Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector

M. Aaboud et al.*
(ATLAS Collaboration)

(Received 27 December 2017; published 9 April 2018)

A search for the associated production of the Higgs boson with a top quark pair ($t\bar{t}H$) is reported. The search is performed in multilepton final states using a data set corresponding to an integrated luminosity of 36.1 fb$^{-1}$ of proton-proton collision data recorded by the ATLAS experiment at a center-of-mass energy $\sqrt{s} = 13$ TeV at the Large Hadron Collider. Higgs boson decays to $WW^*$, $\tau\tau$, and $ZZ^*$ are targeted. Seven final states, categorized by the number and flavor of charged-lepton candidates, are examined for the presence of the Standard Model Higgs boson with a mass of 125 GeV and a pair of top quarks. An excess of events over the expected background from Standard Model processes is found with an observed significance of 4.1 standard deviations, compared to an expectation of 2.8 standard deviations. The best fit for the $t\bar{t}H$ production cross section is $\sigma(t\bar{t}H) = 790^{+230}_{-210}$ fb, in agreement with the Standard Model prediction of $507^{+35}_{-30}$ fb. The combination of this result with other $t\bar{t}H$ searches from the ATLAS experiment using the Higgs boson decay modes to $b\bar{b}$, $\gamma\gamma$ and $ZZ^* \rightarrow 4\ell$, has an observed significance of 4.2 standard deviations, compared to an expectation of 3.8 standard deviations. This provides evidence for the $t\bar{t}H$ production mode.

DOI: 10.1103/PhysRevD.97.072003

I. INTRODUCTION

The study of the origin of electroweak symmetry breaking is one of the key goals of the Large Hadron Collider (LHC) [1]. In the Standard Model (SM) [2–5], the symmetry is broken through the introduction of a complex scalar field doublet, leading to the prediction of the existence of one physical neutral scalar particle, commonly known as the Higgs boson [6–10]. The discovery of a Higgs boson with a mass of approximately 125 GeV by the ATLAS [11] and CMS [12] Collaborations was a crucial milestone. Measurements of its properties performed so far [13–18] are consistent with the predictions for the SM Higgs boson.

These measurements rely primarily on studies of the bosonic decay modes, $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^*$, and $H \rightarrow WW^*$; therefore it is crucial to also measure the Yukawa interactions, which are predicted to account for the fermion masses [3,19]. Thus far, only the Yukawa coupling of the Higgs boson to $\tau$ leptons has been observed [18,20–22] and evidence for the Yukawa coupling of the Higgs boson to $b$ quarks has been found through direct searches [23–25].

The Yukawa coupling of the Higgs boson to the top quark, the heaviest particle in the SM, is expected to be of the order of unity and could be particularly sensitive to effects beyond the SM (BSM). A measurement of the ratio of this coupling to the SM prediction of $0.87 \pm 0.15$ has been obtained from the combined fit of the ATLAS and CMS Higgs boson measurements [18]. This depends largely on the indirect measurement using the top quark contribution to gluon-gluon fusion production and diphoton decay loops for which no BSM contribution is assumed. Therefore, a direct measurement of the coupling of the Higgs boson to top quarks is highly desirable to disentangle any deviation in the top quark’s Yukawa coupling due to couplings to new particles and to significantly reduce the model dependence in the extraction of the top quark’s Yukawa coupling.

A direct measurement can be achieved by measuring the rate of the process in which the Higgs boson is produced in association with a pair of top quarks, $gg/q\bar{q} \rightarrow t\bar{t}H$, which is a tree-level process at lowest order in perturbation theory. Although the $t\bar{t}H$ production cross section at the LHC is 2 orders of magnitude smaller than the total Higgs boson production cross section, the distinctive signature from the top quarks in the final state gives access to many Higgs boson decay modes. The ATLAS and CMS Collaborations have searched for $t\bar{t}H$ production using proton-proton ($pp$) collision data collected during LHC run 1 at center-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV, with analyses mainly sensitive to $H \rightarrow WW^*$, $H \rightarrow \tau^+\tau^-$, $H \rightarrow b\bar{b}$ and $H \rightarrow \gamma\gamma$ [26–30]. The combination of these results yields a
best fit of the ratio of observed and SM cross sections, \( \mu = \sigma / \sigma_{\text{SM}} \) of 2.3^{+0.7}_{-0.6} \ [18].

The ongoing data taking at the LHC at an increased center-of-mass energy of \( \sqrt{s} = 13 \) TeV allows the collection of a larger data set because of an increased \( t\bar{t}H \) production cross section relative to run 1 \ [31–35]. This article reports the results of a search for \( t\bar{t}H \) production using a data set corresponding to an integrated luminosity of 36.1 fb\(^{-1}\) collected with the ATLAS detector at \( \sqrt{s} = 13 \) TeV during 2015 and 2016. Examples of tree-level Feynman diagrams are given in Fig. 1, where the Higgs boson is shown decaying to \( WW^*/ZZ^* \) or \( \tau \tau \). The search uses seven final states distinguished by the number and flavor of charged-lepton (electron, muon and hadronically decaying \( \tau \) lepton) candidates, denoted \( l \). In the following, the term “light lepton,” denoted \( l' \), refers to either electrons or muons and is understood to mean both particle and antiparticle as appropriate. These signatures are primarily sensitive to the decays \( H \rightarrow WW^* \) (with subsequent decay to \( lv\bar{v}l'v' \) or \( llqq \)), \( H \rightarrow \tau^+\tau^- \) and \( H \rightarrow ZZ^* \) (with subsequent decay to \( l\nu\nu \) or \( llqq \)), and their selection is designed to avoid any overlap with the ATLAS searches for \( t\bar{t}H \) production with \( H \rightarrow b\bar{b} \) \ [36], \( H \rightarrow \gamma\gamma \) \ [37] and \( H \rightarrow ZZ^* \rightarrow 4\ell \) \ [38] decays.

Backgrounds to the signal arise from associated production of a top quark pair and a \( W \) or \( Z \) (henceforth \( V \)) boson. Additional backgrounds arise from \( t\bar{t} \) production with leptons from heavy-flavor hadron decays and additional jets (nonprompt leptons) and other processes where the electron charge is incorrectly assigned (labeled as “\( q \) mis-id”) or where jets are incorrectly identified as \( \tau \) candidates. Backgrounds are estimated with a combination of simulation and data-driven techniques (labeled as “prefit”), and then a global fit to the data, in all final states, is used to extract the best estimate for the \( t\bar{t}H \) production rate and adjust the background predictions (labeled as “postfit”).

The article is organized as follows. Section II introduces the ATLAS detector; Sec. III describes the Monte Carlo (MC) simulation samples as well as the recorded data used for this analysis. The reconstruction and identification of the physics objects are discussed in Sec. IV. The event selection and classification are explained in Sec. V. Section VI describes the methods used to estimate the backgrounds. The theoretical and experimental uncertainties are discussed in Sec. VII. The results are presented in Sec. VIII, and the combination with the three other ATLAS searches for \( t\bar{t}H \) production mentioned above is reported in Sec. IX.

II. ATLAS DETECTOR

The ATLAS experiment \ [39] at the LHC is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and a near 4\( \pi \) coverage in solid angle.\(^1\) It consists of an inner tracking detector surrounded by a superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer. The inner tracking detector, covering the pseudorapidity range \( |\eta| < 2.5 \), consists of silicon pixel and silicon microstrip tracking detectors inside a transition-radiation tracker that covers \( |\eta| = 2.0 \). It includes, for the \( \sqrt{s} = 13 \) TeV running period, a newly installed innermost pixel layer, the insertable B layer \ [40]. Lead/liquid-argon

---

\(^1\)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 \( \Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \).
(LAr) sampling calorimeters provide electromagnetic (EM) energy measurements for $|\eta| < 2.5$ with high granularity and longitudinal segmentation. A hadron calorimeter consisting of steel and scintillator tiles covers the central pseudorapidity range ($|\eta| < 1.7$). The end cap and forward regions are instrumented with LAr calorimeters for 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 ($|\eta| < 2.7$) and fast detectors for triggering ($|\eta| < 2.4$). A two-level trigger system is used to select events [41]. 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 about 1 kHz.

III. DATA AND MONTE CARLO SAMPLES

The data were collected by the ATLAS detector during 2015 and 2016 with a peak instantaneous luminosity of $L = 1.4 \times 10^{34}$ cm$^{-2}$ s$^{-1}$. The mean number of $pp$ interactions per bunch crossing in the data set is 24 and the bunch spacing is 25 ns. After the application of beam and data-quality requirements, the integrated luminosity considered corresponds to 36.1 fb$^{-1}$.

