Search for the Associated Production of a Higgs Boson and a Top Quark Pair in Multilepton Final States with the ATLAS Detector

The ATLAS Collaboration

Abstract

The Yukawa coupling of the Higgs boson to the top quark is a key parameter of the Standard Model, and can be constrained using the associated production process $pp \rightarrow t\bar{t}H + X$. A search for this process using final states with multiple leptons, primarily targeting the decays $H \rightarrow WW^{*}$ and $H \rightarrow \tau\tau$, has been performed using 13.2 fb$^{-1}$ of data recorded by the ATLAS detector in 2015 and 2016 at a center of mass energy $\sqrt{s} = 13$ TeV. The best-fit value of the ratio of observed and Standard Model cross sections is $2.5 \pm 0.7 \ (\text{stat}) \ ^{+1.1}_{-0.9} \ (\text{syst})$, and an upper limit on this ratio of 4.9 (2.3 expected) is found at 95% confidence level.

© 2016 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.
1 Introduction

The Yukawa coupling of the Higgs boson to the top quark is a key parameter of the Standard Model (SM). It can be determined from the ratio of the top quark mass and Higgs field vacuum expectation value, from the cross section of $gg \rightarrow H$ production through a top quark loop, or from the cross section of the process $gg/qq \rightarrow t\bar{t}H$, which is a tree-level process at lowest order in perturbation theory. Comparison of these measurements has the potential to identify and disambiguate new physics effects that can modify the $t\bar{t}H$ production cross section relative to the SM expectation.

The ATLAS and CMS collaborations have searched for production of $t\bar{t}H$ in $pp$ collisions at the Large Hadron Collider (LHC) using data collected during LHC Run 1 at a center of mass energy $\sqrt{s}$ of 7 and 8 TeV, with analyses sensitive to $H \rightarrow WW^*$, $\tau\tau$, $b\bar{b}$, and $\gamma\gamma$ decays [1–5]. The combination of ATLAS and CMS Run 1 results yields a best fit of the ratio of observed and SM cross sections, $\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$, of $2.3^{+0.7}_{-0.6}$. The excess over the SM expectation $\mu_{t\bar{t}H} = 1$ is driven primarily by multileptonic final states [6]. Improvement on this measurement requires the addition of more data.

The cross section for $t\bar{t}H$ production in the SM rises by a factor of 3.9 as the center of mass energy is changed from 8 to 13 TeV [7, 8].

This note reports preliminary results of a search for $t\bar{t}H$ production using 13.2 fb$^{-1}$ of data collected with the ATLAS detector at $\sqrt{s} = 13$ TeV during 2015 and 2016. The search uses four final states distinguished by the number and flavor of leptons: two same-charge light leptons ($e$ or $\mu$) and no hadronically-decaying $\tau$ lepton candidates ($2\ell0\tau_{had}$); two same-charge light leptons and one hadronically-decaying $\tau$ lepton candidate ($2\ell1\tau_{had}$); three light leptons ($3\ell$); and four light leptons ($4\ell$). These signatures are primarily sensitive to $H \rightarrow WW^*$ (with subsequent decay to $\ell\nu\ell\nu$ or $\ell\nu j j$) and $H \rightarrow \tau\tau$ decays, and are effective in suppressing $t\bar{t}$ backgrounds. Backgrounds are estimated with a combination of simulation and data-driven techniques, and a global fit to the yields in all final states is used to extract the best estimate for the $t\bar{t}H$ production rate.

2 ATLAS detector

The ATLAS experiment [9] at the LHC is a multi-purpose particle detector with a forward-backward symmetric cylindrical geometry and a near $4\pi$ coverage in solid angle. It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer. The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon micro-strip, and transition radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity and longitudinal segmentation. A hadron (steel/scintillator-tile) calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The end-cap and forward regions are instrumented with LAr calorimeters for both EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large air-core toroid superconducting magnets with eight coils each. It includes a system of precision tracking chambers and fast detectors for triggering.

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

2
Table 1: Configurations used for event generation of signal and background processes. If only one parton distribution function (PDF) is shown, the same one is used for both the matrix element (ME) and parton shower generators; if two are shown, the first is used for the matrix element calculation and the second for the parton shower. “V” refers to production of an electroweak boson (W or Z/γ∗). “Tune” refers to the underlying-event tune of the parton shower generator. “MG5_aMC” refers to MadGraph5_aMC@NLO 2.2.1; “Pythia 6” refers to version 6.427; “Pythia 8” refers to version 2.7. Samples using Pythia 6 and Pythia 8 have heavy flavor hadron decays modeled by EvtGen 1.2.0 [10]. All samples include leading-logarithm photon emission, either modeled by the parton shower generator or by PHOTOS [11].

<table>
<thead>
<tr>
<th>Process</th>
<th>ME Generator</th>
<th>Parton Shower</th>
<th>PDF</th>
<th>Tune</th>
</tr>
</thead>
<tbody>
<tr>
<td>tHqbb</td>
<td>MG5_aMC</td>
<td>Pythia 8</td>
<td>CT10 [17]/NNPDF 2.3 LO</td>
<td>A14</td>
</tr>
<tr>
<td>tW(Z/γ∗)</td>
<td>MG5_aMC</td>
<td>Pythia 8</td>
<td>NNPDF 3.0 NLO/2.3 LO</td>
<td>A14</td>
</tr>
<tr>
<td>t(Z/γ∗)</td>
<td>MG5_aMC</td>
<td>Pythia 6 [22]</td>
<td>CTEQ6L1</td>
<td>Perugia2012 [23]</td>
</tr>
<tr>
<td>tW(Z/γ∗)</td>
<td>MG5_aMC</td>
<td>Pythia 8</td>
<td>NNPDF 2.3 LO</td>
<td>A14</td>
</tr>
<tr>
<td>ttW</td>
<td>MG5_aMC</td>
<td>Pythia 8</td>
<td>NNPDF 2.3 LO</td>
<td>A14</td>
</tr>
<tr>
<td>Wt single top</td>
<td>Powheg-BOX [25, 26]</td>
<td>Pythia 6</td>
<td>CT10/CTEQ6L1</td>
<td>Perugia2012</td>
</tr>
<tr>
<td>VV,qqVV, VVV</td>
<td>Sherpa 2.1.1 [27]</td>
<td>Sherpa</td>
<td>CT10</td>
<td>Sherpa default</td>
</tr>
<tr>
<td>Z → ℓ+ℓ−</td>
<td>Sherpa 2.2</td>
<td>Sherpa</td>
<td>NNPDF 3.0 NLO</td>
<td>Sherpa default</td>
</tr>
</tbody>
</table>

A two-level trigger system is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to a design maximum of 100 kHz. This is followed by a software-based trigger with a sustained average accepted event rate of 1 kHz.

3 Monte Carlo Event Generation

Event generator programs and configurations used for simulating the signal and background processes are shown in Table 1. Detailed descriptions of the generator configurations may be found in Refs. [28–31]. In all cases a Higgs boson mass of 125 GeV is assumed. Generated events are passed through a full GEANT4 [32] simulation of the ATLAS detector. Additional minimum-bias pp interactions (pileup) are modeled with the Pythia 8.1 generator with the MSTW2008LO parton distribution function (PDF) set [33] and the A2 tune [34], and are added to simulated events according to the luminosity profile of the recorded data.

Production of ttH, ttW, and ttZ are simulated with a next-to-leading order (NLO) QCD matrix element computed by MadGraph5_AMC@NLO, matched to the Pythia 8 parton shower generator. In the case of ttZ, the inclusive ttℓℓ matrix element is computed, including off-shell Z and γ∗ contributions with m(ℓℓ) > 5 GeV. For studies of systematic variations, samples with variations of the QCD factorization and renormalization scales by factors of 2 and 0.5 are used. Parton shower effects are studied by comparing the
nominal $t\bar{t}H$ sample with one with the same matrix element calculation but showered using HERWIG++, and by comparing the nominal $t\bar{t}V$ samples with ones with variations in the A14 PYTHIA 8 tune.

The overall cross section for $t\bar{t}H$ production, 507.1 fb, is computed at NLO in QCD and electroweak couplings \([35–41]\) as compiled in Refs. \([7, 8]\). It has uncertainties of $±5.0$% from QCD renormalization/factorization scale choice and $±3.6$% from parton distribution function uncertainties (including $\alpha_s$ uncertainties).

