ATLAS CONF Note
ATLAS-CONF-2019-045
16th October 2019

Analysis of $t\bar{t}H$ and $t\bar{t}W$ production in multilepton final states with the ATLAS detector

The ATLAS Collaboration

A search for the associated production of a top-quark pair with the Higgs boson ($t\bar{t}H$) in multilepton final states is presented. The search is based on a dataset of proton–proton collisions at $\sqrt{s} = 13$ TeV recorded with the ATLAS detector at the CERN Large Hadron Collider and corresponding to an integrated luminosity of 80 fb^{-1}. Six final states, defined by the number and flavour of charged-lepton candidates, and 25 event categories are defined to simultaneously search for the $t\bar{t}H$ signal and constrain several leading backgrounds. The $t\bar{t}W$ background normalisation is left unconstrained in the statistical analysis and the resulting $t\bar{t}W$ normalisation is found to be higher than the theoretical prediction. An excess of events consistent with $t\bar{t}H$ production, over the expected background from Standard Model processes, is found with an observed significance of 1.8 standard deviations, compared to an expectation of 3.1 standard deviations. Assuming Standard Model branching fractions, the best-fit value of the $t\bar{t}H$ production cross section is $\sigma_{t\bar{t}H} = 294^{+182}_{-162}$ fb, which is consistent with the Standard Model prediction. The impact on the $t\bar{t}H$ cross section measurement of the assumptions made on the $t\bar{t}W$ background modelling is discussed.

1 Introduction

Since the discovery of the Higgs boson by the ATLAS [1] and CMS [2] collaborations, its properties have been studied using proton–proton (pp) collision data from the Large Hadron Collider (LHC). Initial measurements of its interactions with other Standard Model (SM) particles focused on measurements of its coupling to gauge bosons [3] followed, more recently, by measurements of its Yukawa interactions with the τ-lepton [4, 5] and the bottom quark [6, 7].

In the SM, the Higgs boson is predicted to couple most strongly to the top quark. The top-quark Yukawa coupling, expected to be of order unity, can be probed directly by measuring the cross section for associated production of a Higgs boson with a top-quark pair ($t\bar{t}H$). The $t\bar{t}H$ process was observed by the ATLAS and CMS collaborations where each experiment combined dedicated searches exploiting the main Higgs boson decay channels and using proton–proton collision data collected at $\sqrt{s} = 13$ TeV [8, 9].

This note reports the results of a search for $t\bar{t}H$ production at $\sqrt{s} = 13$ TeV using a dataset corresponding to an integrated luminosity of 80 fb^{-1} collected with the ATLAS detector during 2015–2017. The measurement uses six final states distinguished by the number and flavour of charged-lepton candidates (electron, muon, and hadronically decaying τ lepton), denoted l. In the following, the term light lepton, denoted ℓ, refers to either electrons or muons. Multilepton signatures are primarily sensitive to the decays $H \rightarrow WW^*$ (with at least one W decaying to leptons), $H \rightarrow \tau^+\tau^-$, and $H \rightarrow ZZ^*$ (with a subsequent decay to $ll\nu\nu$ or $llqq$). The selection avoids overlap with the ATLAS searches for $t\bar{t}H$ production with $H \rightarrow b\bar{b}$ [10], $H \rightarrow \gamma\gamma$ [11], and $H \rightarrow ZZ^* \rightarrow 4\ell$ [12] decays. The main backgrounds to the $t\bar{t}H$ signal arise from $t\bar{t}W$, $t\bar{t}(Z/\gamma^*)$, and diboson (VV) production, as well as from $t\bar{t}$ production with additional light leptons from heavy-flavour hadron decays, misidentified jets, or photon conversions (collectively referred to as "non-prompt leptons"), and other processes where the electron charge is incorrectly assigned or where jets are misidentified as τ_{had} candidates. A binned maximum-likelihood fit is performed in all final states to extract the $t\bar{t}H$ cross-section and the background normalisations. Many aspects of the analysis closely follow Ref. [13] which uses 36 fb^{-1} of data, with key differences being the treatment of the $t\bar{t}W$ background in the statistical analysis, and the background estimation procedure for non-prompt light leptons.

This note is organised as follows. Section 2 introduces the ATLAS detector. Section 3 describes the dataset and Monte Carlo (MC) simulation samples used for this analysis as well as the associated uncertainties. The reconstruction and identification of the physics objects, as well as their associated uncertainties, are discussed in Section 4. The event selection and categorisation are explained in Section 5. Section 6 describes the methods used to estimate the backgrounds and their uncertainties. The statistical model used and the results obtained are presented in Section 7. The cross-checks performed on the robustness of statistical model and the impact of $t\bar{t}W$ modelling are discussed in Section 7.1.

2 ATLAS detector

The ATLAS detector [14] at the LHC covers almost the entire solid angle around the collision point,[1] and consists of an inner tracking detector surrounded by a thin superconducting solenoid producing a 2 T axial magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large toroid magnet assemblies. The inner detector consists of a high-granularity silicon pixel detector, including the insertable B-layer [15], and a silicon microstrip tracker, together providing a precise reconstruction of tracks of charged particles in the pseudorapidity range $|\eta| < 2.5$. The inner detector also includes a transition radiation tracker that provides tracking and electron identification information for $|\eta| < 2.0$. The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic (EM) calorimetry is provided by barrel and endcap high-granularity 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. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements, respectively. The muon spectrometer measures the trajectories of muons with $|\eta| < 2.7$ using multiple layers of high-precision tracking chambers located in a toroidal field of approximately 0.5 T and 1 T in the central and endcap regions of ATLAS, respectively. The muon spectrometer is also instrumented with separate trigger chambers covering $|\eta| < 2.4$. A two-level trigger system [16], consisting of a hardware-based Level-1 trigger followed by a software-based High-Level Trigger (HLT), is used to reduce the event rate to a maximum of around 1 kHz for offline storage.

3 Data and Monte Carlo samples

A dataset of pp collisions at $\sqrt{s} = 13$ TeV collected by the ATLAS experiment during 2015–2017 and corresponding to an integrated luminosity of 79.9 ± 1.6 fb^{-1} is used. The uncertainty in the integrated luminosity [17] is obtained using the LUCID-2 detector [18] for the primary luminosity measurements. The number of additional pp interactions per bunch crossing (pileup) in this dataset ranges from about 8 to 70 interactions, with an average of 34. Only events recorded under stable beam conditions and for which all detector subsystems were known to be in a good operating condition are used. The trigger requirements are discussed in Section 5.

Monte Carlo simulation samples were produced for the different signal and background processes using the configurations shown in Table 1, with the samples used to estimate the systematic uncertainties indicated in parentheses. Pileup is modelled using events from minimum-bias interactions generated with Pythia 8.186 [19] with the A3 set of tuned parameters [20] (referred to as the “tune”), and overlaid onto the simulated hard-scatter events according to the luminosity profile of the recorded data. The generated events were processed through a simulation [21] of the ATLAS detector geometry and response using Geant4 [22], and through the same reconstruction software as the data. Corrections were applied to the

---

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector. The x-axis points from the IP to the centre of the LHC ring, the y-axis points upward, and the z-axis coincides with the axis of the beam pipe. Cylindrical coordinates (r,ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

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

3.1 $t\bar{t}H$ signal

The nominal sample used to model the $t\bar{t}H$ signal was generated using the next-to-leading-order (NLO) generator POWHEG-BOX v2 [23, 24] with the NNPDF3.0 NLO [25] parton distribution function (PDF) set. The renormalisation and factorisation scales, μ_R and μ_F, are set equal to the geometric mean of the transverse energies of the top quark, the antitop quark, and the Higgs boson. The POWHEG-BOX model parameter h_{damp}, which controls matrix element to parton shower matching and effectively regulates the high-p_{T} radiation, is set to $1.5 \times (2m_t + m_H)/2 = 352.5$ GeV. The parton shower (PS) and hadronisation were modelled using PYTHIA 8.2 [26] and the A14 tune [27], and with Higgs decay branching ratios calculated using HDECAY [28, 29]. The simulated sample is normalised using a cross section of 507^{+35}_{-50} fb, which is computed at NLO in quantum chromodynamics (QCD) with the leading NLO electroweak corrections (i.e. $O(\alpha_s^2\alpha^2)$) [30–35]. Uncertainties include $^{+5.8\%}_{-9.2\%}$ estimated by varying the QCD factorisation and renormalisation scales and ±3.6% due to uncertainties on the PDFs and the strong coupling α_s. Uncertainties affecting the modelling of the acceptance and event kinematics include variations in the QCD factorisation and renormalisation scales, the choice of parton shower and hadronisation model, the modelling of initial state radiation (ISR), and PDF uncertainties. The theoretical uncertainties due to the QCD scale choice are estimated varying the renormalisation and factorisation scale independently by a factor of 0.5 and 2.0 with respect to the central value. The largest variation is obtained when varying both scales simultaneously in the same direction. Uncertainties due to the choice of parton shower and hadronisation model are estimated by comparing the nominal prediction with that obtained using an alternative sample generated with POWHEG-BOX interfaced to HERWIG7 [36]. The uncertainty associated with the modelling of ISR is estimated by considering the Var3c A14 tune variation [27], which correspond to a variation of α_s in the A14 tune. The uncertainty due to the choice of PDF set is evaluated using the PDF4LHC15 prescription [37], utilising 30 eigenvector shifts derived from fits to multiple NNLO PDF sets. Finally, the uncertainties associated with the predicted Higgs-boson branching ratios [28] are also considered.

3.2 $t\bar{t}W$ background

The simulated sample for $t\bar{t}W$ production was generated using the SHERPA 2.2.1 [41] generator with the NNPDF3.0 NLO PDF set. The matrix element (ME) was calculated for up to one additional parton at NLO and up to two partons at LO using COMIX [43] and OPENLOOPS [42], and merged with the SHERPA parton shower [44] using the MEPS@NLO prescription [45] and a merging scale of 30 GeV. The choice of renormalisation and factorisation scales is $\mu_R = \mu_F = H_{\mathrm{T}}/2$, where H_{T} is defined as the scalar sum of the transverse masses $\sqrt{p_{\mathrm{T}}^2 + m^2}$ of all final state particles.

The nominal cross section for the simulated $t\bar{t}W$ sample is 601 ± 76 fb, which is computed at NLO in QCD with the leading NLO electroweak corrections (i.e. $O(\alpha_s^2\alpha^2)$) [28, 56, 57]. The uncertainties from QCD scale and PDF+α_{S} variations are ±12% and ±4%, respectively. This value of the cross section was used in the previous result [13]. In this analysis, additional scaling factors to this cross section are applied to account for missing QCD and electroweak corrections. Previous work [38] had shown that NLO QCD corrections to

Table 1: The configurations used for the event generation of signal and background processes. The samples used to estimate the systematic uncertainties are indicated in parentheses. V refers to production of an electroweak boson (W or Z/γ^*). The matrix element (ME) order refers to the order in the strong coupling constant of the perturbative calculation. If only one parton distribution function (PDF) is shown, the same one is used for both the ME and parton shower generators; if two are shown, the first is used for the ME calculation and the second for the parton shower. Tune refers to the underlying-event tune of the parton shower generator. MG5_aMC refers to MadGraph5_aMC@NLO 2.2.X or 2.3.X [38]; Pythia 6 refers to version 6.427 [39]; Pythia 8 refers to version 8.2 [26]; Herwig++ refers to version 2.7 [40]; Herwig7 refers to version 7.0.4 [36]; MePs@Nlo refers to the method used in Sherpa [41–45] to match the matrix element to the parton shower. Samples using Pythia 6 or Pythia 8 have heavy flavour hadron decays modelled by EvtGen 1.2.0 [46]. All samples include leading-logarithm photon emission, either modelled by the parton shower generator or by PHOTOS [47]. The masses of the top quark and SM Higgs boson were set to 172.5 GeV and 125 GeV.

Process	Generator	ME order	Parton shower	PDF	Tune
$t\bar{t}H$	Powheg-BOX [23, 24]	NLO	Pythia 8	NNPDF3.0 NLO [25]/ NNPDF2.3 LO [48]	A14
	(Powheg-BOX)	(NLO)	(Herwig7)	(NNPDF3.0 NLO/ MMHT2014 LO [49])	(H7-UE-MMHT)
$tHqb$	MG5_aMC	LO	Pythia 8	CT10 [50]	A14
tHW	MG5_aMC	NLO	Herwig++	CT10/ CTEQ6L1 [51, 52]	UE-EE-5
$t\bar{t}W$	Sherpa 2.2.1	MePs@Nlo	Sherpa	NNPDF3.0 NNLO	Sherpa default
	(MG5_aMC)	(NLO)	(Pythia 8)	(NNPDF3.0 NLO/ NNPDF2.3 LO)	(A14)
$t\bar{t}(Z/\gamma^*)$	MG5_aMC	NLO	Pythia 8	NNPDF3.0 NLO/ NNPDF2.3 LO	A14
	(Sherpa 2.2.0)	(LO multileg)	(Sherpa)	(NNPDF3.0 NLO)	(Sherpa default)
$t\bar{t} \to W^+bW^-\bar{b}l^+l^-$	MG5_aMC	LO	Pythia 8	NNPDF3.0 LO	A14
tZ	MG5_aMC	LO	Pythia 6	CTEQ6L1	Perugia2012
tWZ	MG5_aMC	NLO	Pythia 8	NNPDF2.3 LO	A14
$t\bar{t}t$, $t\bar{t}t\bar{t}$	MG5_aMC	LO	Pythia 8	NNPDF2.3 LO	A14
$t\bar{t}W^+W^-$	MG5_aMC	LO	Pythia 8	NNPDF2.3 LO	A14
$t\bar{t}$	Powheg-BOX	NLO	Pythia 8	NNPDF3.0 NLO/ NNPDF2.3 LO	A14
Single top (t-, Wt-, s-channel)	Powheg-BOX [53–55]	NLO	Pythia 8	NNPDF3.0 NLO/ NNPDF2.3 LO	A14
VV, $qqVV$, VVV	Sherpa 2.2.2	MePs@Nlo	Sherpa	NNPDF3.0 NNLO	Sherpa default
$Z \to l^+l^-$	Sherpa 2.2.1	MePs@Nlo	Sherpa	NNPDF3.0 NLO	Sherpa default

$t\bar{t}W$+1-jet production can be large. An inclusive scaling factor of 1.11 has been estimated using dedicated samples generated with Sherpa 2.2.5 using the MePs@Nlo prescription, and cross-checked with the NLO generator MadGraph5_aMC@NLO 2.2.1 [38] using the FxFx prescription [58]. In addition, recent work [59] has shown that sub-leading NLO electroweak corrections for $t\bar{t}W$ production are larger than expected, primarily because of the large NLO$_3$ term driven by the $t\bar{t}W$+1-jet diagrams with a Higgs boson exchanged in the t-channel. The corresponding estimated scaling factor is 1.09. Therefore, after applying these two scaling factors, the inclusive cross section, used to normalise the $t\bar{t}W$ sample becomes 727 ± 92 fb.[2] As the $t\bar{t}W$ normalisation will be estimated in data (see Section 7), this uncertainty is not

---

[2] The theoretical uncertainties are not revised based on the additional corrections considered, but are scaled proportionally to the

included as a systematic uncertainty. This cross-section will be referred to as the "updated $t\bar{t}W$ theoretical cross-section".

Systematic uncertainties due to missing higher-order QCD corrections are estimated by varying the factorisation and renormalisation scales in the nominal sample simultaneously by a factor of 0.5 and 2.0 with respect to the central value. Uncertainties associated with the modelling of additional QCD radiation are estimated by comparing the nominal $t\bar{t}W$ prediction with that of an alternative sample that was generated at NLO (thus at lower order than for the nominal sample) with the MadGraph5_aMC@NLO 2.2.1 generator using the same scale choice and PDF set as for the nominal sample, and interfaced to Pythia 8.2 in combination with the A14 tune. This alternative sample was used as nominal sample in the previous result [13]. Finally, the uncertainty due to the choice of PDF set is evaluated using the PDF4LHC15 prescription.