Monte Carlo simulation samples were produced for signal and background processes using the full ATLAS detector simulation [42] based on GEANT4 [43] or, for selected smaller backgrounds, a fast simulation using a parameterization of the calorimeter response and GEANT4 for tracking systems [44]. To simulate the effects of additional $pp$ collisions in the same and nearby bunch crossings (pileup), additional interactions were generated using the low-momentum strong-interaction processes of PYTHIA 8.186 [45,46] with a set of tuned parameters referred to as the A2 tune [47] and the MSTW2008LO set of parton distribution functions (PDFs) [48] and overlaid onto the simulated hard-scatter event. The simulated events are reweighted to match the pileup conditions observed in the data and are reconstructed using the same procedure as for the data. The event generators used for each signal and background sample, together with the program and the set of tuned parameters used for the modeling of the parton shower, hadronization and underlying event are listed in Table I. The simulation samples for $t\bar{t}H$, $t\bar{t}V$, $VV$ and $t\bar{t}$ are described in Refs. [49–51]. The samples used to estimate the systematic uncertainties are indicated in between parentheses in Table I.

<table>
<thead>
<tr>
<th>Process</th>
<th>Event generator</th>
<th>ME order</th>
<th>Parton shower</th>
<th>PDF</th>
<th>Tune</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}H$</td>
<td>MG5_AMC</td>
<td>NLO</td>
<td>PYTHIA 8</td>
<td>NNPDF 3.0 NLO [68]</td>
<td>A14</td>
</tr>
<tr>
<td>($t\bar{t}H$)</td>
<td>(MG5_AMC)</td>
<td>(NLO)</td>
<td>(HERWIG++)</td>
<td>(CT10 [69])</td>
<td>(UE-EE-5)</td>
</tr>
<tr>
<td>$t\bar{t}W$</td>
<td>MG5_AMC</td>
<td>LO</td>
<td>PYTHIA 8</td>
<td>NNPDF 3.0 NLO</td>
<td>A14</td>
</tr>
<tr>
<td>$t\bar{t}(Z/\gamma^{*} \to l\bar{l})$</td>
<td>MG5_AMC</td>
<td>(LO multileg)</td>
<td>(SHERPA)</td>
<td>NNPDF 3.0 NLO</td>
<td>A14</td>
</tr>
<tr>
<td>$t\bar{t}Z$</td>
<td>MG5_AMC</td>
<td>LO</td>
<td>PYTHIA 6</td>
<td>CTEQ6L1</td>
<td>Perugia2012</td>
</tr>
<tr>
<td>$t\bar{t}W$</td>
<td>MG5_AMC</td>
<td>LO</td>
<td>PYTHIA 8</td>
<td>NNPDF 2.3 LO</td>
<td>A14</td>
</tr>
<tr>
<td>$t\bar{t}W'$</td>
<td>MG5_AMC</td>
<td>LO</td>
<td>PYTHIA 8</td>
<td>NNPDF 2.3 LO</td>
<td>A14</td>
</tr>
<tr>
<td>$Z \to l^+l^-$</td>
<td>POWHEG-BOXV2 [70]</td>
<td>NLO</td>
<td>PYTHIA 6</td>
<td>CT10</td>
<td>Perugia2012</td>
</tr>
<tr>
<td>$VV \to llXX$</td>
<td>SHERPA 2.1.1</td>
<td>MEPS NLO</td>
<td>SHERPA</td>
<td>CT10</td>
<td>SHERPA default</td>
</tr>
<tr>
<td>$qqVV, VVV$</td>
<td>SHERPA 2.1.1</td>
<td>MEPS NLO</td>
<td>SHERPA</td>
<td>NNPDF 3.0 NLO</td>
<td>SHERPA default</td>
</tr>
</tbody>
</table>
A Higgs boson mass of 125 GeV, from the combined ATLAS and CMS run 1 measurements [74], and a top quark mass of 172.5 GeV are assumed. The overall $t\bar{t}H$ cross section is 507 fb, which is computed at next-to-leading order (NLO) in quantum chromodynamics (QCD) with NLO electroweak corrections [31–35]. Uncertainties include $+5.8\%$ due to the QCD factorization and renormalization scales and $+3.6\%$ due to the PDFs and the strong coupling $\alpha_S$. The cross sections for $t\bar{t}V$ production, including the process $pp \rightarrow t\bar{t}l^+l^- + X$ over the full $Z/\gamma^* \rightarrow l^+l^-$ mass spectrum, are computed at NLO in QCD and electroweak couplings following Refs. [58,75]. The cross section for $t\bar{t}l^+l^-$, with $m(l^+l^-) > 5$ GeV, is 124 fb, and 601 fb for $t\bar{t}W^\pm$ [31]. The QCD scale uncertainties are $\pm12\%$ and uncertainties from PDF and $\alpha_S$ variations are $\pm4\%$.

Events in the $t\bar{t}$ sample with radiated photons of high transverse momentum ($p_T$) are vetoed to avoid overlap with those from the $t\bar{t}\gamma$ sample. Dedicated samples are included to account for backgrounds from $t\bar{t}(Z/\gamma^*)$, where the $Z/\gamma^*$ has low invariant mass but the leptons enter the analysis phase space via asymmetric internal conversions, or rare $t \rightarrow Wb\ell\bar{\ell}$ radiative decays (referred to as “rare top decay” in the following).

IV. OBJECT RECONSTRUCTION AND IDENTIFICATION

All analysis channels share a common trigger, jet, lepton and overall event preselection. The selections are detailed here and the lepton selection is summarized in Table II. Unless otherwise specified, light leptons are required to pass the loose lepton selection. Further channel-specific requirements are discussed in Sec. V.

The selection of events is based on the presence of light leptons, with either single-lepton or dilepton triggers. For data recorded in 2015, the single-electron (single-muon) trigger required a candidate with transverse momentum $p_T > 24$ (20) GeV [41]; in 2016 the lepton $p_T$ threshold was raised to 26 GeV. The trigger $p_T$ thresholds for the 2015 (2016) data taking were $12 + 12$ $(17 + 17)$ GeV for dielectron and $18 + 8$ $(22 + 8)$ GeV for dimuon triggers. For the electron + muon triggers, they were $17 + 14$ GeV for both data sets. The trigger requirement has an efficiency of 82%–99%, depending on the final state and the data set, for signal events passing the final signal-region selections. The reconstructed light leptons are required to be matched to the trigger signatures. The primary vertex of an event is chosen as the vertex with the highest sum of squared transverse momenta of the associated tracks with $p_T > 400$ MeV [76].

Muon candidates are reconstructed by combining inner detector tracks with track segments or full tracks in the muon spectrometer [77]. 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 $|\eta| < 2.5$ and to pass loose identification requirements [77]. To reduce the nonprompt muon contribution, the track is required to originate from the primary vertex by imposing a requirement on its transverse impact parameter significance $|d_0|/\sigma_{d_0} < 3$ and on its longitudinal impact parameter multiplied by the sine of the polar angle $|z_0 \sin \theta| < 0.5$ mm. Additionally, muons are required to be separated by $\Delta R > \min(0.4, 0.04 + (10 \text{ GeV})/p_T)$ from any selected jets (see below for details of jet reconstruction and selection). The requirement is chosen to maximize the acceptance for prompt muons at a fixed rejection factor for nonprompt and fake muon candidates.

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter that are associated with charged-particle tracks reconstructed in the inner detector [78,79]. They are required to have a transverse momentum $p_T > 10$ GeV and $|\eta_{\text{cluster}}| < 2.47$, and the transition region between the barrel and end cap electromagnetic calorimeters, $1.37 < |\eta_{\text{cluster}}| < 1.52$, is excluded.

