ATLAS CONF Note

ATLAS-CONF-2022-039

11th July 2022

Search for heavy Higgs bosons from a g2HDM in multilepton plus b-jets final states in pp collisions at 13 TeV with the ATLAS detector

The ATLAS Collaboration

A search for new heavy scalars with flavour-violating decays in final states with multiple leptons and b-tagged jets is presented. The results are interpreted in terms of a general two-Higgs-doublet-model involving an additional scalar with couplings to the top-quark and the three up-type quarks (ρ_{tt}, ρ_{tc}, and ρ_{tu}). The targeted signals lead to final states with either a same-sign top-quark pair, three top-quarks, or four top-quarks. The search is based on a dataset of proton-proton collisions at $\sqrt{s}$ = 13 TeV recorded with the ATLAS detector during Run 2 of the Large Hadron Collider, corresponding to an integrated luminosity of 139 fb^{-1}. Events are categorised depending on the multiplicity of light charged leptons (electrons or muons), total lepton charge, and a deep-neural-network-based categorisation to enhance the purity of each of the signals. A mild excess is observed over the Standard Model expectation corresponding to a local significance of 2.81 standard deviations for a signal with m_H = 1000 GeV and ρ_{tt} = 0.32, ρ_{tc} = 0.05, and ρ_{tu} = 0.85. Exclusion limits at 95% confidence are set on the mass and couplings of the heavy Higgs boson. Masses of an additional scalar boson m_H between 200-630 (200-840) GeV with couplings ρ_{tt} = 0.4, ρ_{tc} = 0.2, and ρ_{tu} = 0.2 are observed (expected) to be excluded at 95% confidence level. Additional interpretations are provided in models of R-parity violating supersymmetry, motivated by the recent flavour and $(g-2)_\mu$ anomalies.

ATLAS-CONF-2022-039
13 July 2022

1 Introduction

Several extensions of the Standard Model propose the augmentation of the Higgs sector by the addition of a second complex Higgs doublet [1, 2] (2HDM), giving rise to five Higgs bosons: two CP-even scalar fields h and H, one CP-odd pseudo-scalar A, and two charged fields $H^{\pm}$. The two CP-even scalars are expected to mix, however, the measurement of Higgs boson properties has revealed no deviations from the expectations of the Standard Model. This would mean that extra scalars from 2HDMs have to be either very heavy (decoupling limit) or have a vanishingly small mixing with the SM Higgs (alignment limit). In order to avoid flavour changing neutral Higgs (FCNH) couplings mediated by the SM Higgs, a discrete Z_2 symmetry is usually imposed. Searches for heavy scalars or pseudo-scalars with flavour-conserving decays have been performed extensively in ATLAS [3–10] and CMS [11–23]. However, if the Z_2 symmetry is dropped, alignment automatically emerges when all heavy Higgs quartic couplings are $O(1)$ [24]. Therefore, models without Z_2 symmetry can lead naturally to the alignment limit and predict FCNH couplings in the heavy Higgs sector, while respecting the SM-like nature of the $h(125)$ discovered at the LHC.

The search presented in this paper targets such a general two Higgs doublet model (g2HDM) without Z_2 symmetry, where the heavy Higgs bosons feature FCNH couplings. Only couplings involving top-quarks are considered: ρ_{tt}, ρ_{tc}, and ρ_{tu}. The ρ_{AB} parameters indicate the coupling of the heavy Higgs boson to particles A and B. The notation ρ_{tq} is used to refer to both the ρ_{tc} and ρ_{tu} couplings. These kinds of g2HDMs with extra top Yukawa couplings are phenomenologically interesting since they can explain the generation of the baryon asymmetry through the couplings ρ_{tt} or ρ_{tc} [25]. No distinction is performed between the different chiralities in the coupling, and an effective coupling $\rho_{tq} = \sqrt{\hat{\rho}^2_{t_L q_R} + \hat{\rho}^2_{q_L t_R}}/\sqrt{2}$ is used, where the hat symbol is used to denote the original couplings in the g2HDM lagrangian.

The production and decay modes at tree level considered in the analysis are shown in Figure 1. The presence of the ρ_{tq} coupling opens the possibility of same-sign top production, as shown in Figures 1(a) (sstt) and 1(b) (ttq), and also three-top production, as shown in Figures 1(c) (ttt) and 1(d) (tttq). The three-top signature has been highlighted as a sensitive probe of new physics [26–29]. Additionally, four-top quarks can be produced, as shown in Figure 1(e) (tttt). The targeted final state is characterised by multiple leptons (electrons and muons) and multiple jets containing b-flavoured hadrons (b-jets). Many of the production modes are expected to be charge-asymmetric (with preference to positively charged), and this feature is exploited in the search. The relevance of each production mode depends on the chosen coupling. A benchmark of $\rho_{tt} = 0.4$ and $\rho_{tq} = 0.2$ is chosen to guide the analysis design and optimisation. The values are chosen so that the signal would be responsible for the higher $t\bar{t}W$ and $t\bar{t}t\bar{t}$ yields observed in ATLAS analyses [30–34]. For the chosen couplings, the tH production cross section is two orders of magnitude larger than ttH production, and three orders of magnitude larger than same-sign tops production via t-channel H.

This analysis is the first to target BSM production leading to three-top final states and the first to probe the g2HDM. The production of four-tops in the SM or via heavy scalars has been explored previously by ATLAS [33–37] and CMS [38–41]. Limits on the couplings can be derived from LHC Higgs measurements, B physics, and assuming perturbativity [42], leading to $\rho_{tt} < 2$, $\hat{\rho}_{tc} < 1.5$, and $\hat{\rho}_{ct} < 0.1$, or equivalently $\rho_{tc} < 1.06$. In addition, $K-\bar{K}$ mixing provides the constraint $\hat{\rho}_{ct} < 0.14$ [43], and $D-\bar{D}$ mixing provides the constraint $|\hat{\rho}_{tc}\hat{\rho}^*_{tu}| < 0.02$ [43], which translates to $|\rho_{tc}\rho^*_{tu}| < 0.01$ assuming a negligible $\hat{\rho}_{ut}$. These constraints are derived assuming $m_H \approx m_{H^+} = 500$ GeV, and become weaker for higher masses.

Two additional signal models based on R-parity-violating supersymmetry (RPV SUSY) are used to interpret the results of the search. The models are motivated by the recent flavour anomalies [44–51] and $(g-2)_\mu$

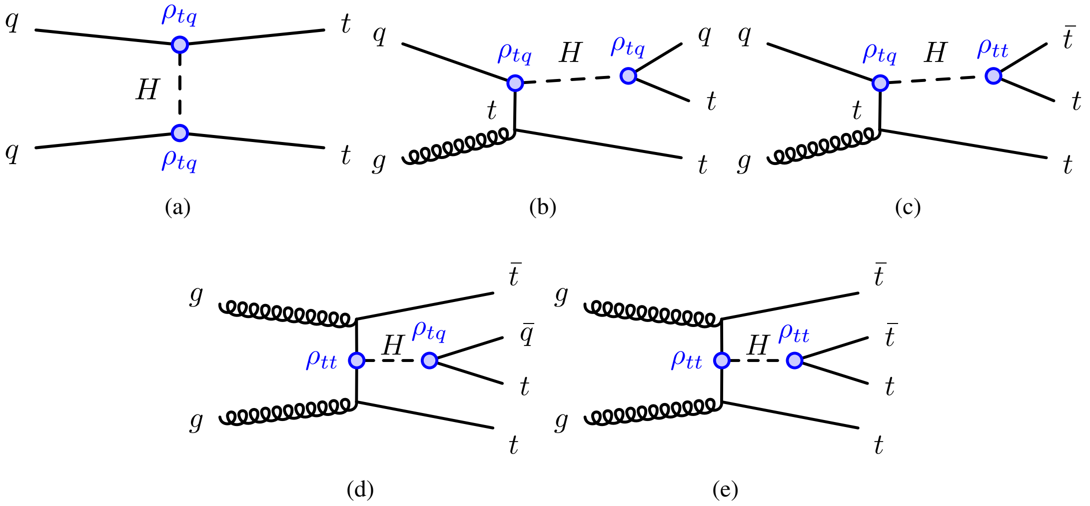


Figure 1: Signal diagrams for the dominant production and decay modes of the heavy scalar considered in the analysis. The subsequent decay can lead to a final state with high multiplicity of leptons and b-jets which is targeted by the search. Single production through gluon fusion is not considered since the decay does not lead to the relevant final state.

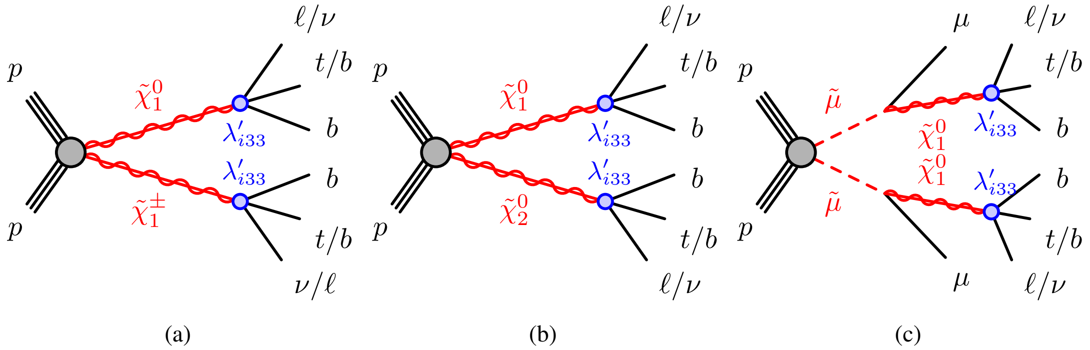


Figure 2: Signal diagrams for the RPV SUSY signals used as additional interpretation in the analysis. The subsequent decay can lead to a final state with high multiplicity of leptons and b-jets which is targeted by the search.

anomaly [52], and can provide a successful explanation with different choices of particles, masses, and couplings [53–62]. The first model features electroweakino production (wino or Higgsino) which decays via a lepton-number-violating RPV coupling of the $LQ\bar{D}$ type to a lepton and third-generation quarks. The corresponding term in the superpotential has the form $\lambda'_{i33}L_iQ_3\bar{D}_3$, where $i \in 2, 3$ is a generation index, and L, Q, $\bar{D}$ are the lepton doublet, quark doublet, and down-type quark singlet superfields, respectively. Relevant diagrams for the production and decay are shown in Figures 2(a) and 2(b). The second model features direct smuon production decaying to a bino-like neutralino, which in turn decays via the same RPV coupling (λ'_{i33}), as shown in Figure 2(c).

2 ATLAS detector

The ATLAS experiment [63] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and nearly 4π coverage in solid angle.[1] It consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic (EM) and hadron calorimeters, and a muon spectrometer (MS). The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip and transition radiation tracking detectors; the innermost layer is 33 mm from the beamline [64, 65]. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic energy measurements with high granularity. A steel/scintillator-tile hadron calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The endcap and forward regions are instrumented with LAr calorimeters for both the EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large superconducting air-core toroidal 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 [66]. The first-level trigger is implemented in hardware and uses a subset of the detector information to keep the accepted rate below approximately 100 kHz. This is followed by a software-based trigger that reduces the accepted event rate to approximately 1 kHz on average depending on the data-taking conditions. An extensive software suite [67] is used for real and simulated data reconstruction and analysis, for operation and in the trigger and data acquisition systems of the experiment.

3 Data and simulated event samples

A dataset of pp collisions at $\sqrt{s} = 13$ TeV collected by the ATLAS experiment during 2015–2018 and corresponding to an integrated luminosity of 139 fb^{-1} is used. The number of additional pp interactions per bunch crossing (pile-up) in this dataset ranges from about 8 to 70, with an average of 34. Only events recorded under stable beam conditions and for which all detector subsystems were known to be in a good operating condition are used. The trigger requirements are discussed in Section 5.

Monte Carlo (MC) simulation samples were produced for the different signal and background processes. Table 1 shows the configurations used in this analysis, with the samples in parentheses and in grey indicating those used to estimate the systematic uncertainties. All simulated samples, except those produced with the Sherpa [68] event generator, utilised EvtGen 1.2.0 [69] to model the decays of heavy-flavour hadrons. All samples showered with Pythia use the A14 tune [70], whereas those showered with Herwig use the H7-UE set of tuned parameters [71]. Pile-up was modelled using events from minimum-bias interactions generated with Pythia 8.186 [72] with the A3 set of tuned parameters [73] (referred to as the 'tune'), and overlaid onto the simulated hard-scatter events according to the luminosity profile of the recorded data. The generated events were processed through either a full simulation of the ATLAS detector geometry and response using Geant4 [74], or a faster simulation where the full Geant4 simulation of the calorimeter response is replaced by a detailed parameterisation of the shower shapes [75]. In both cases, the simulated events were processed through the same reconstruction software as the dataset of pp collisions. Corrections

---

[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 along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, 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 $\eta = -\ln\tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$.

were applied to the simulated events so that the particle candidates' selection efficiencies, energy scales and energy resolutions match those determined from data control samples. The simulated samples are normalised to their cross sections, and generated to the highest order available in perturbation theory.

Samples used to model the g2HDM signal were generated at leading-order (LO) in QCD with MADGRAPH v2.9.3 [76] with the `NNPDF3.1nlo` [77] parton distribution function (PDF) set. The ATLAS Fast Simulation was used for masses in the range of 200 GeV to 1 TeV with a 100 GeV step. All signals were produced with the set of couplings $\rho_{tt} = \rho_{tc} = \rho_{tu} = 0.1$. Each signal process described in Section 1 was generated as a separate MC sample. The cross section obtained from the LO generator is used for the normalisation of the signals. Simulated events for different coupling values are obtained by rescaling the samples to match the target cross section and branching ratio of each subprocess. For a given choice of couplings all the processes are taken into account and rescaled. The RPV SUSY signal samples were generated with MADGRAPH v2.9.3, with up to two extra jets at LO in QCD. The matching scale is set at 1/4 of the mass of the SUSY particle being produced. Supersymmetric particle decays via the RPV coupling are simulated with 25% branching ratio to $\mu/\tau/\nu_\mu/\nu_\tau$ each, a b-quark, and a b- or top-quark depending on the lepton charge. The identical branching ratio to second- and third-generation leptons follows from the choice of $\lambda'_{233} = \lambda'_{333}$, while the balance in charged and neutral leptons originates from the presence of a left-handed lepton superfield in the $LQ\bar{D}$ coupling. All signal events were showered with PYTHIA 8.245 [72] using the `NNPDF2.3lo` [78] PDF set.

The sample used to model the $t\bar{t}W$ ($t\bar{t}Z/\gamma^*$) background was generated using SHERPA-2.2.10 [79] (SHERPA-2.2.11), where the matrix element (ME) was calculated for up to one (zero) additional parton at next-to-leading-order (NLO) in QCD and up to two partons at LO in QCD using COMIX [80] and OPENLOOPS [81] and merged with the SHERPA parton shower (PS) [82] using the MEPS@NLO prescription [83], with a CKKW merging scale of 30 GeV for the $t\bar{t}W$ sample. These samples are generated using the `NNPDF3.0nnlo` [84] PDF set. Both the factorisation and renormalisation scales are set to $\mu_r = \mu_f = m_T/2$ in the $t\bar{t}W$ sample, where m_T is defined as the scalar sum of the transverse masses $\sqrt{m^2 + p_T^2}$ of the particles generated from the ME calculation. The LO $t\bar{t}W$ electroweak (EW) contributions are obtained from a dedicated sample simulated with SHERPA-2.2.10. The invariant mass of the lepton pair ($m_{\ell^+\ell^-}$) in the $t\bar{t}Z/\gamma^*$ sample is set to be greater than 1 GeV. The complete $t\bar{t}W$ simulation is normalised to the total cross section of $\sigma(t\bar{t}W) = 614.7$ fb that comes from the SHERPA configuration outlined above considering NLO QCD and NLO EWK effects, based on a similar strategy as used in Ref. [85]. The $t\bar{t}Z/\gamma^*$ sample is normalised to the cross section $\sigma(t\bar{t}Z/\gamma^*) = 839$ fb, calculated at NLO QCD and NLO EW accuracy using MADGRAPH5_AMC@NLO [86] and scaled by an off-shell correction estimated at one-loop level in α_s.

The production of $t\bar{t}t\bar{t}$ events was modelled using the MADGRAPH5_AMC@NLO v2.6.2 generator which provides matrix elements at NLO in QCD with the `NNPDF3.1nlo`. The functional form of the renormalisation and factorisation scales are set to $\mu_r = \mu_f = m_T/4$. Top quarks are decayed at LO using MADSPIN to preserve all spin correlations. The events are interfaced with PYTHIA 8.230 for the parton shower and hadronisation, using the `NNPDF2.3lo` PDF set. The production of $t\bar{t}t\bar{t}$ events is normalised to a cross-section of 12 fb computed at NLO in QCD including EW corrections [87].

Diboson (VV) background processes are simulated with SHERPA 2.2.2 [79]. The matrix element was calculated using COMIX [80] and OPENLOOPS [81] with NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional partons, and merged with the SHERPA using MEPS@NLO prescription [83]. The NNPDF3.0NNLO set of PDFs was used, along with the dedicated set of tuned parton-shower parameters developed by the SHERPA authors. The cross section of $\sigma(VV) = 104$ pb was computed by SHERPA 2.2.2.

Samples for $t\bar{t}H$, $t\bar{t}$, and single top production were generated using the NLO generator Powheg-Box-v2 [88–93] and interfaced with Pythia 8 for the parton showering and fragmentation. These samples used the `NNPDF3.0nlo` PDF set. The h_{damp} parameter, which controls the transverse momentum of the first additional emission beyond the LO Feynman diagram in the PS and therefore regulates the high-p_{T} radiation, is set to $3/4 \times (m_t + m_{\bar{t}} + m_H)$ in the $t\bar{t}H$ sample and to $1.5 \times m_t$ in the $t\bar{t}$ and single top samples, where m_t (m_H) denotes the mass of the top quark (SM Higgs boson).