3.3 Other backgrounds

The simulated samples for $t\bar{t}(Z/\gamma^*)$, VV, and $t\bar{t}$ production follow Ref. [60–62]. For the $t\bar{t}(Z/\gamma^*)$ sample, the inclusive $t\bar{t}l^+l^-$ ME is computed at NLO, including off-shell Z and γ^* contributions with $m(l^+l^-) > 1$ GeV. A dedicated $t\bar{t}$ sample including rare $t \to Wb\gamma^*(\to l^+l^-)$ radiative decays, $t\bar{t} \to W^+bW^-\bar{b}l^+l^-$, is generated using a LO ME and requiring $m(l^+l^-) > 1$ GeV. In this sample the photon can be radiated from the top quark, the W boson, or the b-quark. Both the $t\bar{t}(Z/\gamma^*)$ and $t\bar{t} \to W^+bW^-\bar{b}l^+l^-$ samples are combined and together form the "$t\bar{t}(Z/\gamma^*)$ (high mass)" sample. The contribution from internal photon conversions ($\gamma^* \to l^+l^-$) with $m(l^+l^-) < 1$ GeV are modelled by QED multiphoton radiation via the PS in an inclusive $t\bar{t}$ sample and is referred to as "$t\bar{t}\gamma^*$ (low mass)". Further details on the generation settings for these samples can be found in Table 1. Care has been taken to avoid both double-counting of contributions and uncovered regions of phase space when combining the different simulated samples. The cross section for inclusive $t\bar{t}l^+l^-$ production, with $m(l^+l^-) > 1$ GeV, is 162 ± 21 fb. This cross section is computed at NLO in QCD and electroweak couplings [28, 38, 57]. The uncertainties from QCD scale and PDF+α_S variations are ±12% and ±4% respectively. The LO cross section from the $t\bar{t} \to W^+bW^-\bar{b}l^+l^-$ sample is scaled by a factor of 1.54, based on comparisons between the NNLO+NLL [63–67] and LO cross sections for $t\bar{t}$ production, and assigned a 50% normalisation uncertainty. Uncertainties affecting the modelling of the acceptance and event kinematics for the $t\bar{t}l^+l^-$ sample include the same QCD scale and tune variations as considered for the $t\bar{t}H$ sample, PDF variations using the PDF4LHC15 prescription, and the comparison to an alternative LO multileg sample (see Table 1). In the case of the inclusive $t\bar{t}$ sample, the uncertainty in the modelling of additional QCD radiation is assessed with two alternative samples generated with settings that increase or decrease the amount of radiation [68]. Diboson backgrounds are normalised using the cross sections computed by Sherpa 2.2.2, and a 50% normalisation uncertainty is assigned and treated as uncorrelated among different subprocesses (WZ+light-jets, WZ+$\geq 1c$, WZ+$\geq 1b$, and ZZ+jets). Rare background contributions (tZ, $t\bar{t}t\bar{t}$, $ttWW$, WtZ, VVV, $tt\bar{t}$, $tHjb$ and WtH) are normalised using their NLO theoretical cross sections, and assigned a 50% normalisation uncertainty.

4 Object identification and uncertainties

Tracks are reconstructed in the inner detector. Interaction vertices from the pp collisions are reconstructed from at least two tracks with transverse momentum (p_T) larger than 400 MeV that are consistent with

---

scaling factors applied.

originating from the beam collision region in the x–y plane. If more than one primary vertex candidate is found, the candidate whose associated tracks form the largest sum of squared p_{T} [69] is selected as the hard-scatter primary vertex.

Muon candidates [70] are reconstructed by matching track segments in different layers of the muon spectrometer to tracks found in the inner detector; the resulting muon candidates are refitted using the complete track information from both detector systems. They are required to satisfy $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$ and to pass loose identification requirements [70].

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter that are associated with inner detector tracks [71]. They are required to have a transverse momentum $p_{\mathrm{T}} > 10$ GeV and $|\eta_{\mathrm{cluster}}| < 2.47$. The transition region between the barrel and endcap calorimeters, $1.37 < |\eta_{\mathrm{cluster}}| < 1.52$, is excluded. The associated track of an electron candidate is required to have at least four hits in the silicon layers, one of them being in the innermost layer (or the next-to-innermost layer if the innermost layer is non-operational), and no association with a vertex from a reconstructed photon conversion [72] in the detector material (denoted “material conversion"). Loose and tight electron identification operating points are used [73], based on a likelihood discriminant employing calorimeter, tracking and combined variables that provide separation between electrons and jets. To further suppress material conversions, additional requirements on the associated track p_{T} and on the ratio of the electron’s calorimeter energy and the track momentum are applied to tight electrons.

Calorimeter- and track-based isolation criteria are applied to electrons and muons that result in a 99% efficiency in $Z \rightarrow \ell^+\ell^-$ events. Calorimetric isolation uses the scalar sum of transverse energies of clusters within $\Delta R = 0.3$ around the light-lepton candidate excluding the electron candidate’s cluster. Track isolation uses the sum of transverse momenta of tracks with $p_{\mathrm{T}} > 1$ GeV consistent with originating from the primary vertex, excluding the light-lepton candidate’s track, within $\Delta R = \min(0.3, 10\ \mathrm{GeV}/p_{\mathrm{T}}(\ell))$.

Non-prompt leptons, i.e. those produced at some displacement from the interaction point, are further rejected using a boosted decision tree (BDT) discriminant based on isolation and b-tagging variables, referred to as the non-prompt lepton BDT [13]. The efficiency at the chosen working point for muons (electrons) satisfying the calorimeter- and track-based isolation criteria is about 80% (65%) for $p_{\mathrm{T}} \sim 20$ GeV and reaches a plateau of 95% (90%) at $p_{\mathrm{T}} \sim 45$ GeV. The corresponding rejection factor against leptons from the decay of b-hadrons is about 3.5 (10), after removing ambiguities between overlapping reconstructed objects.

Tight electrons with incorrect charge assignment are rejected using a BDT discriminant based on calorimeter and tracking quantities [71]. An efficiency of 95% for electrons with correct charge assignment is obtained with a rejection factor of ~17 for electrons with incorrect charge assignment.

The resulting tight electron candidates are further split into three categories: “material conversions", “internal conversions", and "very tight". Material conversion candidates have a reconstructed displaced vertex with radius $r > 20$ mm that includes the track associated with the electron. The invariant mass of the associated track and the closest (in $\Delta\eta$) opposite-charge track reconstructed in the silicon detector, calculated at the conversion vertex, is required to be < 100 MeV. Internal conversion candidates are required to fail the requirements for material conversions, and the di-track invariant mass, this time calculated at the primary vertex, is also required to be < 100 MeV. Very tight electron candidates are tight electrons that fail the internal conversion and material conversion requirements, and have $|\eta| < 2$. The latter requirement rejects a small fraction of electrons with a large charge misidentification rate because of the limited number of hits used in the track reconstruction.

Table 2: The requirements applied to select loose (L), loose and minimally-isolated (L*), tight (T) and very tight (T*) light leptons.

	e				μ		
	L	L*	T	T*	L	L*	T/T*
Identification	Loose		Tight		Loose		Medium
Isolation	No	Yes			No	Yes	
Non-prompt lepton veto	No		Yes		No		Yes
Charge misidentification veto	No		Yes		N/A		
Material/internal conversion veto	No			Yes	N/A		
Lepton $\lvert\eta\rvert$	< 2.47			< 2	< 2.5		
$\lvert d_0\rvert/\sigma_{d_0}$	< 5				< 3		
$\lvert z_0 \sin\theta\rvert$	< 0.5 mm						

Requirements on the transverse (d_0) and longitudinal (z_0) impact parameters are used to select leptons originating from the primary vertex. Electrons and muons are required to satisfy $|z_0 \sin\theta| < 0.5$ mm. Electrons (muons) are required to have the transverse impact parameter significance $|d_0|/\sigma_{d_0} < 5(3)$. Muons are required to be separated by $\Delta R > \min(0.4, 0.04 + (10\text{ GeV})/p_{\text{T},\mu})$ from any selected jets (explained below). If two electrons are closer than $\Delta R = 0.1$, only the one with the higher p_T is considered. An electron lying within $\Delta R = 0.1$ of a selected muon is rejected.

The different light-lepton selections used in the analysis are summarised in Table 2. Uncertainties in light-lepton reconstruction, identification, isolation, and trigger efficiencies are taken into account, but have a negligible impact in the analysis.

Jets are reconstructed from clusters built from energy deposits in the calorimeters [74, 75], using the anti-k_t algorithm with a radius parameter $R = 0.4$ [76, 77]. They are calibrated using simulation with corrections obtained from in situ techniques [78]. Jets are required to satisfy $p_\text{T} > 25$ GeV and $|\eta| < 2.5$. A Jet Vertex Tagger (JVT) is used to remove jets associated to pileup vertices with $p_\text{T} < 60$ GeV and $|\eta| < 2.4$ [79]. Any jets within $\Delta R = 0.3$ of a selected electron or a hadronically decaying τ-lepton candidate are rejected. The uncertainty in the jet energy scale [80, 81] is decomposed into uncorrelated components. The total jet energy scale uncertainty varies from 1% to 2% depending on the jet p_T. Differences in the fractions of jets containing quarks and gluons between physics processes are accounted for in the jet energy scale.

Jets containing b-hadrons are identified (b-tagged) via a multivariate discriminant (MV2c10) combining information from track impact parameters and secondary vertices [82–84]. A working point with an average efficiency in $t\bar{t}$ events of 70% for b-quark jets and rejection factors against light-quark/gluon jets, c-quark jets and hadronically decaying τ leptons of 380, 12 and 55, respectively, is used. Correction factors derived from data are applied to the simulated samples [83]. The uncertainties in the b-tagging efficiencies [82] are also decomposed into uncorrelated components. The approximate relative size of the b-tagging efficiency uncertainty is 2% for b-jets, 10% for c-jets and τ-leptons, and 30% for light jets.

Hadronically decaying τ-lepton candidates (τ_had) are reconstructed from clusters in the calorimeters and associated inner detector tracks [85]. They are required to have either one or three associated tracks, with a total charge of ±1. Candidates must have transverse momentum $p_\text{T} > 25$ GeV and $|\eta| < 2.5$, excluding the EM calorimeter's transition region, and to originate from the primary vertex. A BDT discriminant using calorimeter- and tracking-based variables is used to identify τ_{had} candidates and reject jet backgrounds [86]. The medium working point has a target efficiency of 55% (40%) for one- (three-)prong τ_{had} decays, while the tight working point has an efficiency of 40% (30%) for one- (three-)prong τ_{had} decays. Electrons that are reconstructed as one-prong τ_{had} candidates are removed using a BDT with an efficiency (rejection factor) of 95% (30–100%) for real (fake) τ_{had} candidates depending on the transverse momentum. Additionally, τ_{had} candidates are required to be separated by $\Delta R > 0.2$ from any selected electron or muon candidates. Any τ_{had} candidate that is also b-tagged is rejected. The uncertainty in the identification efficiency for τ_{had} candidates is ~6% [86].

The missing transverse momentum $\vec{p}_{\text{T}}^{\text{miss}}$ (with magnitude $E_{\text{T}}^{\text{miss}}$) is defined as the negative vector sum of the p_{T} of all selected and calibrated objects in the event, including a term to account for momentum from soft particles in the event which are not associated with any of the selected objects [87]. This soft term is calculated from inner-detector tracks matched to the selected primary vertex to make it more resilient to contamination from pileup interactions.

5 Event selection and categorisation

Six final states, termed channels, are analysed, categorised by the number and flavour of loose lepton candidates. Certain channels are further split into categories to gain significance. The selection criteria are orthogonal such that each event only contributes to a single channel.

The six channels are:

- 2ℓSS: two same-charge light leptons and no hadronically decaying τ-lepton candidates;
- 3ℓ: three light leptons and no hadronically decaying τ-lepton candidates;
- 4ℓ: four light leptons;
- $1\ell 2\tau_{\text{had}}$: one light lepton and two opposite-charge hadronically decaying τ-lepton candidates;
- 2ℓSS$1\tau_{\text{had}}$: two same-charge light leptons and one hadronically decaying τ-lepton candidate;
- $3\ell 1\tau_{\text{had}}$: three light leptons and one hadronically decaying τ-lepton candidate.

Events in channels containing at least two light leptons are required to have been selected by dilepton triggers. For the dielectron triggers the p_{T} thresholds on the two electrons were 12 GeV in 2015, 17 GeV in 2016, and 24 GeV in 2017, while for the dimuon triggers the p_{T} thresholds on the leading (sub-leading) muon were 18 GeV (8 GeV) in 2015, and 22 GeV (8 GeV) in 2016 and 2017. For the electron+muon triggers, the p_{T} thresholds on the electron (muon) were 17 GeV (14 GeV) for all datasets. Events in the $1\ell 2\tau_{\text{had}}$ channel are required to have been selected by a single-electron (single-muon) trigger with $p_{\text{T}} >$ 24 (20) GeV in 2015, while for 2016 and 2017, the lepton p_{T} threshold was raised to 26 GeV. Selected electrons or muons are required to match, with $\Delta R < 0.15$, the corresponding leptons reconstructed by the trigger and to have a p_{T} exceeding the trigger p_{T} threshold by 1 GeV or 2 GeV (depending on the lepton trigger and data-taking conditions). The trigger requirement has an efficiency of 80% to 99%, depending on the channel and the dataset, for signal events passing the final selections.

The selection requirements for each channel are summarised in Table 3. The separation of the $t\bar{t}H$ signal from the background is achieved using multivariate techniques in the 2ℓSS, 3ℓ and $1\ell 2\tau_{\text{had}}$ channels, and additional selection criteria in the 4ℓ channel. No further event selection is applied in the 2ℓSS$1\tau_{\text{had}}$ and

Table 3: Offline selection criteria applied to the channels. The common selection criteria for all channels are listed in the first line under the title “Common”. Same-charge (opposite-charge) lepton pairs are also referred to as same-sign (opposite-sign) with abbreviation SS (OS). Same-flavour (SF), OS lepton pairs are referred to as SFOS pairs. In the categories for conversions the selection requirements on one of the leptons are loosened as discussed in Section 4.

Channel	Selection criteria
Common	$N_{\text{jets}} \geq 2$ and $N_{b\text{-jets}} \geq 1$
2ℓSS	Two same-charge (SS) very tight (T*) leptons, $p_{\text{T}} > 20$ GeV No τ_{had} candidates $m(\ell^+\ell^-) > 12$ GeV for all SF pairs **13 categories**: enriched with $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}$, mat. conv, int. conv., split by lepton flavour, charge, jet and b-jet multiplicity
3ℓ	Three loose (L) leptons with $p_{\text{T}} > 10$ GeV; sum of light-lepton charges = ±1 Two SS very tight (T*) leptons, $p_{\text{T}} > 15$ GeV One OS (w.r.t the SS pair) loose-isolated (L*) lepton, $p_{\text{T}} > 10$ GeV No τ_{had} candidates $m(\ell^+\ell^-) > 12$ GeV and $\lvert m(\ell^+\ell^-) - 91.2\text{ GeV}\rvert > 10$ GeV for all SFOS pairs $\lvert m(3\ell) - 91.2\text{ GeV}\rvert > 10$ GeV **7 categories**: enriched with $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, VV, $t\bar{t}$, mat. conv, int. conv
4ℓ	Four loose-isolated (L*) leptons; sum of light lepton charges = 0 $m(\ell^+\ell^-) > 12$ GeV and $\lvert m(\ell^+\ell^-) - 91.2\text{ GeV}\rvert > 10$ GeV for all SFOS pairs $m(4\ell) < 115$ GeV or $m(4\ell) > 130$ GeV **2 categories**: Zenr (Z-enriched;1 or 2 SFOS pairs) or Zdep (Z-depleted; 0 SFOS pairs)
$1\ell 2\tau_{\text{had}}$	One tight (T) lepton, $p_{\text{T}} > 27$ GeV Two OS τ_{had} candidates At least one tight τ_{had} candidate $N_{\text{jets}} \geq 3$
$2\ell\text{SS}1\tau_{\text{had}}$	2ℓSS selection, except: One medium τ_{had} candidate $N_{\text{jets}} \geq 4$
$3\ell 1\tau_{\text{had}}$	3ℓ selection, except: One medium τ_{had} candidate, of opposite charge to the total charge of the light leptons Two SS tight (T) leptons

$3\ell 1\tau_{\text{had}}$ channels. The categories used in each channel are discussed here. Further details about how some of these categories are used to estimate the background (also denoted "control regions") are provided in Section 6. Section 7 discusses the statistical model and the distributions exploited by the fit in each category. The TMVA package [88] is used to train BDTs in the 2ℓSS and $1\ell 2\tau_{\text{had}}$ channels, while XGBoost [89] is used in the 3ℓ channel. Depending on the channel, between 7 and 26 variables are used as input to the BDTs.

