EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Submitted to: JHEP

CERN-EP-2022-109
22nd September 2022

Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector

The ATLAS Collaboration

A search for Higgs boson pair production in events with two b-jets and two τ-leptons is presented, using a proton–proton collision data set with an integrated luminosity of 139 fb^{-1} collected at $\sqrt{s}$ = 13 TeV by the ATLAS experiment at the LHC. Higgs boson pairs produced non-resonantly or in the decay of a narrow scalar resonance in the mass range from 251 to 1600 GeV are targeted. Events in which at least one τ-lepton decays hadronically are considered, and multivariate discriminants are used to reject the backgrounds. No significant excess of events above the expected background is observed in the non-resonant search. The largest excess in the resonant search is observed at a resonance mass of 1 TeV, with a local (global) significance of 3.1σ (2.0σ). Observed (expected) 95% confidence-level upper limits are set on the non-resonant Higgs boson pair-production cross-section at 4.7 ($3.9^{+1.5}_{-1.1}$) times the Standard Model prediction, assuming Standard Model kinematics, and on the resonant Higgs boson pair-production cross-section at between 21 and 900 fb (12 and 840 fb), depending on the mass of the narrow scalar resonance.

1 Introduction

The discovery of a Higgs boson (H) with a mass of about 125 GeV [1, 2] has led to a comprehensive programme of measurements and searches by the ATLAS [3] and CMS [4] collaborations at the Large Hadron Collider (LHC) [5] using proton–proton (pp) collision data. To date, all of the measured properties of the Higgs boson are found to be consistent with their Standard Model (SM) predictions [6–13], and no unexpected particles or Higgs boson decay modes have been observed. The SM predicts non-resonant HH production, with approximately 90% of the total cross-section being due to the gluon–gluon fusion (ggF) process. The leading-order (LO) contributions to ggF HH production are the 'triangle diagram', which includes a Higgs boson self-coupling vertex, and the heavy-quark 'box diagram', which has two fermion–fermion–Higgs vertices, as shown in Figure 1(a) and Figure 1(b), respectively. These diagrams interfere destructively, leading to a small SM ggF non-resonant HH cross-section, which is predicted to be $31.1^{+2.1}_{-7.2}$ fb at next-to-next-to-leading (NNLO) order in α_s, including an approximation of finite top-quark-mass effects, for $m_H = 125$ GeV and $\sqrt{s} = 13$ TeV [14–21]. The vector-boson fusion (VBF) process provides a sub-leading source of HH production in the SM, and has a cross-section of 1.73 ± 0.04 fb, for $m_H = 125$ GeV and $\sqrt{s} = 13$ TeV [22–26]. Due to these small cross-sections, an observation of SM non-resonant HH production is not expected with the currently available LHC data set, although significant non-resonant and resonant enhancements to the HH cross-section are predicted in many beyond-the-SM (BSM) theories.

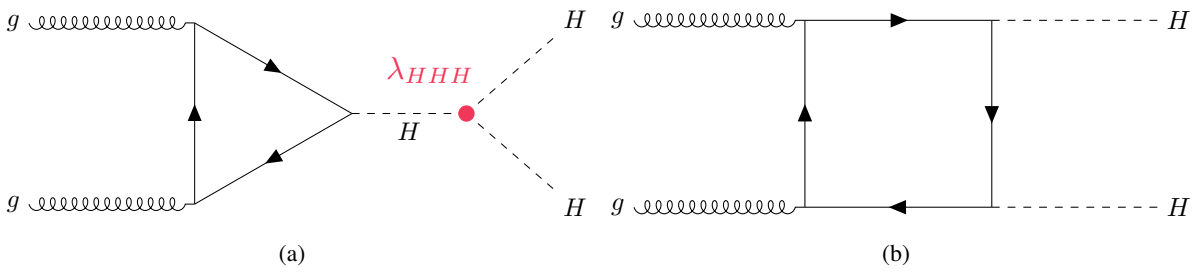


Figure 1: Leading-order Feynman diagrams for ggF non-resonant production of Higgs boson pairs: (a) the 'triangle diagram' and (b) the 'box diagram'. The Higgs boson trilinear self-coupling is denoted by λ_{HHH}.

Due to the diagram shown in Figure 1(a) and its interference with the diagram shown in Figure 1(b), non-resonant HH production is a sensitive probe of the Higgs boson trilinear self-coupling and the shape of the Higgs field potential, which have important implications for the stability of the electroweak vacuum [27, 28] and for baryogenesis [29] and inflation [30, 31]. Modifications to the non-resonant HH cross-section occur in BSM scenarios with new, light, coloured scalars [32], composite Higgs models [33], theoretical scenarios with couplings between pairs of top quarks and pairs of Higgs bosons [34], as well as models with a modified coupling of the Higgs boson to the top quark.

Previous searches for non-resonant HH production were performed by ATLAS and CMS in the $b\bar{b}\tau^+\tau^-$ [35–37], $b\bar{b}\gamma\gamma$ [38–40], $b\bar{b}b\bar{b}$ [41–45], $b\bar{b}\ell^+\nu\ell^-\nu$ [46, 47] decay channels, by ATLAS in the $b\bar{b}qq\ell\nu$ [48], $WW^*\gamma\gamma$ [49] and WW^*WW^* [50] decay channels, and by CMS in the $b\bar{b}\ell^+\ell^-\ell^+\ell^-$ [51] and $WW^*WW^*/WW^*\tau^+\tau^-/\tau^+\tau^-\tau^+\tau^-$ [52] decay channels. In the $b\bar{b}\tau^+\tau^-$ decay channel, ATLAS and CMS set observed (expected) upper limits on the non-resonant HH production cross-section at 12.7 (14.8) [35] and 30 (25) [36] times the SM expectation using 36.1 fb^{-1} and 35.9 fb^{-1} of 13 TeV pp collision data, respectively. Using these data sets, ATLAS and CMS each combined their results from several search channels, improving these observed (expected) limits to 6.9 (10) [53] and 22.2 (12.8) [54] times the SM

expectation, respectively. ATLAS and CMS recently published searches for HH production in the $b\bar{b}\gamma\gamma$ decay mode using their full Run-2 pp collision data sets, which set observed (expected) upper limits on the SM non-resonant HH cross-section at 4.2 (5.7) and 5.2 (7.7) times the SM expectation, respectively [39, 40]. CMS also recently performed a search for HH production in the $b\bar{b}\tau^+\tau^-$ [37] decay mode using 138 fb^{-1} of 13 TeV pp collision data, which set observed (expected) upper limits on the SM non-resonant HH cross-section of 3.3 (5.2) times the SM expectation. ATLAS and CMS have performed combinations of HH searches using their full Run-2 datasets, which set observed (expected) upper limits on the SM non-resonant HH cross-section at 2.4 (2.9) and 3.4 (2.5) times the SM expectation, respectively [13, 55].

Various BSM scenarios predict heavy resonances that can decay into pairs of Higgs bosons. These BSM resonances include heavy Higgs bosons from extended Higgs sectors such as those in two-Higgs-doublet models (2HDMs) [56], the minimal supersymmetric extension of the SM [57, 58], twin Higgs models [59], and composite Higgs models [33, 60]. Heavy resonances that decay into pairs of Higgs bosons also include spin-0 radions and spin-2 gravitons from the Randall–Sundrum model [61–63], and stoponium states in supersymmetric models [64]. Searches for resonant HH production have been performed in many final states by ATLAS and CMS, and no significant excesses have been observed [35, 36, 38, 39, 41, 42, 47–50, 52–54, 65–70].

This paper describes a search for non-resonant and resonant HH production in the final state with two τ-leptons and two jets containing b-hadrons (b-jets). The sizeable fraction of all possible SM decays that result in this final state, $\mathcal{B}(HH \to b\bar{b}\tau^+\tau^-) = 7.3\%$ [71, 72], and relatively low backgrounds make this one of the most sensitive HH search signatures. In the search for non-resonant HH production, the signal kinematics are assumed to follow the SM prediction, and the search is optimised for maximum sensitivity to the cross-section rather than the Higgs boson self-coupling. Additionally, only the ggF and VBF non-resonant HH production modes are considered, because other production modes are not expected to contribute significant additional sensitivity in this search. A narrow CP-even scalar particle (X) with a mass between 251 and 1600 GeV is used as the benchmark model for the resonant signal. Decay modes in which both τ-leptons decay hadronically (τ_{had}), or in which one decays hadronically and the other leptonically (τ_{lep}), are considered; these are referred to as $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$, respectively. The presence of τ_{had} are determined by detector signatures compatible with the expected visible decay products ($\tau_{\text{had-vis}}$). Events are categorised according to the type of trigger that accepted the event, and are required to contain two oppositely charged $\tau_{\text{had-vis}}$ and two b-tagged jets in the $\tau_{\text{had}}\tau_{\text{had}}$ final state, or an electron or muon and an oppositely charged $\tau_{\text{had-vis}}$ and two b-tagged jets in the $\tau_{\text{lep}}\tau_{\text{had}}$ final state. Signal events also contain neutrinos from the decay of τ-leptons and b-hadrons, which manifest themselves as missing momentum transverse to the beamline ($\mathbf{p}_{\text{T}}^{\text{miss}}$). Backgrounds in this search include the production of top-quark pairs ($t\bar{t}$), single top quarks, W and Z bosons in association with jets, dibosons (WW, WZ, ZZ), single Higgs bosons, and multi-jet events. In some background events, quark- or gluon-initiated jets are misidentified as $\tau_{\text{had-vis}}$. The selected events are tested for the presence of the signal by performing profile-likelihood fits to multivariate discriminant distributions. The search is performed using a data set obtained from pp collisions delivered by the LHC during Run 2, between 2015 and 2018, at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The data were collected by the ATLAS detector [3] and correspond to an integrated luminosity of 139 fb^{-1}. Compared to the previous ATLAS search [35], in addition to the greater integrated luminosity, this search profits from improved $\tau_{\text{had-vis}}$ and b-jet reconstruction and identification algorithms, more sophisticated multivariate techniques used to target the resonant signal hypotheses, and new background estimation techniques. In particular, the combined reconstruction and identification efficiencies for $\tau_{\text{had-vis}}$ and b-jets increased by around 25%–38% and 10%, respectively.

The rest of this paper is organised as follows. A brief description of the ATLAS detector is given in Section 2,

and the data and simulation samples used are outlined in Section 3. Overviews of the reconstruction and selection of physics objects are presented in Sections 4 and 5, respectively. The background modelling strategy is described in Section 6. The systematic uncertainties relevant to this search and the statistical interpretation are described in Sections 7 and 8, respectively. The results of the search are given in Section 9. Finally, Section 10 presents the conclusions.

2 ATLAS detector

ATLAS [3] is a general-purpose particle detector covering nearly the entire solid angle[1] around the collision point. It is composed of an inner tracking detector system surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroidal magnets.

The inner detector, located within a 2 T axial magnetic field generated by the superconducting solenoid, is used to measure the trajectories and momenta of charged particles. The inner layers, consisting of high-granularity silicon pixel detectors, instrument a pseudorapidity range $|\eta| < 2.5$. An innermost silicon pixel layer, the insertable B-layer [73, 74], was added to the detector between Run 1 and Run 2. The insertable B-layer improves the ability to identify displaced vertices, which significantly improves the b-jet tagging performance [75]. Silicon strip detectors, which cover the range $|\eta| < 2.5$, surround the pixel detectors. Outside the strip detectors and covering $|\eta| < 2.0$, there are straw-tube tracking detectors, which also provide measurements of transition radiation that are used in electron identification.

The calorimeter system covers the range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel ($|\eta| < 1.475$) and endcap ($1.375 < |\eta| < 3.2$) highly segmented lead/liquid-argon (LAr) electromagnetic calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by a steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters extend the coverage to $|\eta| = 3.2$. The region of $3.2 < |\eta| < 4.9$ is instrumented with copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements, respectively.

The outermost part of the detector is the muon spectrometer, which measures the curved trajectories of muons in the field of three large air-core toroidal magnets. High-precision tracking is performed within the range $|\eta| < 2.7$, and there are chambers for fast triggering within the range $|\eta| < 2.4$.

The ATLAS detector has a two-level trigger system [76] to select events of interest. The first-level (L1) trigger is implemented in custom electronics and, using a subset of the information from the detector, it accepts events from the 40 MHz LHC proton bunch crossings at a rate of about 100 kHz. This is followed by a software-based high-level trigger (HLT) that reduces the accepted event rate to approximately 1 kHz.

An extensive software suite [77] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

[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 $\eta = -\ln\tan(\theta/2)$. The distance in (η, ϕ) coordinates, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, is also used to define cone sizes. Transverse momentum and energy are defined as $p_{\mathrm{T}} = p\sin\theta$ and $E_{\mathrm{T}} = E\sin\theta$, respectively.

3 Data and simulation samples

3.1 Data samples

The data used in this search were collected at a centre-of-mass energy of 13 TeV between 2015 and 2018, using triggers to select events with at least one lepton (where a lepton is defined as an electron or a muon), at least one $\tau_{\text{had-vis}}$, at least one lepton and one $\tau_{\text{had-vis}}$, or at least two $\tau_{\text{had-vis}}$. Details about these triggers are discussed in Section 5.1. Events are selected for analysis only if they are of good quality and if all the relevant detector components are known to be in good operating condition [78]. The total integrated luminosity of the data, after meeting the good-quality criteria, is 139.0 ± 2.4 fb^{-1} [79, 80]. The recorded events contain an average of 34 simultaneous inelastic pp collisions per bunch-crossing.

3.2 Simulated event samples

Monte Carlo (MC) simulated events are used to model SM background production, SM non-resonant HH signal production, and BSM resonant HH signal production. The events were passed through the full ATLAS detector simulation [81] based on GEANT4 [82], with the exception of the BSM resonant HH signal events, which were passed through a fast simulation in which the response of the calorimeters is parameterised rather than fully simulated. The effects of multiple interactions in the same and neighbouring bunch crossings (pile-up) were modelled by overlaying each hard-scatter event with minimum-bias events, simulated using the soft quantum chromodynamics (QCD) processes of PYTHIA 8.186 [83] with a set of tuned parameters called the A3 tune [84] and the NNPDF2.3LO [85] parton distribution functions (PDFs). The EVTGEN program [86] was used to model the decays of bottom and charm hadrons in all samples of simulated events, except those generated using SHERPA [87]. The samples generated with SHERPA used the bottom- and charm-hadron decay model implemented within the generator. The simulated events were processed through the same reconstruction algorithms as the data. For all samples containing a SM Higgs boson, its mass was fixed to 125 GeV. The same mass value is used in the calculation of the Higgs boson decay branching fractions and in the calculation of the single-Higgs-boson and SM non-resonant HH production cross-sections. Unless otherwise specified, the order of the cross-section calculation refers to the expansion in the strong coupling constant (α_s). A summary of the event samples used for the simulation of the signal and background processes is shown in Table 1.

Simulated SM non-resonant HH signal production includes the contributions from the ggF and VBF processes. The simulated ggF events were generated with the POWHEG BOX v2 generator [88] at next-to-leading order (NLO) with finite top-quark mass, and using the PDF4LHC15_NLO_30_PDFAS (code 90400 in the LHAPDF database [89]) PDF set [90]. Parton showers and hadronisation were simulated using PYTHIA 8.244 [83] with the A14 tune [91, 92] and the NNPDF2.3LO PDF set. The cross-section for ggF non-resonant HH production is calculated at next-to-next-to-leading order (NNLO) using FTApprox [20]. The VBF non-resonant HH signal events were generated at LO using the MADGRAPH5_AMC@NLO 2.7.3 [93] generator with the NNPDF3.0NLO [94] PDF set. Parton showering and hadronisation were performed using PYTHIA 8.244 with the A14 tune and the NNPDF2.3LO PDF set. The cross-section for VBF non-resonant HH production is calculated at next-to-next-to-next-to-leading order (N3LO) in QCD in the limit of no partonic exchange between the two protons [26]. The calculated cross-section values for ggF and VBF non-resonant HH production at $\sqrt{s} = 13$ TeV and $m_H = 125$ GeV are given in Section 1. Other non-resonant HH production modes are not considered because their contributions to the analysis sensitivity are expected to be negligible.

Table 1: The generators used to simulate the signal and background processes. If not otherwise specified, the order of the cross-section calculation refers to the expansion in the strong coupling constant (α_s). The acronyms ME, PS and UE are used for matrix element, parton shower and underlying event, respectively. Details of the simulation of the signal and background samples are described in the text. (†) The NNLO(QCD)+NLO(EW) cross-section calculation for the $pp \to ZH$ process already includes the $gg \to ZH$ contribution. The $qq \to ZH$ process is normalised to the NNLO(QCD)+NLO(EW) cross-section for the $pp \to ZH$ process, after subtracting the $gg \to ZH$ contribution.

