy Higgs bosons produced in association with a W boson and decaying into a pair of W bosons. The search uses proton–proton collisions at centre-of-mass energy of 13 TeV corresponding to an integrated luminosity of 139 fb^{-1}. The data were recorded by the ATLAS experiment between 2015 and 2018 at the LHC. The search is performed in the final states with two leptons of the same electric charge, missing transverse momentum and jets. The $W \to qq$ decay is reconstructed from two resolved small-R jets or one boosted large-R jet, and two corresponding signal regions are defined. The data are found to be in good agreement with the estimated backgrounds.

Upper limits on the production of heavy Higgs bosons are derived as a function of the heavy Higgs boson's mass and coupling strengths to vector bosons. For heavy Higgs bosons with masses ranging from 300 GeV to 900 GeV, the 95% CL upper limits on the coupling strengths $|\rho_H f_W/\Lambda^2|$ and $|\rho_H f_{WW}/\Lambda^2|$ are in the range 2.5–2.9 TeV^{-2} and 10–15 TeV^{-2}, respectively. The most stringent exclusion ranges for the coupling strengths, $|\rho_H f_W/\Lambda^2| > 2.5$ TeV^{-2} and $|\rho_H f_{WW}/\Lambda^2| > 10$ TeV^{-2}, are set for the production of heavy Higgs bosons with a mass of 600 GeV or 300 GeV, respectively. The scaling factor ρ_H is set to 0.05 and Λ is set to 5 TeV in the analysis. Heavy Higgs bosons are excluded at 95% CL for masses up to 700 GeV or 900 GeV with anomalous couplings to vector bosons $(\rho_H f_W/\Lambda^2, \rho_H f_{WW}/\Lambda^2)$ fixed at $(0, 12.4\ \text{TeV}^{-2})$ or $(2.7\ \text{TeV}^{-2}, 0)$, respectively.

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].

[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].

[3] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1, arXiv: 1106.0034 [hep-ph].

[4] G. Bhattacharyya and D. Das, *Scalar sector of two-Higgs-doublet models: A minireview*, Pramana **87** (2016) 40, arXiv: 1507.06424 [hep-ph].

[5] R. Aggleton, D. Barducci, N.-E. Bomark, S. Moretti and C. Shepherd-Themistocleous, *Review of LHC experimental results on low mass bosons in multi Higgs models*, JHEP **02** (2017) 035, arXiv: 1609.06089 [hep-ph].

[6] H. Davoudiasl, I. M. Lewis and M. Sullivan, 'Good things to do with extra Higgs doublets', *2022 Snowmass Summer Study*, 2022, arXiv: 2203.01396 [hep-ph].

[7] L. Wang, J. M. Yang and Y. Zhang, *Two-Higgs-doublet models in light of current experiments: a brief review*, Commun. Theor. Phys. **74** (2022) 097202, arXiv: 2203.07244 [hep-ph].

[8] J. Kim, S. Lee, P. Sanyal and J. Song, *CDF W-boson mass and muon g-2 in a type-X two-Higgs-doublet model with a Higgs-phobic light pseudoscalar*, Phys. Rev. D **106** (2022) 035002, arXiv: 2205.01701 [hep-ph].

[9] ATLAS Collaboration, *Search for doubly and singly charged Higgs bosons decaying into vector bosons in multi-lepton final states with the ATLAS detector using proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **06** (2021) 146, arXiv: 2101.11961 [hep-ex].

[10] ATLAS Collaboration, *Search for heavy resonances decaying into a pair of Z bosons in the* $\ell^+\ell^-\ell'^+\ell'^-$ *and* $\ell^+\ell^-\nu\bar{\nu}$ *final states using* $139\,fb^{-1}$ *of proton–proton collisions at* $\sqrt{s} = 13$, *TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2020) 332, arXiv: 2009.14791 [hep-ex].

[11] ATLAS Collaboration, *Search for heavy diboson resonances in semileptonic final states in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Eur. Phys. J. C **80** (2020) 1165, arXiv: 2004.14636 [hep-ex].

[12] ATLAS Collaboration, *Search for doubly charged scalar bosons decaying into same-sign W boson pairs with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 58, arXiv: 1808.01899 [hep-ex].

[13] ATLAS Collaboration, *Search for heavy resonances decaying into WW in the* $e\nu\mu\nu$ *final state in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Eur. Phys. J. C **78** (2018) 24, arXiv: 1710.01123 [hep-ex].

[14] CMS Collaboration, *Search for a heavy Higgs boson decaying to a pair of W bosons in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **03** (2020) 034, arXiv: 1912.01594 [hep-ex].

[15] CMS Collaboration, *Search for a heavy pseudoscalar boson decaying to a Z and a Higgs boson at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 564, arXiv: 1903.00941 [hep-ex].

[16] Y.-P. Kuang, H.-Y. Ren and L.-H. Xia, *Model-independent Probe of anomalous heavy neutral Higgs bosons at the LHC*, Phys. Rev. D **90** (2014) 115002, arXiv: 1404.6367 [hep-ph].

[17] Y.-P. Kuang, H.-Y. Ren and L.-H. Xia, *Further investigation of the model-independent probe of heavy neutral Higgs bosons at LHC Run 2*, Chin. Phys. C **40** (2016) 023101, arXiv: 1506.08007 [hep-ph].

[18] X. Chen et al., *Search for a generic heavy Higgs at the LHC*, Phys. Lett. B **804** (2020) 135358, arXiv: 1905.05421 [hep-ph].