The cross sections for $t\bar{t}V$ production, including the process $pp \rightarrow t\bar{t}\ell^+\ell^- + X$ over the full $Z/\gamma^*$ mass spectrum, are computed at NLO in QCD and electroweak couplings using the configuration of Refs. \([12, 41]\). These have QCD scale uncertainties $\approx 12$% and PDF+$\alpha_s$ uncertainties of 3–4%. The total cross section used for $t\bar{t}\ell^+\ell^-$ (with $M(\ell^+\ell^-) > 5$ GeV) is 123.7 fb, and for $t\bar{t}W^\pm$ is 600.8 fb.

### 4 Object and Event Preselection

All analysis channels share a common trigger, jet, lepton and overall event preselection. These selections are detailed here; further channel-specific requirements are discussed in Section 5.

Events are required to have been selected by single electron or muon triggers. For data recorded in 2015, these achieve maximal efficiency for an isolated $e$ ($\mu$) with $p_T > 25$ (21) GeV [42]; in 2016 these thresholds are 25 GeV for both $e$ and $\mu$. The trigger requirement is 94 to 99% efficient, depending on the final state, for events passing final signal region selections. The primary vertex in an event is chosen as the vertex with the highest $\sum p_T^2$ of associated tracks [43]. Events with significant noise in the calorimeters or data corruption are removed.

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter that are associated with charged particle tracks reconstructed in the inner detector [44]. Only candidates with $p_T > 10$ GeV are considered. They are required to satisfy $|\eta_{\text{cluster}}| < 2.47$. Candidates in the transition region between different electromagnetic calorimeter components, $1.37 < |\eta_{\text{cluster}}| < 1.52$, are rejected. A multivariate likelihood discriminant combining shower shape and track information is used to distinguish real electrons from hadronic showers (fake electrons). Isolation variables are used to reduce the background from non-prompt electrons produced in hadron decays. Calorimetric isolation uses the sum of transverse energies of calorimeter clusters, excluding the electron candidate cluster itself, within a cone of $\Delta R = 0.2$ of the electron candidate. Track isolation uses the sum of transverse momenta of tracks consistent with originating at the primary vertex, excluding the electron candidate track, within a cone of $\Delta R = \min(0.2, 10 \text{ GeV}/p_T(e))$. For the object preselection, a loose electron discriminant working point is used, and an isolation selection tuned to be 99% efficient for prompt electrons in both calorimetric and tracking variables is chosen. To further reduce the non-prompt electron contribution, the track is required to be consistent with originating from the primary vertex; requirements are imposed on the transverse impact parameter significance and the longitudinal impact parameter, as shown in Table 2.

Muon candidates are reconstructed by combining inner detector tracks with track segments or full tracks in the muon spectrometer [46]. In the region $|\eta| < 0.1$, where muon spectrometer coverage is reduced, muon candidates are also reconstructed from inner detector tracks matched to isolated energy deposits in the calorimeters consistent with the passage of a minimum-ionizing particle. Candidates are required to satisfy $p_T > 10$ GeV and $|\eta| < 2.5$. Calorimetric and track isolation are defined for muon candidates similarly as for electron candidates, except that the track isolation uses a larger cone size at low $p_T$ ($\Delta R = \min(0.3, 10 \text{ GeV}/p_T(\mu))$). Calorimeter clusters within $\Delta R = 0.1$ of the muon candidate track are
Table 2: Tight and loose light lepton definitions. The lepton identification working points are documented in Refs. [44, 45]. Selections for tight leptons are applied on top of the selections for loose leptons. “99% eff” refers to isolation working points designed to be 99% efficient for isolated leptons at all $p_T$. (*) An additional “gradient” isolation working point is defined with efficiency and fake/non-prompt lepton rejection intermediate between the loose and tight isolation selections.

<table>
<thead>
<tr>
<th></th>
<th>Loose</th>
<th>Tight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Track isolation</td>
<td>99% eff.</td>
<td>99% eff.</td>
</tr>
<tr>
<td>Calorimeter isolation</td>
<td>99% eff.</td>
<td>99% eff.</td>
</tr>
<tr>
<td>Identification working point</td>
<td>Loose</td>
<td>Loose</td>
</tr>
<tr>
<td>Transverse impact parameter $</td>
<td>d_0</td>
<td>/\sigma_{d_0}$</td>
</tr>
<tr>
<td>$z$ impact parameter $</td>
<td>\Delta z_0 \sin \theta_\ell</td>
<td>$</td>
</tr>
</tbody>
</table>

excluded from the calorimeter isolation computation to avoid counting the energy deposit of the muon. The transverse impact parameter requirement for muon candidates is slightly tighter than for electrons, while the longitudinal impact parameter selection is the same.

Additional tighter isolation working points are applied on top of the loose selections in some cases to reduce fake and non-prompt lepton contributions. The tight working point, used in channels other than $4\ell$, is shown in Table 2; it is more restrictive than the 99% working point for lepton $p_T \leq 40$ GeV and reduces fake and non-prompt lepton backgrounds at lower $p_T$. An additional working point, “gradient,” is designed to give an increase in true isolated lepton efficiency from 87% at $p_T = 10$ GeV to 99% at 60 GeV, and is intermediate between the loose and tight definitions; it is used in the $4\ell$ channel to improve the signal efficiency.

Jets are reconstructed from calibrated topological clusters built from energy deposits in the calorimeters [47], using the anti-$k_t$ algorithm with a radius parameter $R = 0.4$ [48, 49]. Jets with energy contributions likely arising from noise or detector effects are removed from consideration [50], and only jets satisfying $p_T > 25$ GeV and $|\eta| < 2.5$ are used in this analysis. For jets with $p_T < 60$ GeV and $|\eta| < 2.4$, a jet-track association algorithm is used to confirm that the jet originates from the selected primary vertex, in order to reject jets arising from pileup collisions [51]. The average efficiency of this association is 92% per jet.

Hadronically decaying $\tau$ lepton candidates ($\tau_{had}$) are reconstructed from clusters in the calorimeters and associated inner detector tracks [52]. The candidates are required to have either one or three associated tracks, with a total charge of $\pm 1$. Candidates with $p_T > 25$ GeV and $|\eta| < 2.5$, excluding the electromagnetic calorimeter transition region, are considered. A boosted decision tree discriminant using calorimeter and tracking-based variables is used to identify $\tau_{had}$ candidates and reject generic jet backgrounds. The chosen working point has an efficiency of 75% (59%) for one- (three-)prong $\tau_{had}$ decays. Electrons which are reconstructed as one-prong $\tau_{had}$ candidates are removed using a sliding cut on the electron likelihood ID variable; the rejection factor (inverse efficiency) for electrons is $\approx 30–100$ depending on $\eta$ and $p_T$.

Jets containing $b$-hadrons are identified ($b$-tagged) via a multivariate discriminant combining information from the impact parameters of displaced tracks with topological properties of secondary and tertiary decay vertices reconstructed within the jet. The working point used for this search corresponds to an average
efficiency of 70% for $b$-jets with $p_T > 20$ GeV and $|\eta| < 2.5$ in $t\bar{t}$ events. The expected rejection factors against light and $c$-jets are 380 and 12, respectively.

To avoid double counting objects and to remove leptons likely originating from hadron decays, ambiguities are resolved in the following order: any $\tau_{\text{had}}$ candidate within $\Delta R = 0.2$ of an electron or muon candidate is removed; if an electron candidate shares an inner-detector track with a muon, the electron candidate is kept if the muon candidate track has no extension to the muon spectrometer, otherwise the muon candidate is kept; electron candidates which share an inner-detector track with a muon candidate are removed; any jet within $\Delta R = 0.2$ of an electron candidate is removed, then any electron candidate within $\Delta R = 0.4$ of a jet is removed; if a muon candidate and a jet lie within $\Delta R = 0.4$ of each other, the muon is kept and the jet is removed if the jet has two or fewer tracks, else the muon is removed and the jet is kept; and any jet within $\Delta R = 0.2$ of a $\tau_{\text{had}}$ candidate is considered only as a $\tau_{\text{had}}$ candidate in events with two light leptons. This algorithm is applied to the preselected objects and all channel-specific further quality requirements on the leptons and jets start with the surviving candidates after this procedure.

Missing transverse momentum ($E_T^{\text{miss}}$) is used to define certain additional selections for validation of background modeling. The missing transverse momentum vector is defined as minus the vector sum of the transverse momenta of all reconstructed and calibrated physics objects and remaining unclustered energy, the latter of which is estimated from low-$p_T$ tracks associated with the primary vertex but not assigned to a hard object [53]. The value of $E_T^{\text{miss}}$ for an event is the magnitude of this vector.