Process	ME generator	ME QCD order	ME PDF	PS and hadronisation	UE model tune	Cross-section order
Signal						
non-resonant $gg \to HH$ (ggF)	PowHeg Box v2	NLO	PDF4LHC15nlo	Pythia 8.244	A14	NNLO FTApprox
non-resonant $qq \to qqHH$ (VBF)	MadGraph5_aMC@NLO 2.7.3	LO	NNPDF3.0nlo	Pythia 8.244	A14	N3LO(QCD)
resonant $gg \to X \to HH$	MadGraph5_aMC@NLO 2.6.1	LO	NNPDF2.3lo	Herwig 7.1.3	H7.1-Default	–
Top-quark						
$t\bar{t}$	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.230	A14	NNLO+NNLL
t-channel	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.230	A14	NLO
s-channel	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.230	A14	NLO
Wt	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.230	A14	NLO
$t\bar{t}Z$	Sherpa 2.2.1	NLO	NNPDF3.0nnlo	Sherpa 2.2.1	Default	NLO
$t\bar{t}W$	Sherpa 2.2.8	NLO	NNPDF3.0nnlo	Sherpa 2.2.8	Default	NLO
Vector boson + jets						
W/Z+jets	Sherpa 2.2.1	NLO ($\leq$ 2 jets) LO (3,4 jets)	NNPDF3.0nnlo	Sherpa 2.2.1	Default	NNLO
Diboson						
WW, WZ, ZZ	Sherpa 2.2.1	NLO ($\leq$ 1 jet) LO (2,3 jets)	NNPDF3.0nnlo	Sherpa 2.2.1	Default	NLO
Single Higgs boson						
ggF	PowHeg Box v2	NNLO	NNPDF3.0nlo	Pythia 8.212	AZNLO	N3LO(QCD)+NLO(EW)
VBF	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.212	AZNLO	NNLO(QCD)+NLO(EW)
$qq \to WH$	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.212	AZNLO	NNLO(QCD)+NLO(EW)
$qq \to ZH$	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.212	AZNLO	NNLO(QCD)+NLO(EW)$^{(\dagger)}$
$gg \to ZH$	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.212	AZNLO	NLO+NLL
$t\bar{t}H$	PowHeg Box v2	NLO	NNPDF3.0nlo	Pythia 8.230	A14	NLO

The BSM resonant HH signal from the ggF production of a heavy spin-0 resonance and its decay into a pair of SM Higgs bosons, $X \to HH$, was simulated with the MadGraph5_aMC@NLO 2.6.1 generator using the NNPDF2.3lo PDF set at LO accuracy in QCD. The simulated events were interfaced to Herwig 7.1.3 [95, 96] to model the parton shower, hadronisation and underlying event, using the H7.1-Default tune [97] and the NNPDF2.3lo PDF set. The resonant HH signal was simulated for 19 values of the resonance mass, m_X, between 251 GeV and 1.6 TeV. The width of the heavy scalar X was fixed to 10 MeV.

The production of $t\bar{t}$ events, and of single-top-quark events in the Wt-, s- and t-channels, was simulated using the PowHeg Box v2 generator together with the NNPDF3.0nlo PDF set. The simulated events were interfaced to Pythia 8.230 for parton showering and hadronisation using the A14 tune together with the NNPDF2.3lo PDF set. The top-quark spin correlations were preserved for all these simulated top-quark processes. The top-quark mass was set to 172.5 GeV. The $t\bar{t}$ production cross-section is calculated at next-to-next-to-leading-order and next-to-next-to-leading-logarithm (NNLO+NNLL) accuracy [98]. The cross-sections for the three single-top-quark production channels are calculated at NLO [99–101]. The $t\bar{t}$–Wt interference was handled using the diagram removal scheme.

Events containing a W or Z boson produced in association with jets, diboson (WW, WZ and ZZ) production processes, and the $t\bar{t}Z$ production process were simulated with the Sherpa 2.2.1 generator [87], whereas the $t\bar{t}W$ production process was simulated with the Sherpa 2.2.8 generator. These samples used the NNPDF3.0nnlo [94] PDF set with dedicated parton shower tuning developed by the Sherpa authors. For the simulation of W/Z+jets events, the matrix elements were calculated for up to two partons at NLO and up to four partons at LO using the OpenLoops [102] and Comix [103] matrix-element generators. The

expected number of W/Z+jets events is normalised to the NNLO cross-sections [104]. Diboson production was simulated for up to one additional parton at NLO and up to three additional partons at LO using the OpenLoops and Comix programs. The NLO cross-sections from Sherpa are used to normalise the diboson and the $t\bar{t}W/Z$ events.

SM single Higgs boson production is considered as part of the background in this search, and its production modes were simulated using the Powheg Box v2 generator and the NNPDF3.0nlo PDF set. Single Higgs boson production via ggF was simulated at NNLO accuracy in QCD using the Powheg NNLOPS program [105, 106], whereas VBF single Higgs boson production was simulated at NLO accuracy in QCD [107]. Events from both of these production modes were interfaced to Pythia 8.212 for parton showering and hadronisation using the AZNLO tune [108] together with the CTEQ6L1 PDF set [109]. The cross-section for ggF production of single Higgs bosons is based on a computation with N3LO accuracy in QCD, and NLO accuracy in the electroweak (EW) expansion [71, 110–113], whereas the cross-section for VBF production of single Higgs bosons is taken from the NNLO(QCD)+NLO(EW) calculation [71, 114–116]. The $qq \to WH$, $qq \to ZH$ and $gg \to ZH$ simulated events were interfaced to Pythia 8.212 for parton showering and hadronisation using the AZNLO tune together with the CTEQ6L1 PDF set. The cross-sections are taken from the NNLO(QCD)+NLO(EW) calculations for $qq \to WH$ and $qq \to ZH$ [117–123], and from calculations at next-to-leading-order and next-to-leading-logarithm (NLO+NLL) accuracy in QCD for $gg \to ZH$ [124–128]. For Higgs boson production in association with a pair of top quarks ($t\bar{t}H$), the simulated events were interfaced to Pythia 8.230 for parton showering and hadronisation using the A14 tune and the NNPDF2.3lo PDF set. The cross-section for $t\bar{t}H$ production is taken from NLO calculations [71]. SM single Higgs boson production plays a more important role as a background in the non-resonant HH search than in the resonant HH search, due to more similar production kinematics.

4 Object reconstruction

Electrons, muons, $\tau_{\text{had-vis}}$, jets from the hadronisation of quarks and gluons, including b-tagged jets, and $\mathbf{p}_{\text{T}}^{\text{miss}}$ are used in this search. An anti-$\tau_{\text{had-vis}}$ object, defined as a $\tau_{\text{had-vis}}$ with modified identification requirements, is also used to estimate the backgrounds from hadronic jets misidentified as $\tau_{\text{had-vis}}$. Tracks are used in the reconstruction, identification, isolation and vertex compatibility requirements and calibration of many of the physics objects described below, and in vertex reconstruction; they are reconstructed from hits in the inner tracking detectors, and are required to have $p_{\text{T}} > 500$ MeV [129, 130]. Events are required to have at least one collision vertex reconstructed from two or more associated tracks. If multiple vertices are found, the one with the largest $\sum p_{\text{T}}^2$ of the associated tracks is selected as the primary vertex. Finally, an overlap removal procedure is applied to ensure that no detector signature is identified as multiple reconstructed objects.

Electron candidates are reconstructed by matching tracks reconstructed in the inner detector to topological energy clusters in the electromagnetic calorimeter, with a reconstruction efficiency of around 98% [131]. Electron candidates are required to have $p_{\text{T}} > 7$ GeV and $|\eta| < 2.47$, and to be outside the transition region between the calorimeter's barrel and endcaps, $1.37 < |\eta| < 1.52$. They must pass track-quality requirements, followed by a loose likelihood-based selection that requires the shower profile to be compatible with that of an electromagnetic shower. These requirements have an efficiency of around 93%. Isolation requirements are applied; these are based on the presence of tracks in a cone of p_{T}^e-dependent size ΔR around the electron

and of calorimetric energy deposits in a fixed-size cone. Lastly, the electron energy scale is calibrated in data, and the energy resolution is calibrated in simulation, using $Z \to ee$ events [131].

Muon candidates are reconstructed from tracks in the muon spectrometer, matched to tracks in the inner detector where available [132]. In the absence of full tracks in the muon spectrometer, muons with $|\eta| < 0.1$ can be reconstructed from track segments in the muon spectrometer, or energy deposits compatible with that of a minimum-ionising particle in the calorimeters. If an inner-detector track is present, it must match the direction and momentum of the muon spectrometer track for it to be included. The muon momentum is defined by using information from both the muon spectrometer and the inner detector where available. Only muon candidates with $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.7$, and passing loose quality requirements based on the number of hits used to reconstruct the tracks, with an efficiency of around 99%, are considered for further analysis. Lastly, isolation requirements with an efficiency of around 95% that are based on the presence of particle-flow objects [133] in a cone of p_{T}^{μ}-dependent size ΔR around the muon are applied, except for muons used in the b-tagged-jet energy correction described below.

Jets are reconstructed using the anti-k_t jet clustering algorithm [134, 135] with a radius parameter of 0.4, applied to noise-suppressed positive-energy topological energy clusters [136, 137] and charged-particle tracks, processed using a particle-flow algorithm [133]. Jet energies are corrected for contributions from pile-up, calibrated using energy- and η-dependent correction factors determined from comparisons between particle-level objects and reconstructed physics objects in simulated events, and then corrections are applied to account for effects due to the initiating-parton type and hadron composition [138]. In data, a residual in situ correction is applied in order to correct for differences relative to simulation. Jets are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$, or $p_{\mathrm{T}} > 30$ GeV and $2.5 < |\eta| < 4.5$. To reject jets from pile-up, jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.5$ ($20 < p_{\mathrm{T}} < 120$ GeV and $2.5 < |\eta| < 4.5$) are required to pass a 'jet vertex tagger' [139] ('forward jet vertex tagger' [140]) requirement to determine if they originate from the primary vertex [141]. Lastly, quality criteria [142] are applied to jets reconstructed from topological clusters [136, 137] in the calorimeter in order to identify and remove events containing jets from non-collision backgrounds and calorimeter noise.

A multivariate classification algorithm based on a deep neural network, the DL1r tagger [143, 144], is used to distinguish b-jets from the background of light-flavour- and charm-quark-initiated jets using information about the jet kinematics, the impact parameters of tracks associated with the jet, and the presence of displaced vertices. The inputs to the DL1r network include variables based on a recurrent neural network (RNNIP) [145], which can exploit the spatial and kinematic correlations between tracks that are initiated from the same b-hadrons. This analysis uses a b-jet working point that targets an efficiency of 77%. A simulation-based p_{T}-dependent correction is applied to the b-tagged jet momentum [146]. Finally, for b-tagged jets containing a muon, the difference between the reconstructed muon momentum and the momentum corresponding to the energy of its associated cluster in the calorimeter is added to the b-tagged jet momentum [146]. This corrects for the fact that the muon typically deposits only a small fraction of its energy in the calorimeters. These two corrections improve the resolution of the invariant mass of the pair of b-tagged jets by approximately 16% for the non-resonant HH signal.

The reconstructed $\tau_{\text{had-vis}}$ candidates [147] are seeded by jets. The $\tau_{\text{had-vis}}$ energy is calibrated using multivariate methods with information from tracks and calorimeter energy clusters [148]. The $\tau_{\text{had-vis}}$ are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$, excluding $1.37 < |\eta| < 1.52$. Boosted decision trees are used to determine if tracks near a cluster originate from a τ_{had}, and one or three tracks with a total charge of ± 1 are required. The true-$\tau_{\text{had-vis}}$ are discriminated from quark- and gluon-initiated-jet backgrounds by using recurrent neural networks trained to target signatures with either one or three associated tracks [149], and a loose requirement with an efficiency of around 85% (75%) for one-track (three-track) $\tau_{\text{had-vis}}$ candidates is

applied. A separate boosted decision tree is then used to reject $\tau_{\text{had-vis}}$ candidates with one associated track that originate from electrons, with an efficiency of about 95% for τ_{had} [147].

Anti-$\tau_{\text{had-vis}}$ objects are defined in order to estimate the background from jets misidentified as $\tau_{\text{had-vis}}$, as described in Section 6. They are only used for the estimation of these backgrounds, and only one anti-$\tau_{\text{had-vis}}$ per event is considered. Anti-$\tau_{\text{had-vis}}$ objects are reconstructed, and their energy is calibrated, in the same way as for $\tau_{\text{had-vis}}$ candidates, and they must satisfy the nominal $\tau_{\text{had-vis}}$ kinematic and track selection criteria. They are required to pass a looser recurrent neural-network requirement, corresponding to an efficiency of approximately 99% for true-$\tau_{\text{had-vis}}$, but to fail the nominal recurrent neural-network requirement applied to the $\tau_{\text{had-vis}}$ candidates.

Neutrinos are produced in the decay of τ-leptons and in the semileptonic decay of b-hadrons, and contribute to the total $\mathbf{p}_{\text{T}}^{\text{miss}}$ in the event. The $\mathbf{p}_{\text{T}}^{\text{miss}}$ is defined as the negative vector sum of the transverse momenta of all reconstructed and calibrated leptons, jets, and $\tau_{\text{had-vis}}$, and of all tracks matched to the primary vertex but not to other reconstructed objects in the event [150]. The magnitude of $\mathbf{p}_{\text{T}}^{\text{miss}}$ is denoted by $E_{\text{T}}^{\text{miss}}$.

A sequential overlap removal procedure is applied to resolve ambiguities in which multiple electron, muon or jet candidates would otherwise be reconstructed from the same detector signature. This procedure uses a definition of angular distance, $\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$, that is based on the rapidity y of the objects. The following steps are applied in order:

1. if any electrons share a track, all but the highest-p_{T} electron are removed;
2. any $\tau_{\text{had-vis}}$ within $\Delta R_y = 0.2$ of an electron or a muon is removed, except if the $\tau_{\text{had-vis}}$ has $p_{\text{T}} > 50$ GeV, in which case the muon must also be associated with a track in the muon spectrometer for the $\tau_{\text{had-vis}}$ to be removed;
3. if an electron and a muon share a track, the electron is removed if the muon is associated with a signature in the muon spectrometer, otherwise the muon is removed;
4. any jet within $\Delta R_y = 0.2$ of an electron is removed;
5. any electron within $\Delta R_y = 0.4$ of a jet is removed;
6. any jet within $\Delta R_y = 0.2$ of a muon, or which would have the inner-detector track of a muon matched to it using ghost association [151], is removed if it has fewer than three associated tracks;
7. any muon within $\Delta R_y = 0.4$ of a jet is removed;
8. any jet within $\Delta R_y = 0.2$ of a $\tau_{\text{had-vis}}$ is removed;
9. if an anti-$\tau_{\text{had-vis}}$ is within $\Delta R_y = 0.2$ of a jet, it is removed if the jet is b-tagged, otherwise the jet is removed.

5 Event selections

Events are selected in three separate signal categories. The $\tau_{\text{had}}\tau_{\text{had}}$ category targets events with two oppositely charged $\tau_{\text{had-vis}}$ and two b-jets, whereas two $\tau_{\text{lep}}\tau_{\text{had}}$ categories target events with an electron or muon, an oppositely charged $\tau_{\text{had-vis}}$ and two b-jets. A control region (Z + HF CR) is also defined in order to determine the normalisation of the background in which a Z boson is produced in association with one or more jets initiated by heavy-flavour quarks (Z + HF background). Events in the signal regions

(SRs) are analysed for the presence of signal using multivariate discriminants, and discrimination between the Z + HF and $t\bar{t}$ backgrounds is achieved using the dilepton invariant mass ($m_{\ell\ell}$) in the Z + HF CR, as described in Section 5.3.