[19] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, CERN-2013-004 (2013), arXiv: 1307.1347 [hep-ph].

[20] J. F. Gunion, H. E. Haber, G. L. Kane and S. Dawson, *The Higgs Hunter's Guide*, vol. 80, 2000.

[21] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[22] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*, ATLAS-TDR-19; CERN-LHCC-2010-013, 2010, URL: https://cds.cern.ch/record/1291633, Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009, 2012, URL: https://cds.cern.ch/record/1451888.

[23] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, JINST **13** (2018) T05008, arXiv: 1803.00844 [physics.ins-det].

[24] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[25] 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.

[26] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det].

[27] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATL-SOFT-PUB-2021-001, 2021, URL: https://cds.cern.ch/record/2767187.

[28] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph].

[29] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph].

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

[31] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

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

[33] F. Maltoni, K. Mawatari and M. Zaro, *Higgs characterisation via vector-boson fusion and associated production: NLO and parton-shower effects*, Eur. Phys. J. C **74** (2014) 2710, arXiv: 1311.1829 [hep-ph].

[34] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph].

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

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

[37] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: 1111.1220 [hep-ph].

[38] 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].

[39] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231.

[40] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph].

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

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

[43] 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].

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

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

[46] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[47] 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].

[48] 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].

[49] ATLAS Collaboration, *Measurement of the Z/γ^* 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].

[50] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

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

[52] C. Anastasiou, L. Dixon, K. Melnikov and F. Petriello, *High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order*, Phys. Rev. D **69** (2004) 094008, arXiv: hep-ph/0312266.

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

[54] GEANT4 Collaboration, S. Agostinelli et al., *Geant4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[55] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

[56] 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.

[57] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-026, 2015, URL: https://cds.cern.ch/record/2037717.

[58] 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].

[59] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex].

[60] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex].

[61] 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].

[62] ATLAS Collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. D **97** (2018) 072003, arXiv: 1712.08891 [hep-ex].

[63] ATLAS Collaboration, *Evidence for* $t\bar{t}t\bar{t}$ *production in the multilepton final state in proton–proton collisions at* $\sqrt{s} = 13$*TeV with the ATLAS detector*, Eur. Phys. J. C **80** (2020) 1085, arXiv: 2007.14858 [hep-ex].

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

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

[66] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex].

[67] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex].

[68] 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].

[69] 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].

[70] ATLAS Collaboration, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 295, arXiv: 1512.05955 [hep-ex].

[71] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s} = 13$ *TeV*, ATLAS-CONF-2017-029, 2017, URL: https://cds.cern.ch/record/2261772.

[72] 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].

[73] D. Krohn, J. Thaler and L.-T. Wang, *Jet trimming*, JHEP **02** (2010) 084, arXiv: 0912.1342 [hep-ph].

[74] ATLAS Collaboration, *Performance of jet substructure techniques for large-R jets in proton–proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, JHEP **09** (2013) 076, arXiv: 1306.4945 [hep-ex].

[75] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex].

[76] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015–2016 LHC pp collisions*, ATLAS-CONF-2018-023, 2018, URL: https://cds.cern.ch/record/2625233.

[77] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex].

[78] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv: 2004.13447 [hep-ex].

[79] ATLAS Collaboration, *Boosted hadronic vector boson and top quark tagging with ATLAS using Run 2 data*, ATL-PHYS-PUB-2020-017, 2020, URL: https://cds.cern.ch/record/2724149.

[80] ATLAS Collaboration, *Observation of Electroweak Production of a Same-Sign W Boson Pair in Association with Two Jets in pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector*, Phys. Rev. Lett. **123** (2019) 161801, arXiv: 1906.03203 [hep-ex].

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

[82] ATLAS Collaboration, *Measurement of the Inelastic Proton–Proton Cross Section at $\sqrt{s}$ = 13 TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002, arXiv: 1606.02625 [hep-ex].

[83] ATLAS Collaboration, *Jet energy resolution in proton–proton collisions at $\sqrt{s}$ = 7 TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2306, arXiv: 1210.6210 [hep-ex].

[84] ATLAS Collaboration, *Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s}$ = 13 TeV collected with the ATLAS detector*, Eur. Phys. J. C **82** (2021) 95, arXiv: 2109.10627 [hep-ex].

[85] 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.

[86] 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].

[87] ATLAS Collaboration, *A W/Z-boson tagger using Track-CaloCluster jets with ATLAS*, ATLAS-PHYS-PUB-2020-008, 2020, URL: https://cds.cern.ch/record/2718218.

[88] ATLAS Collaboration, *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **07** (2020) 124, arXiv: 2002.07546 [hep-ex].

[89] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072009, arXiv: 1901.03584 [hep-ex].

[90] 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].

[91] ATLAS Collaboration, *Evidence for the production of three massive vector bosons with the ATLAS detector*, Phys. Lett. B **798** (2019) 134913, arXiv: 1903.10415 [hep-ex].

[92] M. Baak, G. J. Besjes, D. Côté, A. Koutsman, J. Lorenz and D. Short, *HistFitter software framework for statistical data analysis*, Eur. Phys. J. C **75** (2015) 153, arXiv: 1410.1280 [hep-ex].

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

[94] ATLAS Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 535, arXiv: 1902.05759 [hep-ex].

[95] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693.

[96] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an], Erratum: Eur. Phys. J. C **73** (2013) 2501.