A dedicated $t\bar{t}$ sample including rare $t \rightarrow Wb\gamma^*(\rightarrow l^+l^-)$ radiative decays, $t\bar{t} \rightarrow W^+bW^-\bar{b}l^+l^-$, is generated using a ME calculated at LO in QCD and requiring $m_{\ell^+\ell^-} > 1$ GeV. In this sample the photon can be radiated from the top quark, the W boson, or the b-quark. Both the $t\bar{t}Z/\gamma^*$ and $t\bar{t} \rightarrow W^+bW^-\bar{b}l^+l^-$ samples are combined and together form the "$t\bar{t}Z$ (high mass)" sample. The contribution from internal photon conversions ($\gamma^* \rightarrow l^+l^-$) with $m_{\ell^+\ell^-} < 1$ GeV are modelled by QED multiphoton radiation via the PS in an inclusive $t\bar{t}$ sample and is referred to as "$t\bar{t}\gamma^*$ (low mass)". Dedicated Z+jets samples containing electrons from material photon conversion ($\gamma \rightarrow e^+e^-$) or internal photon conversion are generated with Powheg-Box and interfaced with Pythia 8 for the parton showering and fragmentation. These samples are used to model the data in control regions enriched in material and internal conversion electrons, as explained in Section 5.

The remaining rare background contributions listed in Table 1 are normalised using their NLO theoretical cross sections, except for the $t\bar{t}t$, $t\bar{t}W^+W^-$, $t\bar{t}ZZ$, $t\bar{t}HH$, and $t\bar{t}WH$ processes, for which a LO cross section is used.

Table 1: The configurations used for event generation of signal and background processes. The samples used to estimate the systematic uncertainties are indicated in parentheses and grey. V refers to production of an electroweak boson (W or Z/γ^*). The matrix element order refers to the order in the strong coupling constant of the perturbative calculation. The "$t\bar{t}W$ (EW)" sample also includes next-to-leading-order electroweak corrections. Tune refers to the underlying-event tune of the parton shower generator. MG5_aMC refers to MadGraph5_aMC@NLO 2.2, 2.3, or 2.6; Pythia 6 refers to version 6.427 [94]; Pythia 8 refers to version 8.2; MePs@Nlo refers to the method used in Sherpa to match the matrix element to the parton shower. All samples include leading-logarithm photon emission, either modelled by the parton shower generator or by Photos [95]. The mass of the top quark (m_t) and SM Higgs boson were set to 172.5 GeV and 125 GeV, respectively.

Process	Generator	ME order	Parton shower	PDF	Tune
g2HDM signal	MG5_aMC	LO	Pythia 8	`NNPDF3.1nlo`	A14
SUSY signal	MG5_aMC	LO	Pythia 8	`NNPDF3.1nlo`	A14
$t\bar{t}W$	Sherpa 2.2.10	MePs@Nlo	Sherpa	`NNPDF3.0nnlo`	Sherpa default
	(MG5_aMC)	(NLO)	(Pythia 8)	(`NNPDF3.0nlo`)	(A14)
$t\bar{t}W$ (EW)	Sherpa 2.2.10	LO	Sherpa	`NNPDF3.0nnlo`	Sherpa default
	(MG5_aMC)	(LO)	(Pythia 8)	(`NNPDF3.0nlo`)	(A14)
$t\bar{t}t\bar{t}$	MG5_aMC	NLO	Pythia 8	`NNPDF3.1nlo`	A14
	(Sherpa 2.2.10)	(MePs@Nlo)	(Sherpa)	(`NNPDF3.0nnlo`)	(Sherpa default)
$t\bar{t}H$	Powheg-BOX	NLO	Pythia 8	`NNPDF3.0nlo`	A14
	(Powheg-BOX)	(NLO)	(Herwig7.0.4)	(`NNPDF3.0nlo`)	(H7-UE-MMHT)
	(MG5_aMC)	(NLO)	(Pythia 8)	(`NNPDF3.0nlo`)	(A14)
$t\bar{t}(Z/\gamma^* \rightarrow l^+l^-)$	Sherpa 2.2.11	MePs@Nlo	Sherpa	`NNPDF3.0nnlo`	Sherpa default
	(MG5_aMC)	(NLO)	(Pythia 8)	(`NNPDF3.0nlo`)	(A14)
$t\bar{t} \rightarrow W^+bW^-\bar{b}l^+l^-$	MG5_aMC	LO	Pythia 8	`NNPDF3.0lo`	A14
$t(Z/\gamma^*)$	MG5_aMC	NLO	Pythia 8	`NNPDF2.3lo`	A14
$tW(Z/\gamma^*)$	MG5_aMC	NLO	Pythia 8	`NNPDF2.3lo`	A14
$t\bar{t}$	Powheg-BOX	NLO	Pythia 8	`NNPDF3.0nlo`	A14
	(Powheg-BOX)	NLO	(Herwig7.1.3)	(`NNPDF3.0nlo`)	(H7-UE-MMHT)
$t\bar{t}t$	MG5_aMC	LO	Pythia 8	`NNPDF2.3lo`	A14
Single top (t-, Wt-, s-channel)	Powheg-Box	NLO	Pythia 8	`NNPDF3.0nlo`	A14
VV, $qqVV$, VVV	Sherpa 2.2.2	MePs@Nlo	Sherpa	`NNPDF3.0nnlo`	Sherpa default
$Z \rightarrow l^+l^-$	Sherpa 2.2.1	MePs@Nlo	Sherpa	`NNPDF3.0nlo`	Sherpa default
$Z \rightarrow l^+l^-(\gamma \rightarrow e^+e^-)$	Powheg-BOX	NLO	Pythia 8	`CTEQ6L1nlo`	A14
$Z \rightarrow l^+l^-(\gamma * \rightarrow e^+e^-)$	Powheg-BOX	NLO	Pythia 8	`CTEQ6L1nlo`	A14
W+jets	Sherpa 2.2.1	MePs@Nlo	Sherpa	`NNPDF3.0nlo`	Sherpa default
VH	Powheg-BOX	NLO	Pythia 8	`NNPDF3.0nlo`	A14
$t\bar{t}W^+W^-$	MG5_aMC	LO	Pythia 8	`NNPDF2.3lo`	A14
$t\bar{t}ZZ$	MG5_aMC	LO	Pythia 8	`NNPDF2.3lo`	A14
$t\bar{t}HH$	MG5_aMC	LO	Pythia 8	`NNPDF2.3lo`	A14
$t\bar{t}WH$	MG5_aMC	LO	Pythia 8	`NNPDF2.3lo`	A14

4 Event reconstruction and object identification

Interaction vertices from the pp collisions are reconstructed from at least two tracks with transverse momentum (p_{T}) larger than 500 MeV that are consistent with originating from the beam collision region in the x–y plane. If more than one primary vertex candidate is found in the event, the candidate for which the associated tracks form the largest sum of squared p_{T} is selected as the hard-scatter primary vertex [96].

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter matched to a track in the ID [97]. They are required to satisfy $p_{\mathrm{T}} > 10$ GeV and $|\eta_{\mathrm{cluster}}| < 2.47$, excluding the transition region between the endcap and barrel calorimeters ($1.37 < |\eta_{\mathrm{cluster}}| < 1.52$). Loose and tight electron identification working points are used [97], based on a likelihood discriminant employing calorimeter, tracking and combined variables that provide separation between electrons and jets. The associated track of an electron candidate is required to have at least two hits in the pixel detector and seven hits total in the pixel and silicon-strip detectors combined. For the tight identification working point, one of these pixel hits must be in the innermost layer, or the next-to-innermost layer if the module traversed in the innermost layer is non-operational, and there must be no association with a vertex from a reconstructed photon conversion [98] in the detector material (denoted as 'material conversion' in this paper).

Muon candidates are reconstructed by combining tracks in the ID with tracks in the MS [99]. The resulting muon candidates are re-fit using the complete track information from both detector systems [100]. They are required to satisfy $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$. Loose and medium muon identification working points are used [100].

Electron (muon) candidates are matched to the primary vertex by requiring that their longitudinal impact parameter, z_0, satisfies $|z_0 \sin\theta| < 0.5$ mm, where θ is the polar angle of the track, and requiring that the significance of their transverse impact parameter, d_0, satisfies $|d_0/\sigma(d_0)| < 5\,(3)$, where $\sigma(d_0)$ is the measured uncertainty in d_0.

To further suppress leptons from heavy-flavour hadron decays, misidentified jets, or photon conversions (collectively referred to as 'non-prompt leptons'), lepton candidates are also required to be isolated in the tracker and in the calorimeter. A track-based lepton isolation criterion is defined by calculating the quantity $I_R = \sum p_{\mathrm{T}}^{\mathrm{trk}}$, where the scalar sum includes all tracks (excluding the lepton candidate itself) within the cone defined by $\Delta R < R_{\mathrm{cut}}$ around the direction of the lepton. The value of R_{cut} is the smaller of r_{min} and $10\ \mathrm{GeV}/p_{\mathrm{T}}^{\ell}$, where r_{min} is set to 0.2 (0.3) for electron (muon) candidates and where p_{T}^{ℓ} is the lepton p_{T}. All lepton candidates must satisfy $I_R/p_{\mathrm{T}}^{\ell} < 0.15$. Additionally, electrons (muons) are required to satisfy a calorimeter-based isolation criterion: the sum of the transverse energy within a cone of size $\Delta R = 0.2$ around the lepton, after subtracting the contributions from pile-up and the energy deposit of the lepton itself, is required to be less than 20% (30%) of p_{T}^{ℓ}. Muons are required to be separated by $\Delta R > 0.2$ from any selected jets (defined below). If two electrons are closer than $\Delta R = 0.1$, only the one with the higher p_{T} is considered. Electrons within $\Delta R = 0.1$ of a selected muon are removed.

The selection criteria described above greatly suppress the contribution from non-prompt leptons. However, several channels considered in this search have additional suppression requirements targeting the main non-prompt lepton types. Non-prompt leptons from hadron decays that contain bottom- and charm-quarks, denoted as 'heavy-flavour (HF) non-prompt leptons', are further rejected using a boosted decision tree (BDT) discriminant, referred to as the non-prompt lepton BDT [101], that uses isolation and lifetime information associated with a track jet that matches the selected light lepton. Three working points (WPs) are used: *Tight*, *VeryTight*, and *Tight-not-VeryTight*. The first two provide a selection of prompt-like leptons with an efficiency for muons (electrons) that satisfy the calorimeter- and track-based isolation

Table 2: Description of the loose inclusive ("L"), medium inclusive ("M"), medium exclusive ("M_{ex}"), and tight ("T") lepton definitions. The electron e^* is required to fulfil, in addition to the corresponding lepton definition requirements, those corresponding to an internal or material conversion candidate.

	e				μ			
	L	M	M_{ex}	T	L	M	M_{ex}	T
Isolation	Yes				Yes			
Non-prompt lepton BDT WP	No	*Tight*	*Tight-not-VeryTight*	*VeryTight*	No	*Tight*	*Tight-not-VeryTight*	*VeryTight*
Identification	Loose	Tight			Loose	Medium		
Electron charge-misassignment veto	No	Yes			N/A			
Electron conversion candidate veto	No	Yes (except e^*)			N/A			
Transverse impact parameter significance $\lvert d_0\rvert/\sigma_{d_0}$	< 5				< 3			
Longitudinal impact parameter $\lvert z_0 \sin\theta\rvert$	< 0.5 mm							

criteria of about 50% (70%) for $p_{\text{T}} \sim 20$ GeV and reaches a plateau of 90% at $p_{\text{T}} \sim 55$ (50) GeV. The corresponding rejection factor[2] against leptons from the decay of b-hadrons ranges from 29 to 50 (6 to 18) depending on the p_{T} and η, after resolving ambiguities between overlapping reconstructed objects. The latter allows to select non-prompt-like leptons and is part of the event selection of the control regions enriched in HF non-prompt lepton background, as described in Section 6.

In order to further suppress electrons with incorrect charge assignment, a BDT discriminant based on calorimeter and tracking quantities [102] is used. An efficiency of 88% for isolated electrons with correct charge assignment is obtained, with a rejection factor of ~3.3 for isolated electrons with incorrect charge assignment. Material and internal conversion candidates are identified based on a combination of requirements on the invariant mass of tracks and the radius from the reconstructed displaced vertex to the primary vertex. Material conversion candidates have a reconstructed displaced vertex with radius $r > 20$ mm that includes the track associated with the electron.[3] The invariant mass of the associated track and the closest (in $\Delta\eta$) opposite-charge track reconstructed in the silicon detector, calculated at the conversion vertex, is required to be < 100 MeV. Internal conversion candidates, which correspond to the internal photon conversions (see Section 3), are required to fail the requirements for material conversions, and the di-track invariant mass, calculated here at the primary vertex, is also required to be < 100 MeV.

The various lepton working points used in this analysis are summarised in Table 2. After the initial categorisation based on loose leptons (corresponding to "L"), the most optimal lepton working point to further optimise the event selection is chosen depending on the main background processes and the number of expected events in each category. The various choices can be seen for the signal and control regions in Section 5. All M, M_{ex}, and T electrons in the analysis are required to not be material nor internal conversion candidates, with the exception of the electron in the control regions enriched with internal and material conversions. Those electrons must be an internal or material conversion candidate, respectively, and is denoted as e^*, regardless of the conversion mechanism (see Section 5).

The constituents for jet reconstruction are identified by combining measurements from both the ID and the calorimeter using a particle flow (PFlow) algorithm [103, 104]. Jet candidates are reconstructed from

---

[2] The rejection factor is defined as the reciprocal of the efficiency.

[3] The beampipe and insertable B-layer inner radii are 23.5 mm and 33 mm, respectively.

these PFlow objects using the anti-k_t algorithm [105, 106] with a radius parameter of $R = 0.4$. They are calibrated using simulation with corrections obtained from in situ techniques in data [104]. Only jet candidates with a $p_{\text{T}} > 25$ GeV and within $|\eta| < 2.5$ are selected. In order to reduce the effect from pile-up, each jet with $p_{\text{T}} < 60$ GeV and $|\eta| < 2.4$ is required to satisfy the “Tight” working point of the Jet Vertex Tagger (JVT) [107] criteria used to identify the jets as originating from the selected primary vertex. A set of quality criteria is also applied to reject events containing at least one jet arising from non-collision sources or detector noise [108].

Jets containing b-hadrons are identified (b-tagged) via the DL1r algorithm [109, 110] that uses a deep-learning neural network based on the distinctive features of the b-hadrons in terms of the impact parameters of tracks and the displaced vertices reconstructed in the ID. Additional input to this network is provided by discriminant variables constructed by a recurrent neural network [111], which exploits the spatial and kinematic correlations between tracks originating from the same b-hadron. For each jet, a value for the multivariate b-tagging discriminant is calculated. A jet is b-tagged if the b-tagging score is above a certain threshold, referred to as a working point (WP). Four WPs are defined with average expected efficiencies for b-jets of 60%, 70%, 77% and 85%, as determined in simulated $t\bar{t}$ events. The b-tagging distribution obtained by ordering the resulting five exclusive bins from the four WPs from higher to lower b-jet efficiency is referred to as “pseudo-continuous” b-tagging score, and it is used as input to the multivariate analysis discriminant described in Section 5. In this search, a jet is considered b-tagged if it passes the WP corresponding to 77% or 60% efficiency to tag a b-jet, with a light-jet[4] rejection factor of about 192 or 2500, and a charm-jet (c-jet) rejection factor of about 6 or 40, as determined for jets with $p_{\text{T}} > 20$ GeV and $|\eta| < 2.5$ in simulated $t\bar{t}$ events. Correction factors derived from dedicated calibration samples enriched in b-jets, c-jets, or light jets, are applied to the simulated samples [109, 112, 113]. The notation $b^{77\%}$ and $b^{60\%}$ is used to denote the number of b-tagged jets with the corresponding WP.

The missing transverse momentum $\vec{p}_{\text{T}}^{\text{miss}}$ (with magnitude $E_{\text{T}}^{\text{miss}}$) is defined as the negative vector sum of the p_{T} of all selected and calibrated objects in the event, including a term to account for the momentum from soft particles in the event that are not associated with any of the selected objects [114]. This soft term is calculated from inner-detector tracks matched to the selected primary vertex, which makes it more resilient to contamination from pile-up interactions. The $E_{\text{T}}^{\text{miss}}$ distribution is used as an input variable to the machine learning training discussed in Section 5.

5 Search strategy

Events are required to pass a minimal preselection and are then categorised into orthogonal signal regions based on different criteria such as number of leptons, total lepton charge (indicated by Q), and a multi-output deep neural network classifier (DNN^{cat}). This categorisation provides a set of regions that are sensitive to all the possible signal production and decay modes considered in this search. A deep neural network is trained in each of the signal regions to discriminate the signal from the backgrounds (DNN^{SB}). Additional orthogonal control regions are defined in order to fit the normalisation of the main backgrounds. Dedicated kinematic selections are applied to the control regions in order to improve the purity of the targeted backgrounds. A maximum-likelihood fit is performed across categories in order to test for a possible signal and constrain in-situ the leading backgrounds simultaneously.

---

[4] ‘Light jet’ refers to a jet originating from the hadronisation of a light quark (u, d, s) or a gluon.

Events were selected for read-out using a combination of single-lepton and di-lepton triggers, requiring the electrons or muons to satisfy identification criteria similar to those used in the offline reconstruction and isolation requirements [115, 116]. For the analysis selection, at least two jets and at least two leptons are required in the event, and leptons are required to match, with $\Delta R < 0.15$, the corresponding leptons reconstructed by the trigger and to have a p_{T} exceeding the trigger p_{T} threshold by 1 GeV. Events are required to contain at least one b-tagged jet with the 60% efficiency working point, or at least two b-tagged jets with the 77% efficiency working point. If events contain pairs of opposite-sign charge and same-flavour leptons (OS-SF), all pairs are required to satisfy a mass requirement on the di-lepton system mass of $m_{\ell^+\ell^-}^{OS-SF} > 12$ GeV and $|m_{\ell^+\ell^-}^{OS-SF} - m_Z| > 10$ GeV. Three disjoint event categories are defined according to the number of loose leptons in the event: same-charge dilepton (2ℓSS), three-lepton (3ℓ), and four-lepton (4ℓ) categories. Leptons are ordered by p_{T} in the 2ℓSS and 4ℓ regions. In the 3ℓ regions the lepton with opposite-sign charge is taken first, followed by the two same-sign leptons in p_{T} order. The p_{T} and identification requirements of each lepton in each category are optimised based on a compromise between non-prompt lepton background suppression and signal acceptance enhancement, and are summarised in Table 3.