5.1 Signal regions

Events in the $\tau_{\text{had}}\tau_{\text{had}}$ category were recorded using a combination of single-$\tau_{\text{had-vis}}$ triggers (STTs) and di-$\tau_{\text{had-vis}}$ triggers (DTTs). The STTs accept events with at least a single $\tau_{\text{had-vis}}$ at the HLT with a minimum p_{T} between 80 GeV and 160 GeV, depending on the data-taking period. The DTTs select events with at least a pair of $\tau_{\text{had-vis}}$ reconstructed at the HLT, with minimum p_{T} of 35 GeV (25 GeV) for the (sub-)leading $\tau_{\text{had-vis}}$, where the (sub-)leading $\tau_{\text{had-vis}}$ is defined as the $\tau_{\text{had-vis}}$ with the (second-)highest p_{T} in the event. From the 2016 data-taking period onward, additional requirements were applied in the L1 trigger to reduce the DTT trigger rates. During 2016 data-taking, an additional jet with $E_{\text{T}} > 25$ GeV was required. For 2017 (2018) data-taking, if two offline[2] jets with $p_{\text{T}} > 45$ GeV are found, then a trigger is used that requires two additional jets with $E_{\text{T}} > 12$ GeV (and $|\eta| < 2.3$) at L1, otherwise another trigger is used that requires one additional jet with $E_{\text{T}} > 25$ GeV and the $\tau_{\text{had-vis}}$ to be reconstructed within $\Delta R = 2.8$ of each other[3]. In order to select events near the trigger efficiency plateaus where the efficiencies are well modelled, the offline $\tau_{\text{had-vis}}$ are required to be within $\Delta R = 0.2$ of the corresponding HLT $\tau_{\text{had-vis}}$ objects and a minimum offline p_{T} requirement is applied to $\tau_{\text{had-vis}}$ and to the jets. The offline p_{T} thresholds for the $\tau_{\text{had-vis}}$ range between 100 GeV and 180 GeV for the STTs, and are 40 GeV (30 GeV) for the (sub-)leading $\tau_{\text{had-vis}}$ for the DTTs. Additional offline requirements for the DTTs are either that two additional jets with $p_{\text{T}} > 45$ GeV are present in the event, or that a jet with $p_{\text{T}} > 80$ GeV is present in the event and the $\tau_{\text{had-vis}}$ are reconstructed within $\Delta R = 2.5$ of each other. For events that pass both the STTs and DTTs, the offline requirements used for the STTs are applied. Events passing the $\tau_{\text{had}}\tau_{\text{had}}$ event selection are analysed together.

Events in the $\tau_{\text{lep}}\tau_{\text{had}}$ categories were recorded using a combination of single-lepton triggers (SLTs) and lepton-plus-$\tau_{\text{had-vis}}$ triggers (LTTs). The SLTs require an electron or muon to be reconstructed at the HLT with an E_{T} threshold that ranges from 24 GeV to 26 GeV for electrons and a p_{T} threshold that ranges from 20 GeV to 26 GeV for muons, depending on the data-taking period. The LTTs require that an electron with $p_{\text{T}} > 17$ GeV or a muon with $p_{\text{T}} > 14$ GeV in addition to a $\tau_{\text{had-vis}}$ with $p_{\text{T}} > 25$ GeV is reconstructed at the HLT. The LTTs used to collect lepton-plus-$\tau_{\text{had-vis}}$ events with $\tau_{\text{had-vis}}$ $p_{\text{T}} < 35$ GeV had additional requirements at the L1 trigger, requiring the presence of either an additional jet with $E_{\text{T}} > 25$ GeV or two additional jets with $E_{\text{T}} > 12$ GeV. The electron-plus-$\tau_{\text{had-vis}}$ triggers that require an additional jet with $E_{\text{T}} > 25$ GeV was only used to select an event if the electron-plus-$\tau_{\text{had-vis}}$ triggers that require two additional jets with $E_{\text{T}} > 12$ GeV did not select the event. For any trigger to select an event, based on the presence of a muon, the muon must have $|\eta| < 2.5$. In order to select events near the trigger efficiency plateaus where the trigger efficiencies are well modelled, the offline electrons, muons and $\tau_{\text{had-vis}}$ objects are required to be within $\Delta R = 0.07$, $\Delta R = 0.1$ and $\Delta R = 0.2$ of the corresponding objects at the HLT, respectively. Minimum p_{T} requirements are applied to the offline objects, and these are 1 GeV above the thresholds for electrons and muons at the HLT, 5 GeV above the thresholds for $\tau_{\text{had-vis}}$ at the HLT, and 80 GeV (45 GeV) for jets with an L1-trigger E_{T} threshold of 25 GeV (12 GeV). Events which pass the offline SLT lepton p_{T} requirements are not considered for the LTT. This ensures that there is no overlap between the SLT and LTT categories. These two categories are analysed separately.

[2] In this paper, offline objects are objects which are reconstructed after the data were collected, as opposed to trigger-level objects.

[3] This dependence of the online selection on the offline requirements ensures that both triggers will not be used for the same events, which avoids the need to study their combined efficiency.

In the $\tau_{\text{had}}\tau_{\text{had}}$ category, in addition to the trigger selection described above, events are required to contain exactly two $\tau_{\text{had-vis}}$ with opposite charges and exactly two b-tagged jets. Events with any additional leptons are vetoed, and the (sub-)leading b-tagged jet is required to have $p_{\text{T}} > 45$ (20) GeV. The invariant mass of the τ-lepton pair ($m_{\tau\tau}^{\text{MMC}}$) is estimated from the four-momenta of the $\tau_{\text{had-vis}}$ and the $\mathbf{p}_{\text{T}}^{\text{miss}}$ using the Missing Mass Calculator (MMC) [152], which assumes that the $\mathbf{p}_{\text{T}}^{\text{miss}}$ is exclusively from the neutrinos produced in the τ-lepton decays. To reject background from low-mass Drell–Yan events, $m_{\tau\tau}^{\text{MMC}}$ is required to be above 60 GeV.

In the $\tau_{\text{lep}}\tau_{\text{had}}$ categories, exactly one 'loose' electron or 'loose' muon, an oppositely charged $\tau_{\text{had-vis}}$, and exactly two b-tagged jets are required. The selected electron (muon) must also pass a tight (medium) identification requirement with an efficiency of around 80% (97%). Events are required to have $m_{\tau\tau}^{\text{MMC}} > 60$ GeV, where $m_{\tau\tau}^{\text{MMC}}$ is calculated using the four-momenta of the electron or muon, the $\tau_{\text{had-vis}}$ and the $\mathbf{p}_{\text{T}}^{\text{miss}}$. The b-tagged jet pair invariant mass (m_{bb}) is required to be less than 150 GeV to reject $t\bar{t}$ background events, and to allow for the definition of a $t\bar{t}$-enriched region which is used in the estimation of $t\bar{t}$ backgrounds, as described in Section 6. A $\tau_{\text{had-vis}}$ with $p_{\text{T}} > 20$ GeV and $|\eta| < 2.3$ is required in the SLT category, and a $\tau_{\text{had-vis}}$ with $p_{\text{T}} > 30$ GeV, or higher if required by the trigger, and $|\eta| < 2.3$ is required in the LTT category. In both categories, the (sub-)leading b-tagged jet must have $p_{\text{T}} > 45$ (20) GeV, in addition to any trigger-dependent requirements.

The full event selection is summarised in Table 2. The acceptance times efficiency for the non-resonant ggF+VBF $HH \to b\bar{b}\tau^+\tau^-$ signal, evaluated with respect to the targeted τ-lepton pair decay mode, is 4.0%, 4.0% and 1.0%, in the $\tau_{\text{had}}\tau_{\text{had}}$ category, and the $\tau_{\text{lep}}\tau_{\text{had}}$ SLT and $\tau_{\text{lep}}\tau_{\text{had}}$ LTT categories, respectively. For the resonant HH signal, the acceptance times efficiency is shown in Figure 2 as a function of the resonance mass. The decrease in acceptance times efficiency for m_X greater than about 1000 GeV is due to the Lorentz boost of the Higgs bosons causing their decay products to become highly collimated more often.

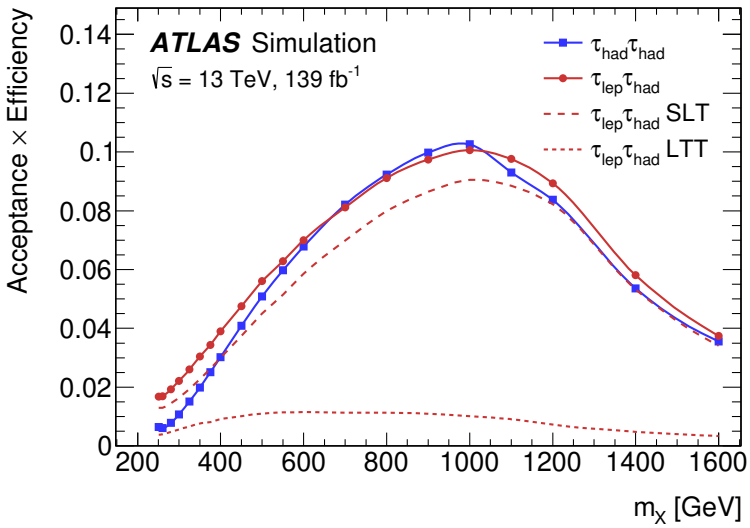


Figure 2: Acceptance times efficiency for the full analysis selections as a function of the resonance mass m_X in the $\tau_{\text{had}}\tau_{\text{had}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ SLT and $\tau_{\text{lep}}\tau_{\text{had}}$ LTT trigger categories, shown in solid line with square markers, dashed and dotted lines, respectively. The solid line with circle markers is the acceptance times efficiency curve for the combined $\tau_{\text{lep}}\tau_{\text{had}}$ category. The acceptance times efficiency for $X \to HH \to b\bar{b}\tau^+\tau^-$ decays is evaluated with respect to the targeted τ-lepton pair decay mode ($\tau_{\text{lep}}\tau_{\text{had}}$ or $\tau_{\text{had}}\tau_{\text{had}}$).

Table 2: Summary of the event selections, shown separately for events that are selected by different triggers. In cases where pairs of reconstructed objects of the same type are required, thresholds for the (sub-)leading p_{T} object are given outside (within) parentheses where different event selection thresholds are applied. When the selection depends on the year of data-taking, the possible values of the requirements are separated by commas, except for the jet selection in the LTT and DTT triggers, which use multiple selection criteria as described in Section 5.1. The trigger p_{T} thresholds shown are applied to the offline physics objects that are matched to the corresponding trigger objects.

$\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ category		$\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ categories	
STT	DTT	SLT	LTT
e/μ selection			
No loose e/μ		Exactly one loose e/μ	
		e (μ) must be tight (medium and have $\lvert\eta\rvert < 2.5$)	
		$p_{\mathrm{T}}^{e} > 25, 27$ GeV	18 GeV $< p_{\mathrm{T}}^{e} <$ SLT cut
		$p_{\mathrm{T}}^{\mu} > 21, 27$ GeV	15 GeV $< p_{\mathrm{T}}^{\mu} <$ SLT cut
$\tau_{\text{had-vis}}$ selection			
Two loose $\tau_{\text{had-vis}}$		One loose $\tau_{\text{had-vis}}$	
		$\lvert\eta\rvert < 2.3$	
$p_{\mathrm{T}} >$ 100, 140, 180 (25) GeV	$p_{\mathrm{T}} > 40$ (30) GeV		$p_{\mathrm{T}} > 30$ GeV
Jet selection			
≥ 2 jets with $\lvert\eta\rvert < 2.5$			
Leading jet $p_{\mathrm{T}} > 45$ GeV	Trigger dependent	Leading jet $p_{\mathrm{T}} > 45$ GeV	Trigger dependent
Event-level selection			
Trigger requirements passed			
Collision vertex reconstructed			
$m_{\tau\tau}^{\mathrm{MMC}} > 60$ GeV			
Opposite-sign electric charges of $e/\mu/\tau_{\text{had-vis}}$ and $\tau_{\text{had-vis}}$			
Exactly two b-tagged jets			
		$m_{bb} < 150$ GeV	

5.2 Multivariate signal extraction

Multivariate discriminants (MVAs) evaluated for events passing the above selections are used to extract possible signals. Parameterised neural networks (PNNs) [153] are used in the search for resonant HH production, and a boosted decision tree (BDT) and neural networks (NNs) are used in the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ category and $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ categories of the non-resonant HH search, respectively. The (P)NNs are trained using Keras [154] with the Tensorflow [155] backend, and the BDT is trained using TMVA [156]. During training, the sum of all backgrounds normalised to their respective cross-sections is used. The backgrounds containing one or more $\tau_{\text{had-vis}}$ from a misidentified quark- or gluon-initiated jet are modelled using simulation, except for the multi-jet background in the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ category, where the data-driven estimate is used. The PNNs are parameterised in the mass of the heavy resonance, providing near-optimal sensitivity and continuity over the range of signal masses considered.

The same choice of MVA input variables is used for the resonant and non-resonant production modes, although different input variables are used in the different analysis categories. These variables are listed in Table 3, and are defined as follows:

- m_{HH} is the invariant mass of the HH system as reconstructed from the τ-lepton pair (calculated using the MMC) and the b-tagged jet pair. In the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ categories the four-momenta of the $b\bar{b}$ and

$\tau^+\tau^-$ systems are multiplied by correction factors (125 GeV/m_{bb} and 125 GeV/$m_{\tau\tau}^{\text{MMC}}$, respectively) in order to improve the m_{HH} mass resolution;

- $\Delta R(\tau,\tau)$ is evaluated between the two $\tau_{\text{had-vis}}$ (the electron or muon and the $\tau_{\text{had-vis}}$) in the $\tau_{\text{had}}\tau_{\text{had}}$ category ($\tau_{\text{lep}}\tau_{\text{had}}$ categories);
- $\Delta R(b,b)$ is evaluated between the b-tagged jets;
- $\Delta p_{\text{T}}(\ell,\tau)$ is the difference between the transverse momenta of the lepton and the $\tau_{\text{had-vis}}$;
- $m_{\text{T}}^{W} = \sqrt{2p_{\text{T}}^{\ell}E_{\text{T}}^{\text{miss}}(1-\cos\Delta\phi_{\ell,\mathbf{p}_{\text{T}}^{\text{miss}}})}$ is the transverse mass of the lepton and the $\mathbf{p}_{\text{T}}^{\text{miss}}$;
- the $\mathbf{p}_{\text{T}}^{\text{miss}}$ ϕ centrality specifies the angular position of the $\mathbf{p}_{\text{T}}^{\text{miss}}$ relative to the $\tau_{\text{had-vis}}$ in the transverse plane [157] and is defined as $(A+B)/\sqrt{A^2+B^2}$, where $A = \sin(\phi_{\mathbf{p}_{\text{T}}^{\text{miss}}} - \phi_{\tau_2})/\sin(\phi_{\tau_1}-\phi_{\tau_2})$, $B = \sin(\phi_{\tau_1} - \phi_{\mathbf{p}_{\text{T}}^{\text{miss}}})/\sin(\phi_{\tau_1}-\phi_{\tau_2})$, and τ_1 and τ_2 represent the two $\tau_{\text{had-vis}}$ (electron or muon and $\tau_{\text{had-vis}}$) in the case of the $\tau_{\text{had}}\tau_{\text{had}}$ category ($\tau_{\text{lep}}\tau_{\text{had}}$ categories);
- $\Delta\phi(\ell\tau, bb)$ is the azimuthal angle between the $\ell + \tau_{\text{had-vis}}$ system and the b-tagged jet pair;
- $\Delta\phi(\ell, \mathbf{p}_{\text{T}}^{\text{miss}})$ is the azimuthal angle between the lepton and the $\mathbf{p}_{\text{T}}^{\text{miss}}$;
- $\Delta\phi(\tau\tau, \mathbf{p}_{\text{T}}^{\text{miss}})$ is the azimuthal angle between the τ-lepton pair system (estimated using the MMC) and the $\mathbf{p}_{\text{T}}^{\text{miss}}$;
- S_{T} is the total transverse energy in the event, summed over all jets, $\tau_{\text{had-vis}}$ and leptons in the event and $E_{\text{T}}^{\text{miss}}$.

The m_{HH}, $m_{\tau\tau}^{\text{MMC}}$ and m_{bb} distributions are shown in Figure 3. For all categories of the non-resonant search, $m_{\tau\tau}^{\text{MMC}}$ and m_{bb} are among the three most important MVA input variables. For the resonant search, five values of m_X were tested in all categories, and m_{HH} was found to be the most important MVA input variable in all cases except at lower values of m_X in the $\tau_{\text{had}}\tau_{\text{had}}$ category. The $m_{\tau\tau}^{\text{MMC}}$ and m_{bb} variables separate resonant $H \to \tau\tau$ and $H \to bb$ signals (respectively) from backgrounds that do not contain these processes, and the m_{HH} variable targets resonant $X \to HH$ decays in the resonant search.