In the 2ℓSS channel, five event categories are defined for the background determination. Two of these categories are defined by requiring the presence of an electron from either an internal conversion (2ℓIntC) or a material conversion (2ℓMatC) as defined in Section 4. The remaining three categories are required to contain either two or three reconstructed jets and are referred to as "low jet multiplicity (LJ)" categories. Events in these three categories are grouped according to the flavour of the sub-leading lepton and those with a sub-leading electron are split into those containing one or two b-jets (2ℓLJ(e1), 2ℓLJ(e2), and 2ℓLJ(μ)). The categorisation according to the flavour of the sub-leading lepton is motivated by the fact that this lepton is more likely to be non-prompt. Two independent BDTs are trained on events with four or more jets: one to separate $t\bar{t}H$ from $t\bar{t}$ (i.e. non-prompt background), and the other to separate $t\bar{t}H$ from $t\bar{t}V$, which is dominated by $t\bar{t}W$. Categories enriched in $t\bar{t}H$, $t\bar{t}W$ and $t\bar{t}$ events are defined by two dimensional cuts on the BDT outputs. The $t\bar{t}$ category is split according to the flavour of the sub-leading leptons. Events with four or more jets are further split into categories according to the lepton charge to exploit the charge asymmetry in the $t\bar{t}W$ background. This results in six additional background categories (2ℓttW+, 2ℓttW-, 2ℓtt(e)+, 2ℓtt(e)-, 2ℓtt(μ)+, and 2ℓtt(μ)-) and two categories for the signal (2ℓttH+, 2ℓttH-). These eight categories are referred to as "high jet multiplicity (HJ)" categories. In total, there are 13 categories in the 2ℓSS channel.

The 3ℓ channel uses a five-dimensional multinomial BDT with the following five classification targets: $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}$ and diboson. The output discriminants are mapped into five categories to maximise the signal significance (3ℓttH, 3ℓttW, 3ℓttZ, 3ℓtt, and 3ℓVV). As for the 2ℓSS channel, two additional categories are defined for conversions (3ℓIntC and 3ℓMatC) by loosening the requirements on one electron following Section 4. In total seven categories are defined.

Cross-check analyses using a cut-and-count categorisation were also developed for the most sensitive 2ℓSS and 3ℓ channels. These follow the same basic event selection and categories are defined using selection criteria based on the most highly ranked BDT variables including jet multiplicity, b-tagged jet multiplicity and lepton flavour.

In the 4ℓ channel, two categories are defined according to the presence (4ℓZenr) or absence (4ℓZdep) of a same-flavour, opposite-charge lepton pair. The purity of the Z-enriched region is improved by applying a requirement on a BDT trained to discriminate the $t\bar{t}H$ signal from the irreducible background arising from $t\bar{t}Z$ and ZZ production.

In the $1\ell 2\tau_{\text{had}}$ channel, three categories are defined based on a BDT trained to discriminate the $t\bar{t}H$ signal from the $t\bar{t}$ background.

In total, 877 events are selected in the data, distributed over the 25 event categories considered in the analysis, which have signal-to-background (S/B) ratios ranging from 0.3% to 104%. The total expected number of reconstructed SM $t\bar{t}H$ events summed over all categories is 170, corresponding to 0.42% of all produced $t\bar{t}H$ events. In the 2ℓSS$1\tau_{\text{had}}$ and $3\ell 1\tau_{\text{had}}$ channels, about 20% of the selected $t\bar{t}H$ events have a non-prompt lepton or fake τ_{had} candidate. These events are treated as signal and are corrected using the dedicated normalisation factors for non-prompt leptons and fake τ_{had} candidates discussed in Section 6.2.

6 Background estimation

Backgrounds are categorised into irreducible and reducible backgrounds. Irreducible backgrounds (Section 6.1) have all selected leptons as prompt, i.e. produced in W/Z boson decays, in leptonic τ-lepton decays, or internal conversions. Reducible backgrounds have prompt leptons with misassigned charge (labeled as "QMisID"; Section 6.2.1), at least one non-prompt light lepton (Section 6.2.2), or jets misidentified as τ_{had} candidates (fake τ_{had}; Section 6.2.3). The QMisID and fake τ_{had} backgrounds are estimated using data-driven techniques, and all other backgrounds are estimated using the simulation. However, the yields of some simulated backgrounds, including the key $t\bar{t}W$ and non-prompt-lepton backgrounds, are adjusted via normalisation factors that are determined simultaneously with the $t\bar{t}H$ cross section by performing a likelihood fit to the data across all categories as discussed in Section 7.

6.1 Irreducible backgrounds

Background contributions with prompt leptons originate from a wide range of physics processes with the relative importance of individual processes varying by channel. The main irreducible backgrounds originate from $t\bar{t}W$ and $t\bar{t}(Z/\gamma^{*})$ production, followed by VV (in particular WZ) production, and have final states and kinematic properties similar to the $t\bar{t}H$ signal. Smaller contributions originate from the following rare processes: tZ, tW, WtZ, $t\bar{t}WW$, VVV, $t\bar{t}t$, and $t\bar{t}t\bar{t}$ production. The associated production of single top quarks with a Higgs boson, tH, contributes at most 2% in any signal region and other Higgs boson production mechanisms contribute negligibly (<0.2%) in any signal region; therefore, they are treated as background processes and fixed to the SM predictions. Backgrounds with prompt leptons are estimated from simulation using the samples described in Section 3, which also discusses the systematic uncertainties in the modelling of these processes.

6.1.1 $t\bar{t}W$ background

The $t\bar{t}W$ background represents the dominant background particularly in the 2ℓSS and 3ℓ channels across multiple event categories, which span a wide range of kinematic regimes. Despite the use of the state-of-art simulations, the accurate modelling of additional QCD radiation in $t\bar{t}W$ production remains challenging. Categories sensitive to the $t\bar{t}W$ background have been introduced to the analysis to study and constrain this background. The jet multiplicity distributions in the 2ℓSS and 3ℓ channels after event selection are shown in Figure 1. Disagreements between the data and the prefit prediction from the simulation are observed. To minimise the dependence of the $t\bar{t}H$ signal extraction on the $t\bar{t}W$ prediction, three independent normalisation factors for the $t\bar{t}W$ background are considered in the likelihood fit: two corresponding to the LJ and HJ categories of the 2ℓSS channel, and one corresponding to the 3ℓ channel categories. The measured normalisation factors are: $\hat{\lambda}^{2\ell\mathrm{LJ}}_{t\bar{t}W} = 1.56^{+0.30}_{-0.28}$, $\hat{\lambda}^{2\ell\mathrm{HJ}}_{t\bar{t}W} = 1.26^{+0.19}_{-0.18}$, and $\hat{\lambda}^{3\ell}_{t\bar{t}W} = 1.68^{+0.30}_{-0.28}$. The agreement is improved after the application of the background corrections resulting from the likelihood fit, in particular the above $t\bar{t}W$ normalisation factors. Additional uncertainties associated with the modelling of the b-jet multiplicity and W-boson charge asymmetry in the $t\bar{t}W$ background are introduced to account for observed discrepancies in the shape of these distributions between data and pre-fit background predictions in the 2ℓSS and 3ℓ channels (see Figure 2). The W-boson charge asymmetry is studied via the distribution of the sign of the sum of lepton charges (referred to as "total charge"). These uncertainties are constructed to affect only the shape of the b-jet multiplicity and total charge distributions, thus preserving the normalisation of the $t\bar{t}W$ background after event selection. The uncertainty associated with the modelling of the b-jet multiplicity distribution is ±25% (∓35%) for events with exactly one (at least two) b-jets. The uncertainty associated with the modelling of the total charge distribution is ±20% (∓35%) for events with positive (negative) total charge. These additional uncertainties are treated as uncorrelated between the 2ℓSS and 3ℓ channels and are referred to as "extrapolation" uncertainties. In total, there are 41 uncertainties that are included in the statistical model to describe the $t\bar{t}W$ background. Further details of the impact of the $t\bar{t}W$ model is provided in Section 7.

6.1.2 Other irreducible backgrounds

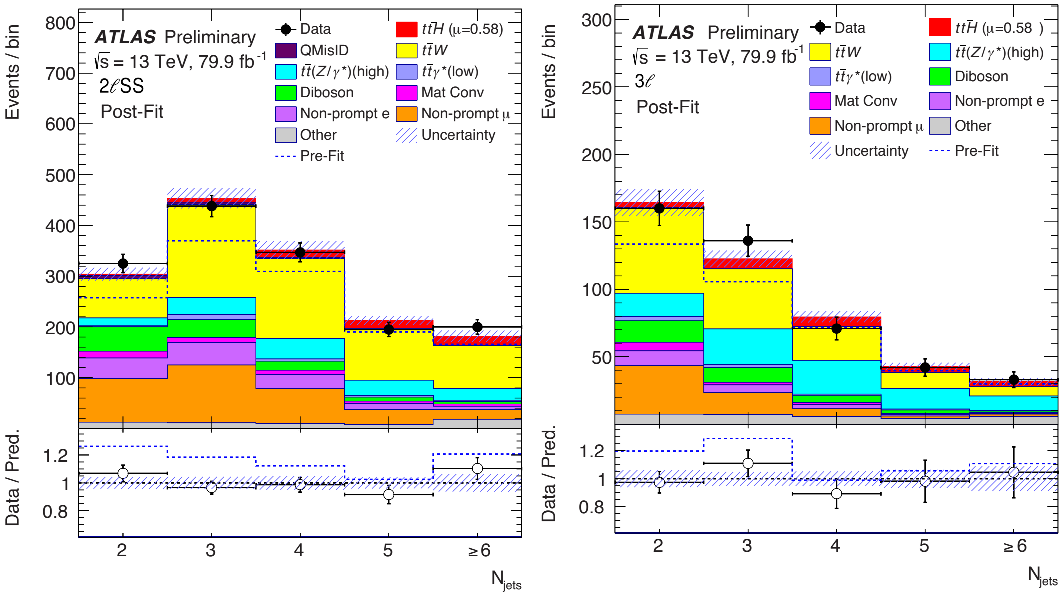
The total yields in the 3ℓttZ and 3ℓVV control regions are used in the likelihood fit to improve the estimation of the background contribution from the $t\bar{t}(Z/\gamma^*)$ and VV processes. The rate of the background from internal conversions with $m(e^+e^-) < 1$ GeV is estimated using the two dedicated control regions (2ℓIntC and 3ℓIntC). The total yield in each category is used in the likelihood fit to determine the following normalisation factor: $\hat{\lambda}_e^{\mathrm{IntC}} = 0.83 \pm 0.32$, where the uncertainty is dominated by the statistical uncertainty. The normalisation of the internal conversion background is validated by comparing data and scaled simulation in a validation region enhanced in $Z \to \mu^+\mu^-\gamma^*(\to e^+e^-)$ candidate events, defined by requiring two opposite-charge muons and one electron that satisfies the internal conversion requirements. The level of agreement found between observed and predicted yields is within 25% (see first bin of distribution in Figure 3(a)), which is assigned as a systematic uncertainty associated with the extrapolation of the estimate from the internal conversion control regions to the other event categories.

6.2 Reducible backgrounds

6.2.1 Charge misassignment

Backgrounds with leptons with the charge incorrectly assigned affect primarily the $2\ell SS$ and $2\ell SS1\tau_{\mathrm{had}}$ channels and predominantly arise from $t\bar{t}$ production, with one electron having undergone a hard bremsstrahlung as well as an asymmetric conversion ($e^\pm \to e^\pm\gamma^* \to e^\pm e^+ e^-$) or a mismeasured track curvature. The muon charge misassignment rate is negligible in the p_{T} range relevant to this analysis. The electron charge misassignment rate is measured in data using samples of $Z \to e^+e^-$ events reconstructed as same-charge pairs and as opposite-charge pairs, with the background subtracted via a sideband method. The charge misassignment rate is measured separately for the three types of tight electrons (internal conversion, material conversion, and very tight) and parameterised as a function of electron p_{T} and $|\eta|$. For very tight electrons, it varies from about 10^{-5} for low-p_{T} electrons ($15 \leq p_{\mathrm{T}} \leq 90$ GeV) at $|\eta| \leq 1.37$, to about 3×10^{-4} for high-p_{T} electrons ($p_{\mathrm{T}} \geq 90$ GeV) at $1.52 \leq |\eta| \leq 2$. The measured charge misassignment rate is then applied to data events satisfying the requirements of the $2\ell SS$ and $2\ell SS1\tau_{\mathrm{had}}$ channels, except that the two leptons are required to be of opposite charge, to estimate the QMisID background in each of the corresponding event categories.

The electron charge misassignment measurement is validated by a closure test in simulation using same-charge electron pairs, with the observed difference between measured and predicted rates taken as the systematic uncertainty. Additional systematic uncertainties include the statistical uncertainty from the data, the variation in the rates when the Z-peak range definition is varied, and the difference between the rates for electrons and positrons in the simulation. The total systematic uncertainty in the charge misassignment background estimate for very tight electrons is about 30%, with the dominant contribution at low p_{T} originating from the closure tests and at high p_{T} from the statistical uncertainty.



(a)

(b)

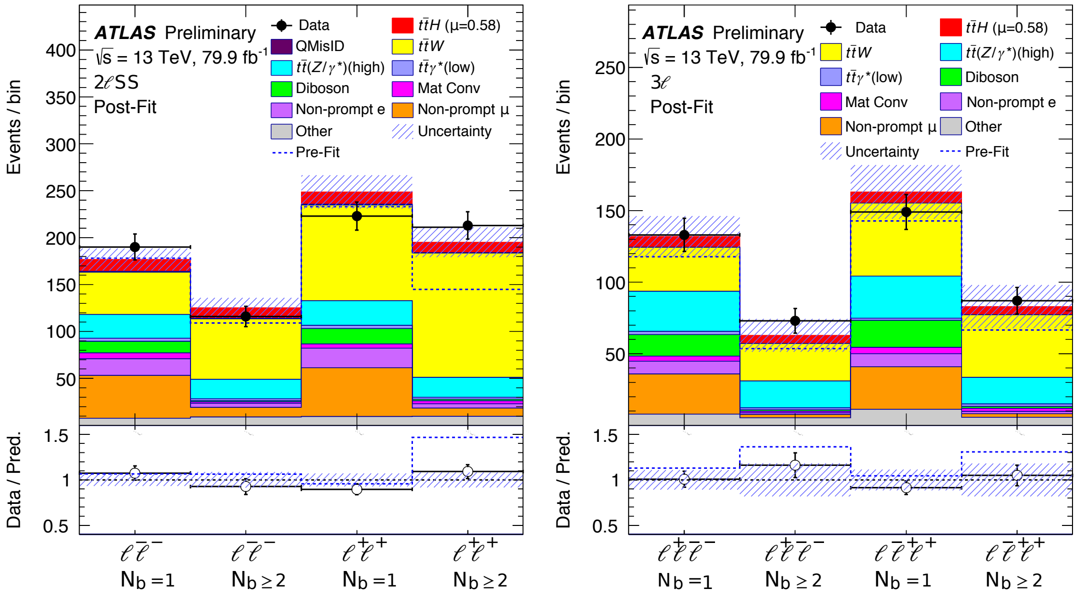
Figure 1: Comparison between data and signal-plus-background prediction for the distribution of jet multiplicity in (a) the 2ℓSS channel and (b) the 3ℓ channel after event selection and before further event categorisation (see Section 5). The background contributions after the likelihood fit (“Post-Fit") are shown as filled histograms. The total signal-plus-background prediction before the fit (“Pre-Fit") is shown as a dashed blue histogram. The $t\bar{t}H$ signal, scaled according to the results of the fit, is shown as a filled red histogram added to the post-fit background. The size of the combined statistical and systematic uncertainty in the signal-plus-background prediction is indicated by the blue hatched band. The ratios of the data to the total pre- and post-fit predictions are shown in the lower panel. The last bin in each figure contains the overflow.