<table>
<thead>
<tr>
<th>Isolation</th>
<th>$e$</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonprompt lepton BDT Identification</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Charge misassignment veto BDT</td>
<td>Loose</td>
<td>No</td>
</tr>
<tr>
<td>Transverse impact parameter significance, $</td>
<td>d_0</td>
<td>/\sigma_{d_0}$</td>
</tr>
<tr>
<td>Longitudinal impact parameter, $</td>
<td>z_0 \sin \theta</td>
<td>$</td>
</tr>
</tbody>
</table>

TABLE II. Loose (L), loose and isolated (L*), loose, isolated and passing the nonprompt BDT (L*), tight (T) and very tight (T*) light-lepton definitions. Selections for the tighter leptons are applied in addition to the looser ones. For the muons, the L*, T and T* lepton definitions are identical.
A multivariate likelihood discriminant combining shower shape and track information is used to distinguish real prompt electrons from electron candidates from hadronic jets, photon conversions and heavy-flavor (HF) hadron decays (fake and nonprompt electrons). Loose and tight electron discriminant working points are used [79], both including the number of hits in the innermost pixel layer to discriminate between electrons and converted photons. The same longitudinal impact parameter selection as for muons is applied, while the transverse impact parameter significance is required to be $|d_{0}|/\sigma_{d_{0}} < 5$. If two electrons closer than $\Delta R = 0.1$ are preselected, only the one with the higher $p_{T}$ is considered. An electron is rejected if, after passing all the above selections, it lies within $\Delta R = 0.1$ of a selected muon.

Hadronically decaying $\tau$-lepton candidates ($\tau_{\text{had}}$) are reconstructed from clusters in the calorimeters and associated inner detector tracks [80]. Candidates are required to have either one or three associated tracks, with a total charge of $\pm 1$. Candidates are required to have a transverse momentum $p_{T} > 25$ GeV and $|\eta| < 2.5$, excluding the electromagnetic calorimeter’s transition region. A boosted decision tree (BDT) discriminant using calorimeter- and tracking-based variables is used to identify $\tau_{\text{had}}$ candidates and reject jet backgrounds. Three types of $\tau_{\text{had}}$ candidates are used in the analysis, referred to as loose, medium and tight: the latter two are defined by working points with a 55% and 45% (40% and 30%) for one- (three-) prong $\tau_{\text{had}}$ decays, respectively [81], while the first one has a more relaxed selection and is only used for background estimates. The corresponding expected rejection factors against light-quark or gluon jets vary from 30 for loose candidates to 300 for tight candidates [80]. Electrons that are reconstructed as one-prong $\tau_{\text{had}}$ candidates are removed via a BDT trained to reject electrons. Additionally, $\tau_{\text{had}}$ candidates are required to be separated by $\Delta R > 0.2$ from any selected electrons and muons. The contribution of fake $\tau_{\text{had}}$ from $b$ jets is removed by vetoing the candidates that are also $b$ tagged, which rejects a large fraction of the $t\bar{t}$ background. The contribution of fake $\tau_{\text{had}}$ from muons is removed by vetoing the candidates that overlap with low-$p_{T}$ reconstructed muons. Finally, the vertex matched to the tracks of the $\tau_{\text{had}}$ candidate is required to be the primary vertex of the event, in order to reject fake candidates arising from pileup collisions.

Jets are reconstructed from three-dimensional topological clusters built from energy deposits in the calorimeters [82,83], using the anti-$k_{T}$ algorithm with a radius parameter $R = 0.4$ [84,85]. Their calibration is based on simulation with additional corrections obtained using in situ techniques [86] to account for differences between simulation and data. Jets are required to satisfy $p_{T} > 25$ GeV and $|\eta| < 2.5$. In order to reject jets arising from pileup collisions, a significant fraction of the total summed scalar $p_{T}$ of the tracks in jets with $p_{T} < 60$ GeV and $|\eta| < 2.4$ must originate from tracks that are associated with the primary vertex [87]. The average efficiency of this requirement is 92% per jet from the hard scatter. The calorimeter energy deposits from electrons are typically also reconstructed as jets; in order to eliminate double counting, any jets within $\Delta R = 0.3$ of a selected electron are not considered. This is also the case for any jets within $\Delta R = 0.3$ of a $\tau_{\text{had}}$ candidate.

Jets containing $b$ hadrons are identified ($b$ tagged) via a multivariate discriminant combining information from algorithms using track impact parameters and secondary vertices reconstructed within the jet [88,89]. These $b$-tagged jets will henceforth be referred to as $b$ jets. The working point used for this search corresponds to an average efficiency of 70% for jets containing $b$ hadrons with $p_{T} > 20$ GeV and $|\eta| < 2.5$ in $t\bar{t}$ events. The expected rejection factors against light-quark or gluon jets, $c$-quark jets and hadronically decaying $\tau$ leptons are 380, 12 and 55, respectively [89,90]. To compensate for differences between data and simulation in the $b$-tagging efficiencies and mistagging rates, correction factors are applied to the simulated samples [89].

The lepton requirements are summarized in Table II. Isolation requirements are applied to all lepton types except the loose definition. Two isolation variables, based on calorimetric and tracking variables, are computed. Calorimetric isolation uses the scalar sum of transverse energies of clusters within a cone of size $\Delta R = 0.3$ around the light-lepton candidate. This excludes the electron candidate’s cluster itself and clusters within $\Delta R = 0.1$ of the muon candidate’s track, respectively, and is corrected for leakage from the electron’s shower and for the ambient energy in the event [91,92]. Track isolation uses the sum of transverse momenta of tracks with $p_{T} > 1$ GeV consistent with originating at the primary vertex, excluding the light-lepton candidate’s track, within a cone of $\Delta R = \min(0.3, 10 \text{ GeV}/p_{T}(\ell))$. Calorimeter- and track-based isolation criteria are applied to electrons and muons to obtain a 99% efficiency in $Z \rightarrow \ell\ell$ events.

Nonprompt leptons are further rejected using a multivariate discriminant, taking as input the energy deposits and charged-particle tracks (including the lepton track) in a cone around the lepton direction, which is referred to as the nonprompt lepton BDT. The jet reconstruction and $b$-tagging algorithms are run on the track collection, and their output is used to train the algorithm together with isolation variables. A reconstructed track jet that is matched to a nonprompt lepton is typically a jet initiated by $b$ or $c$ quarks and may contain a displaced vertex. The most discriminating variables are thus found to be the angular distance between the lepton and the reconstructed jet, the outputs of the $b$-tagging algorithms, the calorimetric and track isolation variables of the lepton, the number of tracks within the jet and the ratio of the lepton $p_{T}$ to the jet $p_{T}$.
The training is performed separately for electrons and muons on prompt and nonprompt leptons from simulated \( t\bar{t} \) events and validated using data in various control regions. The efficiency at the chosen working point to select well-identified prompt muons (electrons) is about 70% (60%) for \( p_T \sim 10\) GeV and reaches a plateau of 98% (96%) at \( p_T \sim 45\) GeV, as shown in Fig. 2, while the rejection factor against leptons from the decay of \( b \) hadrons is about 20. Simulated events are corrected to account for differences between data and simulation for this prompt-lepton isolation efficiency, as well as for the lepton trigger, reconstruction, and identification efficiencies. The corrections were determined using a so-called tag-and-probe method as described in Refs. [77,78] and studied as a function of the number of nearby light- and heavy-flavor jets. This is illustrated in Fig. 2, showing that the corrections for the nonprompt lepton BDT efficiencies are at most 10% at low transverse momentum and decrease with increasing transverse momentum. The largest contribution to the associated systematic uncertainties comes from pileup effects.

There is a small, but non-negligible, probability that electrons and positrons are reconstructed with an incorrect charge. This occurs when an electron (positron) emits a hard bremsstrahlung photon; if the photon subsequently converts to an asymmetric electron-positron pair, and the positron (electron) has high momentum and is reconstructed, the lepton charge can be misidentified. Otherwise it occurs when the curvature of a track is poorly estimated, which typically happens at high momentum. The probability for muons to be reconstructed with incorrect charge is small enough that the charge misassignment is negligible. To reject electrons reconstructed with an incorrect electric charge, a BDT discriminant is built, using the following electron cluster and track properties as input: the electron’s transverse momentum and pseudorapidity, the track curvature significance (defined as the ratio of the electric charge to the track momentum divided by the estimated uncertainty in the measurement) and its transverse impact parameter times the electric charge, the cluster width along the azimuthal direction, and the quality of the matching between the track and the cluster, in terms of both energy/momentum and azimuthal position. The chosen working point achieves a rejection factor of \( \sim 17 \) for electrons passing the tight identification requirements with a wrong charge assignment while providing an efficiency of 95% for electrons with correct charge reconstruction. This requirement is only applied to the very tight electrons. Correction factors to account for differences in the selection efficiency between data and simulation, which are within a few percent for \( |\eta| < 2.4 \) but larger in the forward region, \( 2.4 < |\eta| < 2.47 \), were applied to the selected electrons in the simulation.