The (P)NNs use rectified linear unit and sigmoidal activation functions for the hidden and output layers, respectively, binary cross entropy as the loss function, and stochastic (mini-batch) gradient descent as the optimiser [158]. The PNN used in the $\tau_{\text{had}}\tau_{\text{had}}$ category has 3 hidden layers of 128 nodes, followed by 1 hidden layer of 16 nodes. The (P)NN used in the $\tau_{\text{lep}}\tau_{\text{had}}$ SLT category has 2 hidden layers of 512 nodes, and the (P)NN used in the $\tau_{\text{lep}}\tau_{\text{had}}$ LTT category has 3 hidden layers of 512 (256) nodes. The (P)NN input variables are standardised by subtracting the median value and dividing by the interquartile range. Nesterov momentum and learning-rate decay were used in the training of all (P)NNs, and in the $\tau_{\text{lep}}\tau_{\text{had}}$ categories they used an L2 regularisation ('ridge regression') term in the loss function [158]. The BDT uses 1500 trees with a maximum depth of 2 and a minimum node size of 1% of the training events. Gradient boosting is used with a shrinkage of 0.2.

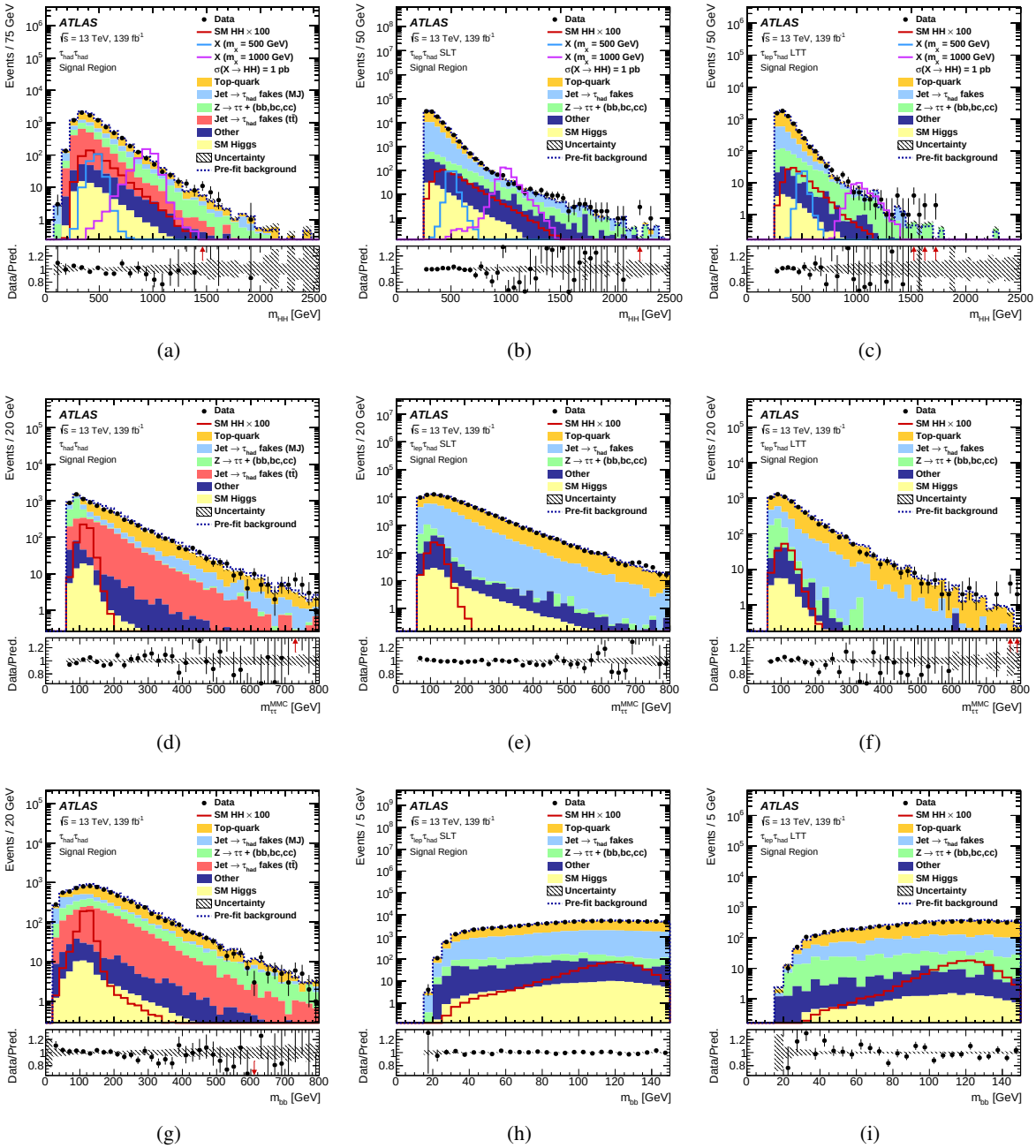


Figure 3: Signal (solid lines), background (filled histograms) and data (dots with error bars) distributions of m_{HH} (top), $m_{\tau\tau}^{\text{MMC}}$ (middle row) and m_{bb} (bottom) for events in the $\tau_{\text{had}}\tau_{\text{had}}$ (left), $\tau_{\text{lep}}\tau_{\text{had}}$ SLT (middle column) and $\tau_{\text{lep}}\tau_{\text{had}}$ LTT (right) categories. The normalisation and shape of the backgrounds and the uncertainty in the total background shown are determined from the likelihood fit (described in Section 8) to data in the non-resonant HH search. The expected non-resonant signal is overlaid with its normalisation scaled by a factor of 100, and the $m_X = 500$ GeV and $m_X = 1000$ GeV resonant signals are overlaid in the m_{HH} distributions with their cross-section set to 1 pb. The dashed histogram shows the total pre-fit background. The size of the combined statistical and systematic uncertainty of the background is indicated by the hatched band. The ratio of the data to the sum of the backgrounds is shown in the lower panels.

Table 3: Variables used as inputs to the MVAs in the three analysis categories. The same choice of input variables is used for the resonant and non-resonant production modes. The variables are defined in the main text.

Variable	$\tau_{\text{had}}\tau_{\text{had}}$	$\tau_{\text{lep}}\tau_{\text{had}}$ SLT	$\tau_{\text{lep}}\tau_{\text{had}}$ LTT
m_{HH}	✓	✓	✓
$m_{\tau\tau}^{\text{MMC}}$	✓	✓	✓
m_{bb}	✓	✓	✓
$\Delta R(\tau,\tau)$	✓	✓	✓
$\Delta R(b,b)$	✓	✓	
$\Delta p_{\text{T}}(\ell,\tau)$		✓	✓
Sub-leading b-tagged jet p_{T}		✓	
m_{T}^{W}		✓	
$E_{\text{T}}^{\text{miss}}$		✓	
$\mathbf{p}_{\text{T}}^{\text{miss}}$ ϕ centrality		✓	
$\Delta\phi(\ell\tau,bb)$		✓	
$\Delta\phi(\ell,\mathbf{p}_{\text{T}}^{\text{miss}})$			✓
$\Delta\phi(\tau\tau,\mathbf{p}_{\text{T}}^{\text{miss}})$			✓
S_{T}			✓

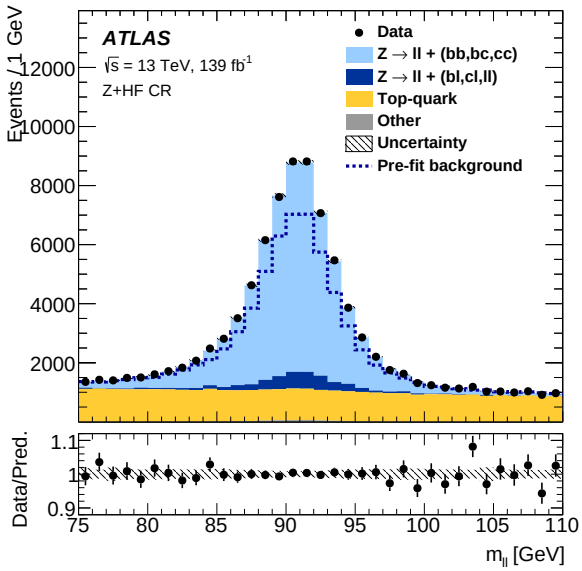
5.3 Z + HF control region

The normalisation of the Z + HF background is determined from data by fitting the $m_{\ell\ell}$ distribution in the Z + HF CR in the likelihood fit (described in Section 8). This is to account for a known discrepancy between the Z + HF production cross-section provided at NLO by Sherpa and the cross-section observed in data. The Z + HF CR targets events containing Z boson decays into electron or muon pairs by using triggers that require either a lepton or a pair of same-flavour leptons. Exactly two oppositely charged same-flavour leptons and exactly two b-tagged jets must be reconstructed offline. The leptons are required to have $p_{\text{T}} > 9$ GeV, pass offline p_{T} thresholds based on the trigger thresholds, be compatible with originating from the primary vertex, and pass medium identification and loose isolation requirements. Lastly, the invariant mass of the lepton pair is required to be between 75 GeV and 110 GeV to select events which include a Z boson decay, and m_{bb} is required to be less than 40 GeV or greater than 210 GeV to ensure orthogonality to the event selection of another search performed by ATLAS. This region also provides constraints on the normalisation of the $t\bar{t}$ background. Figure 4 shows the post-fit background and data $m_{\ell\ell}$ distributions in the Z + HF control region.

6 Background modelling

Backgrounds in this search are estimated using a combination of simulation-based and data-driven techniques.

The main sources of background are top-quark, Z+jets, W+jets, diboson, single Higgs boson and multi-jet production. A reconstructed $\tau_{\text{had-vis}}$, in these background events, can originate either from a τ_{had} decay (true-$\tau_{\text{had-vis}}$), or from a misidentified quark- or gluon-initiated jet (fake-$\tau_{\text{had-vis}}$). Events in which an electron or a muon is misidentified as a $\tau_{\text{had-vis}}$ represent a small additional background.



(a)

Figure 4: Post-fit background and data $m_{\ell\ell}$ distributions in the Z + HF control region. The normalisation and shape of the backgrounds and the uncertainty in the total background are shown as determined from the likelihood fit (described in Section 8) to data in the non-resonant HH search. The uncertainty band includes statistical and systematic uncertainties of the total background. The dashed histogram shows the total pre-fit background.

Most of the background events with fake-$\tau_{\text{had-vis}}$ are from $t\bar{t}$ or multi-jet production. In $t\bar{t}$ events, fake-$\tau_{\text{had-vis}}$ typically originate from quark-initiated jets from the top-quark decay. In multi-jet events, both quark- and gluon-initiated jets may be misidentified as $\tau_{\text{had-vis}}$.

The simulated event samples, summarised in Section 3.2, are used to model background events containing true-$\tau_{\text{had-vis}}$ and events with an electron or a muon misidentified as a $\tau_{\text{had-vis}}$. Events with fake-$\tau_{\text{had-vis}}$ in $t\bar{t}$ or multi-jet production are estimated using techniques relying on both simulated events and data, as detailed in the following subsections. Smaller backgrounds with fake-$\tau_{\text{had-vis}}$ from other production processes are estimated from simulation.

The normalisations of simulated $t\bar{t}$, for both the true- and fake-$\tau_{\text{had-vis}}$ components, and Z + HF backgrounds are determined from data in the likelihood fits of signal and control regions, as outlined in Section 8.

6.1 Fake-$\tau_{\text{had-vis}}$ background in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel

In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, a combined fake-factor method similar to that described in Ref. [35] is used to estimate multi-jet and $t\bar{t}$ backgrounds with fake-$\tau_{\text{had-vis}}$. A schematic depiction of this method is shown in Figure 5. This method employs events in two groups of regions. The events in the identification (ID) regions require one identified $\tau_{\text{had-vis}}$, whereas events in the anti-identification (anti-ID) regions contain one anti-$\tau_{\text{had-vis}}$ candidate.

Prior to the last step of the overlap removal procedure outlined in Section 4, if an event does not contain an identified $\tau_{\text{had-vis}}$, it is checked for reconstructed $\tau_{\text{had-vis}}$ candidates satisfying the anti-$\tau_{\text{had-vis}}$ requirements, as defined in Section 4. If an event contains multiple anti-$\tau_{\text{had-vis}}$ candidates, one is chosen randomly. In the

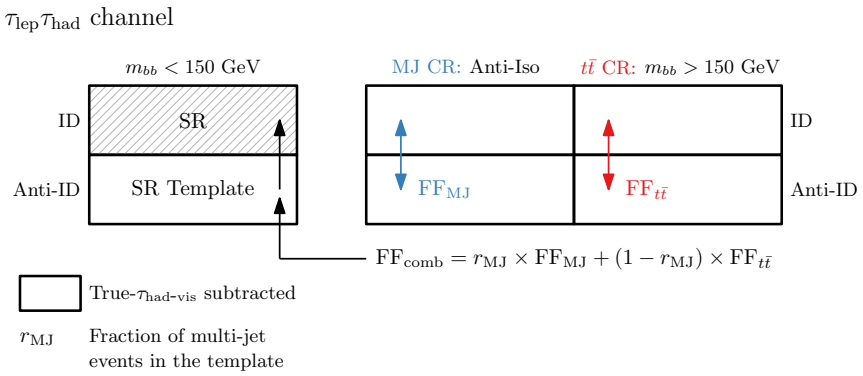


Figure 5: Schematic depiction of the combined fake-factor method used to estimate multi-jet and $t\bar{t}$ backgrounds with fake-$\tau_{\text{had-vis}}$ in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. Backgrounds which are not from events with fake-$\tau_{\text{had-vis}}$ originating from jets are estimated from simulation and are subtracted from data in all control regions. Events in which an electron or a muon is misidentified as a $\tau_{\text{had-vis}}$ are also subtracted, but their contribution is very small. Both sources are indicated by 'True-$\tau_{\text{had-vis}}$ subtracted' in the legend.

LTT category, however, only the anti-$\tau_{\text{had-vis}}$ candidate that is within $\Delta R = 0.2$ of the HLT $\tau_{\text{had-vis}}$ object is considered.

In order to define a fake-$\tau_{\text{had-vis}}$ background template, an anti-ID region is defined using event selection criteria equivalent to the SR selection, with one anti-$\tau_{\text{had-vis}}$ candidate instead of one identified $\tau_{\text{had-vis}}$. This region, which is enriched with fake-$\tau_{\text{had-vis}}$, is defined as the SR Template region. The template for estimating fake-$\tau_{\text{had-vis}}$ in the SR is obtained by subtracting from the data distribution in the SR Template region the distribution of simulated background events in which the $\tau_{\text{had-vis}}$ candidate is not a fake-$\tau_{\text{had-vis}}$ originating from jets. This subtraction is referred to as the true-$\tau_{\text{had-vis}}$ subtraction given that the number of events in which an electron or muon is misidentified as a $\tau_{\text{had-vis}}$, which are also subtracted, is very small. The data and simulated events that are used to build the template are scaled with event weights, referred to as fake-factors (FFs), to estimate the fake-$\tau_{\text{had-vis}}$ background in the SR.