6.2.2 Non-prompt light leptons

Non-prompt leptons originate from material conversions, heavy-flavour hadron decays, or the improper reconstruction of other particles, with an admixture strongly depending on the lepton quality requirements and varying across event categories. The main contribution to the non-prompt-lepton background is from $t\bar{t}$ production, followed by much smaller contributions from V+jets and single-top-quark processes. This background is estimated from simulation, with data-driven corrections, in the $2\ell SS$, 3ℓ, 4ℓ, and $2\ell SS1\tau_{\mathrm{had}}$ channels.

The non-prompt light leptons in the simulated $t\bar{t}$ sample are labelled according to whether they originate from heavy-flavour (HF) or light-flavour (LF) hadron decays, or from a material conversion candidate (MatC). The HF category includes leptons from both bottom and charm decays. Simulated $t\bar{t}$ background events with a selected lepton with mismeasured charge are excluded because the charge misassignment background is estimated from data, as described in Section 6.2.1.

Several of the event categories introduced in Section 5 are used to estimate the non-prompt light lepton background. These control regions were designed to be enriched in specific processes. The 2ℓMatC and 3ℓMatC control regions are enriched in electrons from material conversions. There are eight control regions enriched in contributions from HF non-prompt leptons in $t\bar{t}$ events: seven in the $2\ell SS$ channel and



(a) (b)

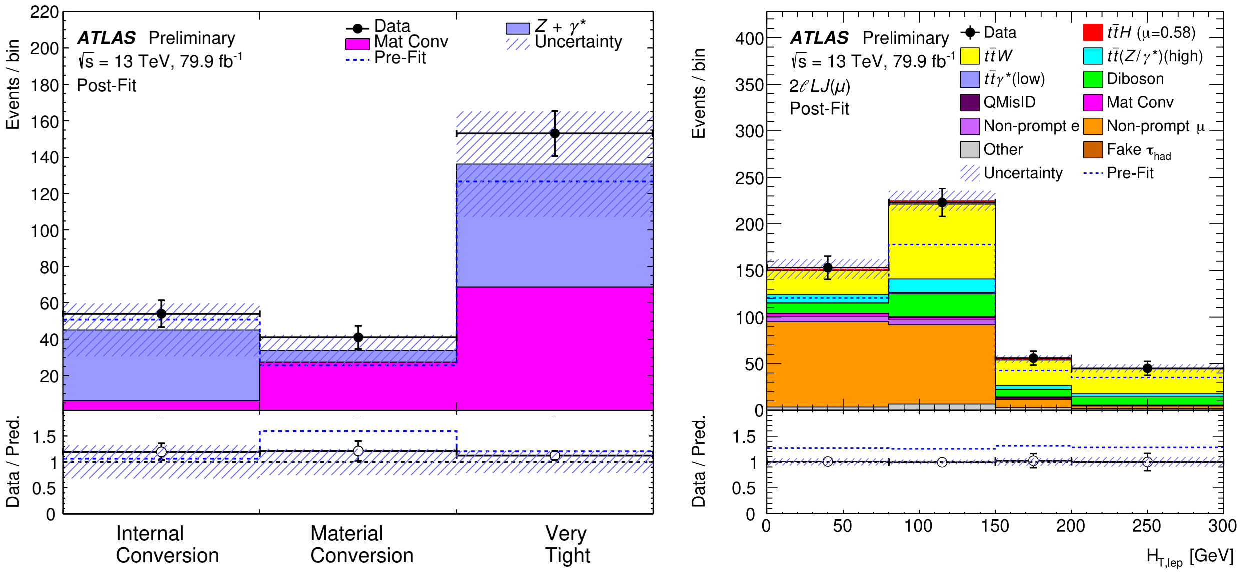
Figure 2: Comparison between data and signal-plus-background prediction for the event yields in (a) the 2ℓSS channel and (b) the 3ℓ channel after event selection and before further event categorisation (see Section 5), split in four separate categories depending on the total charge and b-jet multiplicity. The background contributions after the likelihood fit (“Post-Fit") are shown as filled histograms. The total signal-plus-background prediction before the fit (“Pre-Fit") is shown as a dashed blue histogram. The $t\bar{t}H$ signal, scaled according to the results of the fit, is shown as a filled red histogram added to the post-fit background. The size of the combined statistical and systematic uncertainty in the signal-plus-background prediction is indicated by the blue hatched band. The ratios of the data to the total pre- and post-fit predictions are shown in the lower panel.

one in the 3ℓ channel. Kinematic distributions are used in four control regions to optimise the sensitivity to the HF non-prompt electron and muon contributions. The distribution of the scalar sum of the lepton p_{T} ($H_{\mathrm{T,\,lep}}$), which provides separation against $t\bar{t}W$ (see Figure 3(b)), is used in the 2ℓLJ(μ) and 2ℓLJ(e2) control regions. In the 2ℓLJ(e1) control region, the $\Delta R(\ell,\ell)$ distribution is used as it provides separation against internal conversions, which are characterised by lower values of $\Delta R(\ell,\ell)$ since the virtual photon is usually radiated by one of the reconstructed leptons. The 3ℓtt control region uses the output of the BDT corresponding to the $t\bar{t}$ category. The total event yield is used in all other control regions.

Normalisation factors for three non-prompt-lepton background contributions are estimated from the likelihood fit. The normalisation factor for HF non-prompt leptons is estimated separately for electrons and muons, λ_e^{had} and $\lambda_\mu^{\mathrm{had}}$ respectively. An additional normalisation factor is determined for the material conversion candidates, $\lambda_e^{\mathrm{MatC}}$. The measured normalisation factors are: $\hat{\lambda}_e^{\mathrm{MatC}} = 1.61 \pm 0.48$, $\hat{\lambda}_e^{\mathrm{had}} = 1.12 \pm 0.38$, and $\hat{\lambda}_\mu^{\mathrm{had}} = 1.20 \pm 0.18$, where the uncertainties are dominated by the statistical uncertainty. Among the uncertainties associated with the non-prompt-lepton background estimation, the normalisation factor $\lambda_\mu^{\mathrm{had}}$ has the largest impact on the $t\bar{t}H$ cross section measurement. The systematic uncertainties considered are discussed in the following, although they have a negligible impact on the final result. The background estimation procedure for non-prompt light leptons relies on the simulation to predict the kinematic distributions of the $t\bar{t}$ process, and thus is affected by related modelling uncertainties (see Section 3). Additional uncertainties are estimated by relaxing lepton criteria to enrich the samples in the different types of non-prompt leptons, and comparing the data with the simulation. A 25% uncertainty on material conversions is assigned, based on the comparison between data and simulation in a validation region of selected $Z \rightarrow \mu^+\mu^-\gamma^*(\rightarrow e^+e^-)$ candidate events, with a loose electron that fails the photon conversion veto (see second bin in the distribution of Figure 3(a)). This uncertainty is applied to all categories except for 2ℓMatC and 3ℓMatC as thus acts as an extrapolation uncertainty. A shape uncertainty of up to 15% (6%) is assigned to the HF non-prompt electron (muon) background component from a comparison between data and simulation when the second leading electron (muon) is only required to be loose. The resulting differences between data and scaled simulation, after removal of the overall normalisation effect, is considered as an additional uncertainty correlated bin-to-bin across all distributions and event categories. As the contribution from LF non-prompt leptons is small, about 10% percent of the contribution from HF non-prompt leptons, it is derived from the agreement between data and simulation in a LF enriched region at low values of the non-prompt lepton BDT. The resulting uncertainty is 100%, and is taken to be uncorrelated between the categories enriched in internal and material conversions, and the rest of event categories. An additional uncertainty of 22% is assigned to the non-prompt light lepton background estimate in the $2\ell SS1\tau_{\mathrm{had}}$ channel from the comparison with an alternative data-driven estimation based on the fake factor method [90].

Several studies were performed to validate the non-prompt-lepton background estimates and to confirm that the data excess observed in several event categories is not because this background is mismodelled. In Section 7 it will be shown that this excess is assigned by the fit to $t\bar{t}W$-related fit parameters. A comparison of the data to the total predicted background was made in different event categories of the 2ℓSS channel, split by the flavour of the sub-leading lepton, which has a higher probability to be non-prompt, and the b-jet multiplicity (see Figure 4(a)). There is good agreement with the pre-fit background prediction for events with exactly one b-jet, but a significant data excess in events with at least two b-jets as shown with the dashed blue curve. The excess is shown to be independent of the flavour of the sub-leading lepton and hence uncorrelated with the fractional size of the non-prompt-lepton background. In addition, the b-jet multiplicity in a region strongly enriched in non-prompt leptons, obtained by relaxing the lepton selection requirements, is found to be well modelled. Finally, the score of the BDT which discriminates the $t\bar{t}H$ signal from the $t\bar{t}$ background is well-modelled in the non-prompt-lepton background-dominated region (see Figure 4(b)). This BDT is one of the main discriminators between non-prompt leptons and other backgrounds. In general, regions with large non-prompt backgrounds appear to be well modelled, however regions with large contributions from the $t\bar{t}W$ background and small contributions from non-prompt leptons, show significant discrepancies between data and the background prediction both in normalisation and in shape.

Non-prompt light leptons in the 4ℓ channel are grouped differently. As the contribution of LF electrons in the 4ℓ channel is negligibly small, it is fit together with the electrons from material conversions, while the HF electrons are fit separately. A single normalisation factor is used for non-prompt muons accounting for both HF and LF muons, as the latter are negligible. A control region is defined requiring three loose light leptons, one or two jets, and at least one b-tagged jet, but excluding events passing the 3ℓ selection. The control region is separated into two categories according to the lepton flavour ($ee\mu$ and $e\mu\mu$), each of which is split in three bins of $E_{\mathrm{T}}^{\mathrm{miss}}$. The three normalisation factors obtained from a likelihood fit in this control region are: $\hat{\lambda}_e^{4\ell,\mathrm{MatC}} = 1.4 \pm 1.4$, $\hat{\lambda}_e^{4\ell,\mathrm{had}} = 0.89 \pm 0.41$ and $\hat{\lambda}_\mu^{4\ell,\mathrm{had}} = 1.07 \pm 0.43$, where the uncertainties include statistical and systematic uncertainties, which are applied to each lepton. The systematic uncertainty in each normalisation factor is estimated to be 40% by varying the p_{T} requirements on the leptons.

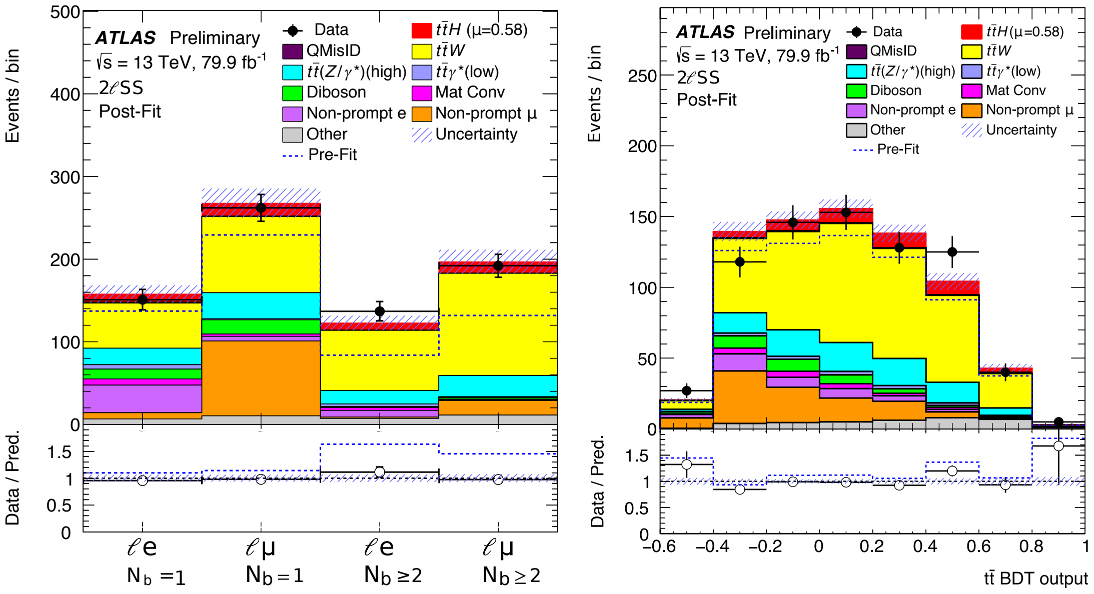


(a) (b)

Figure 3: (a) Comparison between data and prediction for the event yields in the 3ℓ validation region enriched in $Z \to \mu^+\mu^-\gamma^*(\to e^+e^-)$ candidate events, divided intro three categories depending on the requirements satisfied by the electron: internal conversion, material conversion, or very tight. The comparisons are made after correcting the simulation by the normalisation factors for internal conversions and material conversions resulting from the likelihood fit to data in all categories ("Post-Fit"). (b) Comparison between data and prediction for the distribution of the scalar sum of the lepton p_{T} ($H_{\mathrm{T,\,lep}}$) in the $2\ell SS$ $t\bar{t}$ control region at low jet multiplicity with a sub-leading muon, after performing the likelihood fit to data in all categories ("Post-Fit"). The $t\bar{t}H$ signal, scaled according to the results of the fit, is shown as a filled red histogram added to the post-fit background. In both figures the size of the combined statistical and systematic uncertainty in the total prediction is indicated by the blue hatched band. The ratio of the data to the total post-fit prediction is shown in the lower panel. The last bin in each figure contains the overflow.

6.2.3 Fake τ_{had} candidates

In the $2\ell SS1\tau_{\text{had}}$ and $3\ell1\tau_{\text{had}}$ channels, the fake τ_{had} background mainly arises from $t\bar{t}$ and $t\bar{t}V$ events with a jet misidentified as a τ_{had} candidate. A control region is defined requiring two opposite-charge leptons, at least three jets, at least one b-tagged jet, and at least one τ_{had} candidate. This control region is enriched in dileptonic $t\bar{t}$ events, such that the selected τ_{had} candidates primarily originate from jets, and are used to determine a normalisation factor to correct a possible mismodelling of the fake τ_{had} rate in the simulation. The normalisation factor is measured as a function of $p_{\mathrm{T}}^{\tau_{\text{had}}}$ and for one-prong and three-prong τ_{had} candidates separately. In the case of one-prong (three-prong) τ_{had} candidates, the normalisation factors range from 1.05 ± 0.06 (1.25 ± 0.42) for $p_{\mathrm{T}}^{\tau_{\text{had}}}$ in the range of 25–45(25 – 50) GeV, to 0.64 ± 0.12 (0.52 ± 0.71) for $p_{\mathrm{T}}^{\tau_{\text{had}}} \geq 70(75)$ GeV. Systematic uncertainties include the statistical uncertainty in the control region, the uncertainty on the contribution from real τ_{had} candidates that is subtracted in the control region, and the variation in the normalisation factors between the control region and a validation region that is enhanced in Z+jets. The total systematic uncertainty depends on $p_{\mathrm{T}}^{\tau_{\text{had}}}$ and is on average about 13% (60%) for one-prong (three-prong) τ_{had} candidates. Simulated events with both a non-prompt light lepton and a fake τ_{had} candidate are scaled by the product of their corresponding per-lepton normalisation factors. The fraction of fake τ_{had} background with an electron misidentified as a τ_{had} candidate is ~10% and is



(a) (b)

Figure 4: Comparison between data and signal-plus-background prediction in the 2ℓSS channel after event selection for (a) the event yield, split in four separate categories depending on the flavour of the sub-leading lepton and the b-jet multiplicity, and (b) the score of the BDT trained to discriminate $t\bar{t}H$ signal from $t\bar{t}$ background. The background contributions after the likelihood fit ("Post-Fit") are shown as filled histograms. The total background before the fit ("Pre-Fit") is shown as a dashed blue histogram. The $t\bar{t}H$ signal, scaled according to the results of the fit, is shown as a filled red histogram added to the post-fit background. The size of the combined statistical and systematic uncertainty in the signal-plus-background prediction is indicated by the blue hatched band. The ratio of the data to the total post-fit prediction is shown in the lower panel.

estimated with the simulation.