5 Channel Selections

Four final states, categorized by the number and flavor of leptons, are considered in this search: two same-charge light leptons with no hadronically-decaying $\tau$ lepton candidate ($2\ell 0\tau_{\text{had}}$), two same-charge light leptons with one hadronically-decaying $\tau$ lepton candidate ($2\ell 1\tau_{\text{had}}$), three light leptons ($3\ell$), and four light leptons ($4\ell$). The contribution to each of these final states from different Higgs boson decay modes, and the fraction of all $t\bar{t}H$ events captured by the event selections, is shown in Table 3. Based on the number of lepton candidates after the preselection requirements, events are placed into at most one of these categories. Additional requirements beyond the preselection classification are applied depending on the final state. After all object selections, at least one $e$ or $\mu$ candidate in the event is required to correspond to an object identified by the trigger system; this effectively requires one light lepton with $p_T > 25$ GeV (21 GeV for muons in 2015 data) in all channels.

5.1 $2\ell 0\tau_{\text{had}}$ Channel

The two light leptons are required to have the same charge. To reduce fake and non-prompt lepton backgrounds, tighter criteria are applied to the leptons than in the preselection, as detailed in Table 2. Electrons are additionally required to satisfy $|\eta_e| < 1.37$ to reduce the impact of electron charge misreconstruction through the process $e^+ \rightarrow \gamma^* \rightarrow e^+ e^- e^+ e^-$ occurring in detector material. Both leptons are required to have transverse momentum $p_T > 25$ GeV. There must be no $\tau_{\text{had}}$ candidates in the event. Events must have $\geq 5$ jets, of which $\geq 1$ must be $b$-tagged.

After this selection, the events are sorted into three subcategories based on the flavor of the leptons ($ee$, $e\mu$, and $\mu\mu$).
Table 3: Fraction of the expected $t\bar{t}H$ signal arising from different Higgs boson decay modes in each analysis category and acceptance times efficiency ($A \times \epsilon$) for $t\bar{t}H$ signal in each category. The decays contributing to the “other” column are dominantly $H \rightarrow \mu\mu$ and $H \rightarrow bb$. Rows may not add to 100% due to rounding. The acceptance times efficiency includes Higgs boson and top quark branching fractions, detector acceptance, and reconstruction and selection efficiency, and is computed relative to inclusive $t\bar{t}H$ production.

<table>
<thead>
<tr>
<th>Category</th>
<th>Higgs boson decay mode</th>
<th>$WW^*$</th>
<th>$\tau\tau$</th>
<th>$ZZ^*$</th>
<th>Other (\times 10^{-4})</th>
</tr>
</thead>
<tbody>
<tr>
<td>2$\ell$0$\tau_{\text{had}}$</td>
<td>77%</td>
<td>17%</td>
<td>3%</td>
<td>3%</td>
<td>14</td>
</tr>
<tr>
<td>2$\ell$1$\tau_{\text{had}}$</td>
<td>46%</td>
<td>51%</td>
<td>2%</td>
<td>1%</td>
<td>2.2</td>
</tr>
<tr>
<td>3$\ell$</td>
<td>74%</td>
<td>20%</td>
<td>4%</td>
<td>2%</td>
<td>9.2</td>
</tr>
<tr>
<td>4$\ell$</td>
<td>72%</td>
<td>18%</td>
<td>9%</td>
<td>2%</td>
<td>0.88</td>
</tr>
</tbody>
</table>

5.2 2$\ell$1$\tau_{\text{had}}$ Channel

The light leptons are required to pass the tight selections of Table 2 and to have the same charge. No additional $|\eta|$ requirement is placed on the electrons, but in $e^+e^-$ events the invariant mass of the electrons must satisfy $|m(e^+e^-) - 91.2\text{ GeV}| > 10\text{ GeV}$ to remove $Z \rightarrow e^+e^-$ events with a misreconstructed charge. One lepton must have $p_T > 25\text{ GeV}$, and the second must satisfy $p_T > 15\text{ GeV}$. There must be exactly one $\tau_{\text{had}}$ candidate, of opposite charge to the light leptons. There must be $\geq 4$ jets, of which $\geq 1$ must be $b$-tagged.

5.3 3$\ell$ Channel

The total charge of the three leptons in these events must be $\pm 1$. The lepton of opposite charge to that of the other two is designated “lepton 0.” Of the two remaining leptons, the one with smallest $\Delta R(\ell, \ell_0)$ is designated “lepton 1” and the other is “lepton 2.” Backgrounds arising from the addition of a fake or non-prompt lepton to an opposite-charge dilepton event (such as dileptonic $t\bar{t}$ production) will have that additional lepton be lepton 1 or 2, as it will be the same charge as a prompt lepton. Leptons 1 and 2 are required to pass the tight selections of Table 2. Leptons 1 and 2 must additionally satisfy $p_T > 20\text{ GeV}$. Lepton 0 is rarely fake or non-prompt, and so no additional requirements are imposed. To reject the $t\bar{t}Z$ background, all same-flavor $\ell^+\ell^-$ pairs in the event must satisfy $|m(\ell^+\ell^-) - 91.2\text{ GeV}| > 10\text{ GeV}$. To remove leptons from quarkonium decays, all same-flavor $\ell^+\ell^-$ pairs must satisfy $m(\ell^+\ell^-) > 12\text{ GeV}$. To remove potential backgrounds with $Z$ decays to $\ell\ell\gamma$ decays to $\ell\ell'\ell''$, where one lepton has very low momentum and is not reconstructed, the three-lepton invariant mass must satisfy $|m(3\ell) - 91.2\text{ GeV}| > 10\text{ GeV}$. There must be either 3 jets in the event, of which at least two are $b$-tagged, or $\geq 4$ jets, of which at least one must be $b$-tagged. No requirement on additional $\tau_{\text{had}}$ candidates is made, and any jets also reconstructed as $\tau_{\text{had}}$ candidates are treated only as jets.

5.4 4$\ell$ Channel

The sum of the charges of the leptons in these events must be zero. All leptons must pass the gradient isolation selection. To reject $t\bar{t}Z$ and $ZZ$ background, all same-flavor $\ell^+\ell^-$ pairs in the event must satisfy $|m(\ell^+\ell^-) - 91.2\text{ GeV}| > 10\text{ GeV}$. To remove leptons from quarkonium decays, all same-flavor $\ell^+\ell^-$ pairs
must satisfy \(m(\ell^+\ell^-) > 12\) GeV. The four-lepton invariant mass must satisfy \(100 \text{ GeV} < m(4\ell) < 350\) GeV to reduce contamination from \(Z \rightarrow 4\ell\) at low mass and \(t\bar{t}Z\) at high mass. To reduce contamination from other Higgs boson production processes and to ensure statistical independence from dedicated \(H \rightarrow ZZ^* \rightarrow 4\ell\) measurements, a Higgs boson veto \(|m(4\ell) - 125\) GeV| > 5 GeV is applied. There must be at least two jets in the event, of which at least one must be \(b\)-tagged. No requirement on \(\tau_{\text{had}}\) candidates is made, and any jets also reconstructed as \(\tau_{\text{had}}\) candidates are treated only as jets.

The signal region selections are summarized in the top rows of Table 4.

6 Backgrounds

Backgrounds in the signal regions are categorized into those in which all the selected leptons are produced in decays of electroweak bosons or \(\tau\) leptons (prompt leptons) and those in which at least one lepton arises from another source. In the latter case, the leptons arise from hadron decays or photon conversions (non-prompt), other interactions in detector material (charge misreconstruction or fake), or improper reconstruction of other particle species (fake).

6.1 Backgrounds with Prompt Leptons

These backgrounds are estimated using Monte Carlo (MC) simulation. In all final states, \(t\bar{t}V\) production is the largest background with prompt leptons. Diboson production, in particular \(WZ\), is subleading. Rarer processes (\(t(Z/\gamma^*)\), \(tW(Z/\gamma^*)\), \(t\bar{t}t\bar{t}\), \(t\bar{t}W^+W^-\)) also contribute at a total level comparable to diboson production. For the purposes of this analysis associated production of a single top quark and a Higgs boson is considered as a background, and at the SM cross section contributes negligibly.

Systematic uncertainties on the acceptance for the \(t\bar{t}V\) (\(V = W\) or \(Z/\gamma^*)\) backgrounds are derived using MC event simulation. Variations of hard process renormalization/factorization scale and PDF uncertainties are considered, as are uncertainties in the A14 parton shower tune for the \(t\bar{t}V\) samples. For \(t\bar{t}W\) only, a comparison between NLO event generation and a leading-order merged calculation is done to test possible matrix element-parton shower matching effects.