The FFs are derived separately for multi-jet (FF_{MJ}) and $t\bar{t}$ ($\text{FF}_{t\bar{t}}$) events in dedicated control regions. The multi-jet control regions (MJ CRs) and $t\bar{t}$ control regions ($t\bar{t}$ CRs) are defined separately for the ID and the anti-ID regions, depending on whether they contain one identified $\tau_{\text{had-vis}}$ or one anti-$\tau_{\text{had-vis}}$ candidate, respectively. Besides the $\tau_{\text{had-vis}}$ selection, the MJ CRs are defined using the SR selection with an inverted electron or muon isolation requirement (anti-Iso) and without the $m_{bb} < 150$ GeV requirement. The MJ CR's purity in multi-jet production events varies between 65% and 90% depending on the trigger category type (SLT or LTT) and whether the MJ CR is in the ID or anti-ID region. Similarly, the $t\bar{t}$ CRs are defined using the SR selection with an inverted m_{bb} requirement ($m_{bb} > 150$ GeV). The $t\bar{t}$ CR is about 95% pure in the events from $t\bar{t}$ production. Contamination from either non-resonant or resonant HH signal in these CRs is estimated to be negligible. Backgrounds which are not from events with fake-$\tau_{\text{had-vis}}$ originating from jets, are estimated from simulation and are subtracted from the distribution of the data in all the control regions used for the FF measurement. After the subtraction the FFs are derived as the ratio of the number of events in the ID region to the number of events in the anti-ID region. They are parameterised in terms of the $\tau_{\text{had-vis}}$ p_{T}, independently for 1- and 3-prong $\tau_{\text{had-vis}}$ ('1- and 3-prong' refers to the number of tracks associated with a reconstructed $\tau_{\text{had-vis}}$), and separately for the SLT and LTT categories. The FFs corresponding to the individual background processes are combined as

$$\text{FF}_{\text{comb}} = r_{\text{MJ}} \times \text{FF}_{\text{MJ}} + (1 - r_{\text{MJ}}) \times \text{FF}_{t\bar{t}},$$

where r_{MJ} is the expected fraction of multi-jet events in the SR Template. The number of multi-jet events in the SR Template is estimated by taking the number of data events in the SR Template and subtracting the expected number of non-multi-jet background events with both true- and fake-$\tau_{\text{had-vis}}$ as estimated from simulation. The non-multi-jet background events are dominated by $t\bar{t}$ production. The r_{MJ} is parameterised as a function of the $\tau_{\text{had-vis}}$ p_{T}, and it is measured separately for the $\tau_e\tau_{\text{had}}$ and $\tau_\mu\tau_{\text{had}}$ events, for 1- and 3-prong $\tau_{\text{had-vis}}$ categories, and for the SLT and LTT categories. The FF_{comb} is used to scale the events used for the SR Template in order to obtain the fake-$\tau_{\text{had-vis}}$ background prediction in the SR.

The determination of the fake-$\tau_{\text{had-vis}}$ background using the combined FF method is sensitive to the modelling of simulated $t\bar{t}$ events with true-$\tau_{\text{had-vis}}$ given that this is the dominant background that is subtracted from data in the derivation of the FF and r_{MJ}, and when obtaining the SR Template. Additionally, the derivation of r_{MJ} is sensitive to the modelling of simulated $t\bar{t}$ events with fake-$\tau_{\text{had-vis}}$. To improve predictions, the simulated events from $t\bar{t}$ production are differentially reweighted to data distributions depending on the jet multiplicity and the scalar sum of the transverse momentum of all visible final-state objects in the event. These reweighting factors are determined from another $t\bar{t}$ control region ($t\bar{t}$ CR2), which is about 93% pure in events from $t\bar{t}$ production, and is defined using a selection identical to the SR selection, but with an inverted m_{bb} requirement ($m_{bb} > 150$ GeV) and with an additional $m_{\text{T}}^{W} > 40$ GeV requirement. Furthermore, events in this control region are required to have a reconstructed $\tau_{\text{had-vis}}$ candidate, but this candidate is not required to meet the recurrent neural-network identification criteria defined in Section 4. The m_{T}^{W} requirement is introduced to remove any potential contamination from multi-jet events.

Statistical uncertainties in $\text{FF}_{t\bar{t}}$, FF_{MJ} and r_{MJ} are evaluated and propagated to the final result. The difference between the fake-$\tau_{\text{had-vis}}$ background estimates obtained with and without the aforementioned $t\bar{t}$ modelling correction is taken as an uncertainty in the background estimate. A conservative 30% modelling uncertainty is assigned to simulated non-$t\bar{t}$ backgrounds which are subtracted from data. Due to its large dependence on the modelling of simulated $t\bar{t}$ events with fake-$\tau_{\text{had-vis}}$, the obtained values of r_{MJ} are varied by ± 0.5, with the constraint $0 \leq r_{\text{MJ}} \leq 1$. The impact of such a conservative uncertainty is small since the FFs in multi-jet and $t\bar{t}$ events are found to be similar. The total uncertainty in the FF_{comb} value for the SLT category is at most 10%, and at most 25% for the LTT category. The combined FF method is checked for closure in the $t\bar{t}$ CR and it is validated in the 0-b-tagged and 1-b-tagged regions, which are the same as the $\tau_{\text{lep}}\tau_{\text{had}}$ SR except for the requirement on the number of b-tagged jets. The signal contamination in the 0-b-tagged and 1-b-tagged regions is negligible. The estimated background distributions agree well with the observed distributions in all validation regions.

6.2 Fake-$\tau_{\text{had-vis}}$ background in the $\tau_{\text{had}}\tau_{\text{had}}$ channel

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, two separate methods are used to estimate the backgrounds with fake-$\tau_{\text{had-vis}}$ from $t\bar{t}$ and multi-jet production. Multi-jet events can only enter the signal selection when both $\tau_{\text{had-vis}}$ are fake, whereas for $t\bar{t}$ production, usually no more than one reconstructed $\tau_{\text{had-vis}}$ is fake.

6.2.1 Fake-$\tau_{\text{had-vis}}$ background from multi-jet production

In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the fake-$\tau_{\text{had-vis}}$ background from multi-jet production is estimated using a fake-factor method. A schematic depiction of this method is shown in Figure 6. The ID region selection refers to the selection of events with two identified $\tau_{\text{had-vis}}$. In order to define an anti-ID region selection, prior to

the last step of the overlap removal procedure outlined in Section 4, events that have only one identified $\tau_{\text{had-vis}}$ are checked for a reconstructed $\tau_{\text{had-vis}}$ candidate satisfying the anti-$\tau_{\text{had-vis}}$ requirements. The selected anti-$\tau_{\text{had-vis}}$ candidate is required to be within $\Delta R = 0.2$ of an HLT $\tau_{\text{had-vis}}$ object, except in the STT category for events in which the identified $\tau_{\text{had-vis}}$ is already trigger-matched. If multiple anti-$\tau_{\text{had-vis}}$ candidates fulfil the defined criteria, one is selected randomly.

In order to define a multi-jet background template, an anti-ID region is defined by using a selection equivalent to the SR selection, with one identified $\tau_{\text{had-vis}}$ and one anti-$\tau_{\text{had-vis}}$ candidate, instead of two identified $\tau_{\text{had-vis}}$. The template for estimating the multi-jet background in the SR (SR Template) is obtained by subtracting simulated non-multi-jet events from data in the template region. A large fraction of the subtracted non-multi-jet events are from $t\bar{t}$ production, and these simulated $t\bar{t}$ events with fake-$\tau_{\text{had-vis}}$ are corrected with scale factors of the fake-$\tau_{\text{had-vis}}$ misidentification efficiencies in the anti-ID region, which are described in Section 6.2.2. Similarly to the procedure used in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the events that are used to build the template are further scaled with FFs to estimate the multi-jet background in the SR.

A multi-jet-enriched control region is defined by using the $\tau_{\text{had}}\tau_{\text{had}}$ SR selection, but requiring the two $\tau_{\text{had-vis}}$ to have same-sign (SS) charges, as opposed to the SR selection that requires two $\tau_{\text{had-vis}}$ with opposite-sign (OS) charges. Additionally, events in the control region are required to have exactly one b-tagged jet per event (SS CR with 1 b-tagged jet). This control region, with its corresponding anti-ID counterpart, is used for FF measurements. This control region's purity in multi-jet production events varies from 80% to 90% depending on the trigger type (STT or DTT). The FFs are measured as the ratio of the number of events in the ID region to the number of events in the anti-ID region after subtracting all simulated non-multi-jet backgrounds from data. The FFs are determined separately for the STT and DTT categories, and for the different years of data-taking to account for the changes to the $\tau_{\text{had-vis}}$ identification algorithms and event selection topologies used in the trigger. The FFs are derived in the SS CR with 1 b-tagged jet, due to the limited number of selected events and large $t\bar{t}$ background contamination in the SS region with 2 b-tagged jets. For that reason, transfer factors (TFs) are defined to account for the extrapolation from 1-b-tagged-jet events to 2-b-tagged-jet events. In the DTT category, the FFs are binned in p_{T} and η of the anti-$\tau_{\text{had-vis}}$. In the STT category, due to the smaller number of available events, the FFs are measured inclusively in p_{T} and η, but separately according to whether the selected anti-$\tau_{\text{had-vis}}$ is the leading or sub-leading $\tau_{\text{had-vis}}$ candidate in p_{T}. In both categories, the FFs are measured separately for events with 1- or 3-prong anti-$\tau_{\text{had-vis}}$ candidates.

The TFs are defined as ratios of the FFs measured in the SS CR with 2 b-tagged jets to the FFs measured in the SS CR with 1 b-tagged jet, inclusively for the STT and DTT categories. The large contamination from $t\bar{t}$ background in the SS CR with 2 b-tagged jets is removed in the subtraction of all simulated non-multi-jet backgrounds from data. The TFs are also measured inclusively in p_{T} and η of the $\tau_{\text{had-vis}}$, but separately for events with 1- and 3-prong anti-$\tau_{\text{had-vis}}$ candidates, separately according to whether the selected anti-$\tau_{\text{had-vis}}$ is the leading or sub-leading $\tau_{\text{had-vis}}$ candidate in p_{T}, and separately for the different years of data-taking. Their values are compatible with unity within the statistical uncertainty.

The sources of uncertainty in the estimate of the fake-$\tau_{\text{had-vis}}$ background from multi-jet production include the statistical uncertainties in the measured FFs and TFs, and uncertainties in the normalisation and shape of the non-multi-jet backgrounds that are subtracted from data when deriving the SR Template. An uncertainty is also introduced to account for the extrapolation from the SS events to the OS events. The associated systematic uncertainty is estimated by comparing the FFs derived from SS 1-b-tagged-jet events with those derived from OS 1-b-tagged-jet events. To ensure that the sample is dominated by multi-jet events, the OS events are additionally required to satisfy $m_{\tau\tau}^{\text{MMC}} > 110$ GeV and $E_{\text{T}}^{\text{miss}}/\sigma(E_{\text{T}}^{\text{miss}}) < 3$, where $\sigma(E_{\text{T}}^{\text{miss}})$ is the event-based approximation to the resolution of the $E_{\text{T}}^{\text{miss}}$ [159]. The modelling

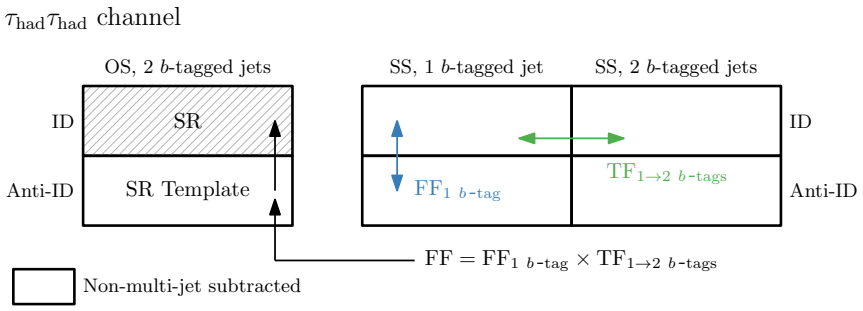


Figure 6: Schematic depiction of the combined fake-factor method to estimate the multi-jet background with fake-$\tau_{\text{had-vis}}$ in the $\tau_{\text{had}}\tau_{\text{had}}$ channel. Backgrounds that are not from multi-jet events are simulated and subtracted from data in all the control regions. This is indicated by 'Non-multi-jet subtracted' in the legend.

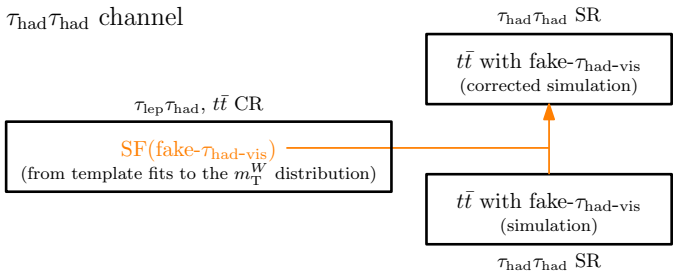


Figure 7: Schematic depiction of the fake-$\tau_{\text{had-vis}}$ scale-factor method to estimate the $t\bar{t}$ background with fake-$\tau_{\text{had-vis}}$ in the $\tau_{\text{had}}\tau_{\text{had}}$ channel.

of the multi-jet background is checked for closure in the SS CRs with 1 and 2 b-tagged jets and good agreement between data and prediction is observed. The FFs are measured in the SS CR with 1 b-tagged jet, and the TFs between the SS CR with 1 b-tagged jet and SS CR with 2 b-tagged jets only correct for the normalisation of the fake-$\tau_{\text{had-vis}}$ prediction. Thus the validation performed on the SS CR with 2 b-tagged jets provides a check of the shape extrapolation from 1 b-tagged jet to 2 b-tagged jets.

6.2.2 Fake-$\tau_{\text{had-vis}}$ background from $t\bar{t}$ production

Background events with fake-$\tau_{\text{had-vis}}$ from $t\bar{t}$ production in the $\tau_{\text{had}}\tau_{\text{had}}$ channel are estimated using simulation. However, the fake-$\tau_{\text{had-vis}}$ misidentification efficiencies are corrected by scale factors (SFs) derived from data. A schematic depiction of this method is shown in Figure 7.

The SFs are determined in the $t\bar{t}$ CR defined within the $\tau_{\text{lep}}\tau_{\text{had}}$ SLT category, as described in Section 6.1. However, in order to harmonise it with the $\tau_{\text{had}}\tau_{\text{had}}$ SR selection, the $t\bar{t}$ CR is redefined to select events with $\tau_{\text{had-vis}}$ $|\eta| < 2.5$. The SFs are extracted as a function of the fake-$\tau_{\text{had-vis}}$ p_{T}, separately for 1- and 3-prong fake-$\tau_{\text{had-vis}}$ objects, by fitting the m_{T}^{W} distribution of simulated events to data using a profile-likelihood fit. The fit of the m_{T}^{W} distribution allows the contributions of the $t\bar{t}$ events with true- and fake-$\tau_{\text{had-vis}}$ to be determined while correcting for the modelling of the $t\bar{t}$ simulation that is common to both contributions. Separate fits are performed for different trigger categories. For 1-prong fake-$\tau_{\text{had-vis}}$ the SFs are close to unity at fake-$\tau_{\text{had-vis}}$ p_{T} below 40 GeV and decrease to SF $\sim$ 0.6 for fake-$\tau_{\text{had-vis}}$ p_{T} above 70 GeV. The SFs for the 3-prong fake-$\tau_{\text{had-vis}}$ are generally about $\sim$ 20% larger than for the 1-prong fake-$\tau_{\text{had-vis}}$ objects.

The $t\bar{t}$ background with fake-$\tau_{\text{had-vis}}$ in the $\tau_{\text{had}}\tau_{\text{had}}$ SR is estimated from the simulated events that pass the SR selection, weighted by the corresponding SFs for each fake-$\tau_{\text{had-vis}}$ in the event.

Uncertainties in the detector response and in the modelling of the $t\bar{t}$ events, as well as of the other minor contributing processes, are taken into account in the likelihood fit when extracting the SFs. The covariance matrix of the measured SFs, which contains all statistical and systematic uncertainties of the measurement, is diagonalised and the resulting eigenvectors are used to define independent nuisance parameters which are propagated to the final signal extraction fit. Theoretical modelling uncertainties in simulated $t\bar{t}$ events to which the SFs are applied are estimated as described in Section 7 and they are also propagated to the final signal extraction fit.

When estimating the fake-$\tau_{\text{had-vis}}$ background from multi-jet production using the fake-factor method (cf. Section 6.2.1), a large fraction of $t\bar{t}$ events containing at least one fake-$\tau_{\text{had-vis}}$ needs to be subtracted from data in the OS 2-b-tagged-jet anti-ID region (SR Template) to estimate the multi-jet contribution in the $\tau_{\text{had}}\tau_{\text{had}}$ SR. The modelling of the simulated $t\bar{t}$ events with fake-$\tau_{\text{had-vis}}$ in the anti-ID region is corrected with SFs that are measured using the same method as described above. These SFs are measured in a control region similar to the $t\bar{t}$ CR except that the $\tau_{\text{had-vis}}$ candidate has to satisfy the anti-$\tau_{\text{had-vis}}$ requirements. The measured SFs for 1-prong fake-$\tau_{\text{had-vis}}$ in the anti-ID region are close to unity at fake-$\tau_{\text{had-vis}}$ p_{T} below 40 GeV, and follow the same trend of decreasing in value with increasing fake-$\tau_{\text{had-vis}}$ p_{T}, as observed in the ID region. The SFs for the 3-prong fake-$\tau_{\text{had-vis}}$ are generally about 10%–20% larger than for the 1-prong fake-$\tau_{\text{had-vis}}$ objects.