Multiple control regions are defined in order to fit the normalisation of the leading backgrounds. These regions are orthogonal to the signal regions and with one another based on different requirements on the lepton working points, dilepton invariant mass, and jet and b-jet multiplicities. Two regions enriched in diboson and $t\bar{t}Z$ are defined by requiring one OS-SF pair compatible with a Z-boson, $|m_{\ell^+\ell^-}^{OS-SF} - m_Z| < 10$ GeV, differing in the jet multiplicity requirement. Two control regions enriched in photon conversions from $Z \rightarrow \mu\mu\gamma^*(\rightarrow ee)$ are defined, according to the identification of the electron as a material conversion or internal conversion candidate. Finally, six control regions are defined enriched in HF non-prompt leptons, making use of the exclusive lepton identification M_{ex} in order to be orthogonal to the signal regions. Events with two same-sign leptons are split according to the criteria (T, M_{ex}), (M_{ex}, T), $(M_{\mathrm{ex}}, M_{\mathrm{ex}})$ for the leading and subleading leptons in p_{T}, and further split according to the subleading lepton flavour. This splitting creates six control regions sensitive to different relative composition of electron and muon non-prompt lepton backgrounds. Additionally, the transverse mass of the leading lepton and the missing transverse energy, $m_T(\ell_0, E_{\mathrm{T}}^{\mathrm{miss}})$, defined as $\sqrt{2 \times E_{\mathrm{T}}^{\mathrm{miss}} \times p_{T,\ell_0} \times (1 - \cos(\phi_{miss} - \phi_{\ell_0}))}$, is required to be lower than 250 GeV in the (T, M_{ex}) and (M_{ex}, T) regions, in order to reduce the $t\bar{t}W$ contribution in these CRs. The full definition of the kinematic selection applied to each control region is given in Table 4. Figure 3 illustrates the categorisation and definition of the signal and control regions being fit simultaneously in this analysis. The signal contamination is found to be at most 3% of the total prediction in the control regions, assuming m_H=400 GeV and $\rho_{tt} = 0.4$, $\rho_{tc} = 0.2$, and $\rho_{tu} = 0.2$.

In order to better target each of the possible signals, a DNN^{cat} is trained to identify each of the five possible production and decay modes of the g2HDM signal. Two DNN^{cat} are trained individually for the 2ℓSS and 3ℓ channels using the Keras library [117] with Tensorflow as a backend [118] and Adam optimiser [119]. Hyperparameters are optimised with the Talos library [120]. The networks consist of nine input features, two dense fully-connected layers of 33 nodes with rectified linear units as activation functions, interleaved with a drop-out layer with 20% rate, and five (three) output nodes with a soft-max activation function for the categorisation of 2ℓSS (3ℓ) events. The output categories correspond to the five production modes considered, ignoring in the 3ℓ category signals that can not produce three leptons. Each event is categorised according to the highest class probability. The nine input features are the number of jets, b-tagging score of the three leading jets, sum of b-tagging score of all jets, sum of all pair-wise angular distances between leptons, scalar sum of jet p_{T}, scalar sum of lepton p_{T}, and the event $E_{\mathrm{T}}^{\mathrm{miss}}$. The network is trained with batch size of 2000 and up to 100 epochs, using all the available signal mass points. To avoid discarding

Table 3: Event selection summary in the signal regions. Leptons are ordered by p_{T} in the 2ℓSS and 4ℓ regions. In the 3ℓ regions the lepton with opposite-sign charge is taken first, followed by the two same-sign leptons in p_{T} order. In the lepton selection, *T, M, L* stand for Tight, Medium and Loose lepton definitions. In the region naming, the "CAT ttX" denotes the category based on the DNN^{cat} output enriched in the signal process "ttX". Each of these regions is split according to the lepton charge of the same-sign lepton pair ("++" or "- -").

Lepton category	2ℓSS	3ℓ	4ℓ
Lepton definition	(T,T) with $\geq 1\ b^{60\%}$ $\|\|$ (T,M) with $\geq 2\ b^{77\%}$	(L,T,M) with $\geq 1\ b^{60\%}$ $\|\|$ (L,M,M) with $\geq 2\ b^{77\%}$	(L,L,L,L)
Lepton p_{T} [GeV]	(20, 20)	(10, 20, 20)	(10, 10, 10, 10)
$m_{\ell^+\ell^-}^{OS-SF}$ [GeV]	–	>12	
$\|m_{\ell^+\ell^-}^{OS-SF} - m_Z\|$ [GeV]	–	>10	
N_{jets}	≥ 2		
$N_{b\mathrm{-jets}}$	$\geq 1\ b^{60\%}$ $\|\|$ $\geq 2\ b^{77\%}$		
Region split	(sstt, ttq, ttt, tttq, tttt) × (Q^{++}, Q^{--})	(ttt, tttq, tttt) × (Q^{+}, Q^{-})	–
Region naming	2ℓSS ++ CAT sstt	3ℓ ++ CAT ttt	4ℓ
	2ℓSS ++ CAT ttq	3ℓ ++ CAT tttq	
	2ℓSS ++ CAT ttt	3ℓ ++ CAT tttt	
	2ℓSS ++ CAT tttq	3ℓ −− CAT ttt	
	2ℓSS ++ CAT tttt	3ℓ −− CAT tttq	
	2ℓSS −− CAT sstt	3ℓ −− CAT tttt	
	2ℓSS −− CAT ttq		
	2ℓSS −− CAT ttt		
	2ℓSS −− CAT tttq		
	2ℓSS −− CAT tttt		

signal events in the evaluation, cross-training is used with the events divided by even/odd event number.

Since several of the probed signal processes are expected to be charge-asymmetric, all the 2ℓSS and 3ℓ regions are further split into two categories each corresponding to the positive and negative total lepton charge selections. Figures 4(a) and 4(b) show the normalized distributions of the targeted signals with a scalar mass of 400 GeV or 1000 GeV, compared to the expected background distribution across the various categories described in Table 3. At high signal mass, a strong migration is observed from the ttt to the tttq category, due to the high probability of additional radiation. Figures 4(c) and 4(d) show the expected fractional signal contribution in each category for the benchmark coupling. The signals originating from top-Higgs associated production (ttq and ttt) are expected to dominate across all regions, including the categories designed to target other processes, due to the much larger production cross section. This contribution is however strongly dependent on the coupling choice. For the benchmark coupling of $\rho_{tt} = 0.4$, $\rho_{tc} = 0.2$, $\rho_{tu} = 0.2$, the decay to top-quark pairs dominates when not suppressed by the

Table 4: Event selection summary in the control regions. The notation e^* is used to denote material conversion or internal conversion candidates, as described in Section 4. In the HF non-prompt lepton region naming, "2ℓSStt(e)" ("2ℓSStt(μ)") refer to the control regions enriched in non-prompt electrons (muons) from semileptonic b-decays originating mostly from $t\bar{t}$ and with the lepton flavours for the leading and subleading leptons corresponding to "ee, μe" ("$\mu\mu$, $e\mu$"). The additional (T, M_{ex}), (M_{ex}, T), and $(M_{\text{ex}}, M_{\text{ex}})$ subscripts refer to the lepton definitions required for the leading and subleading leptons in each region.

Control regions	WZ	$t\bar{t}Z$	Conversions	HF non-prompt
N_{jets}	2 or 3	≥ 4	≥ 0	≥ 2
$N_{b-\text{jets}}$	$\geq 1\ b^{60\%} \parallel \geq 2\ b^{77\%}$		$0\ b^{77\%}$	$1\ b^{77\%}$
Lepton requirement	3ℓ		$\mu\mu e^*$	2ℓSS
Lepton definition	(L, M, M)			$(T, M_{\text{ex}}) \parallel (M_{\text{ex}}, T) \parallel (M_{\text{ex}}, M_{\text{ex}})$
Lepton p_{T} [GeV]	(10, 20, 20)			(20, 20)
$m_{\ell^+\ell^-}^{OS-SF}$ [GeV]	>12		>12	–
$\lvert m_{\ell^+\ell^-}^{OS-SF} - m_Z \rvert$ [GeV]	<10		>10	–
$\lvert m_{\ell\ell\ell} - m_Z \rvert$ [GeV]	–		<10	–
$m_T(\ell_0, E_{\text{T}}^{\text{miss}})$ [GeV]	–			< 250
Region split	–	–	internal / material	subleading $e/\mu \times [(T, M_{\text{ex}}), (M_{\text{ex}}, T), (M_{\text{ex}}, M_{\text{ex}})]$
Region naming	3ℓVV	3ℓttZ	3ℓIntC	2ℓtt(e)$_{(T,M_{\text{ex}})}$, 2ℓtt(e)$_{(M_{\text{ex}},T)}$, 2ℓtt(e)$_{(M_{\text{ex}},M_{\text{ex}})}$
			3ℓMatC	2ℓtt(μ)$_{(T,M_{\text{ex}})}$, 2ℓtt(μ)$_{(M_{\text{ex}},T)}$, 2ℓtt(μ)$_{(M_{\text{ex}},M_{\text{ex}})}$

Table 5: Input variables to the training of the DNN^{cat} and DNN^{SB} discriminants.

Variable	DNN^{cat}	DNN^{SB}
Number of jets (N_{jets})	✓	✓
Sum of pseudo-continuous b-tagging scores of jets	✓	✓
Pseudo-continuous b-tagging score of 1st, 2nd, 3rd leading jet in p_{T}	✓	✓
Sum of p_{T} of the jets and leptons ($H_{T,jets}$, $H_{T,lep}$)	✓	✓
Angular distance of leptons (sum in the case of 3ℓ and 4ℓ)	✓	✓
Missing transverse energy	✓	✓
Leading transverse momentum of jet	-	✓
Invariant mass of leading lepton and missing transverse energy	-	✓
Di/tri/quad-lepton type variable (associated to the number of electrons/muons in event)	-	✓

available phase-space.

A total of 27 analysis regions are defined, with 17 signal regions (10 with 2ℓSS, 6 with 3ℓ, and one 4ℓ) and 10 control regions. In each region, a given kinematic variable is fit to improve the sensitivity to the targeted signal process (signal regions) or to improve the modelling of a particular background process (control regions).

A DNN^{SB} classifier is trained in each signal region to separate the targeted signal from the sum of backgrounds. The networks consist of 12 input features, two dense fully-connected layers of 36 and 48 nodes respectively with sigmoid activation functions, interleaved with a drop-out layer with 40% rate,

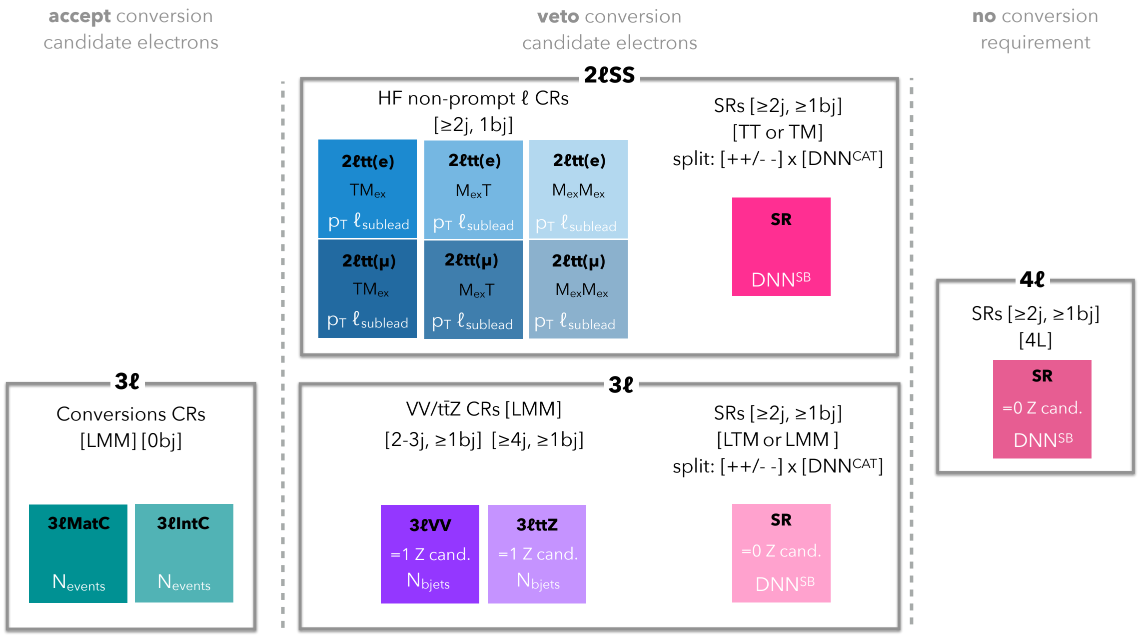
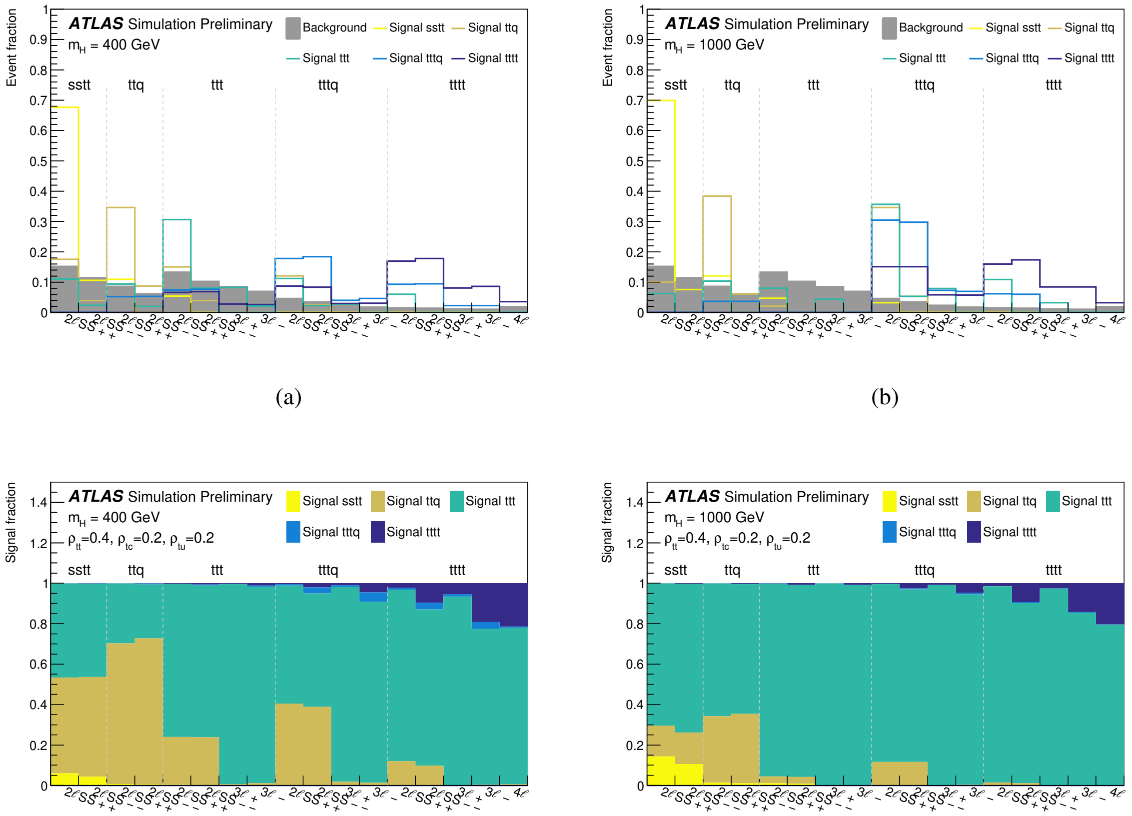


Figure 3: Illustrative sketch of the definition of the signal and control regions. At the bottom of each region box the corresponding observable used in the simultaneous fit as described in Section 8 is shown.

and one output node with a sigmoid activation function. The 12 input features are the leading jet p_{T}, number of muons, transverse mass of leading lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ system, and the nine variables that are used in the DNN^{cat}. Table 5 summarises the input variables used for each multivariate discriminant. In order to achieve good sensitivity over the large range of masses that are tested, the output of the classifier is decorrelated from the signal mass introducing an additional term to the loss function via distance correlation [121, 122]. A hyperparameter λ controls the weight of the additional penalty term, with a value of $\lambda = 0.5$. The value was optimised to achieve a minimal signal mass dependence without compromising the discrimination power. A separate training is performed in each lepton category and signal category. The same DNN^{SB} is used in both positive- and negative-charge regions. Figure 5 shows the DNN^{SB} distribution of the targeted signal in each signal-enriched category, the total signal, and the background in the 2ℓSS ++ CAT sstt, 2ℓSS ++ CAT ttq, 3ℓ ++ CAT ttt, 3ℓ ++ CAT tttq, 3ℓ ++ CAT tttt, and 4ℓ categories.

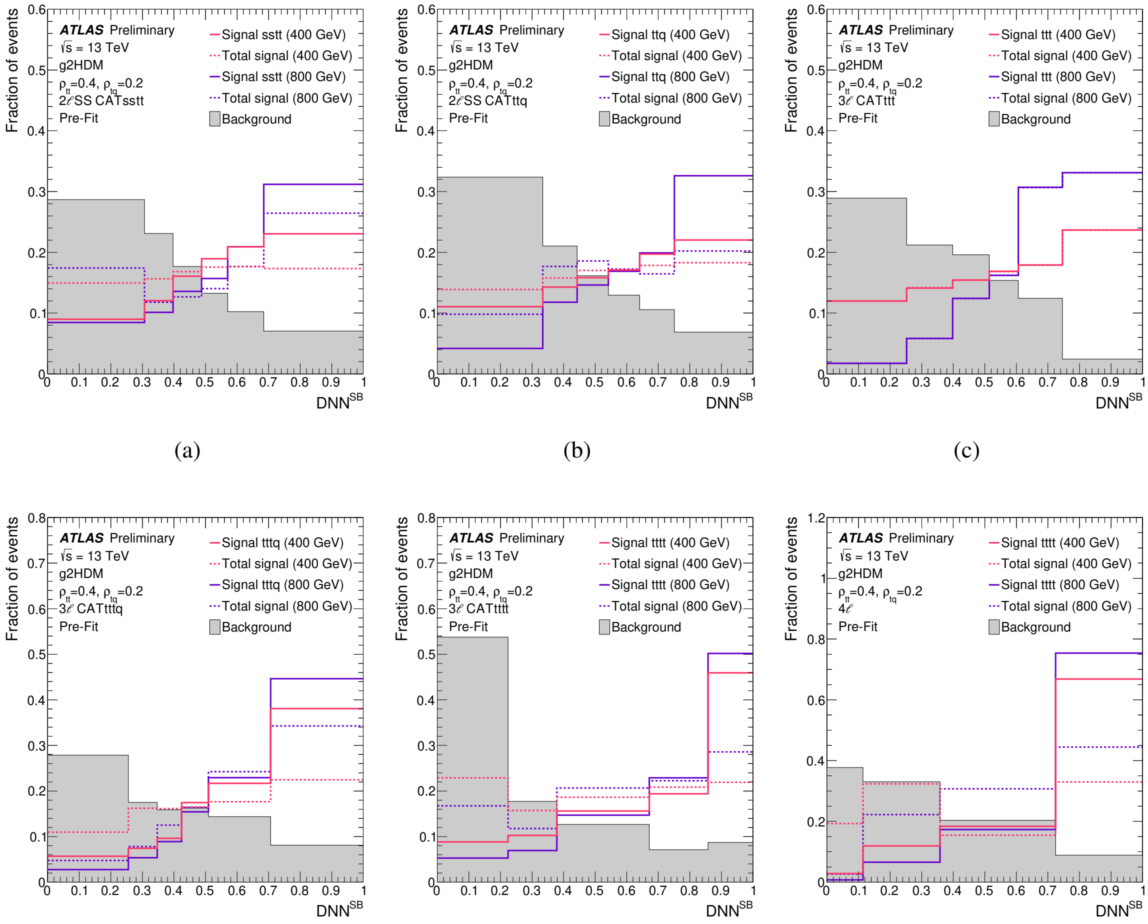
In the diboson and $t\bar{t}Z$ control regions the fitted variable is $N_{b-\mathrm{jets}}$, where the distribution is binned with an upper limit of ≥ 2 b-jets and ≥ 3 b-jets respectively. The subleading lepton p_{T} spectrum is used in the HF non-prompt control regions. Finally, the total event yield is fit in the control regions enriched in electrons from photon conversion.