The largest backgrounds with prompt leptons are studied in validation regions (VR) that are disjoint from the signal regions and enhance specific processes. These permit checks of the overall normalization and jet multiplicity dependence of specific backgrounds. The validation regions are summarized in the lower rows of Table 4 and a comparison of expected and observed yields is shown in Table 5.

A validation region for \(t\bar{t}Z\) is defined by inverting the \(Z\) veto of the \(3\ell\) signal region, and in addition tightening the jet selection by requiring \(\geq 4\) jets of which \(\geq 2\) are \(b\)-tagged. To gain additional events for comparison, a looser selection (\(\geq 4\) jets of which at least one is \(b\)-tagged, or \(3\) jets of which at least 2 are \(b\)-tagged) is also used; the latter includes a larger fraction of \(WZ\) events. Invariant mass plots for these VRs are shown in Figure 1 and show agreement between data and expectation in both regions.

The small number of events in the validation regions means it is not possible to constrain the cross section for production of \(WZ\) or other diboson processes in association with \(c\) and \(b\)-quarks from data with high accuracy. To probe these processes, trilepton events with a \(Z\) candidate and a single \(b\)-tagged jet were studied as a validation region. This VR is expected from simulation to contain similar numbers of mistagged \(WZ + \text{light quark/gluon jet}\), \(WZ + c\), and \(WZ + b\bar{b}\) events, with the fraction of \(WZ + b\bar{b}\)
Table 4: Selections for the signal regions (SR) and validation regions (VR). The variable $H_{T,\text{jets}}$ is the scalar sum of transverse momenta for the considered jets. Same-flavor, opposite-charge lepton pairs are referred to as SFOC pairs. Trigger-matched leptons correspond to an object reconstructed by the trigger, and must have $p_T > 25$ GeV (21 GeV for muons in 2015 data). In all regions at least one selected light lepton is required to be trigger-matched.

<table>
<thead>
<tr>
<th>SR/VR</th>
<th>Channel</th>
<th>Selection criteria</th>
</tr>
</thead>
</table>
| SR    | 2/0$\tau_{\text{had}}$ | Two tight light leptons with $p_T > 25$, 25 GeV  
Sum of light lepton charges $\pm 2$  
Any electrons must have $|\eta_e| < 1.37$  
Zero $\tau_{\text{had}}$ candidates  
$N_{\text{jets}} \geq 5$ and $N_{b\text{-jets}} \geq 1$ |
| SR    | 2$\ell$ $\tau_{\text{had}}$ | Two tight light leptons, with $p_T > 25$, 15 GeV  
Sum of light lepton charges $\pm 2$  
Exactly one $\tau_{\text{had}}$ candidate, of opposite charge to the light leptons  
$|m(ee) - 91.2 \text{ GeV}| > 10 \text{ GeV for } ee \text{ events}$  
$N_{\text{jets}} \geq 4$ and $N_{b\text{-jets}} \geq 1$ |
| SR    | 3$\ell$  | Three light leptons; sum of light lepton charges $\pm 1$  
Two same-charge leptons must be tight and have $p_T > 20$ GeV  
$m(\ell^+\ell^-) > 12 \text{ GeV and } |m(\ell^+\ell^-) - 91.2 \text{ GeV}| > 10 \text{ GeV for all SFOC pairs}$  
$|m(3\ell) - 91.2 \text{ GeV}| > 10 \text{ GeV}$  
$N_{\text{jets}} \geq 4$ and $N_{b\text{-jets}} \geq 1$, or $N_{\text{jets}} = 3$ and $N_{b\text{-jets}} \geq 2$ |
| SR    | 4$\ell$  | Four light leptons; sum of light lepton charges $0$  
All leptons pass “gradient” isolation selection  
$m(\ell^+\ell^-) > 12 \text{ GeV and } |m(\ell^+\ell^-) - 91.2 \text{ GeV}| > 10 \text{ GeV for all SFOC pairs}$  
$100 \text{ GeV} < m(4\ell) < 350 \text{ GeV and } |m(4\ell) - 125 \text{ GeV}| > 5 \text{ GeV}$  
$N_{\text{jets}} \geq 2$ and $N_{b\text{-jets}} \geq 1$ |
| VR    | Tight $t\bar{t}Z$  | 3$\ell$ lepton selection  
At least one $\ell^+\ell^-$ pair with $|m(\ell^+\ell^-) - 91.2 \text{ GeV}| < 10 \text{ GeV}$  
$N_{\text{jets}} \geq 4$ and $N_{b\text{-jets}} \geq 2$ |
| VR    | Loose $t\bar{t}Z$  | 3$\ell$ lepton selection  
At least one $\ell^+\ell^-$ pair with $|m(\ell^+\ell^-) - 91.2 \text{ GeV}| < 10 \text{ GeV}$  
$N_{\text{jets}} \geq 4$ and $N_{b\text{-jets}} \geq 1$, or $N_{\text{jets}} = 3$ and $N_{b\text{-jets}} \geq 2$ |
| VR    | $WZ + 1 \text{ b-tag}$ | 3$\ell$ lepton selection  
At least one $\ell^+\ell^-$ pair with $|m(\ell^+\ell^-) - 91.2 \text{ GeV}| < 10 \text{ GeV}$  
$N_{\text{jets}} \geq 1$ and $N_{b\text{-jets}} = 1$ |
| VR    | $t\bar{t}W$  | 2/0$\tau_{\text{had}}$ lepton selection  
$2 \leq N_{\text{jets}} \leq 4$ and $N_{b\text{-jets}} \geq 2$  
$H_{T,\text{jets}} > 220 \text{ GeV for } ee \text{ and } e\mu \text{ events}$  
$E_T^{|\text{miss}|} > 50 \text{ GeV and } (m(ee) < 75 \text{ or } m(ee) > 105 \text{ GeV}) \text{ for } ee \text{ events}$ |
Table 5: Expected and observed event yields in validation regions (VR). The quoted uncertainties in the expectations include all systematic uncertainties. “Purity” indicates the fraction of events in the VR expected to arise from the targeted process ($t\bar{t}Z$ for the first two VRs, $WZ$ for the third, and $t\bar{t}W$ for the fourth).

<table>
<thead>
<tr>
<th>VR</th>
<th>Purity</th>
<th>Expected</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tight $t\bar{t}Z$</td>
<td>68%</td>
<td>32 ± 4</td>
<td>28</td>
</tr>
<tr>
<td>Loose $t\bar{t}Z$</td>
<td>58%</td>
<td>91 ± 12</td>
<td>89</td>
</tr>
<tr>
<td>$WZ + 1 b$-tag</td>
<td>33%</td>
<td>137 ± 27</td>
<td>147</td>
</tr>
<tr>
<td>$t\bar{t}W$</td>
<td>22%</td>
<td>51 ± 10</td>
<td>55</td>
</tr>
</tbody>
</table>

Figure 1: Invariant mass of leptons 0 and 1 for the a) tight and b) loose $t\bar{t}Z$ validation regions. The leptons are labeled in the same way as for the $3\ell$ signal region. Events away from the $Z$ peak are those satisfying the $Z$ selection with leptons 0 and 2. Non-prompt lepton backgrounds are estimated using data as described in Section 6.2.

increasing with jet multiplicity. The makeup of the signal regions is predicted to be similar to this VR at high jet multiplicity. The jet multiplicity in these events is modeled well as shown in Figure 2. This agreement indicates that the heavy flavor versus light flavor content of these events is simulated reasonably well. A 50% uncertainty on the diboson background normalization is adopted, motivated by the available statistical constraints on $WZ + c$ and $WZ + b\bar{b}$ production from this validation region.

No high-purity, high-yield validation region for $t\bar{t}W$ production is available. A preliminary measurement of $t\bar{t}W$ production with 13 TeV data at the LHC [54] finds a cross section consistent with the SM prediction with a total uncertainty of 56%, dominated by the limited event yield. Here, a $t\bar{t}W$ validation region is defined using the $2\ell0\ell_{\text{had}}$ lepton selection and requiring between two and four jets of which at least two are $b$-tagged. For $ee$ and $e\mu$ events the total scalar sum of jet transverse momenta ($HT_{\text{jets}}$) must exceed 220 GeV; in addition for $ee$ events the dilepton invariant mass must not be in the range 75–105 GeV and the missing transverse energy $E_{\text{T}}^{\text{miss}}$ must exceed 50 GeV. The results are shown in Figure 3; data are consistent with expectation.
Figure 2: Jet multiplicity in the $WZ + 1 \ b$-tag validation region. Non-prompt lepton backgrounds are estimated using data as described in Section 6.2.