7 Systematic uncertainties

The most significant uncertainty in this search is the statistical uncertainty of the data in the signal region. Nonetheless, experimental, theoretical, and modelling uncertainties in the normalisation and shape of the signal and background estimates give a non-negligible contribution to the total uncertainty and are thus evaluated for use in the statistical model (described in Section 8); these are described below. Statistical uncertainties in the predicted background processes are modelled using a simplified version of the Beeston–Barlow technique [160], in which only the uncertainty in the total background content in each bin is considered.

Systematic uncertainties in the detector response are considered in this search. Where relevant, uncertainties in the trigger, reconstruction, identification, and isolation efficiencies, and in the momentum of electrons [131], muons [132] and $\tau_{\text{had-vis}}$ [148] are estimated. Additional uncertainties are estimated for the efficiencies of the electron veto for $\tau_{\text{had-vis}}$ and the track-to-vertex-matching requirements for muons. Jet energy scale and resolution uncertainties [138] and the uncertainty in the efficiency of matching jets to the primary vertex [139] are estimated. These energy scale and resolution uncertainties, in addition to an uncertainty in the tracks matched to the primary vertex but not associated with other reconstructed objects in the event, are propagated to the $E_{\text{T}}^{\text{miss}}$ calculation [150, 161]. Uncertainties in the b-jet tagging efficiency are considered, and these are evaluated as a function of p_{T} for b-tagged jets and c-tagged jets by using $t\bar{t}$ events [143, 162], and as a function of p_{T} for light-flavour jets by using Z + jets events [163]. An uncertainty of 1.7% in the total integrated luminosity [79], obtained using the LUCID-2 detector [80], is assigned to physics processes whose normalisations are taken from simulation. An uncertainty arising from the correction of the pile-up distribution in simulation to that in data is also estimated.

The effects of uncertainties in the parton shower on the acceptances of the resonant (non-resonant) HH signals are assessed by comparing the nominal MC samples with alternative samples which use Pythia 8 (Herwig 7) to model the parton shower. Alternative samples were not generated for some of the resonant HH signals, and uncertainties for these signals are derived by interpolating or extrapolating the uncertainties from signals nearby in m_X for which alternative samples were produced. Independent parton-shower acceptance uncertainties are accounted for in the normalisation of the non-resonant ggF, non-resonant VBF and resonant HH signals, and these range between 0.1% and 19% of the nominal acceptance values. The effects of the renormalisation and factorisation scale, PDF and α_s uncertainties on the ggF and VBF HH signal acceptances are included when they are found to be non-negligible. Separate normalisation uncertainties are applied in the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ LTT categories to account for observed differences in the $\tau_{\text{had-vis}}$ trigger acceptances between fast-simulation and full-simulation resonant HH signal MC samples. The effects of renormalisation and factorisation scale, PDF and α_s uncertainties on the SM ggF and VBF HH cross-sections are considered, and an uncertainty arising from the scheme and value used for the top-quark mass in calculations of the virtual top-quark loop contributions in the SM ggF HH cross-section [21] is included. Lastly, uncertainties in the $H \to b\bar{b}$ and $H \to \tau^+\tau^-$ branching fractions [71] are considered in the HH signal and single-Higgs-boson background samples.

The $t\bar{t}$ normalisation is determined in the likelihood fits, so the analysis is not sensitive to uncertainties in its expected cross-section. However, the relative acceptances of the analysis categories and the distribution shapes of the discriminating variables are still sensitive to modelling uncertainties, and these are evaluated using a method that closely follows Ref. [164]. The $t\bar{t}$ hard-scatter and parton-shower uncertainties are evaluated by comparing the nominal sample with samples generated using MadGraph5_aMC@NLO + Pythia 8 and Powheg Box v2 + Herwig 7, respectively. The hard-scatter uncertainty also accounts for the matrix-element–parton-shower matching and merging uncertainty (henceforth referred to as the matching uncertainty). The impact of the uncertainty in the amount of initial-state QCD radiation is evaluated by co-varying the h_{damp} parameter, the renormalisation and factorisation scales, and the showering tune. The effects of uncertainties in the amount of final-state QCD radiation, the PDF, the α_s value, and the renormalisation and factorisation scales are all evaluated by varying generator parameter values or the PDF used in the generation of the sample. All of these uncertainties affect the relative acceptances of the analysis categories. Additionally, the effects of the hard-scatter and parton-shower uncertainties on the distribution shapes of the MVA output variables are evaluated directly in all categories, but are found to be negligible in the $\tau_{\text{lep}}\tau_{\text{had}}$ LTT category. The impact of the uncertainty in the amount of initial-state QCD radiation on the MVA output variables is considered in the $\tau_{\text{had}}\tau_{\text{had}}$ category.

The Z + HF normalisation is determined in the likelihood fits. The effects of uncertainties in the hard scatter, matching, parton shower, the renormalisation, factorisation and resummation scales, PDF and α_s on the relative acceptances of the analysis categories are all considered. For the renormalisation, factorisation and resummation scales, these uncertainties are evaluated by varying the relevant scales upwards and downwards from their nominal values by a factor of two; both simultaneous and independent variations are considered for the renormalisation and factorisation scales, while the resummation scale is varied independently. In the $\tau_{\text{had}}\tau_{\text{had}}$ category, the impact of the hard-scatter and parton-shower uncertainty is evaluated by comparing the nominal sample with a MadGraph5_aMC@NLO + Pythia 8 sample, and the impact on distribution shapes is parameterised in m_{bb} and propagated to the MVA output variables. This uncertainty is not applied in the $\tau_{\text{lep}}\tau_{\text{had}}$ category and the Z + HF CR. In the $\tau_{\text{had}}\tau_{\text{had}}$ category, the impact of the renormalisation and factorisation scale uncertainties on the shape of the MVA output variables is evaluated directly, and in the $\tau_{\text{lep}}\tau_{\text{had}}$ categories, their impact is parameterised in the p_{T} of the b-tagged jet pair and propagated to the MVA output variables.

The effects of uncertainties in the PDF and the amount of initial- and final-state QCD radiation on the acceptances of the single-top-quark backgrounds are estimated. Cross-section uncertainties are also considered. Despite these backgrounds being small inclusively, their total yields in the most sensitive MVA bins are significant, and the acceptance uncertainties in the different analysis categories range between 14% and 34% of the nominal values. The impact of the uncertainty in the amount of initial-state QCD radiation is evaluated by co-varying the renormalisation and factorisation scales, and the showering tunes. The impact of the uncertainty in the amount of final-state QCD radiation is evaluated by varying the VAR2 parameter [91] in the generation of the samples. The impact of uncertainties in the combined hard scatter and matching (parton shower) of the Wt background is estimated by comparing the nominal sample with a sample generated using MADGRAPH5_AMC@NLO + PYTHIA 8 (POWHEG BOX V2 + HERWIG 7). The impact of the uncertainty in the interference between the Wt and $t\bar{t}$ backgrounds on the acceptance is evaluated by comparing simulated Wt samples produced with the diagram removal and diagram subtraction schemes [165], parameterising this difference in the p_{T} of the b-tagged jet pair, and propagating it to the normalisation and shape of the MVA output distributions.

The effects of uncertainties in the renormalisation and factorisation scales, PDF and α_{s} on the normalisation of the single-Higgs-boson backgrounds are evaluated. Uncertainties of 100% are applied separately to the normalisations of the ggF, VBF and WH single-Higgs-boson backgrounds where the Higgs boson decays into τ-leptons, to account for difficulties in the modelling of these processes in association with heavy-flavour jets [166, 167]. The impact of uncertainties in the parton shower on the acceptances are evaluated for the ZH and $t\bar{t}H$ backgrounds by comparing the nominal samples with alternative samples generated using HERWIG 7 and EVTGEN. Uncertainties in the matching and the amount of initial- and final-state QCD radiation are also considered for the $t\bar{t}H$ backgrounds, and the impact of QCD radiation uncertainties is evaluated by varying the VAR3C parameter [91] and the renormalisation scale. Uncertainties in the shape of the MVA output distributions of all single-Higgs-boson processes are found to be negligible. Cross-section uncertainties are applied to the normalisation of the simulated single-Higgs-boson [71], single-top-quark, Z+light-flavour jets, W+jets, and diboson backgrounds, and acceptance uncertainties are applied to the normalisation of various small backgrounds.

Systematic uncertainties affecting the data-driven background estimates are described in Section 6.

Table 4 shows the breakdown of the relative contributions to the uncertainty in the non-resonant and resonant extracted signal cross-sections, which is obtained from the likelihood fit described in Section 8. Systematic uncertainties in the single-Higgs-boson backgrounds affect the non-resonant search more than these resonant searches, due to the more significant contribution of ggF single-Higgs-boson background in the most signal-like regions of the MVA distributions in the non-resonant search.

8 Statistical interpretation

The HH signal yields are each estimated from data using a simultaneous binned maximum-likelihood fit to the MVA output distributions in the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$, $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ SLT, and $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ LTT event categories, and to the $m_{\ell\ell}$ distribution in the Z + HF CR. The $t\bar{t}$ and Z + HF background normalisations are free parameters in the fit, and are primarily constrained by the background-like MVA output bins and the Z + HF CR. Upper limits are set on the signal normalisations at the 95% confidence level (CL) using the profile-likelihood-ratio test statistic [168] and the modified frequentist $\mathrm{CL_s}$ technique [169] in the asymptotic approximation [168].

Table 4: Breakdown of the relative contributions to the uncertainty in the extracted signal cross-sections, as determined in the likelihood fit (described in Section 8) to data. They are obtained by fixing the relevant nuisance parameters in the likelihood fit, subtracting the square of the obtained uncertainty in the fitted signal cross-section from the square of the total uncertainty, taking the square root, and then dividing by the total uncertainty. The sum in quadrature of the individual components differs from the total uncertainty due to correlations between uncertainties in the different groups.

Uncertainty source	Non-resonant HH	Resonant $X \to HH$		
		300 GeV	500 GeV	1000 GeV
Data statistical + floating normalisation	81%	76%	90%	93%
Data statistical	81%	76%	90%	93%
$t\bar{t}$ and Z + HF normalisations	4%	8%	3%	5%
Systematic	58%	65%	43%	37%
MC statistical	28%	44%	33%	18%
Experimental	12%	31%	8%	12%
Jet and $E_{\mathrm{T}}^{\mathrm{miss}}$	8%	27%	5%	4%
b-jet tagging	5%	5%	3%	7%
$\tau_{\text{had-vis}}$	6%	12%	3%	8%
Electrons and muons	3%	3%	2%	2%
Luminosity and pile-up	3%	2%	2%	5%
Background and signal and modelling	42%	39%	26%	30%
Fake-$\tau_{\text{had-vis}}$	8%	19%	4%	8%
Top-quark	24%	17%	12%	8%
$Z(\to \tau\tau)$ + HF	9%	17%	9%	15%
Single Higgs boson	29%	2%	14%	15%
Other backgrounds	3%	2%	5%	3%
Signal	5%	14%	7%	15%

The systematic uncertainties described in Sections 6 and 7 are represented in the fit as Gaussian- or Poisson-constrained nuisance parameters, which modify the normalisation, relative normalisation between event categories, and/or distribution shape of the discriminating variable for the signal and background processes. Systematic uncertainties are symmetrised and shape uncertainties are smoothed where physically motivated, and then those with a negligible impact are removed from the likelihood fit. Experimental, cross-section, and acceptance uncertainties are correlated across the event categories, except for the parton-shower uncertainty of the $t\bar{t}$ background, since the kinematic and topological properties of this background differ between event categories. Modelling uncertainties in data-driven background estimates are not correlated across different estimation strategies, because different sources of fake-$\tau_{\text{had-vis}}$ are estimated using different procedures.

The binning schemes for the MVA output distributions used in the likelihood fit were chosen to minimise the number of bins, while also maximising the retained expected sensitivity, and ensuring the stability of the fit and the validity of the asymptotic approximation. The binning schemes start from finely binned histograms, and bins are iteratively merged beginning from the most signal-like MVA bins until the following channel-dependent criteria are fulfilled. In the $\tau_{\text{had}}\tau_{\text{had}}$ channel, the bins are required to satisfy $\sigma_{\text{b}}^{\text{MC}} < 0.5 f_{\text{s}} + 1\%$, where $\sigma_{\text{b}}^{\text{MC}}$ is the relative MC statistical uncertainty of the background estimate and f_{s} is the signal fraction in the bin. In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the bins are required to satisfy $10 f_{\text{s}} + 5 f_{\text{b}} > 1$, where f_{s} and f_{b} are the signal and background fractions in the bin, respectively. Bins in all channels must be expected to contain at least five background events to ensure that the asymptotic approximation is valid.

9 Results

The post-fit normalisation factors of the unconstrained $t\bar{t}$ and Z + HF backgrounds for the combined $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, as obtained in the search for non-resonant HH production, are 0.97 ± 0.04 and 1.40 ± 0.11, respectively. These post-fit background normalisation factors are compatible within uncertainties with those obtained for the resonant HH production searches.

The MVA output distributions in the search for non-resonant HH production and for two resonance mass hypotheses (m_X = 500 GeV and 1 TeV) are shown in Figure 8, after performing the fits to the data and assuming a background-only hypothesis. The data, expected signal and post-fit background yields for events entering the most signal-like bin in the non-resonant MVA output histograms in the three event categories are shown in Table 5.

Table 5: Data, expected signal and background yields for events passing the event selection and entering the most signal-like bin in the non-resonant MVA output histograms in the three event categories. The background yields are calculated following a likelihood fit of the background model to data in the non-resonant HH search. The non-resonant HH signal yields are normalised using the SM cross-section. The 'multi-jet fakes' and '$t\bar{t}$ fakes' estimates are only used in the $\tau_{\text{had}}\tau_{\text{had}}$ category, and the 'combined fakes' estimate is only used in the $\tau_{\text{lep}}\tau_{\text{had}}$ categories. All systematic uncertainties are included in the indicated uncertainties.

	$\tau_{\text{had}}\tau_{\text{had}}$	$\tau_{\text{lep}}\tau_{\text{had}}$ SLT	$\tau_{\text{lep}}\tau_{\text{had}}$ LTT
Data	8	7	7
Non-resonant ggF HH	1.58 ± 0.27	0.77 ± 0.13	0.25 ± 0.05
Non-resonant VBF HH	0.0227 ± 0.0019	0.0075 ± 0.0007	0.00455 ± 0.00035
$t\bar{t}$	0.5 ± 0.1	0.44 ± 0.12	1.76 ± 0.32
Single top-quark	0.47 ± 0.20	1.2 ± 0.7	0.61 ± 0.35
Z + HF	2.3 ± 0.4	1.55 ± 0.33	1.7 ± 0.4
Combined fakes	–	1.09 ± 0.22	0.8 ± 0.5
Multi-jet fakes	0.47 ± 0.17	–	–
$t\bar{t}$ fakes	0.29 ± 0.07	–	–
Single Higgs boson	1.7 ± 0.5	1.10 ± 0.28	0.4 ± 0.1
Other backgrounds	0.4 ± 0.1	0.48 ± 0.09	0.33 ± 0.07
Total background	6.1 ± 0.8	6 ± 1	6 ± 1

Table 6 shows the 95% CL upper limits on the cross-sections for non-resonant HH production, and on the ratios of these cross-sections to their SM predictions. The combined observed (expected) limit on the SM non-resonant HH production cross-section via the ggF and VBF processes is 140 fb (110^{+40}_{-30} fb), while the limit on the ratio to the SM prediction is 4.7 ($3.9^{+1.5}_{-1.1}$). The expected limit is calculated assuming no HH production.

Figure 9 shows the data, background and non-resonant HH signal yields, where final-discriminant bins from the $\tau_{\text{had}}\tau_{\text{had}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ SLT and $\tau_{\text{lep}}\tau_{\text{had}}$ LTT categories are combined into bins of $\log_{10}(S/B)$. The fitted background yield B assumes the background-only hypothesis. The signal S that is shown in the overlaid histograms is scaled either to the SM expected cross-section or to the combined expected limit cross-section.