(a) (b)

(c) (d)

Figure 4: Distributions for each signal process category (top row), compared to the total expected background, for (a) 400 GeV and (b) 1000 GeV scalar masses. All distributions are normalised to unity. The vertical dashed grey lines separate categories targeting each of the main signal processes: sstt, ttq, ttt, tttq, and tttt. Signal contributions below 2% in a single bin are omitted for clarity. The expected fractional signal contribution in each category (bottom row) is shown for (c) 400 GeV and (d) 1000 GeV scalar masses for the coupling set $\rho_{tt} = 0.4$, $\rho_{tc} = 0.2$, $\rho_{tu} = 0.2$.



(a) (b) (c)

(d) (e) (f)

Figure 5: Comparison of the DNN^{SB} distribution of the targeted signal (solid line), the total signal (dashed line) and the background (filled grey area) in the (a) 2ℓSS CAT sstt, (b) 2ℓSS CAT ttq, (c) 3ℓ CAT ttt, (d) 3ℓ CAT tttq, (e) 3ℓ CAT tttt, and (f) 4ℓ categories, for an assumed $m_H = 400$ GeV (pink) and $m_H = 800$ GeV (violet), with couplings $\rho_{tt} = 0.4$ and $\rho_{tq} = 0.2$. All distributions are normalised to unity.

6 Background estimation

The background processes passing the signal region selections are categorised into irreducible and reducible backgrounds. Irreducible backgrounds (Section 6.1) produce prompt leptons in their decay, i.e. originating from W/Z boson decays, leptonic τ-lepton decays, or internal conversions. Reducible backgrounds (Section 6.2) have prompt leptons with misassigned charge or at least one non-prompt lepton.

Except for the background from electrons with misassigned charge (denoted as QMisID), all other backgrounds are estimated using the simulated samples described in Section 3. In some cases, the simulation is improved using additional corrections derived from data control samples before the simultaneous fit to data. In particular, the event kinematics of the simulated $t\bar{t}$ and VV backgrounds require dedicated corrections to better describe the data. In addition, the yields of some simulated backgrounds, in particular $t\bar{t}W$, $t\bar{t}Z$, VV and non-prompt-lepton backgrounds, are adjusted via normalisation factors that are determined by performing a likelihood fit to data across all event categories (signal and control regions as defined in Tables 3 and 4) as discussed in Section 8.

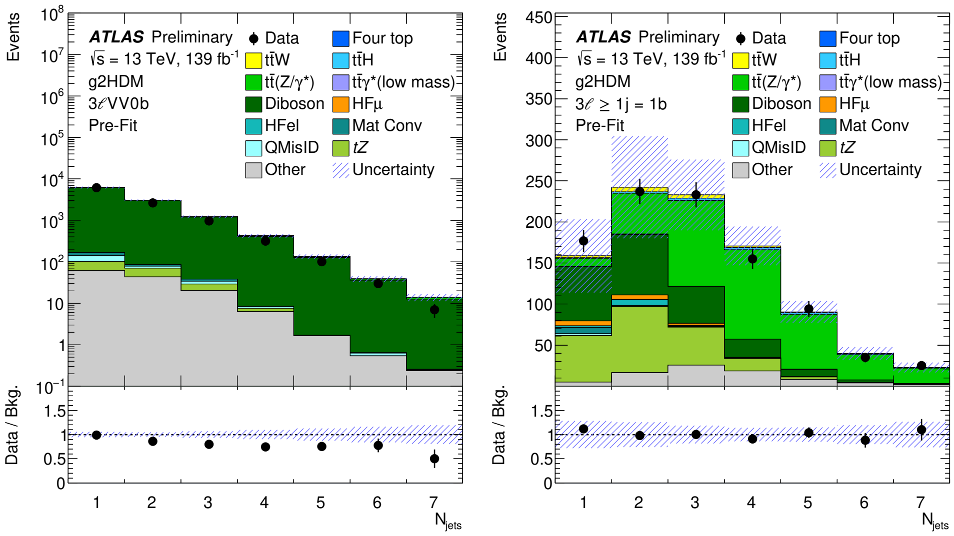
6.1 Irreducible backgrounds

Background contributions with prompt leptons originate from a wide range of physics processes with the relative importance of individual processes varying by channel. The main irreducible backgrounds originate from $t\bar{t}W$, $t\bar{t}t\bar{t}$, and $t\bar{t}Z/\gamma^*$ production, followed by VV (in particular WZ) and $t\bar{t}H$ production, and have final states and kinematic properties similar to the g2HDM signal. Smaller contributions originate from the following rare processes: tZ, tW, tWZ, $t\bar{t}WW$, VVV, and $t\bar{t}t$ production.

6.1.1 $t\bar{t}W$ background

The $t\bar{t}W$ background represents the leading background in several event categories. Despite the use of state-of-the-art simulations, accurate modelling of additional QCD and QED radiation in $t\bar{t}W$ production remains challenging. Given the excellent discriminating power of the DNN^{SB} in the signal regions, the events at lower values of the DNN^{SB} score are enriched in and sensitive to the $t\bar{t}W$ background. Additionally, the signal regions in the 2ℓ and 3ℓ categories are split by the sign of the total lepton charge (Q) to better discriminate some g2HDM signal processes and the $t\bar{t}W$ process, which have a large charge asymmetry, from other SM backgrounds that are charge symmetric. This discrimination improves the modelling of this background in the simultaneous fit.

Disagreement between the data and the prefit prediction from the simulation is observed, which is accommodated by an overall normalisation factor that is assigned to the $t\bar{t}W$ background, and which is determined during the likelihood fit. The measured normalisation factor is $\hat{\lambda}_{t\bar{t}W} = 1.50 \pm 0.14$, which is compatible with that determined in the SM $t\bar{t}t\bar{t}$ analysis [123], and with a previous measurement of the $t\bar{t}W$ production cross section [124].



(a) (b)

Figure 6: Comparison between data and the background prediction for the distribution of (a) the number of jets in the 3ℓVV0b region before the VV jet multiplicity correction and (b) the number of jets in a 3ℓ region with at least one jet and exactly one b-jet defined with the 60% WP after the VV jet multiplicity correction. The ratio of the data to the background prediction ("Pred.") is shown in the lower panel. The size of the combined statistical and systematic uncertainty in the background prediction is indicated by the blue hatched band. The last bin in each figure contains the overflow.

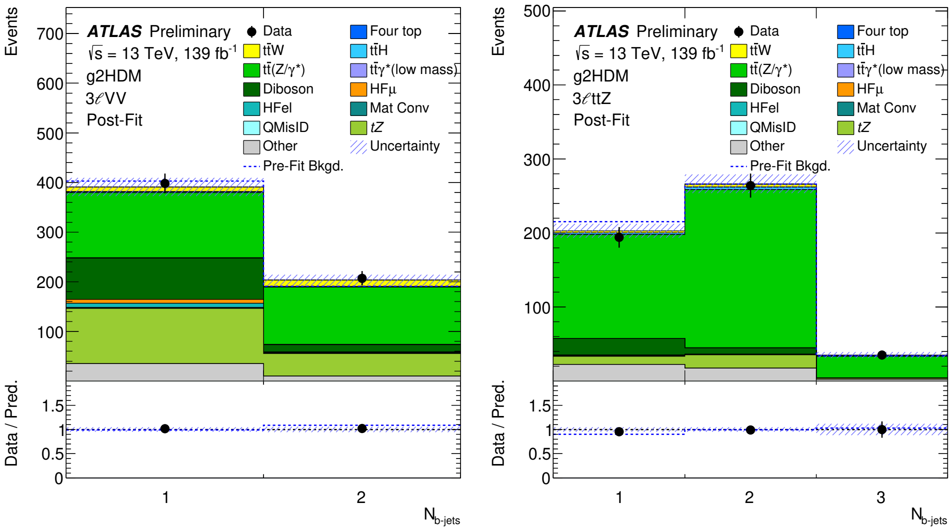
6.1.2 VV and $t\bar{t}Z/\gamma^*$ backgrounds

The VV simulated sample was seen to require additional improvements in the modelling of the jet multiplicity spectrum compared to data. Therefore, a data-driven correction is derived from an inclusive trilepton diboson-enriched region with zero b-jets defined with the 85% WP for b-jet efficiency and at least one jet (denoted as 3ℓVV0b region). The events are required to have three leptons passing the same selection as in the 3ℓVV CR.

Figure 6(a) shows the jet multiplicity distribution in the 3ℓVV0b region before the correction. After the correction is applied to VV, the modelling of the N_{jets} distribution improves significantly in a 3ℓ region with at least one jet and exactly one b-jet defined with the 60% WP, as shown in Figure 6(b).

The 3ℓVV and 3ℓttZ CRs are used in the likelihood fit to improve the prediction of the background contribution from the VV and $t\bar{t}Z/\gamma^*$ processes, respectively. The number of jets and b-jets provide good discrimination between these two processes and are used to build the control regions (number of jets) and as variables used in the fit (number of b-jets). The measured normalisation factors are: $\hat{\lambda}_{VV} = 0.85 \pm 0.30$ and $\hat{\lambda}_{t\bar{t}Z} = 0.97 \pm 0.19$

Figures 7(a) and 7(b) show the b-jet multiplicity distribution in the 3ℓVV and 3ℓttZ CRs after the likelihood fit to data.



(a)

(b)

Figure 7: Comparison between data and the background prediction for the distribution of the b-jet multiplicity in the (a) 3ℓVV and (b) 3ℓttZ CRs after the VV jet multiplicity correction. The background contributions after the likelihood fit to data ("Post-Fit") under the background-only hypothesis are shown as filled histograms. The ratio of the data to the post-fit background prediction ("Pred.") is shown in the lower panel, separately for post-fit background (black points) and pre-fit background (dashed blue line). The size of the combined statistical and systematic uncertainty in the background prediction is indicated by the blue hatched band. The last bin in each figure contains the overflow.

6.1.3 Other irreducible backgrounds

The rate of the background from internal conversions with $m(e^+e^-) < 1$ GeV is estimated using the two dedicated CRs (3ℓIntC and 3ℓMatC). The total yield in each category is used in the likelihood fit to determine the normalisation factor: $\hat{\lambda}_e^{\text{IntC}} = 1.06 \pm 0.23$, where the uncertainty is dominated by the statistical uncertainty.

6.2 Reducible backgrounds

6.2.1 Non-prompt leptons

Non-prompt leptons originate from material conversions, heavy-flavour hadron decays, or the improper reconstruction of other particles, with an admixture strongly depending on the lepton quality requirements and varying across event categories. These backgrounds are in general small in all 2ℓ and 3ℓ SRs and thus are estimated from simulation, with the normalisation determined by the likelihood fit. The non-prompt lepton background contribution in the 4ℓ SR is negligible and is therefore taken from simulation without dedicated data-driven corrections. The main contribution to the non-prompt-lepton background is from $t\bar{t}$ production, with much smaller contributions from V+jets and single-top-quark processes. The non-prompt leptons in the simulated samples are labelled according to whether they originate from heavy-flavour (HF) or light-flavour (LF) hadron decays, or from a material conversion candidate (Mat. Conv.). The HF category includes leptons from both bottom and charm decays.

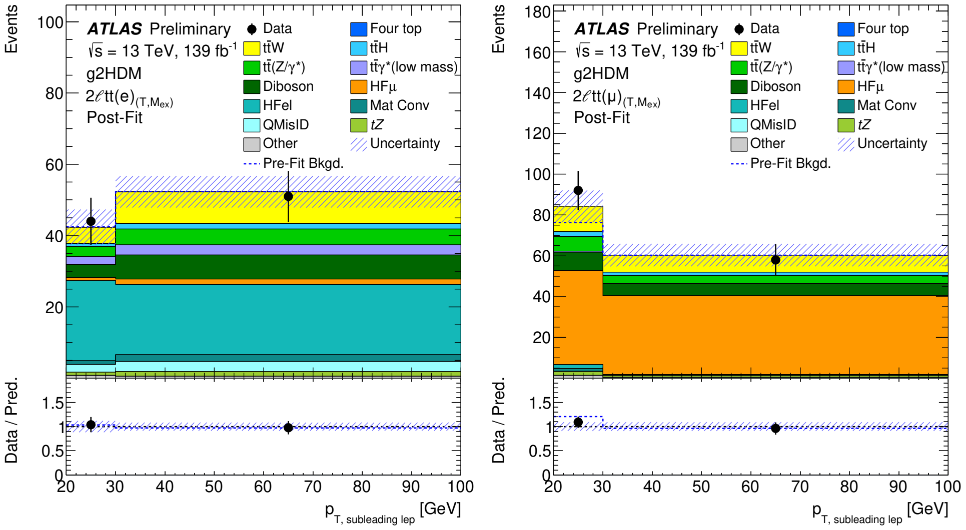
Two corrections are applied to the $t\bar{t}$ and the overall non-prompt lepton background simulation before the fit. First, the $t\bar{t} + \geq 1$ b-jet contribution from simulation is known to be mismodelled and is therefore corrected by a factor of 1.3 as measured by a previous ATLAS analysis sensitive to the in-situ measurement of this contribution in the single- and opposite-sign di-lepton final states [125]. This correction is well motivated since the mismodeling of additional b-jets in $t\bar{t}$ is not expected to depend on the presence of additional non-prompt leptons in the event. Second, the shape of the b-jet multiplicity in the non-prompt lepton background simulation is corrected to match data in an orthogonal 2ℓSS validation region enriched with non-prompt leptons, where one of the leptons is required to pass a looser non-prompt lepton BDT score but not pass the M lepton WP.

Several of the event categories introduced in Section 5 were designed to be enriched in specific processes and are used to derive normalisation factors to improve their modelling by the simulation. The 3ℓMatC CR is enriched in material conversions and only the total event yield is used. There are six 2ℓ CRs enriched in contributions from HF non-prompt leptons in $t\bar{t}$ events, i.e. $2\ell\mathrm{tt(e)}_{(T,M_{\mathrm{ex}})}$, $2\ell\mathrm{tt(e)}_{(M_{\mathrm{ex}},T)}$, $2\ell\mathrm{tt(e)}_{(M_{\mathrm{ex}},M_{\mathrm{ex}})}$, $2\ell\mathrm{tt}(\mu)_{(T,M_{\mathrm{ex}})}$, $2\ell\mathrm{tt}(\mu)_{(M_{\mathrm{ex}},T)}$, and $2\ell\mathrm{tt}(\mu)_{(M_{\mathrm{ex}},M_{\mathrm{ex}})}$. In these CRs, the transverse momentum of the sub-leading lepton, $p_{\mathrm{T,\,sub\text{-}lead\;lep}}$, distribution is used to be able to correct for a possible mismodeling in the p_{T} of the non-prompt lepton. The event requirement to have at least one M_{ex} lepton provides separation from the irreducible backgrounds, in particular $t\bar{t}W$, and thus increases the sensitivity to the HF non-prompt electron and muon contributions. Normalisation factors for three non-prompt-lepton background contributions are estimated from the likelihood fit. The normalisation factor for HF non-prompt leptons is estimated separately for electrons and muons, λ_e^{had} and $\lambda_\mu^{\mathrm{had}}$ respectively. An additional normalisation factor is determined for the material conversions background, $\lambda_e^{\mathrm{Mat\;Conv}}$. The measured normalisation factors are: $\hat{\lambda}_e^{\mathrm{had}} = 1.05 \pm 0.31$, $\hat{\lambda}_\mu^{\mathrm{had}} = 0.92 \pm 0.18$, and $\hat{\lambda}_e^{\mathrm{Mat\;Conv}} = 1.16 \pm 0.29$, where the uncertainties are dominated by the statistical uncertainty.

Figures 8(a) and 8(b) display the sub-leading lepton p_{T} distribution in the $2\ell\mathrm{tt(e)}_{(T,M_{\mathrm{ex}})}$ and $2\ell\mathrm{tt}(\mu)_{(T,M_{\mathrm{ex}})}$ CRs after the likelihood fit to data. As shown in the figures, the purity of HF non-prompt lepton background is 43% and 61%, respectively, which was possible to achieve with the usage of the exclusive M_{ex} lepton working point.

6.2.2 Charge misassignment

Backgrounds with leptons with the charge incorrectly assigned affect primarily the 2ℓ channel and predominantly arise from $t\bar{t}$ production, with one electron having undergone a hard bremsstrahlung as well as an asymmetric conversion ($e^\pm \to e^\pm\gamma^* \to e^\pm e^+ e^-$) or a mismeasured track curvature. The muon charge misassignment rate is negligible in the p_{T} range relevant to this analysis. The electron charge misassignment rate is measured in data using samples of $Z \to e^+e^-$ events reconstructed as same-charge pairs and as opposite-charge pairs, with the background subtracted via a sideband method. The charge misassignment rate is parameterised as a function of electron p_{T} and $|\eta|$, and it varies from about 10^{-5} for low-p_{T} electrons ($17 \leq p_{\mathrm{T}} \leq 50$ GeV) with $|\eta| \leq 1.37$, to about 3×10^{-4} for high-p_{T} electrons ($p_{\mathrm{T}} \geq 100$ GeV) with $1.52 \leq |\eta| \leq 2$. To estimate the QMisID background in each of the corresponding event categories, the measured charge misassignment rate is then applied to data events satisfying the requirements of the 2ℓ channels, except that the two leptons are required to be of opposite charge.