Figure 3: Lepton flavor composition (a) and number of jets (b) for events in the $t\bar{t}W$ validation region. Non-prompt lepton and charge misreconstruction backgrounds (indicated as “QMisReco”) are estimated using data as described in Section 6.2.
6.2 Backgrounds with Non-Prompt Leptons and Fake \( \tau_{\text{had}} \) Candidates

Backgrounds involving non-prompt leptons and fake \( \tau_{\text{had}} \) candidates are estimated using data-driven methods. Monte Carlo studies indicate that, with the lepton identification requirements used, the fraction of \( e \) and \( \mu \) candidates in signal and validation regions that are not real \( e \) and \( \mu \) leptons is negligible. Monte Carlo simulation studies show that the dominant source of non-prompt leptons in all channels is decays of heavy flavor hadrons, primarily in \( t\bar{t} \) events. The background from conversion of high-\( p_T \) photons is tested using dedicated \( t\bar{t} \gamma \) simulation and found to be negligible. The inclusive \( t\bar{t} \) MC sample includes conversions due to photons emitted in initial- or final-state radiation using a leading-logarithm shower [11].

The \( 2\ell_0\tau_{\text{had}}, 2\ell_1\tau_{\text{had}}, 3\ell, \) and \( 4\ell \) final states have different estimation techniques for these backgrounds due to limitations of obtaining enough events in data control regions to estimate rates. In addition the trident process \( e^\pm \rightarrow \gamma (^{(*)} e^\pm \rightarrow e^\pm e^\pm e^\pm \) can result in a high-\( p_T \) electron with opposite charge to the original prompt electron (charge misreconstruction). A large fraction of \( \tau_{\text{had}} \) candidates are quark/gluon-initiated jets, and these are referred to as “fake \( \tau_{\text{had}} \).”

6.2.1 Charge misreconstruction

Charge misreconstruction affects the \( 2\ell_0\tau_{\text{had}} \) categories with electrons and the \( 2\ell_1\tau_{\text{had}} \) category when electrons are present (the impact on the \( \mu\mu + 0\tau_{\text{had}}, 3\ell, \) and \( 4\ell \) channels is negligible): an event with an opposite-charge light lepton pair is made to appear as a same-charge pair. The rates for charge misreconstruction are determined from data using dielectron events with \( m(ee) \) consistent with the \( Z \) mass, comparing the yields of opposite-charge and same-charge events. The charge misreconstruction rate is parametrized as a function of parent electron \( p_T \) and \( |\eta| \) and simulation is used to extrapolate for electron \( p_T > 130 \) GeV. The charge misreconstruction rates vary from \( 2 \times 10^{-4} \) for low-\( p_T \) electrons at small \( |\eta| \) to \( 10^{-2} \) for high-\( p_T \) electrons with \( 1.1 < |\eta| < 1.37 \) (electrons at higher \( |\eta| \) are not used in the \( 2\ell_0\tau_{\text{had}} \) channel). For high-\( p_T \) electrons with \( |\eta| \sim 2.47 \) (relevant for the \( 2\ell_1\tau_{\text{had}} \) channel) the charge misreconstruction rate reaches \( 10^{-1} \). The rates are verified by using them to estimate the yield of \( e^\pm e^\pm \) events compatible with the \( Z \) mass and associated with a \( b \)-jet. The prediction matches the observed data yield within a total uncertainty of 15\%. The charge misreconstruction rates are applied to opposite-charge dilepton events that otherwise pass the same selections as the \( 2\ell_0\tau_{\text{had}} \) subcategories to obtain the background estimate.

Charge misreconstruction for muons arises from improper reconstruction of the curvature of very high \( p_T \) tracks, and is negligible in this analysis.

6.2.2 Non-prompt leptons in the \( 2\ell_0\tau_{\text{had}} \) channels

Backgrounds with non-prompt leptons in the \( 2\ell_0\tau_{\text{had}} \) channels are estimated by extrapolating from control regions dominated by such backgrounds into the signal region. Leptons which pass the loose but not tight selections are termed “anti-tight” leptons and are highly enriched in non-prompt leptons. The ratio of non-prompt leptons passing tight and anti-tight selections is measured from data, and gives transfer factors for non-prompt electrons and muons. These are then applied to events that pass signal region selections except for having one tight and one anti-tight lepton, giving an estimate of the yield of events with non-prompt leptons in the signal region.
The transfer factors are obtained from same-charge $ee$ and $\mu\mu$ control regions with two, three, or four jets, of which at least one is $b$-tagged. The tight–tight regions otherwise follow the $2\ell 0\tau_{\text{had}}$ signal region selections. For the tight–anti-tight regions, one of the leptons must be tight and have caused the event to be accepted by the trigger, while the other must be anti-tight. Both leptons must have $p_T > 25$ GeV and pass the loose isolation and impact parameter selections. Charge-misreconstructed electron and prompt lepton backgrounds in both the tight and anti-tight regions are subtracted based on the data-driven charge misreconstruction rates and simulation. The transfer factors are computed as the ratio of the yields in the tight–tight and tight–anti-tight regions after the prompt lepton background contribution is subtracted. The main non-prompt lepton contribution to the control regions is $t\bar{t} \rightarrow \ell_\nu bqq'b$ production with an additional non-prompt lepton.

The systematic uncertainties on the transfer factors arise from the charge misreconstruction estimate, the extrapolation from the low jet multiplicity region to the signal regions, and statistical uncertainties in the data yields in the control regions. Expected prompt lepton backgrounds are small in all regions used for the transfer factor estimates and so uncertainties on them are not considered.

The systematic uncertainty from extrapolation includes a term from the self-consistency of the transfer factors in $t\bar{t}$ simulation and terms for the potential contamination of the control regions by non-$t\bar{t}$ processes that could cause the transfer factors to change with jet multiplicity. When applied to simulated events, the technique reproduces the yield of non-prompt MC events after the signal region selection within 25% or better. To test possible contamination from other processes, the stability of the non-prompt lepton background estimate in the signal region is checked while varying the minimum $p_T$ of the selected $b$-jet, restricting to two- or three-jet events only, or imposing a minimum $E_{\text{miss}}$ requirement. Good stability is seen and the assigned systematic uncertainty is 19%. Overall, the largest uncertainties in the non-prompt background predictions in the signal regions come from statistical uncertainties in the various control regions (19–24%) and closure and background composition systematic uncertainties (10–25%). The charge misreconstruction subtraction uncertainty on the non-prompt background yield (9–15%) is anticorrelated with the charge misreconstruction prediction in the signal region itself.

As the transfer factors are obtained from $e^\pm e^\pm$ and $\mu^\pm \mu^\pm$ events only, they can be validated in low jet multiplicity $e^\pm \mu^\pm$ events with two tight leptons. Agreement with data is seen within the 16% statistical uncertainty in the prediction.

### 6.2.3 Non-prompt leptons and fake $\tau_{\text{had}}$ candidates in the $2\ell 1\tau_{\text{had}}$ channel

These backgrounds fall into two categories: events that produce prompt same-charge light leptons with an additional fake $\tau_{\text{had}}$ candidate, and those that do not produce prompt same-charge light leptons. The latter processes enter this signal region by adding a non-prompt light lepton.

Processes that produce prompt same-charge lepton pairs include the $t\bar{t}H$ signal and the $t\bar{t}V$ backgrounds. The fraction of events from such processes in the signal region with fake $\tau_{\text{had}}$ candidates can be as large as 50%. The modeling of fake taus by the simulation is tested by selecting events with opposite charge light lepton pairs and one $\tau_{\text{had}}$ candidate that passes a loose selection but fails the nominal identification criterion. This region is dominated by $t\bar{t}$ decays with fake $\tau_{\text{had}}$ candidates. A scale factor of 1.52 is determined to correct the yield of events with a fake $\tau_{\text{had}}$ candidate predicted by MC, with an uncertainty of 14% determined from statistical uncertainties and possible MC generator and process dependence.
Backgrounds that involve a non-prompt light lepton are estimated, similarly to $2\ell 0\tau_{\text{had}}$, using anti-tight/tight light lepton transfer factors derived from a low-multiplicity region requiring two or three jets and at least one $b$-tagged jet. Due to low numbers of events in the control regions, a single overall transfer factor is derived combining both $e$ and $\mu$. Backgrounds with two prompt same-charge light leptons are estimated from simulation and subtracted when computing the transfer factor; the component of such backgrounds with fake $\tau_{\text{had}}$ candidates is adjusted by the above scale factor. The estimated charge misreconstruction background is subtracted as well. Because the fractions of true and fake $\tau_{\text{had}}$ candidates in events with non-prompt light leptons are expected to be similar in all control regions and the signal region, the fake $\tau_{\text{had}}$ component of the non-prompt light lepton background in the signal region is automatically accounted for in this method; a 24% systematic uncertainty is assigned on the fake rate from this extrapolation.