The dominant background in the $1\ell 2\tau_{\text{had}}$ channel is $t\bar{t}$ production with one or two fake τ_{had} candidates. Since in $t\bar{t}$ events there is equal probability for a jet to be reconstructed as a positively or negatively charged τ_{had} candidate, the fake τ_{had} background is estimated from a control region identical to the signal region except that the τ_{had} candidates are required to have the same charge. The selected sample in this control region, after subtraction of the small contribution from $t\bar{t}H$ signal and backgrounds with real τ_{had} candidates estimated with the simulation, is used to predict the fake τ_{had} background in the $1\ell 2\tau_{\text{had}}$ signal region. A systematic uncertainty of 30% for the estimated fake τ_{had} background in the signal region is estimated using a closure test of the method in simulation.

7 Analysis model and results

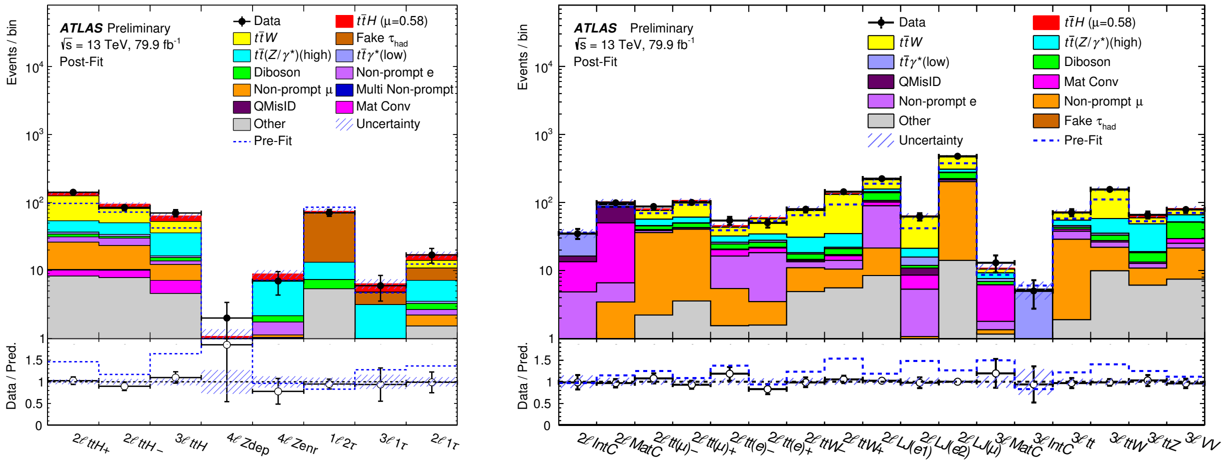
A maximum-likelihood fit is performed on all bins in the 25 event categories defined in Section 5 to determine the $t\bar{t}H$ cross section and the normalisation factors of the $t\bar{t}W$ process and other backgrounds. The $t\bar{t}H$ acceptance in each category is predicted by the simulation assuming the SM. Seventeen categories from the $2\ell SS$ and 3ℓ channels are used as control regions to either determine or constrain different backgrounds (material conversions, internal conversions, $t\bar{t}$ with non-prompt electrons and muons, $t\bar{t}W$,

Table 4: Sources of systematic uncertainty considered in the analysis. “N" means that the uncertainty is taken as normalisation-only for all processes and channels affected. Some of the systematic uncertainties are split into several components, as indicated by the number in the rightmost column. The PDF uncertainties are correlated between the $t\bar{t}H$ signal and the $t\bar{t}W$ background.

Systematic uncertainty	Components
Luminosity (N)	1
Pileup modelling	1
Physics objects	
Electron	8
Muon	11
Tau	7
Jet energy scale and resolution	28
Jet vertex fraction	1
Jet flavour tagging	17
$E_{\mathrm{T}}^{\mathrm{miss}}$	3
Total (Experimental)	77
Data-driven background estimates	
Non-prompt light-lepton estimates (3ℓ, $3\ell 1\tau_{\mathrm{had}}$)	1
Fake τ_{had} estimates	6
Electron charge misassignment	2
Total (Data-driven reducible background)	9
Template fit uncertainties	
Material conversions	1
Internal conversions	1
HF non-prompt leptons	18
LF non-prompt leptons	2
Total (Template fit)	22
$t\bar{t}H$ modelling	
Renormalisation and factorisation scales	3
Parton shower and hadronisation model	1
Higgs boson branching ratio	4
Shower tune	1
PDF	32
$t\bar{t}W$ modelling	
Radiation	1
Generator	1
PDF	32
Extrapolation	4
$t\bar{t}(Z/\gamma^*)$ (high mass) modelling	
Cross section (N)	2
Generator	1
Renormalisation and factorisation scales	3
Shower tune	1
$t\bar{t}$ modelling	
Radiation	1
WZ modelling	
HF composition (N)	3
Shower tune	1
Other background modelling	
Cross section (N)	22
Total (Signal and background modelling)	120
Total (Overall)	218

$t\bar{t}Z$, and VV). In 13 of the control regions the total event yield (i.e. a single bin) is used, but in the remaining four control regions different kinematic variables are used to discriminate between $t\bar{t}$ and $t\bar{t}W$ backgrounds (see Section 6.2.2). The remaining eight categories are used as signal regions to measure the $t\bar{t}H$ cross section. In the $t\bar{t}H$ categories of the $2\ell SS$, 3ℓ, and $1\ell 2\tau_{\mathrm{had}}$ channels a BDT discriminant is used and the total event yield is used in the remaining four signal regions.

The likelihood function $\mathcal{L}(\mu, \vec{\lambda}, \vec{\theta})$ depends on the signal-strength parameter, μ, defined as a multiplicative factor to the yield for the $t\bar{t}H$ signal events, which are normalised to the SM prediction, $\vec{\lambda}$, the normalisation factors for several backgrounds (see Sections 6.1 and 6.2.2), and $\vec{\theta}$, a set of nuisance parameters (NP) encoding systematic uncertainties in the signal and background expectations. Systematic uncertainties can impact the estimated signal and background rates, the migration of events between categories, and the shape of the fitted discriminants; they are summarised in Table 4. Both μ and $\vec{\lambda}$ are treated as free parameters in the likelihood fit. The NPs $\vec{\theta}$ allow variations of the expectations for signal and background according to the systematic uncertainties, subject to Gaussian or log-normal constraints in the likelihood fit. Their fitted values represent the deviations from the nominal expectations that globally provide the best fit to the data. Statistical uncertainties in each bin due to the limited size of the simulated samples are taken into account by dedicated parameters using the Beeston-Barlow technique [91].



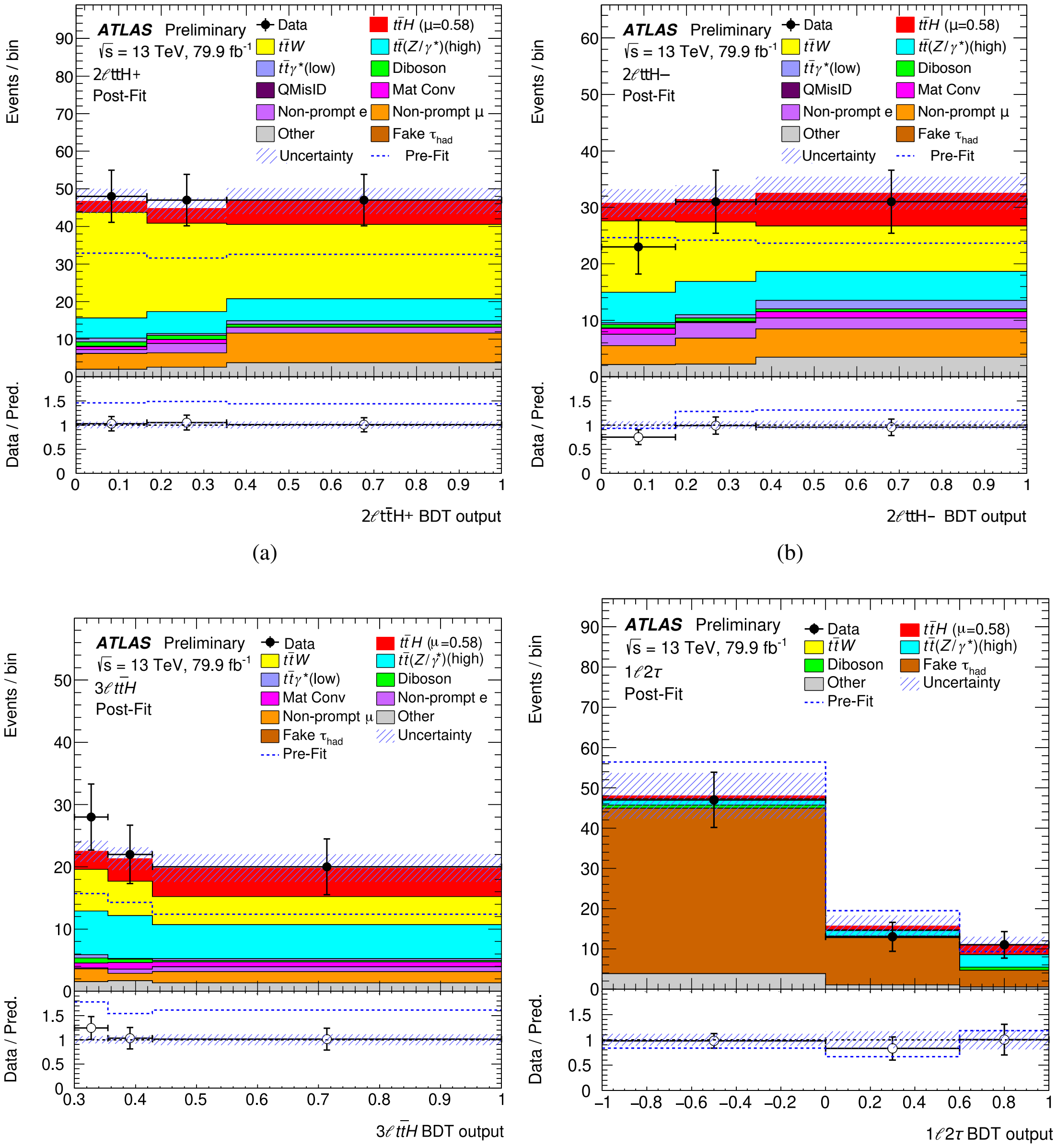
(a) (b)

Figure 5: Comparison between data and prediction for the event yields in (a) the eight $t\bar{t}H$ categories and (b) the 17 control-region categories. The background contributions after the likelihood fit ("Post-Fit") are shown as filled histograms. The total background before the fit ("Pre-Fit") is shown as a dashed blue histogram. The $t\bar{t}H$ signal, scaled according to the results of the fit, is shown as a filled red histogram added to the post-fit background. The size of the combined statistical and systematic uncertainty in the total signal-plus-background prediction is indicated by the blue hatched band. The ratio of the data to the total post-fit prediction is shown in the lower panel.

The test statistic q_0 is defined as the profile likelihood ratio: $q_0 = -2\ln(\mathcal{L}(0, \hat{\hat{\vec{\lambda}}}_0, \hat{\hat{\vec{\theta}}}_0)/\mathcal{L}(\hat{\mu}, \hat{\hat{\vec{\lambda}}}_{\hat{\mu}}, \hat{\hat{\vec{\theta}}}_{\hat{\mu}}))$, where $\hat{\mu}$, $\hat{\hat{\vec{\lambda}}}_{\hat{\mu}}$, and $\hat{\hat{\vec{\theta}}}_{\hat{\mu}}$ are the values of the parameters that maximise the likelihood function, and $\hat{\hat{\vec{\lambda}}}_0$ and $\hat{\hat{\vec{\theta}}}_0$ are the values of the parameters that maximise the likelihood function when fixing μ to zero. The test statistic is evaluated with the RooFit package [92, 93] and is used to quantify how well the observed data agrees with the background-only hypothesis.

The fitted $\hat{\mu}$ value is obtained by maximising the likelihood function with respect to all parameters and the total uncertainty, $\Delta\hat{\mu}$, is obtained from the variation of q_0 by one unit from its minimum. An estimate of the contribution from systematic uncertainties is found by subtracting in quadrature from $\Delta\hat{\mu}$ the statistical uncertainty, which is determined by redoing the fit to data after fixing all NPs to their best-fit values. The contribution from the background normalisation factors is included in the statistical uncertainty. The expected results are obtained in the same way as the observed results by replacing the data in each input bin by the prediction from simulation and the data-driven fake and non-prompt estimates with all NPs set to their best-fit values obtained from the fit to data. The significance is obtained from the test statistic using the asymptotic formulae given in Refs. [94].

Figures 5(a) and 5(b) compare the data to the yields after the predictions were adjusted by the fit in the 25 event categories considered. In all categories, the observed yields agree with the fitted prediction within uncertainties. The background before the fit is shown as a dashed blue histogram. Figure 6 shows the distributions of the BDT discriminants used in selected categories with the bins used in the fit after the fit to the data. In the remaining categories, only the event yields are used in the fit. Differences between the data and the prediction are observed pre-fit, but overall good agreement is observed after the fit. Figure 7 shows the data, background and $t\bar{t}H$ signal yields, where the final-discriminant bins in all event categories are combined into bins of $\log_{10}(S/B)$, where S is the expected $t\bar{t}H$ signal yield and B the fitted background



(a) (b)

(c) (d)

Figure 6: Comparison between data and prediction for the BDT discriminants used in different $t\bar{t}H$ categories: (a) $2\ell ttH+$, (b) $2\ell ttH-$, (c) $3\ell ttH$, and (d) $1\ell 2\tau_{\text{had}}$. The background contributions after the likelihood fit ("Post-Fit") are shown as filled histograms. The total background before the fit is shown as a dashed blue histogram. The total background before the fit ("Pre-Fit") is shown as a dashed blue histogram. The $t\bar{t}H$ signal, scaled according to the results of the fit, is shown as a filled red histogram added to the post-fit background. The size of the combined statistical and systematic uncertainty in the total signal-plus-background prediction is indicated by the blue hatched band. The ratio of the data to the total post-fit prediction is shown in the lower panel.