The missing transverse momentum \( p_T^{\text{miss}} \) (with magnitude \( E_T^{\text{miss}} \)) is defined as the negative vector sum of the transverse momenta of all identified and calibrated leptons and jets and remaining unclustered energy, the latter of which is estimated from low-\( p_T \) tracks associated with the primary vertex but not assigned to any lepton or jet candidate [93,94].

V. EVENT SELECTION AND CLASSIFICATION

The analysis is primarily sensitive to decays of the Higgs boson to \( WW \) or \( \tau \tau \) with a small additional contribution from \( H \rightarrow ZZ \). If the Higgs boson decays to either \( WW \) or \( \tau \tau \), the \( t\bar{t}H \) events typically contain either \( WWWWbb \) or \( \tau\tau WWbb \). In order to reduce the \( t\bar{t} \) background, characterized by a final state of \( WWbb \), final states including three or more charged leptons, or two same-charge light leptons, are selected. Seven final states are analyzed, categorized by the number and flavor of charged-lepton candidates after the preselection requirements, as illustrated in Fig. 3. Each of the seven final

FIG. 2. The efficiency to select well-identified prompt muons (left) and electrons (right) at the chosen nonprompt lepton BDT working point, as a function of the lepton \( p_T \). The muons are required to pass the loose identification requirements, while the electrons are required to pass the tight identification requirements. The measurements in data (simulation) are shown as full black (open red) circles. The bottom panel displays the ratio of data to simulation results, with the blue (yellow) band representing the statistical (total) uncertainty. This ratio is the scale factor that is applied to correct the simulation.
states is called a “channel” and certain channels are further split into categories to gain in significance. Categories include signal and control regions. Additional control regions used for the estimates of the nonprompt backgrounds are discussed in Sec. VI.

The seven channels are

(i) two same-charge light leptons and no hadronically decaying $\tau$ lepton candidates ($2\ell$SS);
(ii) three light leptons and no hadronically decaying $\tau$ lepton candidates ($3\ell$);
(iii) four light leptons ($4\ell$);
(iv) one light lepton and two opposite-charge hadronically decaying $\tau$ lepton candidates ($1\ell + 2\tau_{\text{had}}$);
(v) two same-charge light leptons and one hadronically decaying $\tau$ lepton candidate ($2\ell$SS + $1\tau_{\text{had}}$);

FIG. 3. The channels used in the analysis organized according to the number of selected light leptons and $\tau_{\text{had}}$ candidates. The selection requirements for each channel are in Table III.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Selection criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common</td>
<td>$N_{\text{jets}} \geq 2$ and $N_{b\text{-jets}} \geq 1$</td>
</tr>
</tbody>
</table>
| $2\ell$SS | Two very tight light leptons with $p_T > 20$ GeV  
Same-charge light leptons  
Zero medium $\tau_{\text{had}}$ candidates  
$N_{\text{jets}} \geq 4$ and $N_{b\text{-jets}} < 3$ |
| $3\ell$ | Three light leptons with $p_T > 10$ GeV; sum of light-lepton charges $\pm 1$  
Two same-charge leptons must be very tight and have $p_T > 15$ GeV  
The opposite-charge lepton must be loose, isolated and pass the nonprompt BDT  
Zero medium $\tau_{\text{had}}$ candidates  
m($\ell^+\ell^-) > 12$ GeV and $|m(\ell^+\ell^-) - 91.2$ GeV$| > 10$ GeV for all SFOC pairs  
$m(3\ell) - 91.2$ GeV$| > 10$ GeV |
| $4\ell$ | Four light leptons; sum of light-lepton charges 0  
Third and fourth leading leptons must be tight  
m($\ell^+\ell^-) > 12$ GeV and $|m(\ell^+\ell^-) - 91.2$ GeV$| > 10$ GeV for all SFOC pairs  
$m(4\ell) - 125$ GeV$| > 5$ GeV  
Split two categories: Z-depleted (0 SFOC pairs) and Z-enriched (two or four SFOC pairs) |
| $1\ell + 2\tau_{\text{had}}$ | One tight light lepton with $p_T > 27$ GeV  
Two medium $\tau_{\text{had}}$ candidates of opposite charge, at least one being tight  
$N_{\text{jets}} \geq 3$ |
| $2\ell$SS + $1\tau_{\text{had}}$ | Two very tight light leptons with $p_T > 15$ GeV  
Same-charge light leptons  
One medium $\tau_{\text{had}}$ candidate, with charge opposite to that of the light leptons  
$N_{\text{jets}} \geq 4$  
$|m(ee) - 91.2$ GeV$| > 10$ GeV for $ee$ events |
| $2\ell$OS + $1\tau_{\text{had}}$ | Two loose and isolated light leptons with $p_T > 25, 15$ GeV  
One medium $\tau_{\text{had}}$ candidate  
Opposite-charge light leptons  
One medium $\tau_{\text{had}}$ candidate  
m($\ell^+\ell^-) > 12$ GeV and $|m(\ell^+\ell^-) - 91.2$ GeV$| > 10$ GeV for the SFOC pair  
$N_{\text{jets}} \geq 3$ |
| $3\ell + 1\tau_{\text{had}}$ | 3$\ell$ selection, except:  
One medium $\tau_{\text{had}}$ candidate, with charge opposite to the total charge of the light leptons  
The two same-charge light leptons must be tight and have $p_T > 10$ GeV  
The opposite-charge light lepton must be loose and isolated |
TABLE IV. Summary of the basic characteristics of the seven analysis channels. The lepton selection follows the definition in Table II and is labeled as loose (L), loose and isolated (L†), loose, isolated and passing the nonprompt BDT (L*), tight (T) and very tight (T*) leptons, respectively. The τ had selection is labeled as medium (M) and tight (T).

<table>
<thead>
<tr>
<th>2τSS</th>
<th>3τ</th>
<th>4τ</th>
<th>1τ + 2τ had</th>
<th>2τSS + 1τ had</th>
<th>2τOS + 1τ had</th>
<th>3τ + 1τ had</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light lepton</td>
<td>2T*</td>
<td>1L*, 2T*</td>
<td>2L, 2T</td>
<td>1T</td>
<td>2T*</td>
<td>2L†</td>
</tr>
<tr>
<td>τ had</td>
<td>0M</td>
<td>0M</td>
<td>···</td>
<td>1T, 1M</td>
<td>1M</td>
<td>1M</td>
</tr>
<tr>
<td>N jets, N h-jets</td>
<td>≥ 4, = 1, 2</td>
<td>≥ 2, ≥ 1</td>
<td>≥ 2, ≥ 1</td>
<td>≥ 3, ≥ 1</td>
<td>≥ 4, ≥ 1</td>
<td>≥ 3, ≥ 1</td>
</tr>
</tbody>
</table>

(vi) two opposite-charge light leptons and one hadronically decaying τ lepton candidate (2τ + 1τ had);
(vii) three light leptons and one hadronically decaying τ lepton candidate (3τ + 1τ had).

The selection criteria are designed to be orthogonal to ensure that each event only contributes to a single channel. Channels are made orthogonal through the requirements on the number of loose light leptons and medium τ had candidates. A veto on events containing medium τ had candidates is therefore applied for the 2τSS and 3τ channels, but no veto is applied for the 4τ channel because there is no corresponding τ had channel. In all channels, the light lepton(s) are required to be matched to the lepton(s) selected by either the single-lepton or dilepton triggers. As the 1τ + 2τ had channel has only one light lepton, only single-lepton triggers are used. In order to reduce the diboson background, all channels also require events to include at least two reconstructed jets and that at least one of these jets must be b tagged.

The detailed criteria for each channel are described below and summarized in Table III. In addition, Table IV provides a comparison of the key aspects of the selection used in each channel. After the selection, assuming Standard Model tH production, the total expected number of reconstructed signal events summed over all categories is 91, corresponding to 0.50% of all produced tH events. The breakdown in each channel is given in Table V. In total 332030 events are selected in data. As the background contamination is still large in all channels, except one of the 4τ categories and the 3τ + 1τ had category, further separation of the signal from the background is achieved using multivariate techniques. The TMVA package [95] is used in all channels except for 3τ, which uses XGBoost [96]. Independent cross-check analyses using a simpler cut-and-count categorization were developed for the most sensitive 2τSS, 3τ and 2τSS + 1τ had channels.