(a) (b)

Figure 8: Comparison between data and the background prediction for the distribution of the transverse momentum of the sub-leading lepton ($p_{\text{T, sub-lead lep}}$) in (a) the $2\ell\text{tt(e)}_{(T,M_{\text{ex}})}$ CR and (b) the $2\ell\text{tt}(\mu)_{(T,M_{\text{ex}})}$ CR. The background contributions after the likelihood fit to data ("Post-Fit") under the background-only hypothesis are shown as filled histograms. The ratio of the data to the post-fit background prediction ("Pred.") is shown in the lower panel, separately for post-fit background (black points) and pre-fit background (dashed blue line). The size of the combined statistical and systematic uncertainty in the background prediction is indicated by the blue hatched band. The last bin in each figure contains the overflow.

7 Systematic uncertainties

The signal and background yields in each signal and control region may be affected by several sources of systematic uncertainty, described in the following subsections. Given the low background yields and good signal-to-background separation provided by the final discriminating variable used in the signal-rich event categories, the search sensitivity is determined by the limited number of data events rather than by the systematic uncertainties of the background estimation. The final uncertainty in the background estimate in the SRs is dominated by the uncertainty in the fitted background normalisations, in particular $t\bar{t}W$. A summary of all systematic uncertainties included in the analysis is given in Table 6.

7.1 Experimental uncertainties

The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [126], obtained using the LUCID-2 detector [127] for the primary luminosity measurements.

Uncertainties associated with the lepton selection arise from the trigger, reconstruction, identification and isolation efficiencies, and lepton momentum scale and resolution [97, 99, 102]. Uncertainties associated with the jet selection arise from the jet energy scale (JES), the JVT requirement and the jet energy resolution (JER) [104, 128].

The efficiency of the flavour-tagging algorithm is measured for each jet flavour using control samples in data and in simulation. From these measurements, correction factors are derived to correct the tagging rates in the simulation [109, 113, 129]. These systematic uncertainties are taken as uncorrelated between b-jets, c-jets, and light-flavour jets. An additional uncertainty is assigned to account for the extrapolation of the b-tagging efficiency measurement from the p_{T} region used to determine the correction factors to regions with higher transverse momentum [130]. This uncertainty is the leading experimental uncertainty in the analysis.

The treatment of the uncertainties associated with reconstructed objects is common to all analysis channels, and thus these are considered as fully correlated among different analysis regions.

7.2 Theoretical uncertainties

Modelling uncertainties on the main irreducible backgrounds are assessed through comparison with alternative MC samples, as listed in Table 1. Additional uncertainties are evaluated from renormalisation and factorisation scale variations by a factor of 0.5 and 2, relative to the nominal scales, for the $t\bar{t}W$, $t\bar{t}Z$, and diboson samples. An additional 50% uncertainty is assigned to $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}$ events with additional heavy-flavour jets. This uncertainty is not applied to diboson events with heavy-flavour since its normalisation is fit from data. The statistical uncertainty on the fitted parameters for the N_{jets} VV correction is propagated as an uncertainty on the diboson background. The leading theoretical uncertainties arise from $t\bar{t}W$ modelling and additional heavy-flavour uncertainties.

Finally, additional normalisation uncertainties are included for all processes where their normalisation is not obtained from the fit. The $t\bar{t}t\bar{t}$, $t\bar{t}H$, and tZ processes are assigned an uncertainty of 20% [87], 11% [86], and 5% [131] respectively. A 50% cross section uncertainty is assigned as a conservative estimate to the $t\bar{t}t$, tWZ, $t\bar{t}WW$, and triboson backgrounds, as well as to the additional $t\bar{t}W$ electroweak contribution, which are small backgrounds with low impact on the search.

Uncertainties on the modelling of the signal samples are evaluated through independent variations of the factorisation and renormalisation scales by a factor two. Additional uncertainties due to PDF effects are estimated through an ensemble of eigenvariations of the NNPDF set, and taking the difference to alternative PDF sets [132].

7.3 Non-prompt lepton uncertainties

The normalisation of HF non-prompt leptons is obtained from regions including at least one M_{ex} lepton and extrapolated to the signal regions where the same-sign leptons fulfill the T or M identification requirements. An uncertainty of 20% on the extrapolation from M_{ex} to T leptons is applied from the comparison of the relative efficiency between nominal and alternative $t\bar{t}$ MC samples. An additional 50% uncertainty is assigned to events originating from $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$, decorrelated between flavours. Validation regions with looser lepton requirements and further enriched in non-prompt leptons are defined. A good data/MC agreement is observed in all kinematic variables except for the number of b-jets. Based on this disagreement, an $N_{b-\text{jets}}$-dependent uncertainty is added to the HF non-prompt background ranging from 6%–40% for 1–3 additional b-jets in the non-prompt muon regions, and 10%–80% in the non-prompt electron regions.

Table 6: Sources of systematic uncertainty considered in the analysis. “N” means that the uncertainty is taken as normalisation-only for all processes and channels affected. Some of the systematic uncertainties are split into several components, as indicated by the number in the rightmost column.

Systematic uncertainty	Components
Signal modelling	
Cross section (N)	1
$t\bar{t}W$ modelling	
QCD scale	3
Generator	2
Electroweak cross section	1
Additional heavy-flavour	1
$t\bar{t}Z/\gamma^*$ (high mass) modelling	
QCD scale	2
Generator	2
Additional heavy-flavour	1
$t\bar{t}H$ modelling	
Cross section (N)	1
Parton shower and hadronisation model	1
Generator	1
QCD scale	1
Additional heavy-flavour	1
WZ modelling	
QCD scale	1
Cross section (N)	1
Extra-jets correction	1
$t\bar{t}t\bar{t}$ modelling	
Generator	1
Cross section (N)	1
Other background modelling	
Cross section (N)	6
Total (Signal and background modelling)	29
Luminosity	1
Pile-up reweighting	1
Physics objects	
Electron	6
Muon	15
Electron Non-prompt BDT	14
Muon Non-prompt BDT	20
Jet energy scale	30
Jet energy resolution	12
Jet vertex fraction	1
Jet flavour tagging	62
$E_{\mathrm{T}}^{\mathrm{miss}}$	3
Total (Experimental)	165
Data-driven reducible background estimates	
Material conversions modelling	1
Internal conversions modelling	1
Charge misassignment	1
HF non-prompt	8
$t\bar{t}$ additional heavy-flavour	2
Total (Data-driven reducible background)	13
Total (Overall)	207

The modelling of internal and material conversions is tested in dedicated validation regions with two tight same-sign leptons, requiring one of them to be a conversion candidate. An additional uncertainty of 10% and 50% is assigned to the material and internal conversion backgrounds respectively, evaluated from the data/MC agreement in the validation regions.

A systematic uncertainty of 10–60% is assigned to the background from electrons with misidentified charge. The uncertainty increases with electron p_{T} and decreases with $|\eta|$. The uncertainty is assessed combining the uncertainties from the measurement of the charge misassignment rate, the difference in rates from varying the m_Z window selection, and the different rates measured in data and $Z \to ee$ MC.

8 Results

A maximum-likelihood fit is performed on all bins in the 27 signal and control regions considered in this search to simultaneously determine the background and the g2HDM signal yields that are most consistent with the data. The DNN$^{\mathrm{SB}}$is used as the discriminating variable in the signal regions, whereas the $N_{b-\mathrm{jets}}$, subleading lepton p_{T} and event yields are fit in the control regions. The sum of all the g2HDM signal processes studied here (sstt, ttq, ttt, tttq, tttt) is considered as a single signal template and its acceptance in each category is predicted by the simulation.

The likelihood function $\mathcal{L}(\mu, \vec{\lambda}, \vec{\theta})$ is constructed as a product of Poisson probability terms over all bins considered in the search, and depends on the signal-strength parameter, μ, a multiplicative factor applied to the predicted yield for the g2HDM signal (depending on the coupling configuration ρ_{tt}, ρ_{tc}, ρ_{tu} and on the assumed mass m_H (m_A)), $\vec{\lambda}$, the normalisation factors for several backgrounds (see Section 6), and $\vec{\theta}$, a set of nuisance parameters (NP) encoding systematic uncertainties in the signal and background expectations [133]. Systematic uncertainties can impact the estimated signal and background rates, the migration of events between categories, and the shape of the fitted distributions; they are summarised in Table 6. Both μ and $\vec{\lambda}$ are treated as free parameters in the likelihood fit. The NPs $\vec{\theta}$ allow variations of the expectations for signal and background according to the systematic uncertainties, subject to Gaussian or Poisson constraints in the likelihood fit. Their fitted values represent the deviations from the nominal expectations that globally provide the best fit to the data. Statistical uncertainties in each bin due to the limited size of the simulated samples are taken into account by dedicated parameters using the Beeston–Barlow "lite" technique [134].

The test statistic q_μ is defined as the profile likelihood ratio: $q_\mu = -2\ln(\mathcal{L}(\mu, \hat{\hat{\vec{\lambda}}}_{\mu}, \hat{\hat{\vec{\theta}}}_{\mu})/\mathcal{L}(\hat{\mu}, \hat{\hat{\vec{\lambda}}}_{\hat{\mu}}, \hat{\hat{\vec{\theta}}}_{\hat{\mu}}))$, where $\hat{\mu}$, $\hat{\hat{\vec{\lambda}}}_{\hat{\mu}}$, and $\hat{\hat{\vec{\theta}}}_{\hat{\mu}}$ are the values of the parameters that maximise the likelihood function, and $\hat{\hat{\vec{\lambda}}}_{\mu}$ and $\hat{\hat{\vec{\theta}}}_{\mu}$ are the values of the parameters that maximise the likelihood function for a given value of μ. The test statistic q_μ is evaluated with the RooFit package [135]. A related statistic is used to determine the probability that the observed data are incompatible with the background-only hypothesis (i.e. the discovery test) by setting $\mu = 0$ in the profile likelihood ratio (q_0). The p-value (referred to as p_0) representing the probability of the data being compatible with the background-only hypothesis is estimated by integrating the distribution of q_0 from background-only pseudo-experiments, approximated using the asymptotic formulae given in Ref. [136], above the observed value of q_0. Some model dependence exists in the estimation of the p_0, as a given signal scenario needs to be assumed in the calculation of the denominator of q_0, even if the overall signal normalisation is allowed to float and is fit to data. The observed p_0 is checked for each explored signal scenario. Upper limits on the signal production cross section for each of the signal scenarios considered are derived by using q_μ in the $\mathrm{CL_s}$ method [137, 138]. For a given signal scenario, values of the production cross section (parameterised by μ) yielding $\mathrm{CL_s} < 0.05$, where $\mathrm{CL_s}$ is computed using the asymptotic approximation [136], are excluded at $\geq 95\%$ confidence level (CL).

The smallest p_0 value is observed when assuming a signal with $m_H = 1000$ GeV and (ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85), corresponding to a local significance of 2.81standard deviations. The signal cross section resulting from the fit to data for this g2HDM signal hypothesis is 138 fb, with fractional contributions of 71% ttq, 20% sstt, and 9% ttt. Figure 9 shows the local significance as a function of the three couplings normalised to the sum of the couplings. This normalisation is performed in order to eliminate one degree of freedom related to the total normalisation of the signal, which is not relevant for the computation of the significance.

A comparison of the distributions of observed and expected yields is shown Figure 10(a) for the 17 SRs, and Figure 10(b) for the 10 CRs, after the combined likelihood fit for the signal-plus-background hypothesis. The corresponding post-fit yields for the SRs can be found in Tables 7, 8, and 9 for the 2ℓSS positively-charged, 2ℓSS negatively-charged, and 3ℓ and 4ℓ SRs, respectively. The signal shown in the figures and tables is the g2HDM signal with couplings ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85, and mass of 1000 GeV, which corresponds to the largest observed significance above the background only hypothesis.

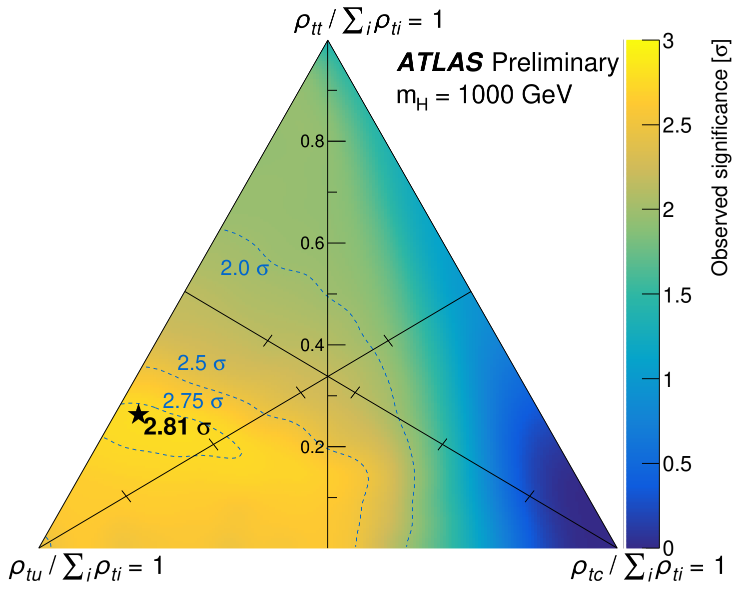


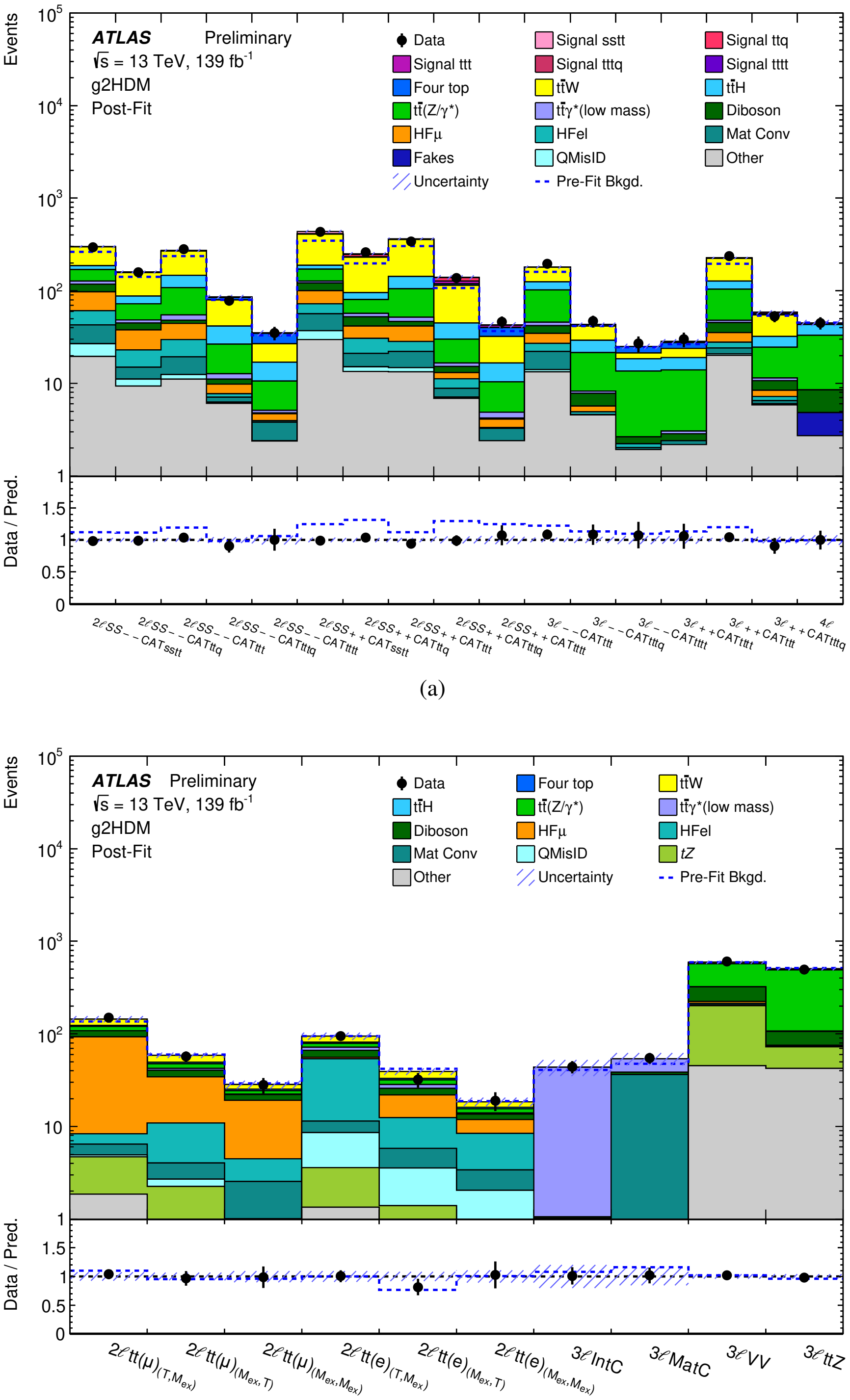
Figure 9: Observed significance for a heavy scalar with a mass of 1000 GeV as a function of the three couplings normalised to the sum of the couplings. This normalisation is performed in order to eliminate one degree of freedom related to the total normalisation of the signal which is not relevant for the computation of the significance. The star indicates the coupling configuration leading to the highest observed significance of 2.81 standard deviations.

In general, good agreement between the data and fitted signal-plus-background yields is found across all event categories.

The systematic uncertainties with the largest impact on the signal strength originate from the modelling of $t\bar{t}W$ with and without additional heavy flavour jets, $t\bar{t}Z$, $t\bar{t}H$, and $t\bar{t}t\bar{t}$ processes. The search is dominated by statistical uncertainties.