The uncertainty in the non-prompt lepton background estimate in the signal region is dominated by the limited number of events and is 76%.

### 6.2.4 Non-prompt leptons in the $3\ell$ channel

The dominant background with non-prompt leptons in this channel is dileptonic $t\bar{t}$ events with an additional non-prompt lepton. This and similar backgrounds are estimated, as in the other final states, with a transfer factor from anti-tight leptons. The transfer factors are taken from the $2\ell 0\tau_{\text{had}}$ analysis.

As described in Section 5.3, the lepton with charge opposite to that of the other two (lepton 0) is rarely non-prompt, and so only leptons 1 and 2 are subject to the non-prompt estimate. The anti-tight leptons in this category are defined as in the $2\ell 0\tau_{\text{had}}$ categories, but the $p_T$ requirement is lowered to 20 GeV and electrons are allowed in the full pseudorapidity range. The transfer factors used in the $2\ell 0\tau_{\text{had}}$ non-prompt background estimate are applied to $3\ell$ with one of either lepton 1 or 2 being anti-tight to obtain the background estimate in this category.

The systematic uncertainties from the $2\ell 0\tau_{\text{had}}$ study are applied to the $3\ell$ background estimate. An additional uncertainty is added for the use of the $p_T > 25$ GeV transfer factor for the $p_T > 20$ GeV leptons in this category; data studies indicate no expected bias from this source. Where relevant, the uncertainties are treated as correlated with the corresponding uncertainties in the $2\ell 0\tau_{\text{had}}$ background estimate.

The nominal non-prompt estimate is cross-checked in this channel by deriving the transfer factors from anti-tight to tight lepton regions using MC simulation. The contribution estimated from this method is consistent with the nominal value within uncertainties. In addition, the transfer factors derived for the $2\ell 0\tau_{\text{had}}$ channel using the alternate anti-tight lepton definitions are also used to derive a non-prompt estimate for $3\ell$, which is in good agreement with the nominal prediction.

### 6.2.5 Non-prompt leptons in the $4\ell$ channel

The rarity of $4\ell$ events means that applying the $2\ell 0\tau_{\text{had}}/3\ell$ non-prompt estimation technique will result in very large statistical uncertainties. It is also found that an appreciable fraction ($\sim 12\%$) of the $t\bar{t}H$ signal enters the signal region in events with at least one non-prompt lepton. Therefore, for this channel, MC simulation of the expected non-prompt lepton yields in the signal region is used, corrected by non-prompt scale factors (NPSFs) determined from data-MC comparison in trilepton control regions. This allows coherent handling of the non-prompt lepton contributions in both background and signal.
The NPSFs are determined separately for electrons and muons. There is also a separation of processes with isolated production of high-$p_T$ b-quarks ($t\bar{t}$-like) and processes with $g \rightarrow b\bar{b}$-dominated production of low-$p_T$ b-quarks ($Z + $jets-like). The $t\bar{t}$-like NPSFs are determined in a trilepton control region with 4$\ell$ lepton quality selections, no same-flavor opposite-charge lepton pairs, and one or two jets of which at least one has $p_T > 30$ GeV. The $Z +$jets-like NPSFs are determined in a trilepton control region with 4$\ell$ lepton quality selections, a Z boson candidate, one or two jets, low $E_T^{\text{miss}}$, and $m_T$ for the lepton not in the $Z$ candidate\(^2\). All together, there are four NPSFs, with values ranging from 0.77 ± 0.12 (muon $t\bar{t}$-like) to 1.35 ± 0.17 (electron $Z +$jets-like), where the quoted uncertainties are statistical only.

The method is verified using trilepton events with $\geq 3$ jets, excluding the $3\ell$ signal region. By requiring or vetoing Z boson candidates, the $t\bar{t}$-like and $Z +$jets-like NPSFs can be tested. The predicted and observed yields in the validation regions agree within statistical uncertainties. Additional cross checks are performed in four-lepton regions with relaxed lepton selections in order to increase yields; these again agree within statistical uncertainties. Systematic uncertainties are estimated from variations in the event selection (20%) and in the MC generators used (50%).

7 Systematic Uncertainties

The systematic uncertainties associated with the generation of background processes and the derivation of non-prompt lepton background estimates are discussed in Sections 3 and 6. The uncertainties in the 2$\ell$0$\tau_{\text{had}}$, 2$\ell$1$\tau_{\text{had}}$, and 3$\ell$ non-prompt estimates have the largest effects on the background estimates. The most important uncertainty arising from theoretical predictions is on the modeling of the acceptance for $t\bar{t}W$ events with high jet multiplicity requirements.

The value of $\mu_{t\bar{t}H}$ depends on the assumed SM cross section and acceptance for $t\bar{t}H$ production. The systematic uncertainties on the $t\bar{t}H$ signal process are obtained similarly to those for the $t\bar{t}V$ backgrounds. The uncertainties in these quantities have an impact on $\mu_{t\bar{t}H}$ comparable to the equivalent uncertainties in $t\bar{t}W$ and $t\bar{t}Z$ production.

The most important detector-related systematic uncertainty arises from the efficiency of the jet-to-vertex association method of approximately 2.5% per jet with $p_T < 60$ GeV, which becomes important in high-jet-multiplicity final states such as the ones considered here. The uncertainties in the modeling of pileup interactions and the jet energy scale also contribute significantly. The effect of uncertainties in the modeling of pileup are determined by varying the assumed inelastic cross section by $^{+16}_{-6}$%. Uncertainties in lepton reconstruction and trigger efficiencies have negligible impact.

The preliminary uncertainty on the combined 2015+2016 integrated luminosity is 2.9%. It is derived, following a methodology similar to that detailed in Refs. [55] and [56], from a preliminary calibration of the luminosity scale using x–y beam-separation scans performed in August 2015 and May 2016. Because this impacts both the estimation of the prompt backgrounds and the conversion from observed yields to an effective $t\bar{t}H$ cross section, it has an impact on $\mu_{t\bar{t}H}$ of nearly 20%.

A global fit to the observed and expected yields in all signal regions is performed to extract $\mu_{t\bar{t}H}$, as described in the following section. The systematic uncertainties are included in this fit with appropriate correlations between signal regions. The impact of the most important systematic uncertainties on $\mu_{t\bar{t}H}$, evaluated after the global fit, is shown in Table 6.

---

\(^2\) The transverse mass $m_T$ for a given lepton $\ell$ is defined as $\sqrt{2E_T^{\text{miss}}p_T,\ell\left(1 - \cos\Delta\phi(\ell, E_T^{\text{miss}})\right)}$. 

15
Table 6: Summary of the effects of the systematic uncertainties on \( \mu \). Due to correlations between the different sources of uncertainties, the total systematic uncertainty can be different from the sum in quadrature of the individual sources. The impact of the systematic uncertainties is evaluated after the fit described in Section 8.

<table>
<thead>
<tr>
<th>Uncertainty Source</th>
<th>( \Delta \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-prompt leptons and charge misreconstruction</td>
<td>+0.56 -0.64</td>
</tr>
<tr>
<td>Jet-vertex association, pileup modeling</td>
<td>+0.48 -0.36</td>
</tr>
<tr>
<td>( t\bar{t}W ) modeling</td>
<td>+0.29 -0.31</td>
</tr>
<tr>
<td>( t\bar{t}H ) modeling</td>
<td>+0.31 -0.15</td>
</tr>
<tr>
<td>Jet energy scale and resolution</td>
<td>+0.22 -0.18</td>
</tr>
<tr>
<td>( t\bar{t}Z ) modeling</td>
<td>+0.19 -0.19</td>
</tr>
<tr>
<td>Luminosity</td>
<td>+0.19 -0.15</td>
</tr>
<tr>
<td>Diboson modeling</td>
<td>+0.15 -0.14</td>
</tr>
<tr>
<td>Jet flavor tagging</td>
<td>+0.15 -0.12</td>
</tr>
<tr>
<td>Light lepton (( e, \mu )) and ( \tau ) had ID, isolation, trigger</td>
<td>+0.12 -0.10</td>
</tr>
<tr>
<td>Other background modeling</td>
<td>+0.11 -0.11</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>+1.1 -0.9</td>
</tr>
</tbody>
</table>

Table 7: Expected and observed yields in the six signal region categories in 13.2 fb\(^{-1}\) of data at \( \sqrt{s} = 13 \) TeV. Uncertainties in the background expectations due to systematic effects and MC statistics are shown. “Other” backgrounds include \( tZ, tWZ, tHqb, tHW, t\bar{t}l, t\bar{t}WW \), and triboson production. Values are obtained pre-fit, i.e., using the initial values of background systematic uncertainty nuisance parameters.