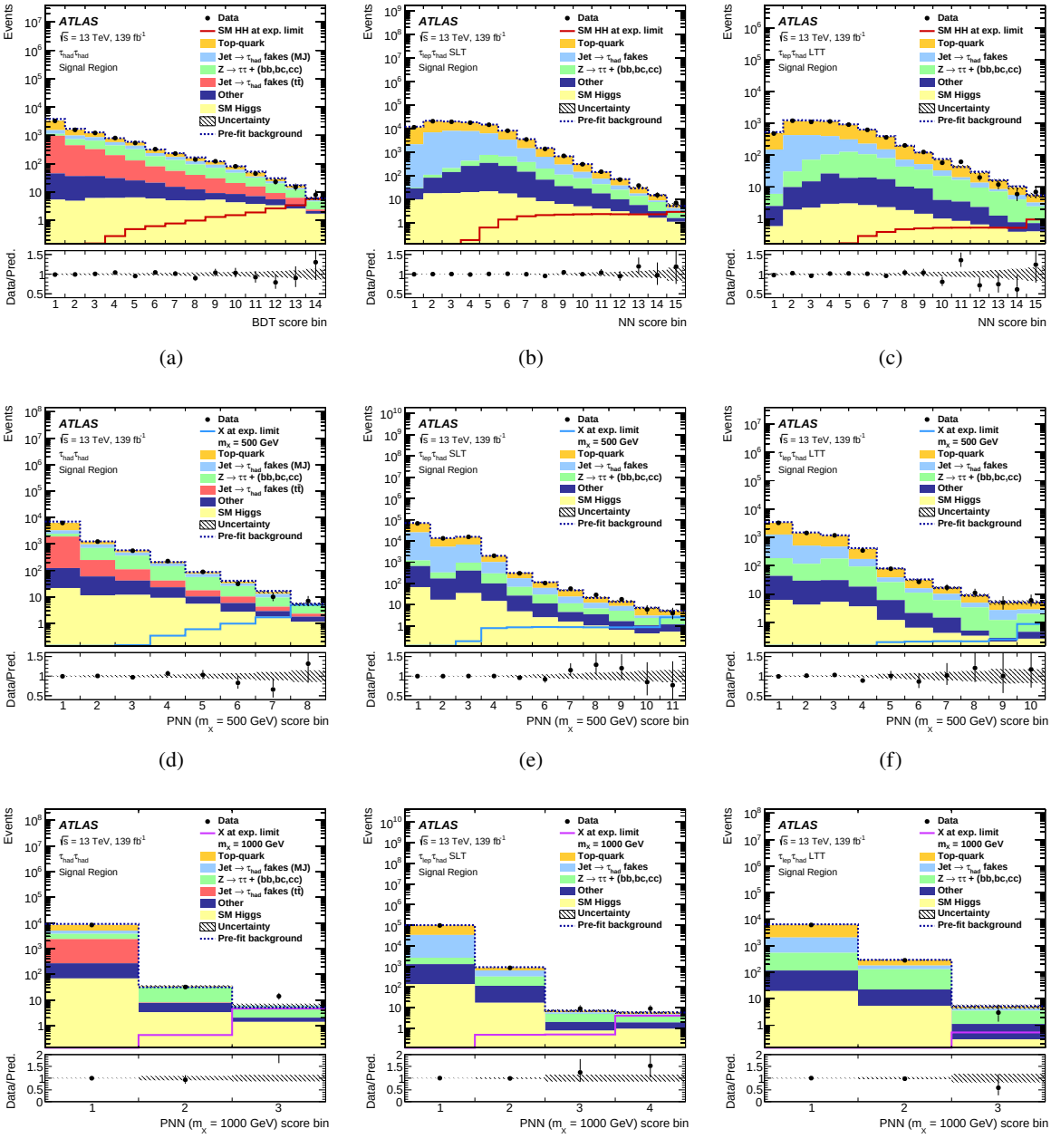


Figure 8: The MVA output distributions in the search for non-resonant HH signal (top) and in the search for resonant HH signal with $m_X = 500$ GeV (middle row) and $m_X = 1000$ GeV (bottom), in the $\tau_{\text{had}}\tau_{\text{had}}$ (left), $\tau_{\text{lep}}\tau_{\text{had}}$ SLT (middle column) and $\tau_{\text{lep}}\tau_{\text{had}}$ LTT (right) categories. The distributions are shown after performing the fits to data and assuming the background-only hypothesis. The signal is overlaid and scaled to the combined expected limit. The dashed histogram shows the total pre-fit background. The lower panels show the ratio of data to the total post-fit background, where the hatched band shows the statistical and systematic uncertainties of that background. For visualisation purposes, these histograms are displayed using uniform bin widths instead of the bin edges used in the fit, although the bin contents correspond to those used in the fit. Indices are used to label the bins.

Table 6: Observed and expected 95% CL upper limits on the cross-section for non-resonant HH production, assuming SM kinematics, and on the cross-section for non-resonant HH production divided by the SM prediction. The $\pm1\sigma$ and $\pm2\sigma$ variations around the expected limit are also shown.

		Observed	-2σ	-1σ	Expected	$+1\sigma$	$+2\sigma$
$\tau_{\text{had}}\tau_{\text{had}}$	$\sigma_{\text{ggF+VBF}}$ [fb]	150	70	95	130	180	240
	$\sigma_{\text{ggF+VBF}}/\sigma^{\text{SM}}_{\text{ggF+VBF}}$	5.0	2.4	3.2	4.4	6.1	8.2
$\tau_{\text{lep}}\tau_{\text{had}}$	$\sigma_{\text{ggF+VBF}}$ [fb]	280	120	170	230	320	430
	$\sigma_{\text{ggF+VBF}}/\sigma^{\text{SM}}_{\text{ggF+VBF}}$	9.7	4.2	5.6	7.8	11	15
Combined	$\sigma_{\text{ggF+VBF}}$ [fb]	140	62	83	110	160	210
	$\sigma_{\text{ggF+VBF}}/\sigma^{\text{SM}}_{\text{ggF+VBF}}$	4.7	2.1	2.8	3.9	5.4	7.2

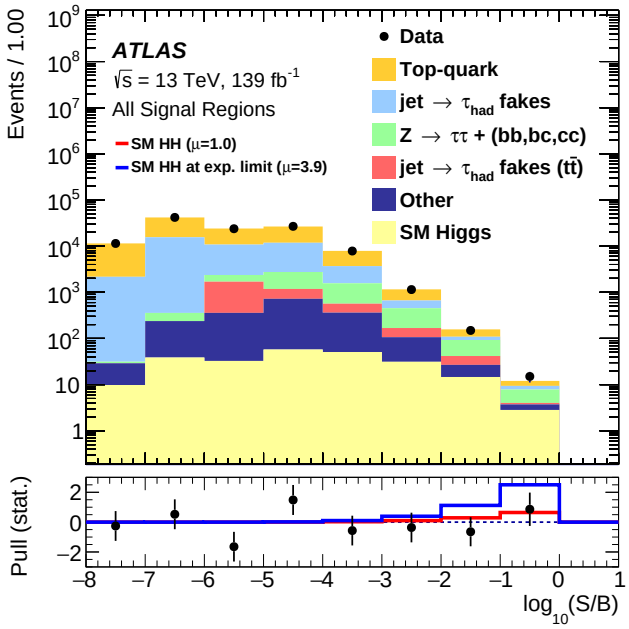


Figure 9: Event yields as a function of $\log_{10}(S/B)$ for data, background and non-resonant HH signal. Final-discriminant bins from the $\tau_{\text{had}}\tau_{\text{had}}$, $\tau_{\text{lep}}\tau_{\text{had}}$ SLT and $\tau_{\text{lep}}\tau_{\text{had}}$ LTT categories are combined into bins of $\log_{10}(S/B)$. The fitted background yield B assumes the background-only hypothesis. In the lower panel, the pull of the data relative to the background (the statistical significance of the difference between data and fitted background) is shown with statistical uncertainties only. The full lines indicate the pull expected from the sum of the fitted background and the signal, which is scaled to either the SM expected cross-section (red) or the combined expected limit cross-section (blue), relative to the fitted background.

For resonant HH production, the exclusion limits on the cross-section are shown as a function of m_X in Figure 10. The $\tau_{\text{had}}\tau_{\text{had}}$ channel has better sensitivity than the $\tau_{\text{lep}}\tau_{\text{had}}$ channel below a resonance mass of 1 TeV, while their sensitivities become comparable at higher masses. A broad excess is observed in both channels in the mass range 700 GeV $< m_X <$ 1.2 TeV. The most significant excess for the $\tau_{\text{had}}\tau_{\text{had}}$ ($\tau_{\text{lep}}\tau_{\text{had}}$) channel is found for the signal mass hypothesis of 1 TeV (1.1 TeV) with a local significance of 2.8σ (1.6σ). The most significant combined excess is at a signal mass hypothesis of 1 TeV with a local significance of 3.1σ and a global significance of 2.0σ.

The p-value defining the global significance is given by the probability of finding a maximum local excess larger than the observed value under the background-only hypothesis, regardless of the resonance mass value m_X where the largest excess is found. It is estimated using fits of 'toy' experiments drawn from a model providing a joint description of all observables in the statistical interpretation. The set of observables comprises the number of events observed in a bin entering either a signal or control region, separately for all bins, as well as the central values of auxiliary measurements used to propagate systematic uncertainties to the results. This model differs from the probabilistic model used to define the likelihood functions since these describe only one PNN discriminant at a time, targeting a single signal mass hypothesis. However, non-trivial correlations between observables are expected since events are subject to the same selection criteria regardless of the resonance mass value m_X to be probed, but different multivariate discriminants are used for each of these signal hypotheses. Observables describing the number of events in a given bin are generated from a multivariate Poisson distribution with the dependence structure between observables given by the copula [170] of a centred multivariate normal distribution with covariance matrix given by the expected linear correlations estimated from a simulation- and data-driven background model. Observables describing the statistical precision of the background estimate are generated using a resampling approach [171]. Observables associated with all other experimental and theoretical uncertainties are varied coherently for all hypothesis tests.

10 Conclusion

A search for non-resonant and resonant Higgs boson pair production in $b\bar{b}\tau^+\tau^-$ events is conducted, where the non-resonant signal is assumed to be produced with SM kinematics, and the resonant signal corresponds to a narrow scalar resonance with a mass m_X in the range 251 to 1600 GeV. The 13 TeV pp collision data set used was collected at the LHC by the ATLAS experiment between 2015 and 2018, and corresponds to an integrated luminosity of 139 fb^{-1}. The sensitivity of this search to the non-resonant signal hypothesis improves on the previous ATLAS search in this channel by around a factor of four. Roughly half of this improvement is due to the larger data set, while most of the remaining sensitivity gain is due to significant improvements in the $\tau_{\text{had-vis}}$ and b-jet reconstruction and identification. Analysis-level improvements include the use of more sophisticated multivariate techniques to target the resonant signal hypotheses, and new fake-$\tau_{\text{had-vis}}$ estimation methods. The data are found to be compatible with the background-only hypothesis, with the largest deviation being found in the search for resonant HH production at a mass of 1 TeV and corresponding to a local (global) significance of 3.1σ (2.0σ). The observed (expected) upper limit on the non-resonant Higgs boson pair-production cross-section, set at the 95% confidence level, is 4.7 ($3.9^{+1.5}_{-1.1}$) times the SM expectation. Observed (expected) upper limits are placed at the 95% confidence level on resonant Higgs boson production and exclude cross-sections above 21–900 fb (12–840 fb), depending on the mass of the resonance. This search provides the highest expected sensitivity to non-resonant HH production of any individual search to date, and provides limits on resonant HH production that are more

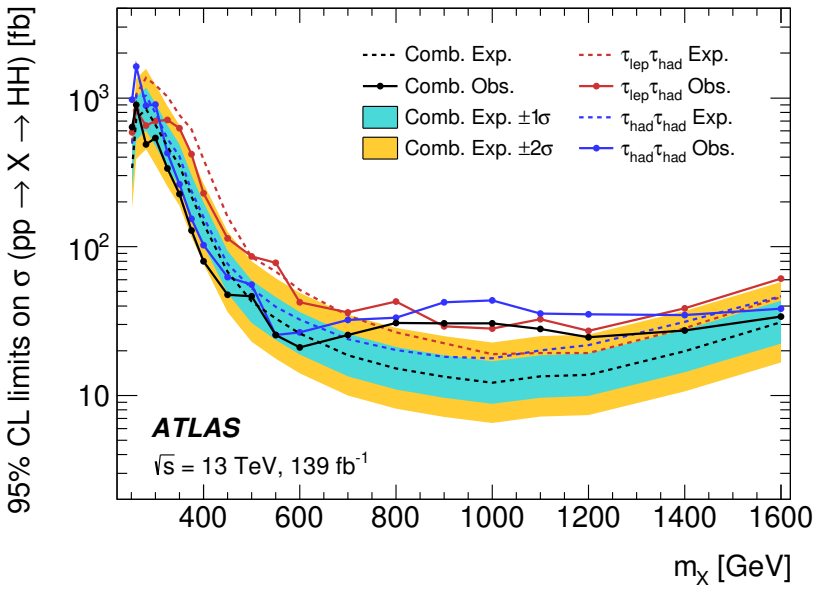


Figure 10: Observed and expected limits at 95% CL on the cross-section for resonant HH production as a function of the scalar resonance mass m_X. The dashed lines show the expected limits while the solid lines show the observed limits. The blue and red lines are the limits for the $\tau_{\text{had}}\tau_{\text{had}}$ channel and $\tau_{\text{lep}}\tau_{\text{had}}$ channel, respectively. The black lines are the combined limits of the two channels. The $\pm 1\sigma$ and $\pm 2\sigma$ variations around the expected combined limit are indicated by the turquoise and yellow bands, respectively. The limits are obtained using the profile-likelihood test statistic and the modified frequentist CL_s technique.

stringent than, or competitive with, the most recently published ATLAS and CMS HH resonant search combinations over much of the m_X range explored.

References

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

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

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

[4] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[5] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[6] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803, arXiv: 1503.07589 [hep-ex].

[7] CMS Collaboration, *A measurement of the Higgs boson mass in the diphoton decay channel*, Phys. Lett. B **805** (2020) 135425, arXiv: 2002.06398 [hep-ex].

[8] ATLAS Collaboration,
Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector, Eur. Phys. J. C **75** (2015) 476, arXiv: 1506.05669 [hep-ex],
Erratum: Eur. Phys. J. C **76** (2016) 152.

[9] ATLAS Collaboration, *CP Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and tH Processes Using $H \to \gamma\gamma$ with the ATLAS Detector*, Phys. Rev. Lett. **125** (2020) 061802, arXiv: 2004.04545 [hep-ex].

[10] ATLAS Collaboration, *Constraints on off-shell Higgs boson production and the Higgs boson total width in $ZZ \to 4\ell$ and $ZZ \to 2\ell 2\nu$ final states with the ATLAS detector*, Phys. Lett. B **786** (2018) 223, arXiv: 1808.01191 [hep-ex].

[11] CMS Collaboration, *Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state*, Phys. Rev. D **99** (2019) 112003, arXiv: 1901.00174 [hep-ex].

[12] ATLAS Collaboration,
A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery, Nature **607** (2022) 52, arXiv: 2207.00092 [hep-ex].

[13] CMS Collaboration,
A portrait of the Higgs boson by the CMS experiment ten years after the discovery, Nature **607** (2022) 60, arXiv: 2207.00043 [hep-ex].

[14] S. Dawson, S. Dittmaier and M. Spira,
Neutral Higgs boson pair production at hadron colliders: QCD corrections, Phys. Rev. D **58** (1998) 115012, arXiv: hep-ph/9805244.

[15] S. Borowka et al., *Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence*, Phys. Rev. Lett. **117** (2016) 012001, arXiv: 1604.06447 [hep-ph], Erratum: Phys. Rev. Lett. **117** (2016) 079901.

[16] J. Baglio et al., *Gluon fusion into Higgs pairs at NLO QCD and the top mass scheme*, Eur. Phys. J. C **79** (2019) 459, arXiv: 1811.05692 [hep-ph].

[17] D. de Florian and J. Mazzitelli,
Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD, Phys. Rev. Lett. **111** (2013) 201801, arXiv: 1309.6594 [hep-ph].

[18] D. Y. Shao, C. S. Li, H. T. Li and J. Wang,
Threshold resummation effects in Higgs boson pair production at the LHC, JHEP **07** (2013) 169, arXiv: 1301.1245 [hep-ph].

[19] D. de Florian and J. Mazzitelli,
Higgs pair production at next-to-next-to-leading logarithmic accuracy at the LHC, JHEP **09** (2015) 053, arXiv: 1505.07122 [hep-ph].

[20] M. Grazzini et al., *Higgs boson pair production at NNLO with top quark mass effects*, JHEP **05** (2018) 059, arXiv: 1803.02463 [hep-ph].

[21] J. Baglio et al., *$gg \to HH$: Combined uncertainties*, Phys. Rev. D **103** (2021) 056002, arXiv: 2008.11626 [hep-ph].