A. 2τSS channel

Selected events are required to include exactly two reconstructed light leptons with the same electric charge. To reduce the background from fake and nonprompt leptons as well as electrons reconstructed with incorrect electric charge, the very tight selection requirements described in Sec. IV are applied and the leptons are required to satisfy \( p_T > 20 \) GeV. Events must include at least four reconstructed jets to suppress tt and tttW backgrounds, among which either one or two are required to be b tagged. A slight disagreement is observed between the Standard Model prediction and the data for events containing two same-charge light leptons and three or more b jets. To avoid any potential systematic bias, these events are vetoed, at no expense in sensitivity.

Two independent BDTs are trained using the selected events. The first aims to separate the signal from the nonprompt and fake background, while the second aims to separate the signal from the tttV background. The data-driven estimate of the nonprompt and fake background described in Sec. VI B 1 is used in the training, which is performed for both BDTs with the nine variables listed in Table VI. The outputs of the two BDT classifiers are combined to maximize the signal significance.

A cross-check is provided by an independent cut-and-count analysis using 12 categories, which places requirements on the jet multiplicity, b-tagged jet multiplicity and the lepton flavor.

B. 3τ channel

Selected events are required to include exactly three reconstructed light leptons with the total charge equal to ±1. The lepton of opposite charge to the other two is found to be prompt in 97% of the selected events in tt simulated samples and therefore only required to be loose, isolated and pass the nonprompt BDT selection requirements, as

TABLE V. Acceptance times efficiency (A × e) for tH signal in each analysis channel. This includes Higgs boson and top quark branching fractions, detector acceptance, and reconstruction and selection efficiency and is computed relative to inclusive tH production considering all Higgs boson and top decays. In the 4τ channel, the two numbers correspond to the Z-enriched and the Z-depleted categories.

<table>
<thead>
<tr>
<th>2τSS</th>
<th>3τ</th>
<th>4τ</th>
<th>1τ + 2τ had</th>
<th>2τSS + 1τ had</th>
<th>2τOS + 1τ had</th>
<th>3τ + 1τ had</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A × e[10^{-4}]</td>
<td>23</td>
<td>13</td>
<td>0.6 + 0.1</td>
<td>2.3</td>
<td>1.7</td>
<td>7.8</td>
<td>0.8</td>
</tr>
</tbody>
</table>
TABLE VI. Variables used in the multivariate analysis (denoted by ×) for the $2\ell SS$, $3\ell$, $4\ell$ (Z-enriched category), $1\ell + 2\tau_{\text{had}}$, $2\ell SS + 1\tau_{\text{had}}$ and $2\ell OS + 1\tau_{\text{had}}$ channels. For $2\ell SS$ and $2\ell SS + 1\tau_{\text{had}}$, lepton 0 and lepton 1 are the leading and subleading leptons, respectively. For $3\ell$, lepton 0 is the lepton with charge opposite to that of the same-charge pair, while the same-charge leptons are labeled with increasing index (lepton 1 and lepton 2) as $p_T$ decreases. The best Z-candidate dilepton invariant mass is the mass of the dilepton pair closest to the Z boson mass. The variables also used in the cross-check analyses are indicated by an *. 

<table>
<thead>
<tr>
<th>Variable</th>
<th>$2\ell SS$</th>
<th>$3\ell$</th>
<th>$4\ell$</th>
<th>$1\ell + 2\tau_{\text{had}}$</th>
<th>$2\ell SS + 1\tau_{\text{had}}$</th>
<th>$2\ell OS + 1\tau_{\text{had}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lepton properties</td>
<td>Leading lepton $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Second leading lepton $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Third lepton $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Dilepton invariant mass (all combinations)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Three-lepton invariant mass</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Four-lepton invariant mass</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Best Z-candidate dilepton invariant mass</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Other Z-candidate dilepton invariant mass</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Scalar sum of all leptons $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Second leading lepton track isolation</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Maximum $</td>
<td>\eta</td>
<td>$ (lepton 0, lepton 1)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Lepton flavor</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Lepton charge</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>Jet properties</td>
<td>Number of jets</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Number of $b$-tagged jets</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Leading jet $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Second leading jet $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Leading $b$-tagged jet $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Scalar sum of all jets $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Scalar sum of all $b$-tagged jets $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Has leading jet highest $b$-tagging weight?</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$b$-tagging weight of leading jet</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$b$-tagging weight of second leading jet</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$b$-tagging weight of third leading jet</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Pseudorapidity of fourth leading jet</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$\tau_{\text{had}}$</td>
<td>Leading $\tau_{\text{had}}$ $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Second leading $\tau_{\text{had}}$ $p_T$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Di-$\tau_{\text{had}}$ invariant mass</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Invariant mass $\tau_{\text{had}}$ -furthest lepton</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>Angular distances</td>
<td>$\Delta R$ (lepton 0, lepton 1)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R$ (lepton 0, lepton 2)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R$ (lepton 0, closest jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R$ (lepton 0, leading jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R$ (lepton 0, closest $b$-jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R$ (lepton 1, closest jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R$ (lepton 2, closest jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Smallest $\Delta R$ (lepton, jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Smallest $\Delta R$ (lepton, $b$-tagged jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Smallest $\Delta R$ (non-tagged jet, $b$-tagged jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Smallest $\Delta R$ (lepton, $b$-tagged jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Smallest $\Delta R$ (non-tagged jet, $b$-tagged jet)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$p_T^\text{miss}$</td>
<td>Missing transverse momentum $E_T^\text{miss}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Azimuthal separation</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$\Delta \phi$ (leading jet, $p_T^\text{miss}$)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>Transverse mass leptons ($H/Z$ decay)</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td></td>
<td>$\vec{p}_T^\text{miss}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
</tbody>
</table>

Pseudo-matrix-element | $\times$ | $\times$ | $\times$ | $\times$ | $\times$ |
Invariant mass within 10 GeV of the $Z$ suppresses background from resonances that decay to light lepton pairs. A $Z$ veto is applied, excluding events containing a same-flavor opposite-charge lepton pair with an invariant mass within 10 GeV of the $Z$ mass to suppress the $t\bar{t}Z$ background. Finally, to eliminate potential backgrounds with $Z$ decays to $\ell\ell\ell'\ell''$ ($\ell\ell$), where one lepton has very low momentum and is not reconstructed, the three-lepton invariant mass must satisfy $|m(3\ell')-91.2$ GeV$| > 10$ GeV.

Selected events are classified using a five-dimensional multinomial boosted decision tree. The five classification targets used in the training are: $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}$, and diboson. In total, 28 variables based on topological aspects of the events as listed in Table VI are used in the training. The output discriminants are mapped into five categories that maximize the signal significance. The output discriminants are used in the signal region to further reduce the $t\bar{t}$ background. Seven variables listed in Table VI are used in the training, including a pseudo-matrix-element discriminator exploiting partially reconstructed resonances ($t$, $H$ and $Z$) [98]. A requirement on the discriminant is then imposed to define the $Z$-enriched signal region.

### D. $1\ell^+ + 2\ell^-$ channel

Selected events are required to include exactly one light lepton and exactly two medium $\tau_{\text{had}}$ candidates of opposite charge. At least one of the $\tau_{\text{had}}$ candidates is required to be tight. In order to suppress the $t\bar{t}$ and $t\bar{t}V$ backgrounds, events must include at least three reconstructed jets. A BDT is trained to further reduce the main $t\bar{t}$ background, in which events had one or two fake $\tau_{\text{had}}$ candidates. Seven variables listed in Table VI are used in the training, including the invariant mass of the visible decay products of the $\tau_{\text{had}}$ system.

### E. $2\ell^0\ell^0 + 1\tau_{\text{had}}$ channel

Selected events are required to contain exactly one medium $\tau_{\text{had}}$ candidate but otherwise to meet the requirements for the $2\ell^0\ell^0$ channel discussed in Sec. VA, except that the light-lepton $p_T$ threshold is lowered from 20 to 15 GeV and that events with three or more $b$ jets are included. The reconstructed charge of the $\tau_{\text{had}}$ candidate must be opposite to that of the light leptons. The $Z$ veto is applied to dielectron events to suppress $Z + \tau$ events with a misassigned charge. A BDT is trained using the 13 variables listed in Table VI on events with relaxed selection requirements: the light leptons are required to be loose instead of tight and the requirement on the number of jets is reduced to two. This BDT is used to further reduce the $t\bar{t}$ background.

A cross-check is provided by an independent cut-and-count analysis using three categories, which places requirements on the number of jets and on the maximum $|\eta|$ of the two light leptons.