<table>
<thead>
<tr>
<th>Source</th>
<th>( 2\ell\tau_{\text{had}} ) ee</th>
<th>( 2\ell\tau_{\text{had}} ) e( \mu )</th>
<th>( 2\ell\tau_{\text{had}} ) ( \mu\mu )</th>
<th>( 3\ell )</th>
<th>( 4\ell )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t\bar{t}W )</td>
<td>2.9 \pm 0.7</td>
<td>9.1 \pm 2.5</td>
<td>6.6 \pm 1.6</td>
<td>0.8 \pm 0.4</td>
<td>6.1 \pm 1.3</td>
</tr>
<tr>
<td>( t\bar{t}(Z/\gamma^*) )</td>
<td>1.55 \pm 0.29</td>
<td>4.3 \pm 0.9</td>
<td>2.6 \pm 0.6</td>
<td>1.6 \pm 0.4</td>
<td>11.5 \pm 2.0</td>
</tr>
<tr>
<td>Diboson</td>
<td>0.38 \pm 0.25</td>
<td>2.5 \pm 1.4</td>
<td>0.8 \pm 0.5</td>
<td>0.20 \pm 0.15</td>
<td>1.8 \pm 1.0</td>
</tr>
<tr>
<td>Non-prompt leptons</td>
<td>12 \pm 6</td>
<td>12 \pm 5</td>
<td>8.7 \pm 3.4</td>
<td>1.3 \pm 1.2</td>
<td>20 \pm 6</td>
</tr>
<tr>
<td>Charge misreconstruction</td>
<td>6.9 \pm 1.3</td>
<td>7.1 \pm 1.7</td>
<td>—</td>
<td>0.24 \pm 0.03</td>
<td>—</td>
</tr>
<tr>
<td>Other</td>
<td>0.81 \pm 0.22</td>
<td>2.2 \pm 0.6</td>
<td>1.4 \pm 0.4</td>
<td>0.63 \pm 0.15</td>
<td>3.3 \pm 0.8</td>
</tr>
<tr>
<td>Total background</td>
<td>25 \pm 6</td>
<td>38 \pm 6</td>
<td>20 \pm 4</td>
<td>4.8 \pm 1.4</td>
<td>43 \pm 7</td>
</tr>
<tr>
<td>( t\bar{t}H ) (SM)</td>
<td>2.0 \pm 0.5</td>
<td>4.8 \pm 1.0</td>
<td>2.9 \pm 0.6</td>
<td>1.43 \pm 0.31</td>
<td>62 \pm 1.1</td>
</tr>
<tr>
<td>Data</td>
<td>26</td>
<td>59</td>
<td>31</td>
<td>14</td>
<td>46</td>
</tr>
</tbody>
</table>

8 Results

The expected backgrounds, \( t\bar{t}H \) signal, and observed data yields in each category are shown in Figure 4 and Table 7. The best-fit value of the \( t\bar{t}H \) signal strength \( \mu_{t\bar{t}H} \) is obtained using a maximum likelihood fit to the data yields of the six categories, which are treated as distinct Poisson terms in the likelihood function. The fit uses the profile-likelihood approach, in which systematic uncertainties are implemented as nuisance parameters with prior uncertainties which the fit is allowed to further constrain [57]. In practice the changes in the central values and uncertainties of the nuisance parameters after the fit are small. The likelihood function is also used to obtain 95\% confidence level (CL) upper limits on \( \mu_{t\bar{t}H} \) using the CL\(_{s}\) method [57, 58], where the test statistic (the profile likelihood ratio) is computed against the \( \mu_{t\bar{t}H} = 0 \) hypothesis. For the \( 4\ell \) channel the limits are obtained with the use of pseudo-Monte Carlo events, while for other channels and the combination the limits are obtained using asymptotic approximations for the
Figure 4: Pre-fit background and signal predictions and observed data yields for each signal region. The $t\bar{t}H$ prediction corresponds to the SM expectation ($\mu_{t\bar{t}H} = 1$). Charge misreconstruction backgrounds are indicated as “QMisReco.”

The best-fit value of $\mu_{t\bar{t}H}$, combining all channels, is $2.5 \pm 0.7 \text{ (stat)} ^{+1.1}_{-0.2} \text{ (syst)}$. The best-fit value of and 95% CL upper limit on $\mu_{t\bar{t}H}$ for each individual channel and the combination of all channels are shown in Figures 5 and 6 and Table 8. For the 4$\ell$ channel, the observation of zero events makes it difficult to quote a best-fit result with meaningful uncertainties, and a 68% confidence level CL$_s$ upper limit is shown instead. In the presence of the SM $t\bar{t}H$ signal, the fit is expected to return $\mu_{t\bar{t}H} = 1.0 ^{+0.7}_{-0.6} \text{ (stat)} ^{+0.9}_{-0.8} \text{ (syst)}$.

The $p$-value associated with the no-$t\bar{t}H$ hypothesis is 0.015 (2.2$\sigma$), and the $p$-value associated with the SM expectation $\mu_{t\bar{t}H} = 1$ is 0.09 (1.3$\sigma$).

Figures 7, 8, and 9 show the lepton flavor composition, jet, and $b$-tagged jet multiplicity of the events in the $2\ell 0_{\text{had}}$, $2\ell 1_{\text{had}}$, and $3\ell$ signal regions.
Figure 5: Best fit values of the $t\bar{t}H$ signal strength $\mu_{t\bar{t}H}$ by final state category and combined. The SM prediction is $\mu_{t\bar{t}H} = 1$. For the $4\ell$ category, as zero events are observed, a 68% CL upper limit is shown instead.

Figure 6: Upper limits on the $t\bar{t}H$ signal strength $\mu_{t\bar{t}H}$ at 95% CL by final state category and combined. The SM prediction is $\mu_{t\bar{t}H} = 1$. The median upper limit that would be set in the presence of a SM $t\bar{t}H$ signal ($\mu = 1$) is also shown.
<table>
<thead>
<tr>
<th>Category</th>
<th>Best fit ( \mu_{iH} )</th>
<th>Observed (expected)</th>
<th>Signal-injected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>95% CL upper limit</td>
<td>95% CL upper limit</td>
</tr>
<tr>
<td>2(\ell) (0\tau_{had})</td>
<td>4.0 (+1.2, +1.7) (-1.1, -1.3)</td>
<td>7.8 ((3.5, +1.7) (-1.0)</td>
<td>4.2</td>
</tr>
<tr>
<td>2(\ell) (1\tau_{had})</td>
<td>6.2 (+2.8, +2.3) (-2.3, -1.4)</td>
<td>12.9 ((5.9, +2.9) (-1.6)</td>
<td>6.3</td>
</tr>
<tr>
<td>3(\ell)</td>
<td>0.5 (+1.2, +1.2) (-1.0, -1.3)</td>
<td>3.9 ((3.5, +1.5) (-1.0)</td>
<td>4.3</td>
</tr>
<tr>
<td>4(\ell)</td>
<td>(&lt;2.2) (68% CL)</td>
<td>5.2 ((6.6, +2.9) (-1.4)</td>
<td>7.4</td>
</tr>
<tr>
<td>Combined</td>
<td>2.5 (+0.7, +1.1) (-0.7, -0.9)</td>
<td>4.9 ((2.3, +1.1) (-0.6)</td>
<td>3.1</td>
</tr>
</tbody>
</table>

Figure 7: Characteristics of events in the 2\(\ell\)0\(\tau_{had}\) signal region: (a) lepton flavor composition; (b) 10\(\times\) the number of \(b\)-tagged jets plus the total number of jets. The signal is set to the SM expectation (\(\mu_{iH} = 1\)) and the background expectation is pre-fit (using initial values of the background systematic uncertainty nuisance parameters). The hatched region shows the total uncertainty on the background plus SM signal prediction in each bin. Charge misreconstruction backgrounds are indicated as “QMisReco.”
Figure 8: Characteristics of events in the $2\ell 1\tau_{\text{had}}$ signal region: (a) lepton flavor composition; (b) $10\times$ the number of $b$-tagged jets plus the total number of jets. The signal is set to the SM expectation ($\mu_{tH} = 1$) and the background expectation is pre-fit (using initial values of the background systematic uncertainty nuisance parameters). The hatched region shows the total uncertainty on the background plus SM signal prediction in each bin. Charge misreconstruction backgrounds are indicated as “QMisReco.”