[22] J. Baglio et al., *The measurement of the Higgs self-coupling at the LHC: theoretical status*, JHEP **04** (2013) 151, arXiv: 1212.5581 [hep-ph].

[23] R. Frederix et al., *Higgs pair production at the LHC with NLO and parton-shower effects*, Phys. Lett. B **732** (2014) 142, arXiv: 1401.7340 [hep-ph].

[24] L.-S. Ling et al., *NNLO QCD corrections to Higgs pair production via vector boson fusion at hadron colliders*, Phys. Rev. D **89** (2014) 073001, arXiv: 1401.7754 [hep-ph].

[25] F. A. Dreyer and A. Karlberg, *Fully differential vector-boson fusion Higgs pair production at next-to-next-to-leading order*, Phys. Rev. D **99** (2019) 074028, arXiv: 1811.07918 [hep-ph].

[26] F. A. Dreyer and A. Karlberg, *Vector-boson fusion Higgs pair production at N^3LO*, Phys. Rev. D **98** (2018) 114016, arXiv: 1811.07906 [hep-ph].

[27] V. Branchina, H. Faivre and V. Pangon, *Effective potential and vacuum stability*, J. Phys. G **36** (2009) 015006, arXiv: 0802.4423 [hep-ph].

[28] D. Buttazzo et al., *Investigating the near-criticality of the Higgs boson*, JHEP **12** (2013) 089, arXiv: 1307.3536 [hep-ph].

[29] M. Reichert et al., *Probing baryogenesis through the Higgs boson self-coupling*, Phys. Rev. D **97** (2018) 075008, arXiv: 1711.00019 [hep-ph].

[30] F. L. Bezrukov and M. Shaposhnikov, *The Standard Model Higgs boson as the inflaton*, Phys. Lett. B **659** (2008) 703, arXiv: 0710.3755 [hep-th].

[31] I. Masina and A. Notari, *Higgs mass range from standard model false vacuum inflation in scalar-tensor gravity*, Phys. Rev. D **85** (2012) 123506, arXiv: 1112.2659 [hep-ph].

[32] G. D. Kribs and A. Martin, *Enhanced di-Higgs production through light colored scalars*, Phys. Rev. D **86** (2012) 095023, arXiv: 1207.4496 [hep-ph].

[33] R. Gröber and M. Mühlleitner, *Composite Higgs boson pair production at the LHC*, JHEP **06** (2011) 020, arXiv: 1012.1562 [hep-ph].

[34] R. Contino et al., *Anomalous couplings in double Higgs production*, JHEP **08** (2012) 154, arXiv: 1205.5444 [hep-ph].

[35] ATLAS Collaboration, *A search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. Lett. **121** (2018) 191801, arXiv: 1808.00336 [hep-ex], Erratum: Phys. Rev. Lett. **122** (2019) 089901.

[36] CMS Collaboration, *Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **778** (2018) 101, arXiv: 1707.02909 [hep-ex].

[37] CMS Collaboration, *Search for nonresonant Higgs boson pair production in final state with two bottom quarks and two tau leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, (2022), arXiv: 2206.09401 [hep-ex].

[38] ATLAS Collaboration, *Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV pp collision data collected by the ATLAS experiment*, JHEP **11** (2018) 040, arXiv: 1807.04873 [hep-ex].

[39] 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].

[40] CMS Collaboration, *Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **03** (2021) 257, arXiv: 2011.12373 [hep-ex].

[41] ATLAS Collaboration, *Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **01** (2019) 030, arXiv: 1804.06174 [hep-ex].

[42] ATLAS Collaboration, *Search for the $HH \to b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **07** (2020) 108, arXiv: 2001.05178 [hep-ex], Erratum: JHEP **01** (2021) 145.

[43] CMS Collaboration, *Search for nonresonant Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state at $\sqrt{s}$ = 13 TeV*, JHEP **04** (2019) 112, arXiv: 1810.11854 [hep-ex].

[44] CMS Collaboration, *Search for Higgs boson pair production in the four b quark final state in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, (2022), arXiv: 2202.09617 [hep-ex].

[45] CMS Collaboration, *Search for nonresonant pair production of highly energetic Higgs bosons decaying to bottom quarks*, (2022), arXiv: 2205.06667 [hep-ex].

[46] ATLAS Collaboration, *Search for non-resonant Higgs boson pair production in the $bb\ell\nu\ell\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s}$ = 13 TeV*, Phys. Lett. B **801** (2020) 135145, arXiv: 1908.06765 [hep-ex].

[47] CMS Collaboration, *Search for resonant and nonresonant Higgs boson pair production in the $b\bar{b}\ell\nu\ell\nu$ final state in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **01** (2018) 054, arXiv: 1708.04188 [hep-ex].

[48] ATLAS Collaboration, *Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **04** (2019) 092, arXiv: 1811.04671 [hep-ex].

[49] ATLAS Collaboration, *Search for Higgs boson pair production in the $\gamma\gamma WW^*$ channel using pp collision data recorded at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Eur. Phys. J. C **78** (2018) 1007, arXiv: 1807.08567 [hep-ex].

[50] ATLAS Collaboration, *Search for Higgs boson pair production in the $WW^{(*)}WW^{(*)}$ decay channel using ATLAS data recorded at $\sqrt{s}$ = 13 TeV*, JHEP **05** (2019) 124, arXiv: 1811.11028 [hep-ex].

[51] CMS Collaboration, *Search for nonresonant Higgs boson pair production in the four leptons plus two b jets final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, (2022), arXiv: 2206.10657 [hep-ex].

[52] CMS Collaboration, *Search for Higgs boson pairs decaying to WWWW, WW$\tau\tau$, and $\tau\tau\tau\tau$ in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, (2022), arXiv: 2206.10268 [hep-ex].

[53] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Lett. B **800** (2020) 135103, arXiv: 1906.02025 [hep-ex].

[54] CMS Collaboration, *Combination of Searches for Higgs Boson Pair Production in Proton–Proton Collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. Lett. **122** (2019) 121803, arXiv: 1811.09689 [hep-ex].

[55] ATLAS Collaboration, *Constraining the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using* pp *collisions at* $\sqrt{s} = 13$ *TeV*, CERN-EP-2022-149, 2022.

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

[57] H. E. Haber and G. L. Kane, *The search for supersymmetry: Probing physics beyond the standard model*, Phys. Rept. **117** (1985) 75.

[58] S. Dimopoulos and H. Georgi, *Softly broken supersymmetry and SU(5)*, Nucl. Phys. B **193** (1981) 150.

[59] Z. Chacko, Y. Nomura, M. Papucci and G. Perez, *Natural little hierarchy from a partially goldstone twin Higgs*, JHEP **01** (2006) 126, arXiv: hep-ph/0510273.

[60] J. Mrazek et al., *The other natural two Higgs doublet model*, Nucl. Phys. B **853** (2011) 1, arXiv: 1105.5403 [hep-ph].

[61] L. Randall and R. Sundrum, *Large Mass Hierarchy from a Small Extra Dimension*, Phys. Rev. Lett. **83** (1999) 3370, arXiv: hep-ph/9905221.

[62] Y. Tang, *Implications of LHC searches for massive graviton*, JHEP **08** (2012) 078, arXiv: 1206.6949 [hep-ph].

[63] K. Cheung, *Phenomenology of the radion in the Randall-Sundrum scenario*, Phys. Rev. D **63** (2001) 056007, arXiv: hep-ph/0009232.

[64] N. Kumar and S. P. Martin, *LHC search for di-Higgs decays of stoponium and other scalars in events with two photons and two bottom jets*, Phys. Rev. D **90** (2014) 055007, arXiv: 1404.0996 [hep-ph].

[65] CMS Collaboration, *Search for Higgs boson pair production in the* $\gamma\gamma b\bar{b}$ *final state in* pp *collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Lett. B **788** (2019) 7, arXiv: 1806.00408 [hep-ex].

[66] ATLAS Collaboration, *Search for resonant pair production of Higgs bosons in the* $b\bar{b}b\bar{b}$ *final state using* pp *collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **105** (2022) 092002, arXiv: 2202.07288 [hep-ex].

[67] 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].

[68] CMS Collaboration, *Search for resonant pair production of Higgs bosons decaying to bottom quark-antiquark pairs in proton–proton collisions at* 13 *TeV*, JHEP **08** (2018) 152, arXiv: 1806.03548 [hep-ex].

[69] 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].

[70] ATLAS Collaboration, *Reconstruction and identification of boosted di-τ systems in a search for Higgs boson pairs using 13 TeV proton-proton collision data in ATLAS*, JHEP **11** (2020) 163, arXiv: 2007.14811 [hep-ex].

[71] 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].

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

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

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

[75] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022, 2015, URL: https://cds.cern.ch/record/2037697.

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

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

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

[79] 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.

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

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

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

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

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

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

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

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

[88] 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].

[89] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, Eur. Phys. J. C **75** (2015) 132, arXiv: 1412.7420 [hep-ph].

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

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

[92] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: https://cds.cern.ch/record/1474107.

[93] 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].

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

[95] M. Bähr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: 0803.0883 [hep-ph].

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

[97] S. Gieseke, C. Röhr and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. C **72** (2012) 2225, arXiv: 1206.0041 [hep-ph].

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

[99] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, arXiv: 1001.5034 [hep-ph].

[100] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503, arXiv: 1103.2792 [hep-ph].

[101] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, Phys. Rev. D **82** (2010) 054018, arXiv: 1005.4451 [hep-ph].

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

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

[104] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini, *Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order*, Phys. Rev. Lett. **103** (2009) 082001, arXiv: 0903.2120 [hep-ph].

[105] K. Hamilton, P. Nason, E. Re and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, JHEP **10** (2013) 222, arXiv: 1309.0017 [hep-ph].

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

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

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

[109] J. Pumplin et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **07** (2002) 012, arXiv: hep-ph/0201195 [hep-ph].

[110] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, JHEP **05** (2016) 058, arXiv: 1602.00695 [hep-ph].

[111] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, Phys. Rev. Lett. **114** (2015) 212001, arXiv: 1503.06056 [hep-ph].

[112] F. Dulat, A. Lazopoulos and B. Mistlberger, *iHixs 2 – Inclusive Higgs cross sections*, Comput. Phys. Commun. **233** (2018) 243, arXiv: 1802.00827 [hep-ph].

[113] S. Actis, G. Passarino, C. Sturm and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron colliders*, Phys. Lett. B **670** (2008) 12, arXiv: 0809.1301 [hep-ph].

[114] M. Ciccolini, A. Denner and S. Dittmaier, *Strong and Electroweak Corrections to the Production of a Higgs Boson + 2 Jets via Weak Interactions at the Large Hadron Collider*, Phys. Rev. Lett. **99** (2007) 161803, arXiv: 0707.0381 [hep-ph].

[115] M. Ciccolini, A. Denner and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC*, Phys. Rev. D **77** (2008) 013002, arXiv: 0710.4749 [hep-ph].

[116] P. Bolzoni, F. Maltoni, S.-O. Moch and M. Zaro, *Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD*, Phys. Rev. Lett. **105** (2010) 011801, arXiv: 1003.4451 [hep-ph].

[117] M. L. Ciccolini, S. Dittmaier and M. Krämer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys. Rev. D **68** (2003) 073003, arXiv: hep-ph/0306234 [hep-ph].

[118] O. Brein, A. Djouadi and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. B **579** (2004) 149, arXiv: hep-ph/0307206.

[119] G. Ferrera, M. Grazzini and F. Tramontano, *Associated Higgs-W-Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at NNLO*, Phys. Rev. Lett. **107** (2011) 152003, arXiv: 1107.1164 [hep-ph].

[120] O. Brein, R. V. Harlander, M. Wiesemann and T. Zirke,
Top-quark mediated effects in hadronic Higgs-Strahlung, Eur. Phys. J. C **72** (2012) 1868,
arXiv: 1111.0761 [hep-ph].

[121] G. Ferrera, M. Grazzini and F. Tramontano,
Higher-order QCD effects for associated WH production and decay at the LHC,
JHEP **04** (2014) 039, arXiv: 1312.1669 [hep-ph].

[122] G. Ferrera, M. Grazzini and F. Tramontano,
Associated ZH production at hadron colliders: The fully differential NNLO QCD calculation,
Phys. Lett. B **740** (2015) 51, arXiv: 1407.4747 [hep-ph].

[123] J. M. Campbell, R. K. Ellis and C. Williams, *Associated production of a Higgs boson at NNLO*,
JHEP **06** (2016) 179, arXiv: 1601.00658 [hep-ph].

[124] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak and T. J. E. Zirke,
Gluon-induced Higgs-strahlung at next-to-leading order QCD, JHEP **02** (2013) 078,
arXiv: 1211.5015 [hep-ph].

[125] B. Hespel, F. Maltoni and E. Vryonidou,
Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM,
JHEP **06** (2015) 065, arXiv: 1503.01656 [hep-ph].

[126] R. V. Harlander, A. Kulesza, V. Theeuwes and T. Zirke,
Soft gluon resummation for gluon-induced Higgs Strahlung, JHEP **11** (2014) 082,
arXiv: 1410.0217 [hep-ph].

[127] R. V. Harlander, S. Liebler and T. Zirke,
Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-doublet model, JHEP **02** (2014) 023,
arXiv: 1307.8122 [hep-ph].

[128] O. Brein, R. V. Harlander and T. J. E. Zirke, *vh@nnlo – Higgs Strahlung at hadron colliders*,
Comput. Phys. Commun. **184** (2013) 998, arXiv: 1210.5347 [hep-ph].

[129] ATLAS Collaboration,
Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2,
Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex].

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

[131] 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].

[132] 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].

[133] 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].

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

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

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

[137] ATLAS Collaboration, *Properties of jets and inputs to jet reconstruction and calibration with the ATLAS detector using proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-036, 2015, URL: https://cds.cern.ch/record/2044564.

[138] 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].

[139] 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].

[140] ATLAS Collaboration, *Forward Jet Vertex Tagging: A new technique for the identification and rejection of forward pileup jets*, ATL-PHYS-PUB-2015-034, 2015, URL: https://cds.cern.ch/record/2042098.

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

[142] 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.

[143] 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].

[144] 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.

[145] 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.

[146] ATLAS Collaboration, *Evidence for the* $H \to b\bar{b}$ *decay with the ATLAS detector*, JHEP **12** (2017) 024, arXiv: 1708.03299 [hep-ex].

[147] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in* pp *collisions at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. C **75** (2015) 303, arXiv: 1412.7086 [hep-ex].

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

[149] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, ATL-PHYS-PUB-2019-033, 2019, URL: https://cds.cern.ch/record/2688062.

[150] 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].

[151] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. B **659** (2008) 119, arXiv: 0707.1378 [hep-ph].

[152] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A new mass reconstruction technique for resonances decaying to* $\tau\tau$, Nucl. Instrum. Meth. A **654** (2011) 481, arXiv: 1012.4686 [hep-ex].

[153] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski and D. Whiteson, *Parameterized neural networks for high-energy physics*, Eur. Phys. J. C **76** (2016) 235, arXiv: 1601.07913 [hep-ex].

[154] F. Chollet et al., *Keras*, 2015, URL: https://keras.io.

[155] Martín Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015, URL: https://www.tensorflow.org/.

[156] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007, arXiv: physics/0703039 [physics.data-an].

[157] ATLAS Collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, JHEP **04** (2015) 117, arXiv: 1501.04943 [hep-ex].

[158] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, http://www.deeplearningbook.org, MIT Press, 2016.

[159] ATLAS Collaboration, *Object-based missing transverse momentum significance in the ATLAS Detector*, ATLAS-CONF-2018-038, 2018, URL: https://cds.cern.ch/record/2630948.

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

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

[162] 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. 39 p, arXiv: 2109.10627.

[163] 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.

[164] ATLAS Collaboration, *Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector*, JHEP **01** (2021) 033, arXiv: 2006.09274 [hep-ex].

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

[166] ATLAS Collaboration, *Study of heavy flavor quarks produced in association with top quark pairs at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, Phys. Rev. D **89** (2014) 072012, arXiv: 1304.6386 [hep-ex].

[167] ATLAS Collaboration, *Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **06** (2013) 084, arXiv: 1302.2929 [hep-ex].

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

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

[170] R. B. Nelsen, *An Introduction to Copulas*, Springer, New York (2006) 1.

[171] ATLAS Collaboration, *Evaluating statistical uncertainties and correlations using the bootstrap method*, ATL-PHYS-PUB-2021-011, 2021, URL: https://cds.cern.ch/record/2759945.