Comparisons between data and the background prediction for the DNN^{SB} distributions used in the different SRs is shown in Figures 11 and 12. The binning used for the DNN^{SB} distributions in the different SRs represents a compromise between preserving enough discrimination in the fit between the background and the signal for the different values of the heavy H mass considered and keeping the statistical uncertainty of the background prediction per bin well below 30%. The signal regions with the largest pre-fit tension between data and the background yields (shown in the blue dashed line) at high values of the DNN^{SB} are the 2ℓSS ++ CAT tttq, the 2ℓSS ++ CAT tttt, the 2ℓSS ++ CAT sstt, and the 2ℓSS ++ CAT ttq regions. Within this model, the $\hat{\lambda}_{t\bar{t}W}$ remains higher than 1, as observed by other analyses. However, charge-asymmetric tensions both at low and high jet and b-jet multiplicities are accommodated by the best fit g2HDM signal, where the largest signal contributions in the 2ℓSS ++ CAT tttq and the 2ℓSS ++ CAT tttt regions originate mainly from ttq and ttt processes (excess at high jet multiplicities), and the largest signal contributions in the 2ℓSS ++ CAT sstt and the 2ℓSS ++ CAT ttq regions originate mainly from the corresponding targeted signals sstt and ttq processes (excess at low jet multiplicities).

Exclusion limits on the heavy Higgs boson mass are set for different choices of the couplings, as shown in Figure 13. Masses of an additional scalar boson m_H between 200-630 (200-840) GeV with couplings $\rho_{tt} = 0.4$, $\rho_{tc} = 0.2$, and $\rho_{tu} = 0.2$ are observed (expected) to be excluded at 95% confidence level. Limits on the mass heavy Higgs boson mass are also set for a scenario without coupling to two top quarks, $\rho_{tt} = 0$,



(a)

(b)

Figure 10: Comparison between data and the background prediction for the event yields in (a) the 17 signal region categories and (b) the 10 control region categories. The expected signal for m_H = 1000 GeV and (ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85) as well as the background contributions are shown after the likelihood fit to data ("Post-Fit") under the signal-plus-background hypothesis. The total background prediction before the likelihood fit to data ("Pre-Fit") is shown as a dashed blue histogram in the upper panel. The ratio of the data to the total prediction is shown in the lower panel, separately for post-fit signal-plus-background (black points) and pre-fit background (dashed blue line). The size of the combined statistical and systematic uncertainty in the background prediction is indicated by the blue hatched band.

ρ_{tc} = 0.2, ρ_{tu} = 0.2, resulting on an observed (expected) limit of 200-350 (200-580) GeV on the heavy Higgs boson mass. No limits can be set on scenarios without off-diagonal couplings, leading only to four-top final states with a coupling set $\rho_{tt} = 1$, $\rho_{tc} = 0$, $\rho_{tu} = 0$. The sensitivity of the analysis on the four-top final state is similar to previous ATLAS analyses [37]. The excluded mass is also presented as a function of two couplings under different assumptions, as shown in Figure 14.

The search is also used to set limits on RPV SUSY models using the existing DNNs that were trained for the g2HDM model. Figures 15(a) and 15(b) show the exclusion limits obtained on the higgsino and wino models respectively. Higgsinos (winos) with masses up to 585 (670) GeV are excluded. Figure 16 shows limits on the smuon-bino model. Smuon masses up to 460 GeV are excluded, with weaker exclusion limits for small mass splittings between the smuon and the LSP, or for LSP masses close to the top-quark threshold.

Table 7: Post-fit yields of the 2ℓSS positively-charged signal regions. The best-fit signal for m_H = 1000 GeV and (ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85) is shown for a $\mu = 1.8 \pm 0.8$.

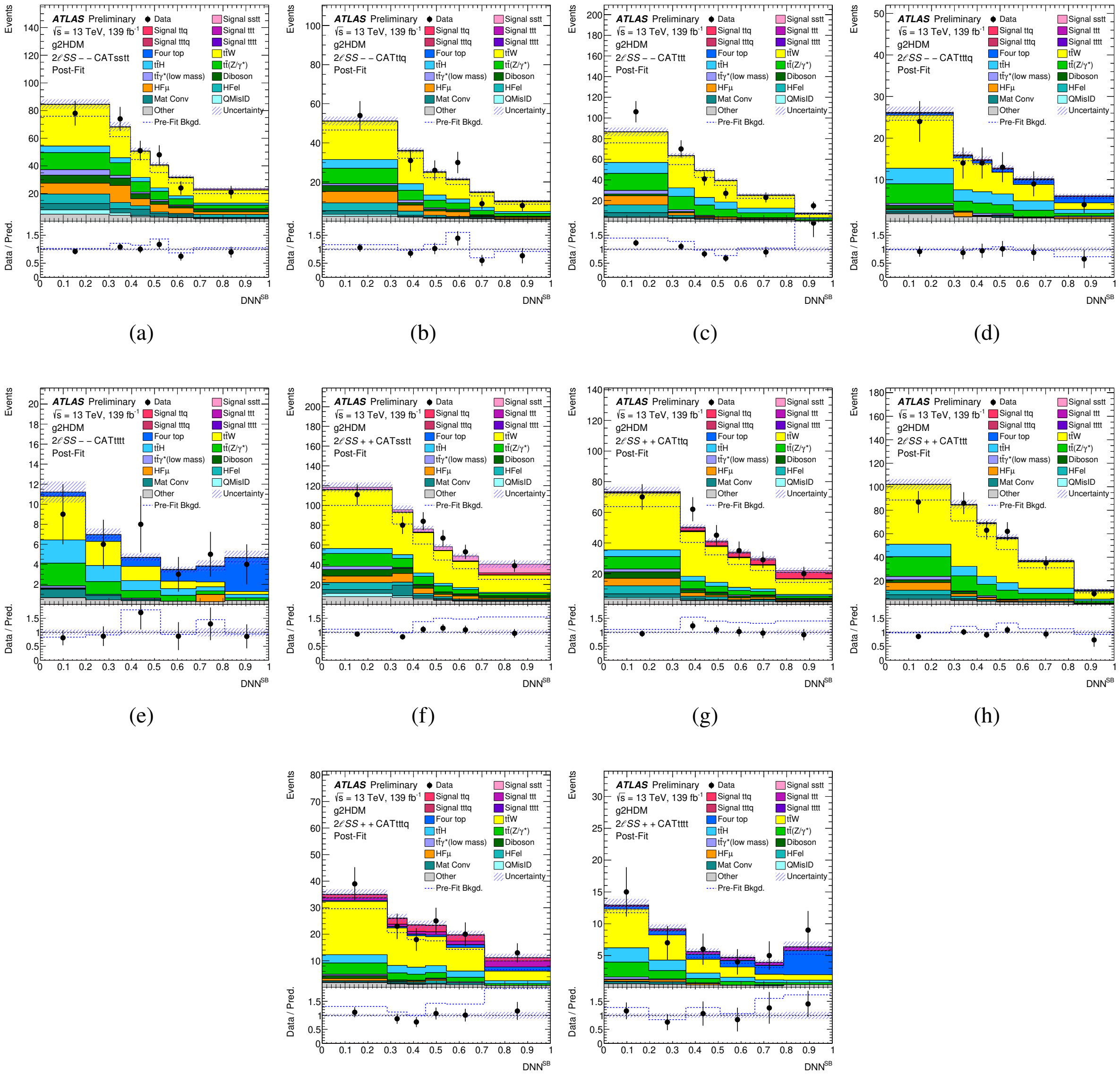
	2ℓSS ++ CAT sstt	2ℓSS ++ CAT ttq	2ℓSS ++ CAT ttt	2ℓSS ++ CAT tttq	2ℓSS ++ CAT tttt
Signal	8 ±4	16 ±7	3.8 ±1.7	23 ±10	4.2 ±1.9
$t\bar{t}t\bar{t}$	0.2 ±0.1	0.4 ±0.1	2.3 ±0.5	4.6 ±0.9	7.8 ±1.4
$t\bar{t}W$	231 ±19	139 ±10	219 ±16	68 ±7	15 ±2
$t\bar{t}H$	17 ±3	15 ±2	37 ±5	14 ±3	6.1 ±1.4
$t\bar{t}(Z/\gamma*)$	43 ±5	24 ±3	53 ±5	14 ±1	5.5 ±0.8
$t\bar{t}\gamma*$ (low mass)	5.3 ±3.0	4.0 ±2.4	4.9 ±2.8	1.3 ±0.8	0.6 ±0.4
VV	22 ±7	11 ±4	5.2 ±1.9	2.2 ±0.9	0.2 ±0.1
tZ	24 ±1	9.3 ±0.6	5.2 ±0.4	0.7 ±0.1	0.01 ±0.01
Non-prompt ℓ	45 ±17	20 ±7	19 ±6	4.2 ±1.9	0.8 ±0.7
Mat Conv	19 ±5	5.9 ±1.6	7.2 ±2.1	1.7 ±0.7	0.9 ±0.2
QMisID	7.4 ±2.7	1.7 ±0.6	1.4 ±0.5	0.2 ±0.1	0.02 ±0.01
Other	5.8 ±2.0	4.0 ±1.3	8.0 ±2.2	6.1 ±1.9	2.4 ±0.6
Total	428 ±15	250 ±9	367 ±12	141 ±9	43 ±3
Data	434	261	342	138	46

Table 8: Post-fit yields of the 2ℓ negatively-charged signal regions. The best-fit signal for m_H = 1000 GeV and (ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85) is shown for a $\mu = 1.8 \pm 0.8$.

	2ℓSS −− CAT sstt	2ℓSS −− CAT ttq	2ℓSS −− CAT ttt	2ℓSS −− CAT tttq	2ℓSS −− CAT tttt
Signal	1.0 ±0.4	2.1 ±1.0	0.5 ±0.2	3.0 ±1.3	0.6 ±0.3
$t\bar{t}t\bar{t}$	0.2 ±0.1	0.5 ±0.1	2.4 ±0.6	4.7 ±0.9	7.8 ±1.4
$t\bar{t}W$	115 ±9	72 ±5	126 ±9	37 ±4	9.4 ±1.4
$t\bar{t}H$	17 ±3	15 ±2	38 ±6	15 ±3	6.1 ±1.4
$t\bar{t}Z/\gamma*$	42 ±5	23 ±3	53 ±5	14 ±2	5.5 ±0.8
$t\bar{t}\gamma*$ (low mass)	9 ±5	3.3 ±1.9	6.2 ±3.5	1.4 ±0.9	0.3 ±0.3
VV	21 ±6	7.7 ±2.7	3.9 ±1.5	1.4 ±0.5	0.1 ±0.1
tZ	13 ±1	5.5 ±0.3	3.2 ±0.3	0.27 ±0.02	0.04 ±0.02
Non-prompt ℓ	54 ±16	23 ±8	25 ±8	2.6 ±0.8	0.8 ±0.5
Mat Conv	16 ±4	3.9 ±1.1	6.7 ±1.8	0.8 ±0.5	1.4 ±0.6
QMisID	7.4 ±2.7	1.7 ±0.6	1.4 ±0.5	0.2 ±0.1	0.02 ±0.01
Other	6.4 ±2.1	3.8 ±1.2	7.7 ±2.3	5.7 ±1.8	2.3 ±0.6
Total	302 ±10	161 ±5	273 ±8	86 ±4	34 ±2
Data	296	158	282	78	35

Table 9: Post-fit yields of the 3ℓ and 4ℓ signal regions. The best-fit signal for $m_H = 1000$ GeV and (ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85) is shown for a $\mu = 1.8 \pm 0.8$.

	3ℓ ++ CAT tttt	3ℓ ++ CAT ttt	3ℓ ++ CAT tttq	3ℓ $--$ CAT tttt	3ℓ $--$ CAT ttt	3ℓ $--$ CAT tttq	4ℓ
Signal	1.2 ±0.5	1.6 ±0.7	2.9 ±1.3	0.2 ±0.1	0.2 ±0.1	0.4 ±0.2	0.05 ±0.02
$t\bar{t}t\bar{t}$	3.7 ±0.7	0.8 ±0.2	1.7 ±0.3	3.7 ±0.7	0.9 ±0.2	1.7 ±0.3	1.5 ±0.9
$t\bar{t}W$	4.6 ±0.6	100 ±7	23 ±3	2.7 ±0.4	55 ±4	12 ±2	0.5 ±0.1
$t\bar{t}H$	4.8 ±1.0	23 ±3	7.5 ±1.5	4.8 ±1.0	23 ±3	7.6 ±1.4	10 ±3
$t\bar{t}Z/\gamma*$	11 ±1	55 ±6	13 ±2	11 ±1	56 ±6	13 ±2	24 ±3
$t\bar{t}\gamma*$ (low mass)	0.2 ±0.2	2.6 ±1.5	0.7 ±0.4	0 ±0	3.2 ±1.8	0.5 ±0.3	0 ±0
VV	0.5 ±0.2	11 ±4	2.4 ±0.9	0.4 ±0.2	7.9 ±3.0	2.1 ±0.6	3.6 ±1.3
tZ	0.27 ±0.02	13 ±1	2.5 ±0.2	0.11 ±0.01	7.2 ±0.4	1.2 ±0.1	0 ±0
Non-prompt ℓ	0 ±0	11 ±4	1.9 ±0.8	0.2 ±0.1	12 ±6	1.1 ±0.6	2.1 ±0.3
Mat Conv	0.2 ±0.1	3.2 ±1.1	0.5 ±0.3	0.1 ±0.1	8 ±5	0.04 ±0.01	0 ±0
Other	1.9 ±0.5	7.4 ±2.3	3.3 ±1.0	1.8 ±0.5	5.9 ±2.0	3.3 ±1.0	2.7 ±0.9
Total	28 ±2	229 ±6	60 ±3	25 ±2	180 ±7	43 ±2	45 ±3
Data	30	236	53	27	195	47	45

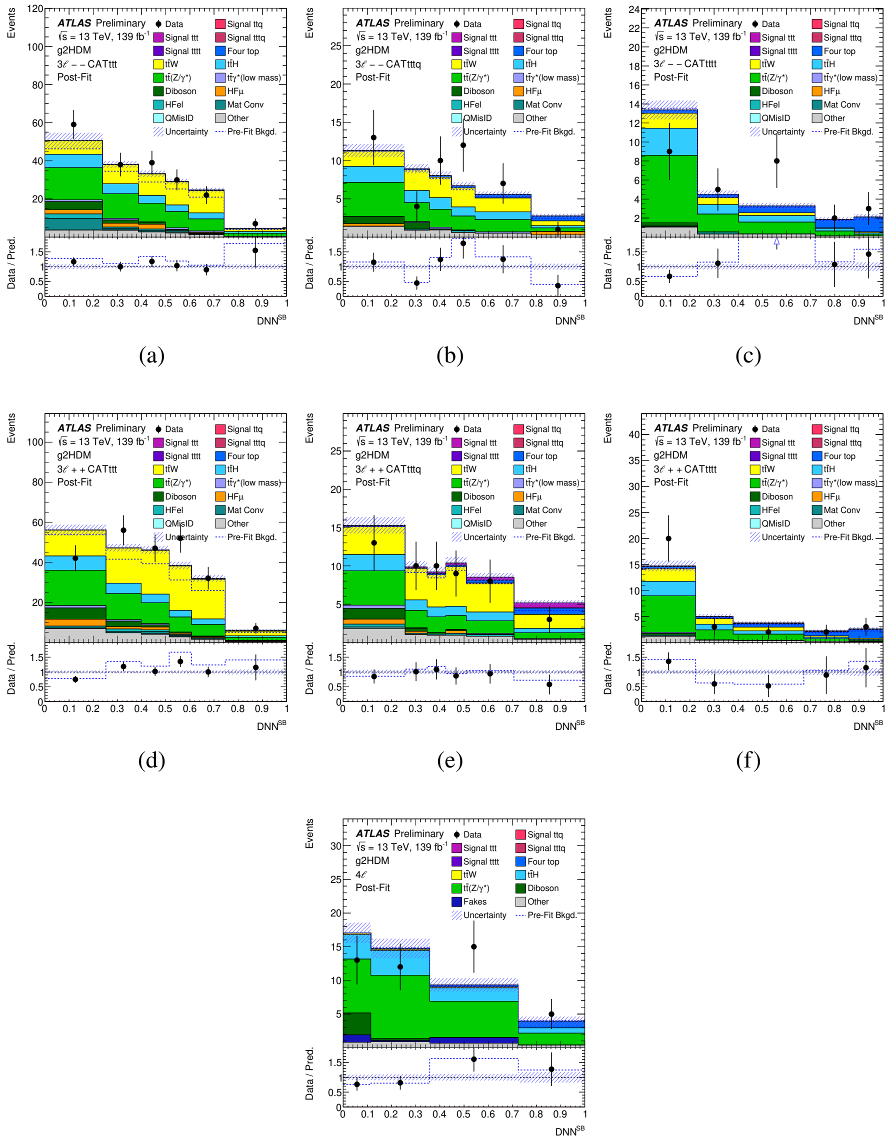


(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j)

Figure 11: Comparison between data and prediction for the DNN^{SB} distribution used in different signal region categories of the 2ℓSS channel: (a) 2ℓSS $--$ CAT sstt, (b) 2ℓSS $--$ CAT ttq, (c) 2ℓSS $--$ CAT ttt, (d) 2ℓSS $--$ CAT tttq, (e) 2ℓSS $--$ CAT tttt, (f) 2ℓSS $++$ CAT sstt, (g) 2ℓSS $++$ CAT ttq, (h) 2ℓSS $++$ CAT ttt, (i) 2ℓSS $++$ CAT tttq, and (j) 2ℓSS $++$ CAT tttt. The expected signal for $m_H = 1000$ GeV and (ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85) as well as the background contributions are shown after the likelihood fit to data (“Post-Fit”) under the signal-plus-background hypothesis. The ratio of the data to the prediction (“Pred.”) is shown in the lower panel, separately for post-fit signal-plus-background (black points) and pre-fit background (dashed blue line). The size of the combined statistical and systematic uncertainty in the signal-plus-background prediction is indicated by the blue hatched band. The last bin in each Figure contains the overflow.