Figure 9: Characteristics of events in the $3\ell$ signal region: (a) lepton flavor composition; (b) $10\times$ the number of $b$-tagged jets plus the total number of jets. The signal is set to the SM expectation ($\mu_{tH} = 1$) and the background expectation is pre-fit (using initial values of the background systematic uncertainty nuisance parameters). The hatched region shows the total uncertainty on the background plus SM signal prediction in each bin.
9 Conclusion

A search for $t\bar{t}H$ production in multilepton final states has been performed using $13.2\, \text{fb}^{-1}$ of proton-proton collision data at $\sqrt{s} = 13\, \text{TeV}$ recorded by the ATLAS experiment at the LHC. The best-fit result of the ratio $\mu_{t\bar{t}H}$ of the $t\bar{t}H$ production rate to the Standard Model expectation is $2.5 \pm 0.7$ (stat) $^{+1.1}_{-0.9}$ (syst), which is consistent with the Standard Model expectation. At 95% confidence level, $\mu_{t\bar{t}H}$ is found to be less than 4.9 (2.3 expected).
### Appendix

Table 9: Post-fit background, signal, and observed yields in the six signal region categories in 13.2 fb$^{-1}$ of data at $\sqrt{s} = 13$ TeV. Uncertainties in the background expectations due to systematic effects and MC statistics are shown. “Other” backgrounds include $t\bar{t}Z$, $t\bar{t}WZ$, $t\bar{t}Hqb$, $t\bar{t}HW$, $t\bar{t}t\bar{t}$, $t\bar{t}t\bar{t}W$, and triboson production. Background expectations have been updated to reflect the values of systematic uncertainty nuisance parameters after the fit to data. The prediction and uncertainties for $t\bar{t}H$ reflect the best-fit production rate of $2.5^{+1.3}_{-1.1}$ times the Standard Model expectation. The uncertainty on the total background estimation is smaller than for the pre-fit values due to anticorrelations between the nuisance parameters obtained during the fit.

<table>
<thead>
<tr>
<th></th>
<th>2/$0\tau_{\text{had}} e\mu</th>
<th>2/0\tau_{\text{had}} e\mu</th>
<th>2/0\tau_{\text{had}} \mu\mu</th>
<th>2/1\tau_{\text{had}}</th>
<th>3\ell</th>
<th>4\ell</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}W$</td>
<td>3.2 ± 0.9</td>
<td>10.4 ± 2.9</td>
<td>7.4 ± 1.8</td>
<td>1.0 ± 0.5</td>
<td>6.5 ± 1.5</td>
<td>—</td>
</tr>
<tr>
<td>$t\bar{t}(Z/\gamma^*)$</td>
<td>1.53 ± 0.29</td>
<td>4.3 ± 0.9</td>
<td>2.6 ± 0.6</td>
<td>1.7 ± 0.4</td>
<td>11.3 ± 1.9</td>
<td>1.08 ± 0.20</td>
</tr>
<tr>
<td>Diboson</td>
<td>0.40 ± 0.26</td>
<td>2.6 ± 1.5</td>
<td>0.8 ± 0.5</td>
<td>0.21 ± 0.15</td>
<td>1.9 ± 1.0</td>
<td>0.04 ± 0.04</td>
</tr>
<tr>
<td>Non-prompt leptons</td>
<td>9 ± 4</td>
<td>11 ± 4</td>
<td>8.9 ± 3.3</td>
<td>1.9 ± 1.6</td>
<td>15 ± 4</td>
<td>0.17 ± 0.10</td>
</tr>
<tr>
<td>Charge misreconstruction</td>
<td>7.2 ± 1.4</td>
<td>7.6 ± 1.8</td>
<td>—</td>
<td>0.25 ± 0.03</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Other</td>
<td>0.83 ± 0.16</td>
<td>2.3 ± 0.6</td>
<td>1.5 ± 0.4</td>
<td>0.66 ± 0.16</td>
<td>3.4 ± 0.8</td>
<td>0.12 ± 0.05</td>
</tr>
<tr>
<td>Total background</td>
<td>22.2 ± 3.4</td>
<td>39 ± 5</td>
<td>21 ± 4</td>
<td>5.7 ± 1.7</td>
<td>39 ± 5</td>
<td>1.42 ± 0.24</td>
</tr>
<tr>
<td>$t\bar{t}H$ (2.5 × SM)</td>
<td>5.3 ± 1.8</td>
<td>13 ± 4</td>
<td>7.6 ± 2.5</td>
<td>4.0 ± 1.2</td>
<td>16 ± 5</td>
<td>1.5 ± 0.5</td>
</tr>
<tr>
<td>Data</td>
<td>26</td>
<td>59</td>
<td>31</td>
<td>14</td>
<td>46</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 10: Post-fit background and signal predictions and observed data yields for each signal region. Background expectations have been updated to reflect the values of systematic uncertainty nuisance parameters after the fit to data. The $t\bar{t}H$ prediction corresponds to the best-fit value ($\mu_{t\bar{t}H} = 2.5^{+1.2}_{-1.1}$) and the displayed total uncertainties reflect the uncertainty in $t\bar{t}H$ as well as the backgrounds. Charge misreconstruction backgrounds are indicated as “QMisReco.”
Figure 11: Effect of the fifteen most important systematic uncertainty nuisance parameters \( \theta \) on the signal strength \( \mu = \mu_{ttH} \) and constraints on the nuisance parameters from the fit. The blue and cyan bars show the \( \pm 1\sigma \) impact of the nuisance parameter on the signal strength (shown on the top axis). The points and associated error bars show the best-fit values of the nuisance parameters and post-fit uncertainties on the nuisance parameters (shown on the bottom axis). The open bars show the effect of the systematic uncertainties on \( \mu_{ttH} \) before the fit, and the solid bars show the effect after the fit. The nuisance parameters are initially normalized to \( 0 \pm 1 \). The dotted vertical lines show \( \pm 1\sigma \) excursions of the nuisance parameters from their initial values.
Figure 12: Event display for a candidate $ee$ event in the $2\ell 0\tau_{\text{had}}$ category. The blue tracks are the two selected electrons. Green and yellow bars indicate energy deposits in the electromagnetic (liquid argon) and hadronic (tile) calorimeters, respectively. In the inset display, the three azure cones are $b$-tagged jets and the yellow cones are the six non-$b$-tagged jets.
Figure 13: Event display for a candidate $e\mu_{\text{had}}$ event in the $2\ell 1_{\text{had}}$ category. The blue track is the selected electron; the red track is the selected muon; and the white cone is the $\tau_{\text{had}}$ candidate. The azure cone is the selected $b$-tagged jet, and the three yellow cones are the non-$b$-tagged jets. Green and yellow bars indicate energy deposits in the electromagnetic (liquid argon) and hadronic (tile) calorimeters, respectively.
Figure 14: Event display for a candidate $3\mu$ event in the $3\ell$ category. The red tracks are the selected muons. The two blue cones are the selected $b$-tagged jets, and the five yellow cones are the non-$b$-tagged jets. Green and yellow bars indicate energy deposits in the electromagnetic (liquid argon) and hadronic (tile) calorimeters, respectively.
Figure 15: Characteristics of events in the 2ℓ0τ_{had} signal region, by flavor category. The variable plotted is 10× the number of b-tagged jets plus the total number of jets for (a) ee events, (b) eμ events, (c) μμ events. The signal is set to the SM expectation (μttH = 1) and the background expectation is pre-fit (using initial values of the background systematic uncertainty nuisance parameters). The hatched region shows the total uncertainty on the background plus SM signal prediction in each bin. Charge misreconstruction backgrounds are indicated as “QMisReco.”
Figure 16: Additional characteristics of events in the $2\ell 1\tau_{\text{had}}$ signal region: (a) number of jets; (b) number of tracks in the $\tau_{\text{had}}$ candidate. The signal is set to the SM expectation ($\mu_{t\bar{t}}H = 1$) and the background expectation is pre-fit (using initial values of the background systematic uncertainty nuisance parameters). The hatched region shows the total uncertainty on the background plus SM signal prediction in each bin. Charge misreconstruction backgrounds are indicated as “QMisReco.”

Figure 17: Expected contribution to the background in each channel from various sources, using values of the background estimates before the fit. Charge misreconstruction backgrounds are indicated as “QMisReco.”
References


[29] 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.


[54] ATLAS Collaboration, *Measurement of the \( t\bar{t}Z \) and \( t\bar{t}W \) production cross sections in multilepton final states using 3.2 fb\(^{-1}\) of pp collisions at 13 TeV at the LHC*, (2016), ATLAS-CONF-2016-003, url: https://cds.cern.ch/record/2138947.