### F. $2\ell^0 + 1\tau_{\text{had}}$ channel

Selected events are required to include exactly two reconstructed loose and isolated leptons of opposite charge with leading (subleading) $p_T > 25$ (15) GeV and exactly one medium $\tau_{\text{had}}$ candidate. In order to reduce the $t\bar{t}$, $Z$ +
jets and $t\bar{t}V$ backgrounds, events must include at least three reconstructed jets. The $Z$ veto is applied to same-flavor lepton pairs to suppress the $Z + jets$ background with a fake $\tau_{\text{had}}$ candidate. To suppress background from resonances that decay to light leptons, events containing a same-flavor lepton pair with an invariant mass below 12 GeV are also removed. A BDT is trained using the 13 variables listed in Table VI on the selected events, with the aim of further reducing the main $t\bar{t}$ background with a fake $\tau_{\text{had}}$ candidate.

G. $3\ell + 1\tau_{\text{had}}$ channel

Selected events are required to contain exactly one medium $\tau_{\text{had}}$ candidate but otherwise to meet the requirements for the $3\ell$ channel discussed in Sec. VB, except that the two same-charge leptons must be tight and have $p_T > 10$ GeV and the opposite-charge lepton must be loose and isolated. The reconstructed charge of the $\tau_{\text{had}}$ candidate must be opposite to the total charge of the light leptons. Due to the high purity of the signal, no further selection is required and only the event yields are used in the fit.

FIG. 4. Left: The fraction of the expected $t\bar{t}H$ signal arising from different Higgs boson decay modes in each signal region. The decays labeled as “other” are mostly $H \rightarrow \mu\mu$ and $H \rightarrow b\bar{b}$. Right: Prefit $S/B$ (black line) and $S/\sqrt{B}$ (red dashed line) ratios for each of the 12 analysis categories including the four $3\ell$ control regions. The background prediction methods are described in Sec. VI.

FIG. 5. The fractional contributions of the various backgrounds to the total predicted background in each of the 12 analysis categories. The background prediction methods are described in Sec. VI: “Nonprompt,” “Fake $\tau_{\text{had}}$” and “$q$ mis-id” refer to the data-driven background estimates (largely $t\bar{t}$ but also include other electroweak processes), and rare processes ($tZ$, $tW$, $tWZ$, $tWW$, triboson production, $t\bar{t}t$, $t\bar{t}t\bar{t}$, $tH$, rare top decay) are labeled as “Other.”
H. Channel summary

Twelve categories are defined in the previous subsections: eight signal regions and four control regions (CR) from the 3ℓ channel. The fraction of the expected signal arising from different Higgs boson decay modes in each signal region is shown in Fig. 4 (left). The signal-to-background ratio $S/B$ for each signal and control region is shown in Fig. 4 (right). This ranges from 0.014 to almost 2. The ratio $S/\sqrt{B}$ is also indicated. The acceptance for each channel is shown in Table V. The background composition in each region is shown in Fig. 5. The background prediction methods are described in the next section. Multivariate techniques have been applied in most channels to improve the discrimination between the signal and the background. The variables used in each channel are indicated in Table VI. The modeling of each variable was checked and no significant disagreement between data and simulation was found.

VI. BACKGROUND ESTIMATION

The irreducible backgrounds all have selected light leptons produced in $W$ or $Z/\gamma^*$ boson decays or leptonic $τ$ decays (prompt leptons, Sec. VI A). The reducible backgrounds have at least one lepton arising from another source (Sec. VI B). In the latter case, light leptons originate from heavy-flavor hadron decays, photon conversions, improper reconstruction of other particles such as hadronic jets, or prompt leptons whose charge is misassigned. Such misidentified and nonprompt light leptons are collectively referred to as nonprompt leptons in the following, as this is the dominant source. The fake $\tau$ had candidates are typically jets, including HF jets.

A. Backgrounds with prompt leptons

Background contributions with prompt leptons originate from a wide range of processes and the relative importance of individual processes varies by channel. The largest backgrounds with prompt leptons are from top production in association with a vector boson, $t\bar{t}W$ and $t\bar{t}(Z/\gamma^*)$, and diboson production, $VV$. These background estimates are a crucial part of the analysis, because their final state and kinematics are similar to the signal. In addition, there are contributions from a number of rare processes: rare top decay, $tZ$, $tW$, $tWZ$, $t\bar{t}WW$, $VVV$, $t\bar{t}t$ and $t\bar{t}\bar{t}$ production. The associated production of single top quarks with a Higgs boson, which contributes at most 2% in any signal region, is also considered as a background process. All other Higgs boson production mechanisms contribute negligibly (<0.2%) in any signal region.

All these backgrounds are estimated from simulation using the samples described in Sec. III. The systematic uncertainties in the modeling of these processes by the simulation are discussed in Sec. VII. The prompt-lepton estimates were validated in various regions, as illustrated in Fig. 6 for the 3ℓ $t\bar{t}Z$ and $t\bar{t}W$ control regions.
TABLE VII. Selection criteria applied to define the control regions used for the nonprompt lepton (top part) and fake \( \tau \) had (bottom part) estimates. The 2\( \ell \)SS CR is used for both the 2\( \ell \)SS and 3\( \ell \) channels, as indicated by putting 3\( \ell \) in parentheses. Same-flavor, opposite-charge (same-charge) lepton pairs are referred to as SFOC (SFSC) pairs.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Region</th>
<th>Selection criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>2( \ell )SS</td>
<td>2 ( \leq N_{\text{jets}} \leq 3 ) and ( N_{b\text{-jets}} \geq 1 )</td>
<td>2 ( \leq N_{\text{jets}} \leq 3 ) and ( N_{b\text{-jets}} \geq 1 )</td>
</tr>
<tr>
<td>(3( \ell ))</td>
<td>One very tight, one loose light lepton with ( p_T &gt; 20(15) ) GeV</td>
<td>One very tight, one loose light lepton with ( p_T &gt; 20(15) ) GeV</td>
</tr>
<tr>
<td></td>
<td>( \epsilon_{\text{real}} )</td>
<td>Zero ( \tau ) had candidates</td>
</tr>
<tr>
<td>4( \ell )</td>
<td>( 1 \leq N_{\text{jets}} \leq 2 )</td>
<td>( 1 \leq N_{\text{jets}} \leq 2 )</td>
</tr>
<tr>
<td></td>
<td>Three loose light leptons; sum of light lepton charges ( \pm 1 )</td>
<td>Three loose light leptons; sum of light lepton charges ( \pm 1 )</td>
</tr>
<tr>
<td></td>
<td>Subleading same-charge lepton must be tight</td>
<td>Subleading same-charge lepton must be tight</td>
</tr>
<tr>
<td></td>
<td>Veto on 3( \ell ) selection</td>
<td>Veto on 3( \ell ) selection</td>
</tr>
<tr>
<td></td>
<td>( \epsilon_{\text{real}} )</td>
<td>( \epsilon_{\text{real}} )</td>
</tr>
<tr>
<td>2( \ell )SS + 1( \tau ) had</td>
<td>2 ( \leq N_{\text{jets}} \leq 3 ) and ( N_{b\text{-jets}} \geq 1 )</td>
<td>2 ( \leq N_{\text{jets}} \leq 3 ) and ( N_{b\text{-jets}} \geq 1 )</td>
</tr>
<tr>
<td></td>
<td>One very tight, one loose light lepton with ( p_T &gt; 15 ) GeV</td>
<td>One very tight, one loose light lepton with ( p_T &gt; 15 ) GeV</td>
</tr>
<tr>
<td></td>
<td>A SFSC pair (</td>
<td>m(\ell^+\ell^-) - 91.2 \text{ GeV}</td>
</tr>
<tr>
<td></td>
<td>Zero or one medium ( \tau ) had candidate, opposite in charge to the light leptons</td>
<td>Zero or one medium ( \tau ) had candidate, opposite in charge to the light leptons</td>
</tr>
<tr>
<td></td>
<td>( \epsilon_{\text{fake}} )</td>
<td>( \epsilon_{\text{fake}} )</td>
</tr>
<tr>
<td>1( \ell ) + 2( \tau ) had</td>
<td>( N_{\text{jets}} \geq 3 ) and ( N_{b\text{-jets}} \geq 1 )</td>
<td>( N_{\text{jets}} \geq 3 ) and ( N_{b\text{-jets}} \geq 1 )</td>
</tr>
<tr>
<td></td>
<td>One tight light lepton, with ( p_T &gt; 27 ) GeV</td>
<td>One tight light lepton, with ( p_T &gt; 27 ) GeV</td>
</tr>
<tr>
<td></td>
<td>Two ( \tau ) had candidates of same charge</td>
<td>Two ( \tau ) had candidates of same charge</td>
</tr>
<tr>
<td></td>
<td>Veto on 3( \ell ) selection</td>
<td>Veto on 3( \ell ) selection</td>
</tr>
<tr>
<td></td>
<td>( \epsilon_{\text{real}} )</td>
<td>( \epsilon_{\text{real}} )</td>
</tr>
<tr>
<td>2( \ell )OS + 1( \tau ) had</td>
<td>Two loose and isolated light leptons, with ( p_T &gt; 25, 15 ) GeV</td>
<td>Two loose and isolated light leptons, with ( p_T &gt; 25, 15 ) GeV</td>
</tr>
<tr>
<td></td>
<td>One loose ( \tau ) had candidate (</td>
<td>m(\ell^+\ell^-) - 91.2 \text{ GeV}</td>
</tr>
</tbody>
</table>