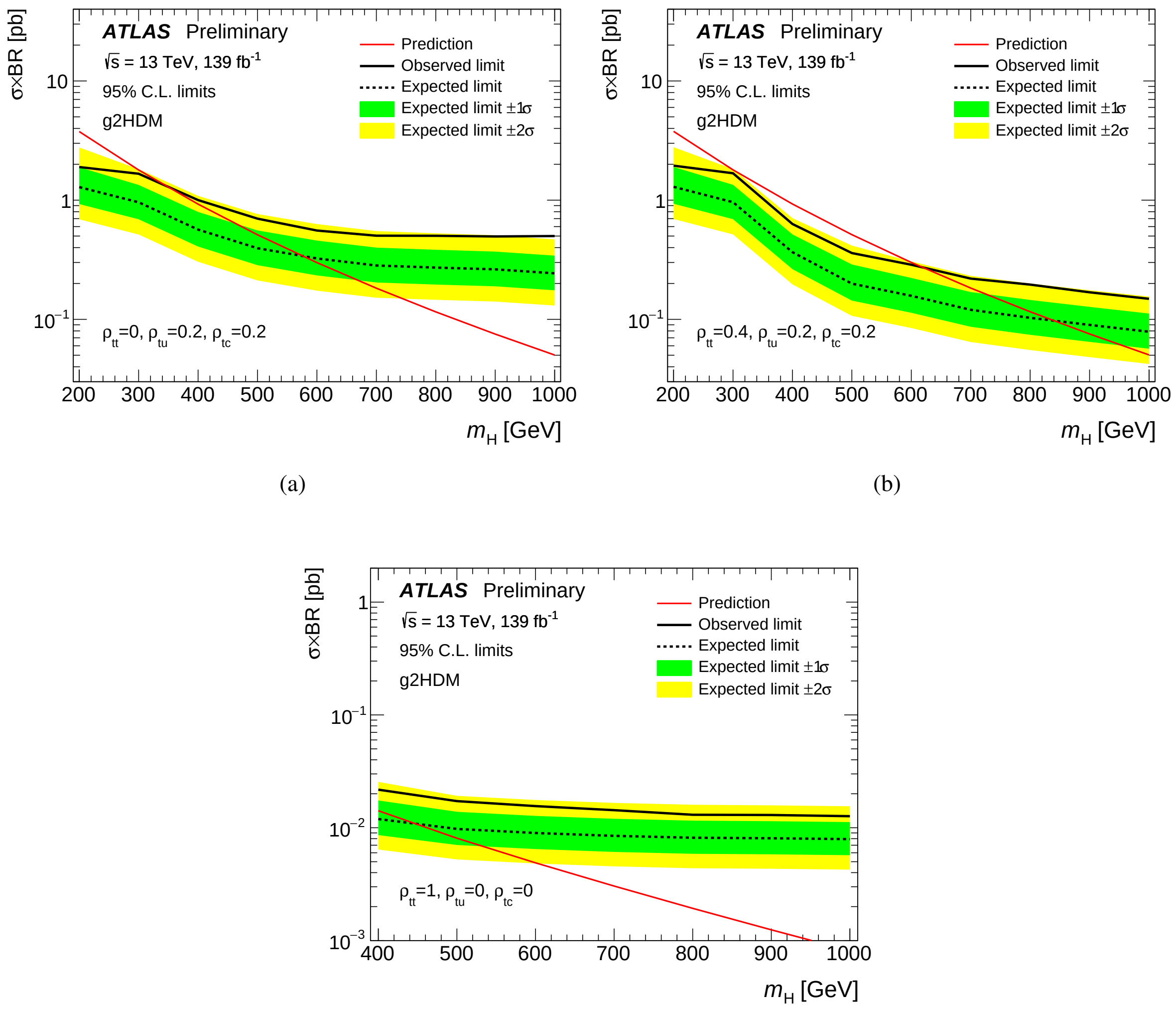


(a) (b) (c)

(d) (e) (f)

(g)

Figure 12: Comparison between data and prediction for the DNN^{SB} distribution used in different signal region categories of the 3ℓ and 4ℓ channels: (a) 3ℓ −− CAT ttt, (b) 3ℓ −− CAT tttq, (c) 3ℓ −− CAT tttt, (d) 3ℓ ++ CAT ttt, (e) 3ℓ ++ CAT tttq, (f) 3ℓ ++ CAT tttt, and (g) 4ℓ. The expected signal for $m_H = 1000$ GeV and (ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85) as well as the background contributions are shown after the likelihood fit to data ("Post-Fit") under the signal-plus-background hypothesis. The ratio of the data to the prediction ("Pred.") is shown in the lower panel, separately for post-fit signal-plus-background (black points) and pre-fit background (dashed blue line). The size of the combined statistical and systematic uncertainty in the signal-plus-background prediction is indicated by the blue hatched band. The last bin in each Figure contains the overflow.

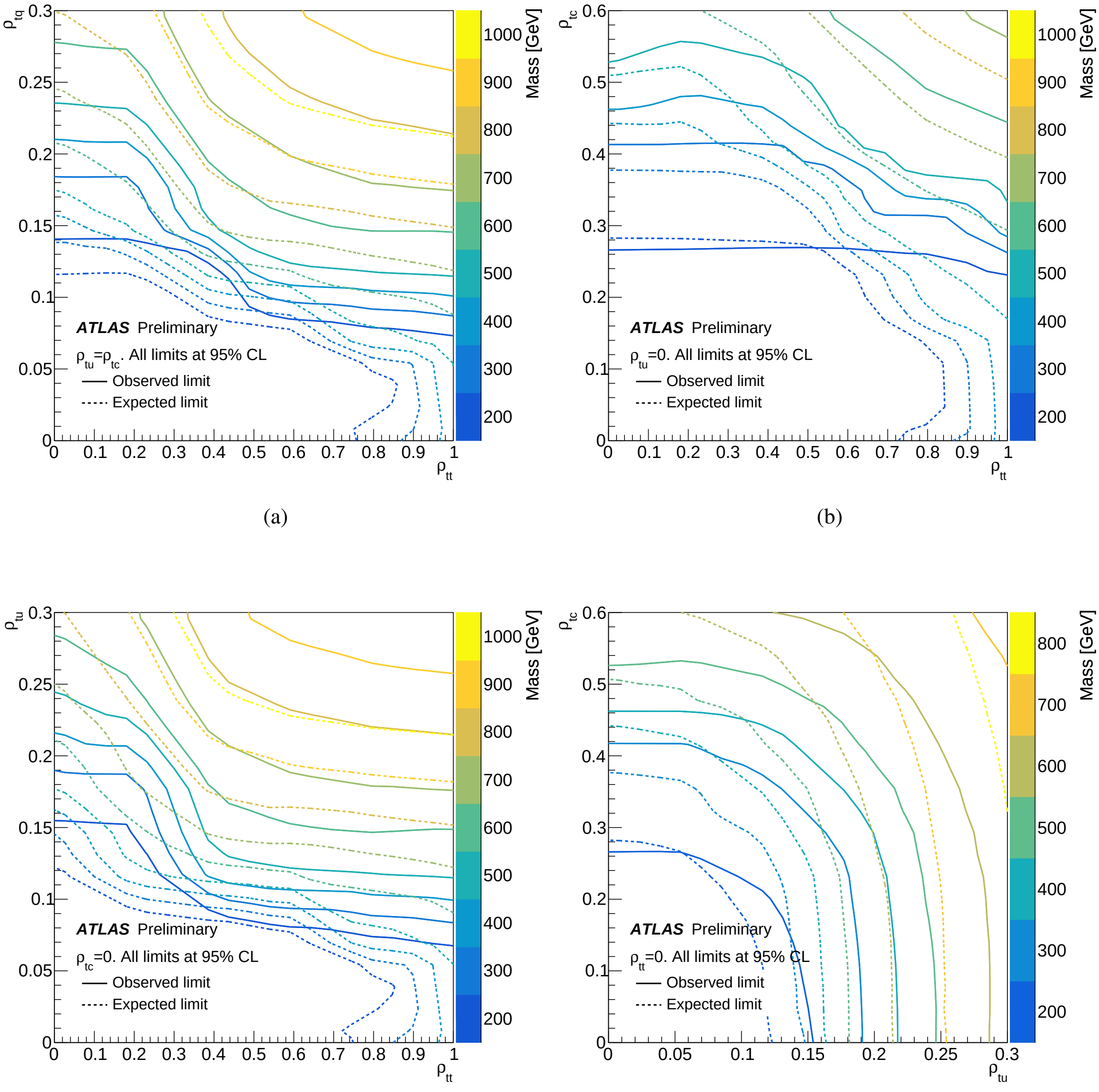


(a)

(b)

(c)

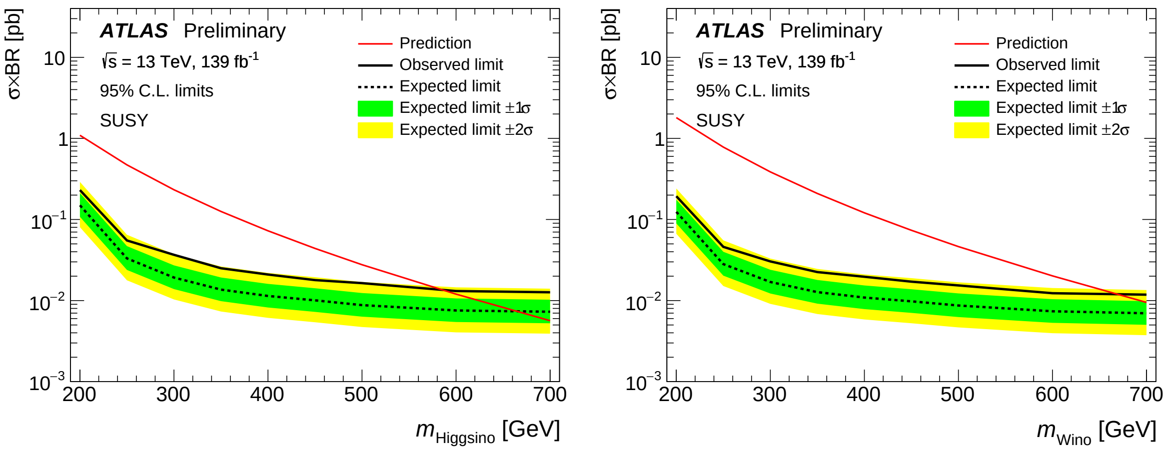
Figure 13: Observed and expected exclusion limits at 95% confidence level on the heavy Higgs boson mass for the g2HDM signal model for different couplings choices: (a) $\rho_{tt} = 0$, $\rho_{tc} = 0.2$, $\rho_{tu} = 0.2$, (b) $\rho_{tt} = 0.4$, $\rho_{tc} = 0.2$, $\rho_{tu} = 0.2$, and (c) $\rho_{tt} = 1$, $\rho_{tc} = 0$, $\rho_{tu} = 0$. The yellow and green contours of the band around the expected limit are the $\pm1\sigma$ and $\pm2\sigma$ variations including all uncertainties, respectively. The theoretical prediction for the signal production cross section is also shown as a red line. The production cross section is the sum of the five production modes considered in the search.



(a) (b)

(c) (d)

Figure 14: Observed (solid line) and expected (dashed line) exclusion limits on the scalar mass as a function of the coupling under different assumptions: (a) $\rho_{tc} = \rho_{tu}$, (b) $\rho_{tu} = 0$, (c) $\rho_{tc} = 0$, and (d) $\rho_{tt} = 0$.



(a)

(b)

Figure 15: Expected and observed exclusion limits on the cross section × branching ratio of electroweakino production in RPV SUSY models as a function of the sparticle masses: (a) higgsino model, and (b) wino model. The yellow and green contours of the band around the expected limit are the $\pm 1\sigma$ and $\pm 2\sigma$ variations including all uncertainties, respectively. The theoretical prediction for the signal production cross section is also shown as a red line.

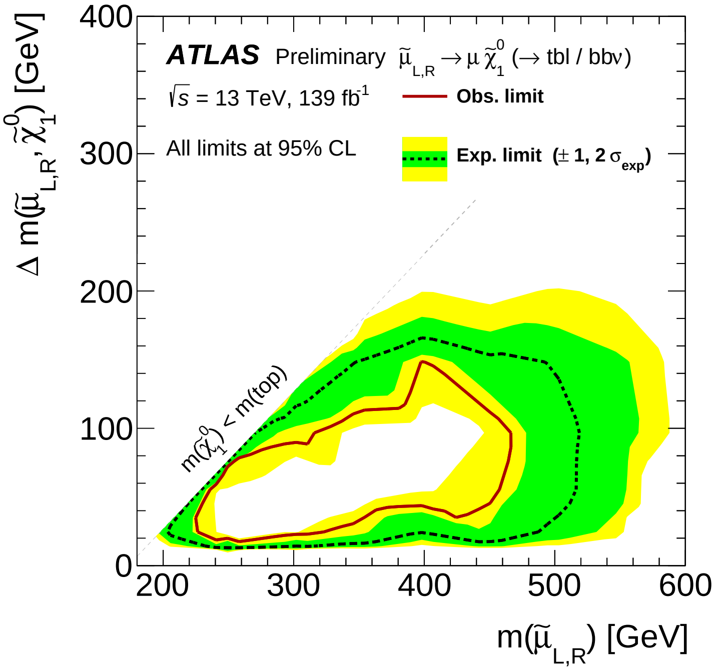


Figure 16: Expected and observed exclusion limits on the smuon plus bino RPV SUSY model. The yellow and green contours of the band around the expected limit are the $\pm 1\sigma$ and $\pm 2\sigma$ variations including all uncertainties, respectively. The diagonal grey lines indicates the allowed kinematic limit for the decays. The neutralino is assumed to decay to tbl or $bb\nu$ with equal probability.

9 Conclusion

A search for a general two Higgs doublet model is presented, where the heavy Higgs bosons feature flavour changing couplings. Such couplings allow for same-sign top and three-top production among others, with a sizeable charge asymmetry. The targeted final state is characterised by multiple leptons and multiple b-jets. In order to improve the sensitivity of the search, events are categorised according to the lepton multiplicity, total lepton charge, and a multi-output deep neural network classifier. The dominant backgrounds originate from $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}$, and are estimated from Monte-Carlo simulation and normalised to data. The analysis is performed with proton-proton collision data at $\sqrt{s} = 13$ TeVcollected from 2015 to 2018 with the ATLAS detector at the LHC, corresponding to an integrated luminosity of 139 fb^{-1}. This search is the first collider result on general two Higgs doublet model with flavour violation. It also represents the first search to target explicitly beyond-the-standard-model production of three top quarks.

A mild excess is observed over the Standard Model expectation corresponding to a local significance of 2.81standard deviations for a signal with $m_H = 1000$ GeV and (ρ_{tt}=0.32, ρ_{tc}=0.05, and ρ_{tu}=0.85). The observed charge-asymmetric tensions both at low and high jet and b-jet multiplicities are accommodated by the best fit g2HDM signal, where the largest signal contributions in the 2ℓSS ++ CAT tttq and the 2ℓSS ++ CAT tttt regions originate mainly from ttq and ttt processes (excess at high jet multiplicities), and the largest signal contributions in the 2ℓSS ++ CAT sstt and the 2ℓSS ++ CAT ttq regions originate mainly from the corresponding targeted signals sstt and ttq processes (excess at low jet multiplicities). Since the excess is not significant, 95% exclusion limits are also set on the mass and couplings of the heavy Higgs bosons. An additional scalar boson with couplings $\rho_{tt} = 0.4$, $\rho_{tc} = 0.2$, and $\rho_{tu} = 0.2$ is observed (expected) to be excluded at 95% confidence level for masses m_H between 200-630 (200-840) GeV. Additional mass limits are set for different coupling choices. For a fixed mass of $m_H = 400$ GeV exclusion limits are set on the allowed coupling strengths, as low as $\rho_{tt} = 0.3$, $\rho_{tc} = \rho_{tu} = 0.16$. Different assumptions are tested to set 2-dimensional exclusion limits on the three couplings. Additional models based on R-parity-violating supersymmetry with the lepton-number-violating coupling λ'_{i33} (with $i \in 2, 3$), are used to further interpret the results of the search. Scenarios with direct electroweak production of higgsinos (winos) are excluded for masses between 200-585 (200-670) GeV. Smuons with masses between 225 and 600 GeVare excluded in a model with direct smuon production, decaying to a bino-like neutralino, which in turn decays via the λ'_{i33} coupling.

References

[1] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1, arXiv: `1106.0034 [hep-ph]` (cit. on p. 2).

[2] J. F. Gunion and H. E. Haber, *The CP conserving two Higgs doublet model: The Approach to the decoupling limit*, Phys. Rev. D **67** (2003) 075019, arXiv: `hep-ph/0207010` (cit. on p. 2).

[3] ATLAS Collaboration, *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, (2021), arXiv: `2112.11876 [hep-ex]` (cit. on p. 2).

[4] ATLAS Collaboration, *Search for dark matter produced in association with a Standard Model Higgs boson decaying into b-quarks using the full Run 2 dataset from the ATLAS detector*, JHEP **11** (2021) 209, arXiv: `2108.13391 [hep-ex]` (cit. on p. 2).

[5] ATLAS Collaboration, *Search for resonances decaying into photon pairs in* $139\,fb^{-1}$ *of pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Lett. B **822** (2021) 136651, arXiv: `2102.13405 [hep-ex]` (cit. on p. 2).

[6] ATLAS Collaboration, *Search for charged Higgs bosons decaying into a top quark and a bottom quark at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, JHEP **06** (2021) 145, arXiv: `2102.10076 [hep-ex]` (cit. on p. 2).

[7] ATLAS Collaboration, *Search for a heavy Higgs boson decaying into a Z boson and another heavy Higgs boson in the* $\ell\ell bb$ *and* $\ell\ell WW$ *final states in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2020) 396, arXiv: `2011.05639 [hep-ex]` (cit. on p. 2).

[8] 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]` (cit. on p. 2).

[9] 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]` (cit. on p. 2).

[10] ATLAS Collaboration, *Search for Heavy Higgs Bosons Decaying into Two Tau Leptons with the ATLAS Detector Using pp Collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. Lett. **125** (2020) 051801, arXiv: `2002.12223 [hep-ex]` (cit. on p. 2).

[11] CMS Collaboration, *Search for heavy resonances decaying to WW, WZ, or WH boson pairs in a final state of a lepton and a large jet in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. D **105** (2021) 032008, arXiv: `2109.06055 [hep-ex]` (cit. on p. 2).

[12] CMS Collaboration, *Search for a heavy Higgs boson decaying into two lighter Higgs bosons in the* $\tau\tau bb$ *final state at* 13 *TeV*, JHEP **11** (2021) 057, arXiv: `2106.10361 [hep-ex]` (cit. on p. 2).

[13] CMS Collaboration, *Search for charged Higgs bosons produced in vector boson fusion processes and decaying into vector boson pairs in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **81** (2021) 723, arXiv: `2104.04762 [hep-ex]` (cit. on p. 2).

[14] CMS Collaboration, *Search for resonant pair production of Higgs bosons in the bbZZ channel in proton–proton collisions at* $\sqrt{s} = 13\,TeV$, Phys. Rev. D **102** (2020) 032003, arXiv: `2006.06391 [hep-ex]` (cit. on p. 2).