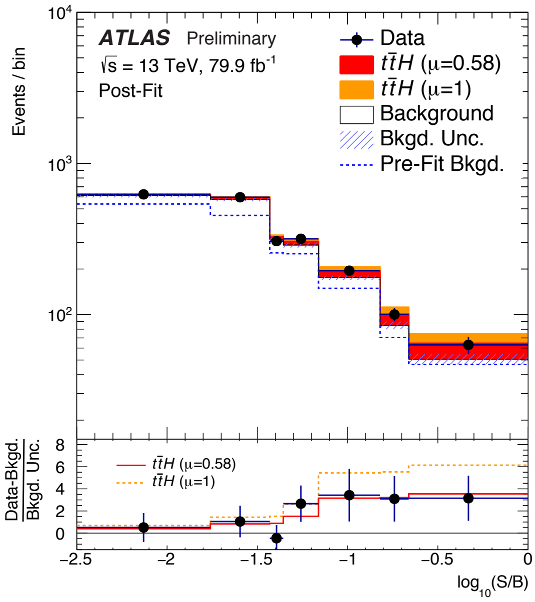


Figure 7: Event yields as a function of $\log_{10}(S/B)$ for data, background and a Higgs boson signal with $m_H = 125$ GeV. The discriminant bins in all categories are combined into bins of $\log_{10}(S/B)$, where S is the expected $t\bar{t}H$ signal yield and B the background yield from the unconditional fit. The background yields are shown as the fitted values, while the $t\bar{t}H$ signal yields are shown for the fitted value ($\mu = 0.58$) and the SM prediction ($\mu = 1$). The total background before the fit is shown as a dashed blue histogram. The size of the combined statistical and systematic uncertainty in the background prediction is indicated by the blue hatched band. The pull (residual divided by its uncertainty) of the data relative to the background-only prediction is shown in the lower panel, where the full red line (dashed orange line) indicates the pull of the prediction for signal with $\mu = 0.58$ ($\mu = 1$) and background relative to the background-only prediction.

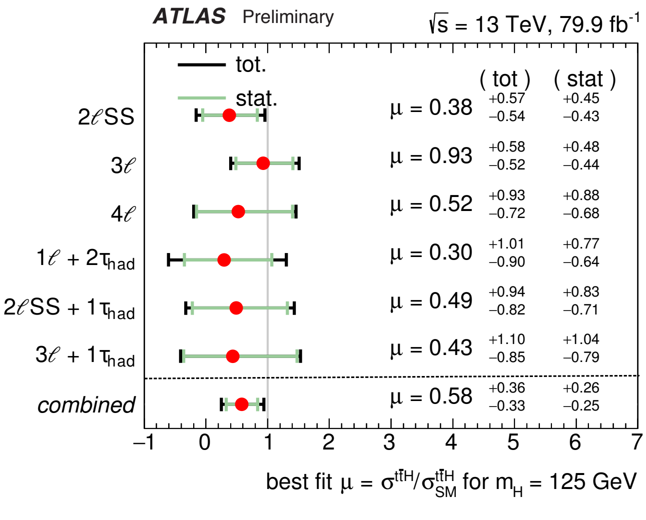


Figure 8: The observed best-fit values of the $t\bar{t}H$ signal strength μ and their uncertainties by analysis channel and combined. The individual μ values for the channels are obtained from a simultaneous fit with the signal-strength parameter for each channel floating independently. The SM prediction corresponds to μ=1.

yield. The total background before the fit is shown as a dashed blue histogram. The significance of the observed (expected) excess above the background-only expectation ($\mu = 0$) is 1.8 (3.1) standard deviations. The best-fit value of μ is:

$$\hat{\mu} = 0.58^{+0.26}_{-0.25}\ (\text{stat.})^{+0.19}_{-0.15}\ (\text{exp.})^{+0.13}_{-0.11}\ (\text{bkg. th.})^{+0.08}_{-0.07}\ (\text{sig. th.}) = 0.58^{+0.36}_{-0.33}. \tag{1}$$

The best-fit value of μ for each individual channel and the combination of all channels are shown in Figure 8. The individual channel results are extracted from the full fit but with a separate parameter of interest for each channel. The probability that the six fitted signal strengths are compatible with a single value is 98%.

Normalisation factors for several important irreducible and reducible backgrounds are determined by the fit (see Section 6). Of particular interest are the three measured normalisation factors for the $t\bar{t}W$ background in the 2ℓSS and 3ℓ event categories: $\hat{\lambda}^{2\ell\text{LJ}}_{t\bar{t}W} = 1.56^{+0.30}_{-0.28}$, $\hat{\lambda}^{2\ell\text{HJ}}_{t\bar{t}W} = 1.26^{+0.19}_{-0.18}$, and $\hat{\lambda}^{3\ell}_{t\bar{t}W} = 1.68^{+0.30}_{-0.28}$. They are consistent with each other and systematically above unity, indicating a preference of the data for a higher value of the $t\bar{t}W$ cross section than the updated $t\bar{t}W$ theoretical cross section (see Section 3). Because the $t\bar{t}W$ modelling uncertainties are constructed to only affect the shapes of distributions while the total yield is fixed, the normalisation factors represent a scaling factor for $t\bar{t}W$ events selected in this analysis. Uncertainties to extrapolate the $t\bar{t}W$ scaling factor to the inclusive phase space are not included.

The contributions from the different sources of uncertainty on the measured signal strength is shown in Table 5. The leading systematic uncertainty is in the jet energy scale and resolution. The most important systematic uncertainties arising from theoretical predictions are in the modelling of $t\bar{t}W$ production and the normalisation of the $t\bar{t}(Z/\gamma^*)$ background. The uncertainty associated with the τ_{had} background estimate is also significant. Only few NPs in the fit showed significant adjustments and/or constraints (see Figure 13). Among them are the NPs associated with the b-jet multiplicity and total charge extrapolation uncertainties on the $t\bar{t}W$ background in the 2ℓSS channel. They were adjusted by +0.33 and +0.75 pre-fit standard deviations, respectively, and their uncertainties reduced by factors of 3 and 2, respectively. The NP associated with the uncertainty in the closure of the non-prompt lepton estimate in $1\ell 2\tau_{\text{had}}$ channel was adjusted by −0.56 pre-fit standard deviations and its uncertainty reduced by a factor of 1.7; however, this uncertainty does not significantly impact the $t\bar{t}H$ cross section measurement.

An extrapolation to the inclusive phase space, assuming SM $t\bar{t}H$ kinematics, is made and the measured $t\bar{t}H$ production cross section is:

$$\hat{\sigma}(t\bar{t}H) = 294^{+132}_{-127}\ (\text{stat.})^{+94}_{-74}\ (\text{exp.})^{+73}_{-56}\ (\text{bkg. th.})^{+41}_{-39}\ (\text{sig. th.})\ \text{fb} = 294^{+182}_{-162}\ \text{fb}. \tag{2}$$

The predicted SM cross section is $\sigma(t\bar{t}H) = 507^{+35}_{-50}$ fb computed at NLO in QCD and electroweak couplings [28]. The measured cross section is consistent with the SM prediction within uncertainties.

7.1 Cross-checks

A number of cross-checks of the assumptions in the statistical model were performed. The measured signal strength was found to be robust under these cross-checks, provided that the $t\bar{t}W$ normalisation was not fixed. We report here six key cross-checks to test the robustness of the model and the consistency of the results with other analyses.

A comparison was made between the results from the combination of the 2ℓSS and 3ℓ categories in the nominal multivariate analysis and those in the cross-check cut-and-count analysis, which has lower expected

Table 5: Breakdown of the contributions to the uncertainties in $\hat{\mu}$. The contribution of the different sources of uncertainty is evaluated after the fit described in Section 7. The total statistical uncertainty is evaluated, as described in the text, by fixing all the nuisance parameters in the fit except for the free-floating background normalisation factors. The contribution from the uncertainty in those normalisation factors is then included in the quoted total statistical uncertainty rather than in the systematic uncertainty component. The statistical uncertainty evaluated after also fixing the background normalisation factors is then indicated as "intrinsic statistical uncertainty". Statistical uncertainties from data-driven background estimates are included within the experimental uncertainties. The other quoted numbers are obtained by repeating the fit after having fixed a certain set of nuisance parameters corresponding to a group of systematic uncertainty sources, and subtracting in quadrature the resulting total uncertainty of μ from the uncertainty from the full fit. The same procedure is followed for quoting the individual effects of background normalisation factors. Due to rounding effects and small correlations between the different sources of uncertainty, the total systematic uncertainty is different from the sum in quadrature of the individual sources.

Uncertainty source	$\Delta\hat{\mu}$	
Jet energy scale and resolution	+0.13	−0.13
$t\bar{t}(Z/\gamma^*)$ (high mass) modelling	+0.09	−0.09
$t\bar{t}W$ modelling (radiation, generator, PDF)	+0.08	−0.08
Fake τ_{had} background estimate	+0.07	−0.07
$t\bar{t}W$ modelling (extrapolation)	+0.05	−0.05
$t\bar{t}H$ cross section	+0.05	−0.05
Simulation sample size	+0.05	−0.05
$t\bar{t}H$ modelling	+0.04	−0.04
Other background modelling	+0.04	−0.04
Jet flavour tagging and τ_{had} identification	+0.04	−0.04
Other experimental uncertainties	+0.03	−0.03
Luminosity	+0.03	−0.03
Diboson modelling	+0.01	−0.01
$t\bar{t}\gamma^*$ (low mass) modelling	+0.01	−0.01
Charge misassignment	+0.01	−0.01
Template fit (non-prompt leptons)	+0.01	−0.01
Total systematic uncertainty	+0.25	−0.22
Intrinsic statistical uncertainty	+0.23	−0.22
$t\bar{t}W$ normalisation factors	+0.10	−0.10
Non-prompt leptons normalisation factors (HF, material conversions)	+0.05	−0.05
Total statistical uncertainty	+0.26	−0.25
Total uncertainty	+0.36	−0.33

sensitivity relative to the background-only hypothesis (2.4 vs 1.4 standard deviations). This analysis included the same $t\bar{t}W$ control regions as the nominal analysis. The resulting best-fit signal strengths were found to be $\hat{\mu} = 0.67^{+0.44}_{-0.41}$ and $\hat{\mu} = 0.43^{+0.66}_{-0.65}$ for the 2ℓSS and 3ℓ categories in the nominal and the cut-and-count analyses, respectively. In both cases the best-fit $t\bar{t}W$ normalisation factors were found to be high and consistent with those from the nominal analysis.

The combined fit was performed separately on the 2015–2016 dataset (36 fb^{-1}) and the 2017 dataset (44 fb^{-1}), which were recorded under different pileup conditions (average number of pp interactions per crossing of 25 and 38, respectively). The resulting best-fit signal strengths were found to be consistent between both datasets: $\hat{\mu} = 0.68^{+0.50}_{-0.45}$ and $\hat{\mu} = 0.52^{+0.45}_{-0.40}$, respectively. Similarly, the best-fit $t\bar{t}W$ normalisation factors were found to be high and consistent between both datasets.

An alternative scenario with a single $t\bar{t}W$ normalisation factor instead of three was tested. In this case the best-fit values obtained are $\hat{\mu} = 0.70^{+0.36}_{-0.33}$ and $\hat{\lambda}_{t\bar{t}W} = 1.39^{+0.17}_{-0.16}$, with only minor changes to the rest of the fitted parameters. The compatibility of this alternative fit with the nominal one is 28%, corresponding to 0.59 standard deviations.

The value of $\hat{\mu}$ obtained by this analysis was compared to that published using 36 fb^{-1} of data [13], which was obtained with the $t\bar{t}W$ normalisation constrained within uncertainties to its NLO QCD+EW theoretical cross section [28]. When the fit is performed in the current analysis with the $t\bar{t}W$ normalisation forced to the same theoretical prediction, and without additional extrapolation uncertainties, a value of $\hat{\mu}$ consistent with that of Ref. [13] is obtained on the same dataset. Compared to Ref. [13], this analysis has a factor of three improvement in the $t\bar{t}W$ sensitivity, as a result of the reduced non-prompt-lepton background, the improved event categorisation, and the higher integrated luminosity used. This allows fewer assumptions in the background estimates to be made in the statistical analysis, and a larger value for the fitted $t\bar{t}W$ normalisation is obtained.

Validation studies of the non-prompt-lepton background estimate (see Section 6.2.2) disfavour mismodelling of this background being the explanation for the enhanced $t\bar{t}W$ normalisation factors. These studies, along with the preference of the fit model to assign the observed data excess to $t\bar{t}W$-related fit parameters, are consistent with the data requiring a larger $t\bar{t}W$ normalisation compared to theoretical predictions and the current estimate of the non-prompt-lepton background.

The value of $\hat{\lambda}_{t\bar{t}W}$ obtained by this analysis in the alternative scenario of a single $t\bar{t}W$ normalisation factor (see above) was compared with that obtained by a previous measurement of the $t\bar{t}W$ cross section using 36 fb^{-1} of data [95]. Such measurement yielded $\hat{\lambda}_{t\bar{t}W} = 1.19 \pm 0.26$ with respect to the updated $t\bar{t}W$ theoretical cross section (see Section 3).[3] While both results are similar, a detailed assessment of their compatibility is beyond the scope of this study.

8 Conclusion

A search for $t\bar{t}H$ production in multilepton final states using a dataset corresponding to an integrated luminosity of 80 fb^{-1} of proton–proton collision at $\sqrt{s} = 13$ TeV recorded by the ATLAS experiment at the LHC is presented. Six final states, targeting Higgs boson decays to WW^*, $\tau\tau$, and ZZ^*, categorised by the number and flavour of charged-lepton candidates, are analysed. An excess of events over the expected background from SM processes is found, which is interpreted as an observed significance of 1.8 standard deviations for a SM Higgs boson of mass 125 GeV. The expected significance for a SM Higgs boson is 3.1 standard deviations. The best-fit result of the observed production cross section is $\sigma(t\bar{t}H) = 294^{+182}_{-162}$ fb, in agreement with the SM prediction of 507^{+35}_{-50} fb. The normalisation factors obtained for the $t\bar{t}W$ background in the phase space selected by this analysis are in the range 1.3–1.7 above the updated theoretical predictions. In addition, modelling issues are observed in regions dominated by $t\bar{t}W$ production. An improved description of the $t\bar{t}W$ background is needed to reach greater precision in the future.

---

[3] The number reported has been obtained by dividing the result of Ref. [95] by a factor of 1.21.

References

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

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

[3] ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at* $\sqrt{s} = 7$ *and* 8 *TeV*, JHEP **08** (2016) 045, arXiv: `1606.02266 [hep-ex]` (cit. on p. 2).

[4] CMS Collaboration, *Observation of the Higgs boson decay to a pair of* τ *leptons*, Phys. Lett. B **779** (2018) 283, arXiv: `1708.00373 [hep-ex]` (cit. on p. 2).

[5] ATLAS Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of* τ*-leptons in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072001, arXiv: `1811.08856 [hep-ex]` (cit. on p. 2).

[6] ATLAS Collaboration, *Observation of* $H \rightarrow b\bar{b}$ *decays and* VH *production with the ATLAS detector*, Phys. Lett. B **786** (2018) 59, arXiv: `1808.08238 [hep-ex]` (cit. on p. 2).

[7] CMS Collaboration, *Observation of Higgs Boson Decay to Bottom Quarks*, Phys. Rev. Lett. **121** (2018) 121801, arXiv: `1808.08242 [hep-ex]` (cit. on p. 2).

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

[9] CMS Collaboration, *Observation of* $t\bar{t}H$ *Production*, Phys. Rev. Lett. **120** (2018) 231801, arXiv: `1804.02610 [hep-ex]` (cit. on p. 2).

[10] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a* $b\bar{b}$ *pair in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, (2017), CERN-EP-2017-291 (cit. on p. 2).

[11] ATLAS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with* 36.1 fb^{-1} *pp collision data at the center-of-mass energy of* 13 *TeV with the ATLAS detector*, (2017), CERN-EP-2017-288 (cit. on p. 2).

[12] ATLAS Collaboration, *Measurement of the Higgs boson coupling properties in the* $H \rightarrow ZZ^{*} \rightarrow 4\ell$ *decay channel at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, (2017), CERN-EP-2017-206 (cit. on p. 2).

[13] ATLAS Collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. D **97** (2018) 072003, arXiv: `1712.08891 [hep-ex]` (cit. on pp. 2, 4, 6, 7, 25).

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

[15] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19 (2010), `https://cds.cern.ch/record/1291633` (cit. on p. 3).

[16] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: `1611.09661 [hep-ex]` (cit. on p. 3).

[17] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, (2019), ATLAS-CONF-2019-021, URL: `https://cds.cern.ch/record/2677054` (cit. on p. 3).