FIG. 7. The composition from simulation of (a) the fake and nonprompt light leptons and (b) the fake \( \tau \) had in selected analysis regions. The light-lepton composition is shown separately depending on the lepton flavor in the regions used in the estimate of the nonprompt contribution. The control regions labeled 2I\( \ell \)Sx are used for the 2\( \ell \)SS and 3\( \ell \) channels; those labeled 3lx are used for the 4\( \ell \) channel, where x denotes the flavor of the lowest-\( p_T \) lepton, and those labeled 2I\( \ell \)Sx + 1\( \tau \) are used for the 2\( \ell \)SS + 1\( \tau \) had channel. The nonprompt lepton background has been separated into the components from \( b \) jets, \( c \) jets, other jets, \( J/\psi \), photon conversions and other contributions. The latter includes pion, kaon and nonprompt tau decays and cases where reconstructed leptons cannot be assigned unambiguously to a particular source. The \( \tau \) had composition is shown both in the control regions used in the estimates and in the signal regions of each channel. The \( \tau \) had background has been separated into the components from \( b \) jets, \( c \) jets, light-quark jets, gluon jets, electrons and other contributions. The latter includes muons, hadrons and cases where reconstructed leptons cannot be assigned unambiguously to a particular source.
TABLE VIII. Summary of the nonprompt lepton and fake $\tau_{\text{had}}$ background estimate strategies of the seven analysis channels. DD means data-driven background estimates and the techniques used are the matrix method (MM) and the fake-factor method (FF). The scale factor method (SF), which scales the estimate from simulation by a correction factor measured in data, is partially data driven. The lower half of the table lists the selection requirements used to define the control regions. The lepton selection follows the same convention as in Table II and is labeled as loose (L), loose and isolated (L$^\ast$), loose, isolated and passing the nonprompt BDT (L$^\ast$), tight (T) and very tight (T$^*$), respectively. Analogously, the $\tau_{\text{had}}$ selection is labeled as medium (M) and tight (T).

<table>
<thead>
<tr>
<th>Nonprompt lepton strategy</th>
<th>2$\ell$SS</th>
<th>3$\ell$</th>
<th>4$\ell$</th>
<th>1$\ell$ + 2$\tau_{\text{had}}$</th>
<th>2$\ell$SS + 1$\tau_{\text{had}}$</th>
<th>2$\ell$OS + 1$\tau_{\text{had}}$</th>
<th>3$\ell$ + 1$\tau_{\text{had}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light lepton</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
</tr>
<tr>
<td>$\tau_{\text{had}}$</td>
<td>DD</td>
<td>1T$^*$, 1L</td>
<td>3L</td>
<td>1T</td>
<td>1T$^*$, 1L</td>
<td>2L$^\dagger$</td>
<td>\hspace{1cm}</td>
</tr>
<tr>
<td>$N_{\text{jets}}$</td>
<td>0M</td>
<td>1 ≤ $N_{\text{jets}}$ ≤ 2</td>
<td>\hspace{1cm}</td>
<td>1 ≤ $N_{\text{jets}}$ ≤ 2</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
</tr>
<tr>
<td>$N_{b\text{-jets}}$</td>
<td>≥ 1</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
<td>\hspace{1cm}</td>
</tr>
</tbody>
</table>

\section*{B. Backgrounds with nonprompt leptons and fake $\tau_{\text{had}}$ candidates}

Data-driven methods are used to estimate the backgrounds with nonprompt light leptons and fake $\tau_{\text{had}}$ candidates, defining control regions enriched in such backgrounds and extrapolating the observed yields to the signal regions. The control regions used for this purpose are summarized in Table VII. They are orthogonal to the signal regions. Figure 7 summarizes the origin of the nonprompt leptons and fake $\tau_{\text{had}}$ candidates in these control regions and some signal regions based on predictions from simulation, where the statistical uncertainties of the absolute fractions can be as large as 7%.

Table VIII summarizes the strategies used to estimate the nonprompt lepton and fake $\tau_{\text{had}}$ backgrounds in each of the channels, motivated by the different event topologies and the statistical power available in the control regions. The matrix method and fake-factor method are largely similar but differ in that the fake-factor method estimates the prompt contribution from simulation, while the matrix method uses the measured prompt lepton efficiency from data.

\subsection*{1. Nonprompt leptons in the 2$\ell$SS and 3$\ell$ channels}

The nonprompt lepton background in the 2$\ell$SS and 3$\ell$ channels is a mixture of leptons from semileptonic HF decays and conversions. These backgrounds are estimated using a matrix method similar to that described in Refs. [99,100]. The matrix method estimates the number of nonprompt leptons in the signal region by selecting events passing all selection requirements except the tight-lepton requirements and splitting the events into four categories. The four categories contain exactly two tight leptons, one tight and one loose-but-not-tight lepton, one loose-but-not-tight and one tight lepton, and two loose-but-not-tight leptons (where the leptons are ordered according to their $p_T$). The probabilities for both the loose prompt and nonprompt leptons to be tight are measured in control regions independent from the signal regions. These are used to estimate the number of nonprompt events in the signal regions via the following formula: $f_{\text{SR}} = w_{TT}N_{TT} + w_{TT}N_{TT} + w_{TT}N_{TT} + w_{TT}N_{TT}$. The $w$ weights depend on the measured prompt and nonprompt lepton efficiencies, T and T$^\ast$ denote leptons passing the tight and loose-but-not-tight lepton selections, respectively.

In the 2$\ell$SS channel, the method allows either of the candidate leptons to be nonprompt, while in the 3$\ell$ channel, the opposite-charge lepton is assumed to always be prompt, as is seen in the simulation for 97% of the cases. The efficiencies are measured separately for electrons and muons.

The control regions used to measure the prompt ($\epsilon_{\text{real}}$) and nonprompt ($\epsilon_{\text{fake}}$) lepton efficiencies are defined in Table VII. They have lower jet multiplicity than the signal regions. The lepton efficiencies are parameterized as a function of $p_T$. The nonprompt electron efficiency is additionally parameterized as a function of the angular distance between the lepton and the closest jet to account for effects of nearby jets. The residual prompt background in the control regions is subtracted using the prediction from simulation, while the background from charge misassignment is subtracted using the estimate described in Sec. VII B 4.

The efficiency for electrons from conversions is significantly higher than that for electrons from HF decays; therefore the change in the fraction of conversions when going from the control to the signal regions is estimated from simulation and used to correct $\epsilon_{\text{fake}}$. Systematic uncertainties in this correction are estimated to be 40%. They include a 15% uncertainty in the modeling of conversions in the simulation [101], a 20% uncertainty from a measurement of $\gamma t$ [102], a 50% uncertainty in the modeling of semileptonic $b$ decays and the uncertainties in the nonprompt lepton efficiencies.
The performance of the matrix method was tested in simulation using a closure test by comparing the prediction from the method to the results from the simulation. Closure tests were performed for each channel using \( t\bar{t} \) simulation and the level of the nonclosure is found to be at most \((11\pm8)\%\) and \((9\pm8)\%\) for the \(2\ell SS\) and \(3\ell\) channels, respectively, which is accounted for as a systematic uncertainty. Additional systematic uncertainties due to the subtraction of the prompt backgrounds in the control regions are included. The total uncertainty in the nonprompt lepton estimate varies from 20% for \(e^\pm\mu^\pm\) to 30% for \(3\ell\). The ratio for the nonprompt background yield in

![Graph](image_url)
data to the predictions from simulation is found to be $2.0 \pm 0.5$ for $ee$, $1.5 \pm 0.5$ for $\mu\mu$ and $1.7 \pm 0.4$ for $ee\mu$ in the $2\ell$SS signal region. It is $1.8 \pm 0.8$ for $3\ell$ in the signal region and $2.2 \pm 0.5$ in the $t\bar{t}$ control region. The nonprompt lepton estimates were validated in various regions, as illustrated in Figs. 8(a) and 8(b) in a region identical to the $2\ell$SS signal region except for being orthogonal in the $N_{\text{jets}}$ requirement (low multiplicity $N_{\text{jets}} = 2, 3$).