[15] CMS Collaboration, *Search for a light charged Higgs boson in the* $H^{\pm} \to cs$ *channel in proton–proton collisions at* $\sqrt{s} = 13\,TeV$, Phys. Rev. D **102** (2020) 072001, arXiv: `2005.08900 [hep-ex]` (cit. on p. 2).

[16] CMS Collaboration, *Search for charged Higgs bosons decaying into a top and a bottom quark in the all-jet final state of pp collisions at* $\sqrt{s} = 13\,TeV$, JHEP **07** (2020) 126, arXiv: `2001.07763 [hep-ex]` (cit. on p. 2).

[17] 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]` (cit. on p. 2).

[18] CMS Collaboration, *Search for new neutral Higgs bosons through the* $H \to ZA \to \ell^{+}\ell^{-}b\bar{b}$ *process in pp collisions at* $\sqrt{s} = 13\,TeV$, JHEP **03** (2020) 055, arXiv: `1911.03781 [hep-ex]` (cit. on p. 2).

[19] CMS Collaboration, *Search for a heavy pseudoscalar Higgs boson decaying into a* 125 *GeV Higgs boson and a Z boson in final states with two tau and two light leptons at* $\sqrt{s} = 13\,TeV$, JHEP **03** (2020) 065, arXiv: `1910.11634 [hep-ex]` (cit. on p. 2).

[20] CMS Collaboration, *Search for a charged Higgs boson decaying into top and bottom quarks in events with electrons or muons in proton–proton collisions at* $\sqrt{s} = 13\,TeV$, JHEP **01** (2020) 096, arXiv: `1908.09206 [hep-ex]` (cit. on p. 2).

[21] CMS Collaboration, *Search for heavy Higgs bosons decaying to a top quark pair in proton–proton collisions at* $\sqrt{s} = 13\,TeV$, JHEP **04** (2020) 171, arXiv: `1908.01115 [hep-ex]` (cit. on p. 2).

[22] CMS Collaboration, *Search for charged Higgs bosons in the* $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ *decay channel in proton–proton collisions at* $\sqrt{s} = 13\,TeV$, JHEP **07** (2019) 142, arXiv: `1903.04560 [hep-ex]` (cit. on p. 2).

[23] 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]` (cit. on p. 2).

[24] W.-S. Hou and M. Kikuchi, *Approximate Alignment in Two Higgs Doublet Model with Extra Yukawa Couplings*, EPL **123** (2018) 11001, arXiv: `1706.07694 [hep-ph]` (cit. on p. 2).

[25] K. Fuyuto, W.-S. Hou and E. Senaha, *Electroweak baryogenesis driven by extra top Yukawa couplings*, Phys. Lett. B **776** (2018) 402, arXiv: `1705.05034 [hep-ph]` (cit. on p. 2).

[26] W. Altmannshofer et al., *Collider Signatures of Flavorful Higgs Bosons*, Phys. Rev. D **94** (2016) 115032, arXiv: `1610.02398 [hep-ph]` (cit. on p. 2).

[27] M. Kohda, T. Modak and W.-S. Hou, *Searching for new scalar bosons via triple-top signature in* $cg \to tS^{0} \to tt\bar{t}$, Phys. Lett. B **776** (2018) 379, arXiv: `1710.07260 [hep-ph]` (cit. on p. 2).

[28] W.-S. Hou, M. Kohda and T. Modak, *Implications of Four-Top and Top-Pair Studies on Triple-Top Production*, Phys. Lett. B **798** (2019) 134953, arXiv: `1906.09703 [hep-ph]` (cit. on p. 2).

[29] G. W. S. Hou, *Top-Higgs Associated Production involving A^0, H^0 with Mass at 300 GeV*, PoS **EPS-HEP2019** (2020) 337, arXiv: `1910.08002 [hep-ph]` (cit. on p. 2).

[30] ATLAS Collaboration,
Analysis of $t\bar{t}H$ and $t\bar{t}W$ production in multilepton final states with the ATLAS detector, ATLAS-CONF-2019-045, 2019, URL: `https://cds.cern.ch/record/2693930` (cit. on p. 2).

[31] 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]` (cit. on p. 2).

[32] 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]` (cit. on p. 2).

[33] ATLAS Collaboration, *Measurement of the $t\bar{t}t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **11** (2021) 118, arXiv: `2106.11683 [hep-ex]` (cit. on p. 2).

[34] ATLAS Collaboration, *Search for R-parity-violating supersymmetry in a final state containing leptons and many jets with the ATLAS experiment using $\sqrt{s} = 13$ TeV proton–proton collision data*, Eur. Phys. J. C **81** (2021) 1023, arXiv: `2106.09609 [hep-ex]` (cit. on p. 2).

[35] ATLAS Collaboration, *Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 052009, arXiv: `1811.02305 [hep-ex]` (cit. on p. 2).

[36] ATLAS Collaboration, *Search for heavy resonances in four-top-quark final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2021-048, 2021, URL: `https://cds.cern.ch/record/2781173` (cit. on p. 2).

[37] ATLAS Collaboration, *Search for $t\bar{t}H/A \to t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2022-008, 2022, URL: `https://cds.cern.ch/record/2805212` (cit. on pp. 2, 27).

[38] CMS Collaboration, *Search for standard model production of four top quarks with same-sign and multilepton final states in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **78** (2018) 140, arXiv: `1710.10614 [hep-ex]` (cit. on p. 2).

[39] CMS Collaboration, *Search for the production of four top quarks in the single-lepton and opposite-sign dilepton final states in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **11** (2019) 082, arXiv: `1906.02805 [hep-ex]` (cit. on p. 2).

[40] CMS Collaboration, *Search for production of four top quarks in final states with same-sign or multiple leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **80** (2020) 75, arXiv: `1908.06463 [hep-ex]` (cit. on p. 2).

[41] CMS Collaboration, *Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **77** (2017) 578, arXiv: `1704.07323 [hep-ex]` (cit. on p. 2).

[42] B. Altunkaynak, W.-S. Hou, C. Kao, M. Kohda and B. McCoy,
Flavor Changing Heavy Higgs Interactions at the LHC, Phys. Lett. B **751** (2015) 135, arXiv: `1506.00651 [hep-ph]` (cit. on p. 2).

[43] A. Crivellin, A. Kokulu and C. Greub, *Flavor-phenomenology of two-Higgs-doublet models with generic Yukawa structure*, Phys. Rev. D **87** (2013) 094031, arXiv: `1303.5877 [hep-ph]` (cit. on p. 2).

[44] Belle Collaboration, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, Phys. Rev. Lett. **124** (2020) 161803, arXiv: `1910.05864 [hep-ex]` (cit. on p. 2).

[45] Belle Collaboration, *Test of Lepton-Flavor Universality in $B \to K^* \ell^+ \ell^-$ Decays at Belle*, Phys. Rev. Lett. **126** (2021) 161801, arXiv: `1904.02440 [hep-ex]` (cit. on p. 2).

[46] Belle Collaboration, *Test of lepton flavor universality and search for lepton flavor violation in $B \to K\ell\ell$ decays*, JHEP **03** (2021) 105, arXiv: `1908.01848 [hep-ex]` (cit. on p. 2).

[47] LHCb Collaboration, *Angular Analysis of the $B^+ \to K^{*+} \mu^+ \mu^-$ Decay*, Phys. Rev. Lett. **126** (2021) 161802, arXiv: `2012.13241 [hep-ex]` (cit. on p. 2).

[48] LHCb Collaboration, *Measurement of CP-Averaged Observables in the $B^0 \to K^{*0} \mu^+ \mu^-$ Decay*, Phys. Rev. Lett. **125** (2020) 011802, arXiv: `2003.04831 [hep-ex]` (cit. on p. 2).

[49] LHCb Collaboration, *Measurement of the ratio of the $B^0 \to D^{*-} \tau^+ \nu_\tau$ and $B^0 \to D^{*-} \mu^+ \nu_\mu$ branching fractions using three-prong τ-lepton decays*, Phys. Rev. Lett. **120** (2018) 171802, arXiv: `1708.08856 [hep-ex]` (cit. on p. 2).

[50] LHCb Collaboration, *Test of lepton universality with $B^0 \to K^{*0} \ell^+ \ell^-$ decays*, JHEP **08** (2017) 055, arXiv: `1705.05802 [hep-ex]` (cit. on p. 2).

[51] LHCb Collaboration, *Test of lepton universality in beauty-quark decays*, (2021), arXiv: `2103.11769 [hep-ex]` (cit. on p. 2).

[52] Muon g-2 Collaboration, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, Phys. Rev. Lett. **126** (2021) 141801, arXiv: `2104.03281 [hep-ex]` (cit. on p. 3).

[53] A. Chakraborty and S. Chakraborty, *Probing $(g-2)_\mu$ at the LHC in the paradigm of R-parity violating MSSM*, Phys. Rev. D **93** (2016) 075035, arXiv: `1511.08874 [hep-ph]` (cit. on p. 3).

[54] M.-D. Zheng and H.-H. Zhang, *Studying the $b \to s\ell^+\ell^-$ Anomalies and $(g-2)_\mu$ in RPV-MSSM Framework with Inverse Seesaw*, (2021), arXiv: `2105.06954 [hep-ph]` (cit. on p. 3).

[55] W. Altmannshofer, P. S. B. Dev, A. Soni and Y. Sui, *Addressing $R_{D^{(*)}}$, $R_{K^{(*)}}$, muon $g-2$ and ANITA anomalies in a minimal R-parity violating supersymmetric framework*, Phys. Rev. D **102** (2020) 015031, arXiv: `2002.12910 [hep-ph]` (cit. on p. 3).

[56] D. Das, C. Hati, G. Kumar and N. Mahajan, *Scrutinizing R-parity violating interactions in light of $R_{K^{(*)}}$ data*, Phys. Rev. D **96** (2017) 095033, arXiv: `1705.09188 [hep-ph]` (cit. on p. 3).

[57] K. Earl and T. Grégoire, *Contributions to $b \to s\ell\ell$ Anomalies from R-Parity Violating Interactions*, JHEP **08** (2018) 201, arXiv: `1806.01343 [hep-ph]` (cit. on p. 3).

[58] N. G. Deshpande and X.-G. He, *Consequences of R-parity violating interactions for anomalies in $\bar{B} \to D^{(*)}\tau\bar{\nu}$ and $b \to s\mu^+\mu^-$*, Eur. Phys. J. C **77** (2017) 134, arXiv: `1608.04817 [hep-ph]` (cit. on p. 3).

[59] Q.-Y. Hu, Y.-D. Yang and M.-D. Zheng,
Revisiting the B-physics anomalies in R-parity violating MSSM, Eur. Phys. J. C **80** (2020) 365,
arXiv: `2002.09875 [hep-ph]` (cit. on p. 3).

[60] Q.-Y. Hu and L.-L. Huang,
Explaining $b \to s\ell^+\ell^-$ *data by sneutrinos in the R -parity violating MSSM*,
Phys. Rev. D **101** (2020) 035030, arXiv: `1912.03676 [hep-ph]` (cit. on p. 3).

[61] W. Altmannshofer, P. S. Bhupal Dev and A. Soni,
$R_{D^{(*)}}$ *anomaly: A possible hint for natural supersymmetry with R-parity violation*,
Phys. Rev. D **96** (2017) 095010, arXiv: `1704.06659 [hep-ph]` (cit. on p. 3).

[62] S. Trifinopoulos,
Revisiting R-parity violating interactions as an explanation of the B-physics anomalies,
Eur. Phys. J. C **78** (2018) 803, arXiv: `1807.01638 [hep-ph]` (cit. on p. 3).

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

[64] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*,
ATLAS-TDR-19; CERN-LHCC-2010-013, 2010,
URL: `https://cds.cern.ch/record/1291633` (cit. on p. 4),
Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009, 2012, URL:
`https://cds.cern.ch/record/1451888`.

[65] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*,
JINST **13** (2018) T05008, arXiv: `1803.00844 [physics.ins-det]` (cit. on p. 4).

[66] ATLAS Collaboration,
The performance of the jet trigger for the ATLAS detector during 2011 data taking,
Eur. Phys. J. C **76** (2016) 526, arXiv: `1606.07759 [hep-ex]` (cit. on p. 4).

[67] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*,
ATL-SOFT-PUB-2021-001, 2021, URL: `https://cds.cern.ch/record/2767187`
(cit. on p. 4).

[68] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007,
arXiv: `0811.4622 [hep-ph]` (cit. on p. 4).

[69] D. J. Lange, *The EvtGen particle decay simulation package*,
Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 4).

[70] 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. 4).

[71] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196,
arXiv: `1512.01178 [hep-ph]` (cit. on p. 4).

[72] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A brief introduction to PYTHIA 8.1*,
Comput. Phys. Commun. **178** (2008) 852, arXiv: `0710.3820 [hep-ph]` (cit. on pp. 4, 5).

[73] ATLAS Collaboration, *Further ATLAS tunes of* PYTHIA *6 and Pythia 8*,
ATL-PHYS-PUB-2011-014, 2011, URL: `https://cds.cern.ch/record/1400677`
(cit. on p. 4).

[74] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250
(cit. on p. 4).

[75] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: `1005.4568 [physics.ins-det]` (cit. on p. 4).

[76] 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]` (cit. on p. 5).

[77] R. D. Ball et al., *Parton distributions from high-precision collider data*, Eur. Phys. J. C **77** (2017) 663, arXiv: `1706.00428 [hep-ph]` (cit. on p. 5).

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

[79] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: `1905.09127 [hep-ph]` (cit. on p. 5).

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

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

[82] 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).

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

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

[85] R. Frederix and I. Tsinikos, *On improving NLO merging for* $t\bar{t}W$ *production*, JHEP **11** (2021) 029, arXiv: `2108.07826 [hep-ph]` (cit. on p. 5).

[86] 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. 5, 22).

[87] R. Frederix, D. Pagani and M. Zaro, *Large NLO corrections in* $t\bar{t}W^{\pm}$ *and* $t\bar{t}t\bar{t}$ *hadroproduction from supposedly subleading EW contributions*, JHEP **02** (2018) 031, arXiv: `1711.02116 [hep-ph]` (cit. on pp. 5, 22).

[88] 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. 6).

[89] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: `hep-ph/0409146` (cit. on p. 6).

[90] 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 p. 6).

[91] 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 p. 6).

[92] 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, [Erratum: JHEP **02** (2010) 011], arXiv: `0907.4076 [hep-ph]` (cit. on p. 6).

[93] 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. 6).

[94] T. Sjöstrand et al., *High-energy-physics event generation with PYTHIA 6.1*,
Comput. Phys. Commun. **135** (2001) 238, arXiv: `hep-ph/0010017` (cit. on p. 7).

[95] P. Golonka and Z. Was,
PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays,
Eur. Phys. J. C **45** (2006) 97, arXiv: `hep-ph/0506026` (cit. on p. 7).

[96] ATLAS Collaboration, *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*,
ATL-PHYS-PUB-2019-015, 2019, URL: `https://cds.cern.ch/record/2670380`
(cit. on p. 8).

[97] 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]` (cit. on pp. 8, 21).

[98] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, Eur. Phys. J. C **79** (2019) 205,
arXiv: `1810.05087 [hep-ex]` (cit. on p. 8).

[99] 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]` (cit. on pp. 8, 21).

[100] 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. 8).

[101] 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]`
(cit. on p. 8).

[102] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13\,TeV$*,
Eur. Phys. J. C **79** (2019) 639, arXiv: `1902.04655 [hep-ex]` (cit. on pp. 9, 21).

[103] 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]` (cit. on p. 9).

[104] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13\,TeV$ with the ATLAS detector*, Eur. Phys. J. C **81** (2020) 689,
arXiv: `2007.02645 [hep-ex]` (cit. on pp. 9, 10, 21).

[105] 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. 10).

[106] 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. 10).

[107] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: `https://cds.cern.ch/record/1700870` (cit. on p. 10).

[108] 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. 10).

[109] 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. 10, 22).

[110] 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. 10).

[111] ATLAS Collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, ATL-PHYS-PUB-2017-003, 2017, URL: `https://cds.cern.ch/record/2255226` (cit. on p. 10).

[112] 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. 10).

[113] 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 pp. 10, 22).

[114] 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]` (cit. on p. 10).

[115] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv: `2004.13447 [hep-ex]` (cit. on p. 11).

[116] 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]` (cit. on p. 11).

[117] F. Chollet et al., *Keras*, `https://keras.io`, 2015 (cit. on p. 11).

[118] *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, Software available from tensorflow.org, 2015, URL: `https://www.tensorflow.org/` (cit. on p. 11).

[119] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 2017, arXiv: `1412.6980 [cs.LG]` (cit. on p. 11).

[120] *Autonomio Talos [Computer software]*, 2020, URL: `https://autonomio.github.io/talos` (cit. on p. 11).

[121] G. J. Székely, M. L. Rizzo and N. K. Bakirov, *Measuring and testing dependence by correlation of distances*, Ann. Statist. **35** (2007) 2769 (cit. on p. 14).

[122] G. Kasieczka and D. Shih, *Robust Jet Classifiers through Distance Correlation*, Phys. Rev. Lett. **125** (2020) 122001, arXiv: `2001.05310 [hep-ph]` (cit. on p. 14).

[123] 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]` (cit. on p. 17).

[124] 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]` (cit. on p. 17).

[125] ATLAS Collaboration, *Measurement of Higgs boson decay into b-quarks in associated production with a top-quark pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2021), arXiv: `2111.06712 [hep-ex]` (cit. on p. 20).

[126] 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. 21).

[127] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 21).

[128] 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. 21).

[129] 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]` (cit. on p. 22).

[130] ATLAS Collaboration, *Simulation-based extrapolation of b-tagging calibrations towards high transverse momenta in the ATLAS experiment*, ATL-PHYS-PUB-2021-003, 2021, URL: `https://cds.cern.ch/record/2753444` (cit. on p. 22).

[131] 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]` (cit. on p. 22).

[132] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: `1510.03865 [hep-ph]` (cit. on p. 22).

[133] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN-OPEN-2012-016, 2012, URL: `https://cds.cern.ch/record/1456844` (cit. on p. 24).

[134] J. S. Conway, *Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*, (2011), URL: `https://arxiv.org/abs/1103.0354` (cit. on p. 24).

[135] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv: `physics/0306116 [physics.data-an]` (cit. on p. 24).

[136] 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, [Erratum: Eur. Phys. J. C **73** (2013) 2501], arXiv: `1007.1727 [physics.data-an]` (cit. on p. 24).

[137] 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. 24).

[138] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693 (cit. on p. 24).