[18] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 3).

[19] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: `0710.3820 [hep-ph]` (cit. on p. 3).

[20] ATLAS Collaboration, *Further ATLAS tunes of PYTHIA6 and Pythia 8*, (2011), ATL-PHYS-PUB-2011-014, URL: `https://cds.cern.ch/record/1400677` (cit. on p. 3).

[21] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: `1005.4568 [physics.ins-det]` (cit. on p. 3).

[22] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250 (cit. on p. 3).

[23] S. Frixione, G. Ridolfi and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv: `0707.3088 [hep-ph]` (cit. on pp. 4, 5).

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

[25] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: `1410.8849 [hep-ph]` (cit. on pp. 4, 5).

[26] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: `1410.3012 [hep-ph]` (cit. on pp. 4, 5).

[27] ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes*, ATL-PHYS-PUB-2014-021, 2014, URL: `https://cds.cern.ch/record/1966419` (cit. on p. 4).

[28] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: `1610.07922 [hep-ph]` (cit. on pp. 4, 6, 23, 25).

[29] 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 [hep-ph]` (cit. on p. 4).

[30] W. Beenakker et al., *Higgs radiation off top quarks at the Tevatron and the LHC*, Phys. Rev. Lett. **87** (2001) 201805, arXiv: `hep-ph/0107081 [hep-ph]` (cit. on p. 4).

[31] W. Beenakker et al., *NLO QCD corrections to t anti-t H production in hadron collisions*, Nucl. Phys. B **653** (2003) 151, arXiv: `hep-ph/0211352 [hep-ph]` (cit. on p. 4).

[32] S. Dawson, L. H. Orr, L. Reina and D. Wackeroth, *Associated top quark Higgs boson production at the LHC*, Phys. Rev. D **67** (2003) 071503, arXiv: `hep-ph/0211438 [hep-ph]` (cit. on p. 4).

[33] S. Dawson, C. Jackson, L. H. Orr, L. Reina and D. Wackeroth, *Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections*, Phys. Rev. D **68** (2003) 034022, arXiv: `hep-ph/0305087 [hep-ph]` (cit. on p. 4).

[34] Y. Zhang, W.-G. Ma, R.-Y. Zhang, C. Chen and L. Guo, *QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider*, Phys. Lett. B **738** (2014) 1, arXiv: `1407.1110 [hep-ph]` (cit. on p. 4).

[35] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao and M. Zaro, *Weak corrections to Higgs hadroproduction in association with a top-quark pair*, JHEP **09** (2014) 065, arXiv: `1407.0823 [hep-ph]` (cit. on p. 4).

[36] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: `1512.01178 [hep-ph]` (cit. on pp. 4, 5).

[37] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: `1510.03865 [hep-ph]` (cit. on p. 4).

[38] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: `1405.0301 [hep-ph]` (cit. on pp. 4–6).

[39] T. Sjöstrand et al., *High-energy-physics event generation with Pythia 6.1*, Comput. Phys. Commun. **135** (2001) 238, arXiv: `hep-ph/0010017 [hep-ph]` (cit. on p. 5).

[40] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: `0803.0883 [hep-ph]` (cit. on p. 5).

[41] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: `0811.4622 [hep-ph]` (cit. on pp. 4, 5).

[42] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: `1111.5206 [hep-ph]` (cit. on pp. 4, 5).

[43] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: `0808.3674 [hep-ph]` (cit. on pp. 4, 5).

[44] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: `0709.1027 [hep-ph]` (cit. on pp. 4, 5).

[45] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv: `1207.5030 [hep-ph]` (cit. on pp. 4, 5).

[46] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. Phys. Res. A **462** (2001) 152 (cit. on p. 5).

[47] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv: `hep-ph/0506026 [hep-ph]` (cit. on p. 5).

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

[49] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, arXiv: `1412.3989 [hep-ph]` (cit. on p. 5).

[50] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv: `1007.2241 [hep-ph]` (cit. on p. 5).

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

[52] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. D **78** (2008) 013004, arXiv: `0802.0007 [hep-ph]` (cit. on p. 5).

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

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

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

[56] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{\pm}$ production and decay at NLO*, JHEP **07** (2012) 052, arXiv: `1204.5678 [hep-ph]` (cit. on p. 4).

[57] S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, JHEP **06** (2015) 184, arXiv: `1504.03446 [hep-ph]` (cit. on pp. 4, 6).

[58] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, JHEP **12** (2012) 061, arXiv: `1209.6215 [hep-ph]` (cit. on p. 5).

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

[60] ATLAS Collaboration, *Modelling of the $t\bar{t}H$ and $t\bar{t}V$ $(V = W, Z)$ processes for $\sqrt{s} = 13$ TeV ATLAS analyses*, (2015), ATL-PHYS-PUB-2016-005, URL: `https://cds.cern.ch/record/2120826` (cit. on p. 6).

[61] ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, (2015), ATL-PHYS-PUB-2016-002, URL: `https://cds.cern.ch/record/2119986` (cit. on p. 6).

[62] ATLAS Collaboration, *Simulation of top quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV*, (2015), ATL-PHYS-PUB-2016-004, URL: `https://cds.cern.ch/record/2120417` (cit. on p. 6).

[63] M. Cacciari, M. Czakon, M. Mangano, A. Mitov and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. B **710** (2012) 612, arXiv: `1111.5869 [hep-ph]` (cit. on p. 6).

[64] P. Bärnreuther, M. Czakon and A. Mitov, *Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv: `1204.5201 [hep-ph]` (cit. on p. 6).

[65] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **12** (2012) 054, arXiv: `1207.0236 [hep-ph]` (cit. on p. 6).

[66] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, JHEP **01** (2013) 080, arXiv: `1210.6832 [hep-ph]` (cit. on p. 6).

[67] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv: `1303.6254 [hep-ph]` (cit. on p. 6).

[68] ATLAS Collaboration, *Simulation of top-quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2016-004, 2016, URL: `https://cds.cern.ch/record/2120417` (cit. on p. 6).

[69] 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` (cit. on p. 7).

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

[71] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. (2019), arXiv: `1902.04655 [hep-ex]` (cit. on p. 7).

[72] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, Eur. Phys. J. C **79** (2019) 205, arXiv: `1810.05087 [hep-ex]` (cit. on p. 7).

[73] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, submitted to JINST (2009), arXiv: `1908.00005 [hep-ex]` (cit. on p. 7).

[74] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017), arXiv: `1603.02934 [hep-ex]` (cit. on p. 8).

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

[76] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: `0802.1189 [hep-ph]` (cit. on p. 8).

[77] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: `1111.6097 [hep-ph]` (cit. on p. 8).

[78] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: `1703.09665 [hep-ex]` (cit. on p. 8).

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

[80] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 17, arXiv: `1406.0076 [hep-ex]` (cit. on p. 8).

[81] ATLAS Collaboration, *Jet energy resolution in proton–proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2306, arXiv: `1210.6210 [hep-ex]` (cit. on p. 8).

[82] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, JINST **11** (2016) P04008, arXiv: `1512.01094 [hep-ex]` (cit. on p. 8).

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

[84] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016, URL: `https://cds.cern.ch/record/2160731` (cit. on p. 8).

[85] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, ATL-PHYS-PUB-2015-045, 2015, URL: `https://cds.cern.ch/record/2064383` (cit. on p. 8).

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

[87] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: `1802.08168 [hep-ex]` (cit. on p. 9).

[88] A. Hoecker et al., *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv: `physics/0703039` (cit. on p. 11).

[89] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, CoRR **abs/1603.02754** (2016), URL: `http://arxiv.org/abs/1603.02754` (cit. on p. 11).

[90] ATLAS Collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. D **97** (2018) 072003, arXiv: `1712.08891 [hep-ex]` (cit. on p. 16).

[91] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219 (cit. on p. 19).

[92] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv: `physics/0306116 [physics.data-an]` (cit. on p. 20).

[93] W. Verkerke and D. Kirkby, *RooFit Users Manual*, URL: `http://roofit.sourceforge.net` (cit. on p. 20).

[94] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, Erratum: Eur. Phys. J. C **73** (2013) 2501, arXiv: `1007.1727 [physics.data-an]` (cit. on p. 20).

[95] ATLAS Collaboration, *Measurement of the* $t\bar{t}Z$ *and* $t\bar{t}W$ *cross sections in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072009, arXiv: `1901.03584 [hep-ex]` (cit. on p. 25).

Appendix

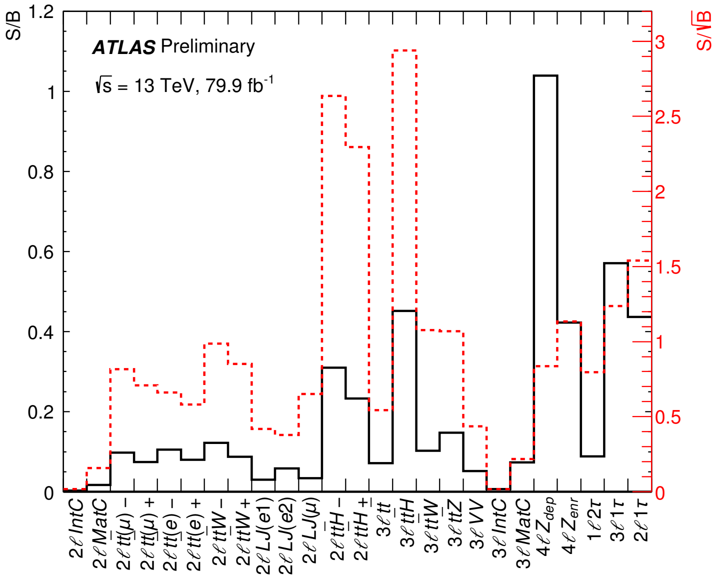


Figure 9: Pre-fit S/B (black line) and $S/\sqrt{B}$ (red dashed line) ratios for each analysis category. The background prediction methods are described in Section 6.

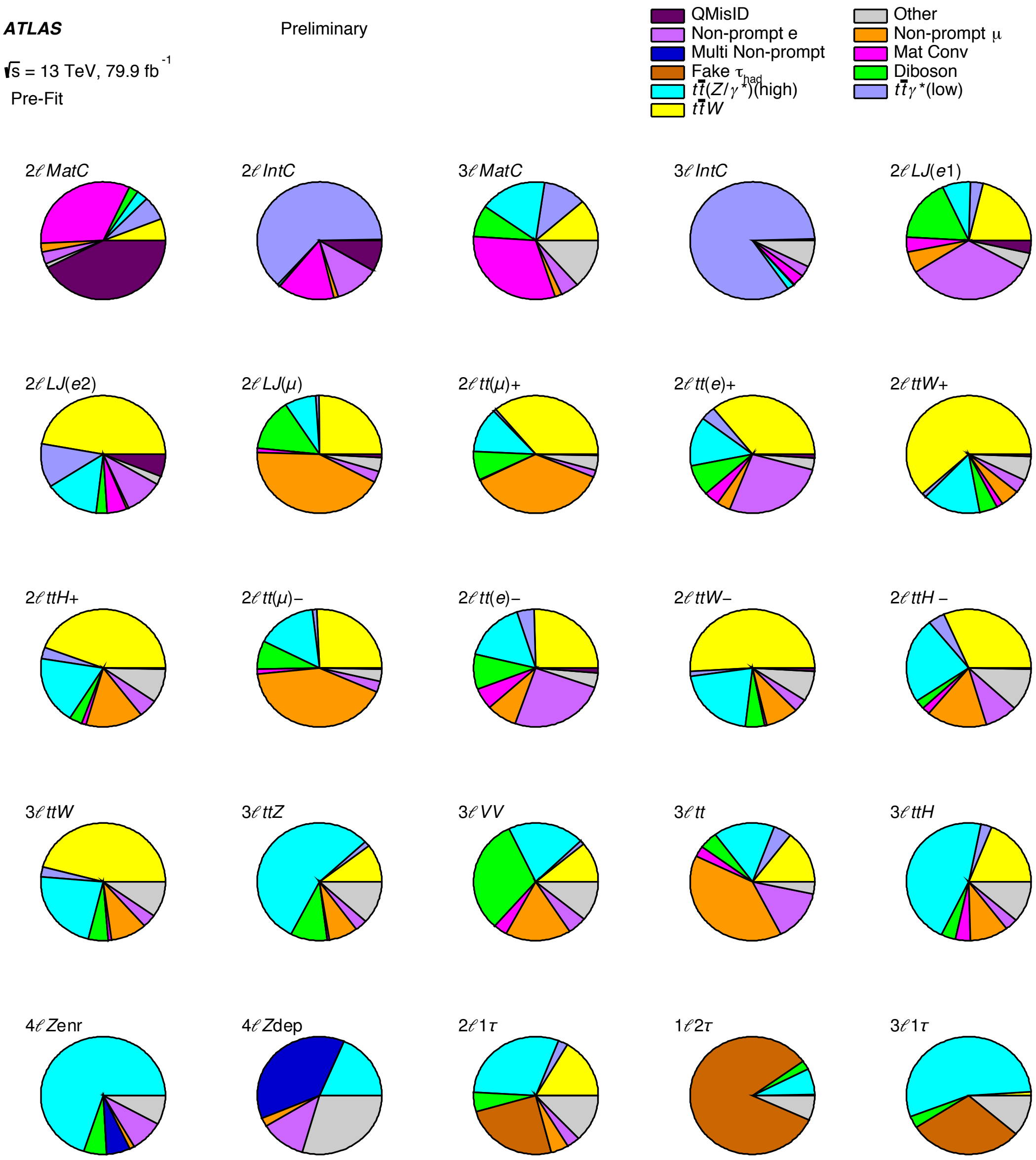
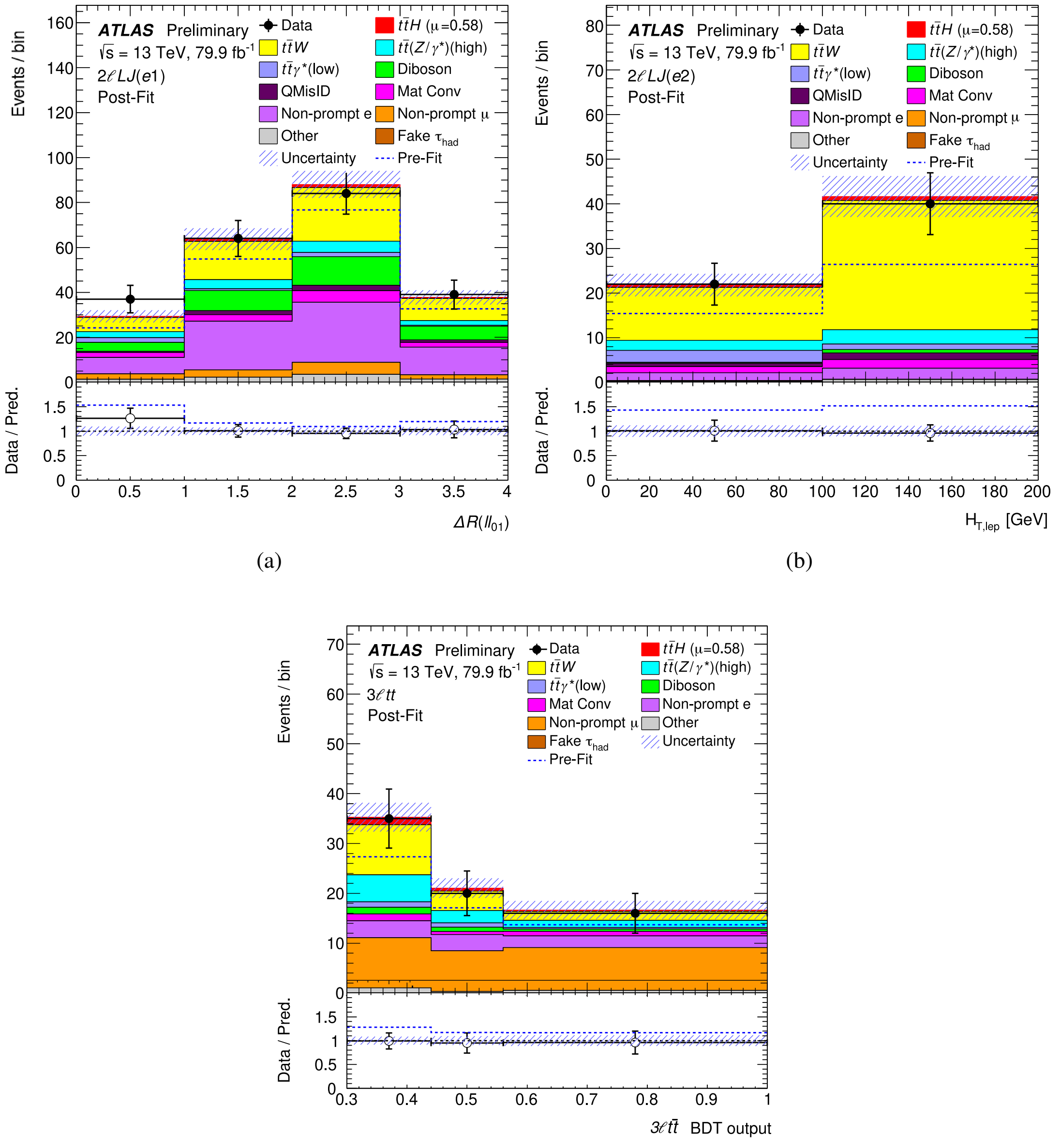


Figure 10: The fractional contributions of the various backgrounds to the total predicted background in each of the 25 event categories. The background estimation methods are described in Sect. 6: “Non-prompt (e/μ)", “Mat Conv", and “QMisID" refer to the data-driven background estimates (largely $t\bar{t}$ but also include other electroweak processes), and rare processes (tZ, tW, WtZ, $t\bar{t}WW$, VVV, $t\bar{t}t$, $t\bar{t}t\bar{t}$, and tH) are labelled as “Other".

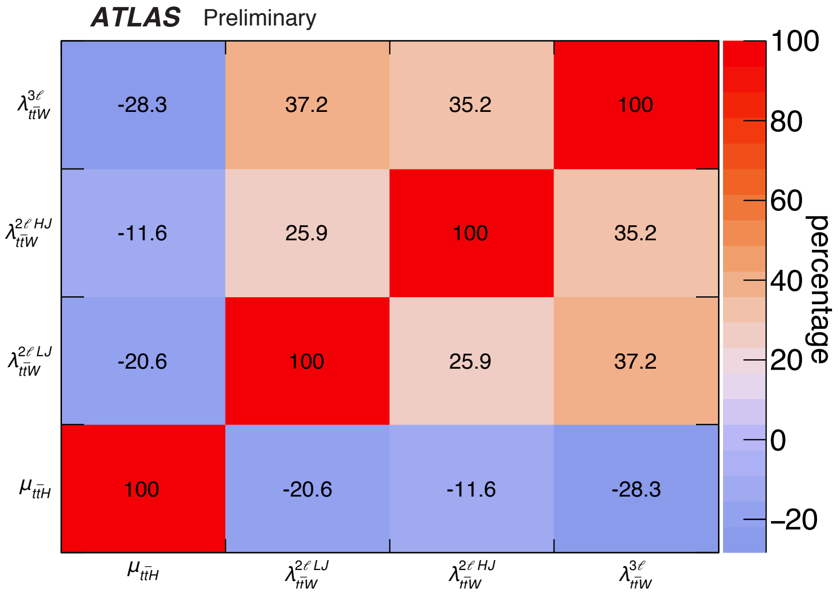


(a)

(b)

(c)

Figure 11: The distributions of the kinematic variables used in the three out of the four binned control regions: (a) $\Delta R(\ell,\ell)$ in the 2ℓLJ(e1) control region, (b) scalar sum of the lepton p_{T} ($H_{\mathrm{T,\,lep}}$) in the 2ℓLJ(e2) control region (see Figure 3(b) for the corresponding distribution in the 2ℓLJ(μ) control regions), and (c) BDT score for the $t\bar{t}$ category in the 3ℓtt control region. The background contributions after the likelihood fit ("Post-Fit") are shown as filled histograms. The total background before the fit ("Pre-Fit") is shown as a dashed blue histogram. The $t\bar{t}H$ ($t\bar{t}W$) signal, scaled according to the results of the fit, is shown as a filled red (yellow) histogram added to the post-fit background. The size of the combined statistical and systematic uncertainty in the total signal-plus-background prediction is indicated by the blue hatched band. The ratio of the data to the total post-fit prediction is shown in the lower panel. The last bin in each figure contains the overflow.



$\lambda_{t\bar{t}W}^{3\ell}$	-28.3	37.2	35.2	100
$\lambda_{t\bar{t}W}^{2\ell\,HJ}$	-11.6	25.9	100	35.2
$\lambda_{t\bar{t}W}^{2\ell\,LJ}$	-20.6	100	25.9	37.2
$\mu_{t\bar{t}H}$	100	-20.6	-11.6	-28.3
	$\mu_{t\bar{t}H}$	$\lambda_{t\bar{t}W}^{2\ell\,LJ}$	$\lambda_{t\bar{t}W}^{2\ell\,HJ}$	$\lambda_{t\bar{t}W}^{3\ell}$

Figure 12: Observed correlations between the signal strength μ and the normalisation factors for the $t\bar{t}W$ background in the profile likelihood fit to the data.

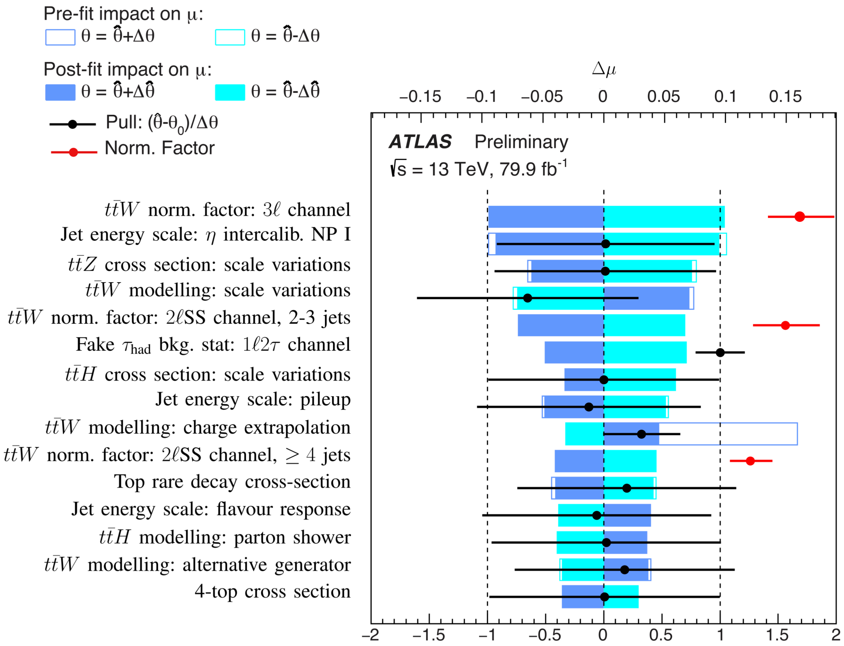


Figure 13: Ranking of the parameters included in the fit according to their impact on the signal strength μ. Only the 15 most highly ranked parameters are shown. The empty blue rectangles correspond to the pre-fit impact on μ and the filled blue ones to the post-fit impact on μ, both referring to the upper x-axis scale. The impact of each nuisance parameter (NP), $\Delta\mu$, is computed by comparing the nominal best-fit value of μ with the result of the fit when fixing the considered NP to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). The black points show the pulls of the NPs relative to their nominal values, θ_0. The nominal value for all NPs is $\theta_0 = 0$, with the exception of the NP associated with the limited sample size in the estimation of the fake τ_{had} background in the $1\ell 2\tau_{\mathrm{had}}$ channel, for which the nominal value is $\theta_0 = 1$. These pulls and their relative post-fit errors, $\Delta\hat{\theta}/\Delta\theta$, refer to the lower x-axis scale. The $t\bar{t}W$ normalisation factors (red points) also refer to the lower x-axis scale, and correspond to the floating normalisations of the $t\bar{t}W$ background, for which the pre-fit impact on μ is not defined. The nominal value of the $t\bar{t}W$ normalisation factors is 1, which corresponds to the $t\bar{t}W$ prediction based on the “updated theoretical cross section" discussed in Section 3. For experimental uncertainties that are decomposed into several independent sources, “NP I" corresponds to the first nuisance parameter, ordered by its impact on μ.

Table 6: Offline selection criteria applied to the 2ℓSS and 3ℓ channels in the cut-and-count analysis, together with the event categories defined. The common selection criteria for all channels are listed in the first line under the title “Common”. Same-charge (opposite-charge) lepton pairs are also referred to as same-sign (opposite-sign) with abbreviation SS (OS). In the 2ℓSS channel the leading and trailing (in p_{T}) SS leptons are denoted ℓ_0 and ℓ_1 respectively. In the 3ℓ channel, the OS lepton (w.r.t the SS pair) is denoted ℓ_0, but is not necessarily the one with highest p_{T}; the remaining SS leptons are denoted ℓ_1 (closest in ΔR to ℓ_0) and ℓ_2 (the remaining one). Same-flavour (SF), OS lepton pairs are referred to as SFOS pairs.

Channel	Selection criteria	
Common	$N_{\mathrm{jets}} \geq 2$ and $N_{b\text{-jets}} \geq 1$	
2ℓSS	Two SS very tight (T*) leptons, $p_{\mathrm{T}} > 20$ GeV	
	No τ_{had} candidates	
	$m(\ell_0\ell_1) > 12$ GeV	
	12 categories based on the following criteria:	
	• Number of jets: $N_{\mathrm{jets}} = 4$ or $N_{\mathrm{jets}} > 4$	
	• Number of b-tagged jets: $N_{b\text{-jets}} = 1$ or $N_{b\text{-jets}} > 1$	
	• Flavour of SS leptons: ee, $\mu\mu$ or opposite flavour (OF)	
3ℓ	Three light (L) leptons with $p_{\mathrm{T}} > 10$ GeV; sum of light-lepton charges = ±1	
	Two SS very tight (T*) leptons, $p_{\mathrm{T}} > 15$ GeV	
	One OS (w.r.t the SS pair) loose-isolated (L*) lepton, $p_{\mathrm{T}} > 10$ GeV	
	No τ_{had} candidates	
	$m(\ell^+\ell^-) > 12$ GeV for all SFOS pairs	
	$\lvert m(3\ell) - 91.2\ \mathrm{GeV}\rvert > 10$ GeV	
	12 categories based on the following criteria:	
	LjZPeak	$3 \leq N_{\mathrm{jets}} \leq 5$; 1 SFOS pair, $m(\ell^+\ell^-) \in Z_{\mathrm{win}}$
	HjZPeak	$N_{\mathrm{jets}} \geq 6$; 1 SFOS pair, $m(\ell^+\ell^-) \in Z_{\mathrm{win}}$
	LjHmZenr	$3 \leq N_{\mathrm{jets}} \leq 5$; $m(\ell_0\ell_1) > 70$ GeV; 1 SFOS pair, $m(\ell^+\ell^-) \notin Z_{\mathrm{win}}$
	HjHmZenr	$N_{\mathrm{jets}} \geq 6$; $m(\ell_0\ell_1) > 70$ GeV; 1 SFOS pair, $m(\ell^+\ell^-) \notin Z_{\mathrm{win}}$
	LjHmZdep_pp	$3 \leq N_{\mathrm{jets}} \leq 5$; $m(\ell_0\ell_1) > 70$ GeV; 0 SFOS pair; ℓ_1 and ℓ_2 positively charged
	LjHmZdep_mm	$3 \leq N_{\mathrm{jets}} \leq 5$; $m(\ell_0\ell_1) > 70$ GeV; 0 SFOS pair; ℓ_1 and ℓ_2 negatively charged
	LjLm1bZenr	$3 \leq N_{\mathrm{jets}} \leq 5$; $N_{b\text{-jets}} = 1$; $m(\ell_0\ell_1) < 70$ GeV; 1 SFOS pair, $m(\ell^+\ell^-) \notin Z_{\mathrm{win}}$
	LjLm1bZdep	$3 \leq N_{\mathrm{jets}} \leq 5$; $N_{b\text{-jets}} = 1$; $m(\ell_0\ell_1) < 70$ GeV; 0 SFOS pair
	LjLm2bZenr	$3 \leq N_{\mathrm{jets}} \leq 5$; $N_{b\text{-jets}} \geq 2$; $m(\ell_0\ell_1) < 70$ GeV; 1 SFOS pair, $m(\ell^+\ell^-) \notin Z_{\mathrm{win}}$
	LjLm2bZdep	$3 \leq N_{\mathrm{jets}} \leq 5$; $N_{b\text{-jets}} \geq 2$; $m(\ell_0\ell_1) < 70$ GeV; 0 SFOS pair
	HjLmZenr	$N_{\mathrm{jets}} \geq 6$; $m(\ell_0\ell_1) < 70$ GeV; 1 SFOS pair, $m(\ell^+\ell^-) \notin Z_{\mathrm{win}}$
	HjLmZdep	$N_{\mathrm{jets}} \geq 6$; $m(\ell_0\ell_1) < 70$ GeV; 0 SFOS pair

$Z_{\mathrm{win}} = [M_Z \pm 10\ \mathrm{GeV}]$, where M_Z denotes the Z-boson pole mass.

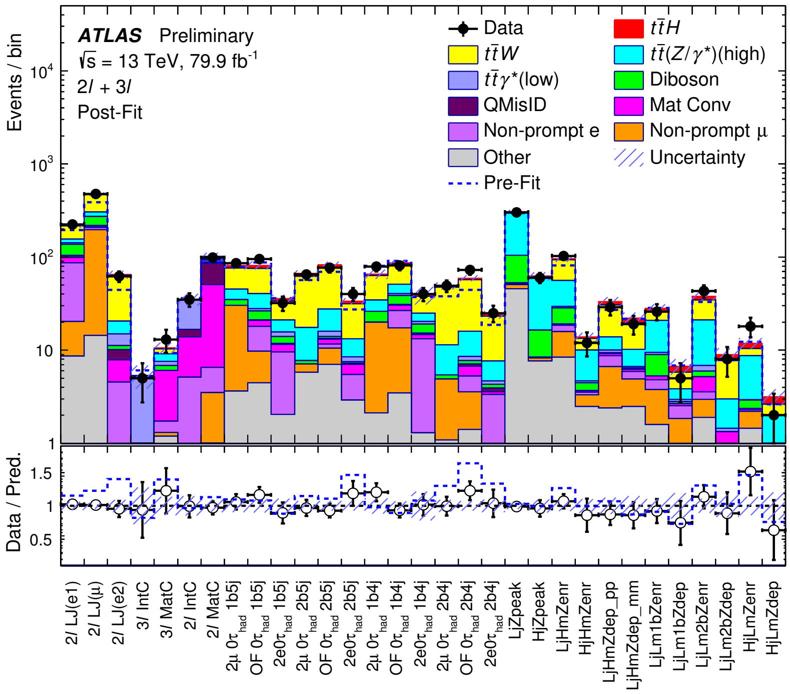


Figure 14: Comparison between data and prediction for the event yields in the categories in the 2ℓSS and 3ℓ channels from the cut-and-count cross-check analysis. The background contributions after the likelihood fit in the cut-and-count analysis (“Post-Fit") are shown as filled histograms. The total background before the fit (“Pre-Fit") is shown as a dashed blue histogram. The $t\bar{t}H$ signal, scaled according to the results of the fit, is shown as a filled red histogram added to the post-fit background. The size of the combined statistical and systematic uncertainty in the total signal-plus-background prediction is indicated by the blue hatched band. The ratio of the data to the total post-fit prediction is shown in the lower panel.