2. Nonprompt leptons in the $4\ell$ channel

A semi-data-driven estimate of the nonprompt leptons is used in the $4\ell$ channel. Leptons are separated according to their origin: prompt, heavy-flavor and light-flavor, with the latter designation including leptons from photon conversions. As the rate of nonprompt muons originating from light-flavor hadrons is extremely low, the muons of heavy- and light-flavor origin are treated together. The control region defined in Table VII for the nonprompt lepton estimate in the $4\ell$ channel, where three light leptons are required, is used. It is composed of roughly 50% $Z+\text{jets}$, 30% diboson and 20% $t\bar{t}$ events. The control region is separated into four categories according to the flavor of the leptons ($eee$, $e\mu\mu$, $e\mu\mu$ and $\mu\mu\mu$) and a fit to the leading jet $p_T$ distribution is performed to extract three normalization factors: $\lambda_{\text{heavy}} = 1.48 \pm 0.22$, $\lambda_{\text{light}} = 0.72 \pm 0.53$ and $\lambda_{\ell} = 0.66 \pm 0.19$, where the errors are statistical. The normalization factors are applied to all events containing nonprompt leptons to correct the yields from the simulation in each category to data. The composition of the nonprompt leptons in the control region is shown in Fig. 7(a). The systematic uncertainty in each normalization factor is estimated to be 30% by varying the $p_T$ requirements on the leptons. The nonprompt lepton estimates were validated in various regions, as illustrated in Fig. 8(c) in a region identical to the control region used to extract the normalization factors except for being orthogonal in the $N_{\text{jets}}$ requirement (higher multiplicity $N_{\text{jets}} > 2$).

3. Nonprompt leptons and fake $\tau_{\text{had}}$ candidates in other channels

In the $3\ell+1\tau_{\text{had}}$, $2\ell OS+1\tau_{\text{had}}$ and $1\ell+2\tau_{\text{had}}$ channels, the background from nonprompt light leptons is a few percent and is estimated from simulation, but the fake $\tau_{\text{had}}$ background, mainly arising from $t\bar{t}$ and $t\bar{t}V$, is estimated from data. In the $2\ell SS+1\tau_{\text{had}}$ channel, both backgrounds are significant and hence are estimated from data. In the $2\ell OS+1\tau_{\text{had}}$ channel, the fake-factor method is used to estimate the background from events containing a fake $\tau_{\text{had}}$ candidate. The method assumes that the real contribution is described well by simulation. The fake factors are estimated using the control region defined in Table VII, which applies the nominal $2\ell OS+1\tau_{\text{had}}$ selection but requires at least three jets and vetoes events containing $b$ jets. The fake factors are parameterized as a function of $p_T^{\tau_{\text{had}}}$ and no significant dependence on other key event properties was found. Systematic uncertainties include the statistical uncertainty in the control regions, differences in the fake composition between the control and signal regions and the variation in the fake factors between different control regions. The total systematic uncertainty in the fake $\tau_{\text{had}}$ background estimate in this channel is 11%. Figure 8(d) illustrates a validation of this estimate in the $2\ell OS+1\tau_{\text{had}}$ selection region, which is largely dominated by events with a fake $\tau_{\text{had}}$.

As the origin of the $\tau_{\text{had}}$ fakes is very similar between the channels, as demonstrated in Fig. 7(b), an extrapolation is made to the $2\ell SS+1\tau_{\text{had}}$ and $3\ell+1\tau_{\text{had}}$ channels. The fake factors derived in the $2\ell OS+1\tau_{\text{had}}$ control region are then applied in the respective signal regions. Its dependence on $p_T$ was found to be negligible. Uncertainties in the scale factor are derived by comparing the value in the nominal control region to those obtained in control regions enriched in $t\bar{t}$ and $Z$ boson events, respectively. The final scale factor is $1.36 \pm 0.16$ including statistical and systematic uncertainties.

In the $2\ell SS+1\tau_{\text{had}}$ channel, this scale factor is applied only to backgrounds containing prompt leptons and fake $\tau_{\text{had}}$ candidates. An additional fake-factor method is used to estimate the background from events containing nonprompt light leptons. This fake factor is derived in a control region defined in Table VII, which differs from the signal region by looser lepton requirements and lower jet multiplicity. As in the $2\ell SS$ and $3\ell$ nonprompt lepton estimates, the change in the fraction of conversions from the control to the signal region is taken into account, with the same associated uncertainties. The total systematic uncertainty in the nonprompt lepton estimate in this channel is 55%, dominated by the statistical uncertainty in the closure test of the method found in simulation.

The dominant background in the $1\ell+2\tau_{\text{had}}$ signal region is $t\bar{t}$ production where one or two $\tau_{\text{had}}$ are fakes from $t\bar{t}$ decays. As there is equal probability for a jet to be reconstructed as a positively or negatively charged $\tau_{\text{had}}$, the fakes are estimated from a control region identical to the signal region except that the $\tau_{\text{had}}$ candidates are required to have the same charge, as shown in Table VII. This region contains almost entirely fakes from $t\bar{t}$ decays. The estimate is extrapolated to the signal region after using simulation to subtract the contribution from real $\tau_{\text{had}}$ in the control region. Using simulation, the nonclosure of this method was found to be below 30%, which is included as a systematic uncertainty.

4. Charge misassignment

The electron charge misassignment rate is measured in data, and the corresponding background is taken into account in the $2\ell SS$, $2\ell SS+1\tau_{\text{had}}$ channels and,
indirectly, in the $3\ell$ channel via the nonprompt background estimate, by scaling opposite-charge data events by this rate. The measurement is performed within a sample of $Z \rightarrow ee$ events reconstructed as same-charge pairs and as opposite-charge pairs. Six bins in $|\eta|$ and four bins in $p_T$ are used. The bins were chosen in accord with the size of the event sample and the variation of the rate with $|\eta|$ and $p_T$. The background is subtracted using a sideband method. The charge misassignment rate varies from $5 \times 10^{-5}$ for low-$p_T$ electrons ($p_T \approx 10$ GeV) at small $|\eta|$ to $10^{-2}$ for high-$p_T$ electrons ($p_T \gtrsim 100$ GeV) with $|\eta| > 2$.

The electron charge misassignment measurement is validated by a closure test in simulation using same-charge pairs, with the observed difference between measured and predicted rates being taken as the systematic uncertainty. An additional validation is performed in data by comparing the measured and estimated numbers of same-charge events. The results are found to agree

<table>
<thead>
<tr>
<th>Systematic uncertainty</th>
<th>Type</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>Pileup reweighting</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Physics Objects</td>
<td>SN</td>
<td>6</td>
</tr>
<tr>
<td>Electron</td>
<td>SN</td>
<td>15</td>
</tr>
<tr>
<td>Muon</td>
<td>SN</td>
<td>10</td>
</tr>
<tr>
<td>$\tau_{\text{had}}$</td>
<td>SN</td>
<td>28</td>
</tr>
<tr>
<td>Jet energy scale and resolution</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Jet vertex fraction</td>
<td>SN</td>
<td>126</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$</td>
<td>SN</td>
<td>3</td>
</tr>
<tr>
<td>Total (experimental)</td>
<td>SN</td>
<td>191</td>
</tr>
<tr>
<td>Data-driven nonprompt or fake leptons and charge misassignment</td>
<td>SN</td>
<td>38</td>
</tr>
<tr>
<td>Control region statistics</td>
<td>SN</td>
<td>22</td>
</tr>
<tr>
<td>Light-lepton efficiencies</td>
<td>SN</td>
<td>5</td>
</tr>
<tr>
<td>Nonprompt light-lepton estimates: nonclosure</td>
<td>N</td>
<td>5</td>
</tr>
<tr>
<td>$\gamma$-conversion fraction</td>
<td>N</td>
<td>5</td>
</tr>
<tr>
<td>Fake $\tau_{\text{had}}$ estimates</td>
<td>N/SN</td>
<td>12</td>
</tr>
<tr>
<td>Electron charge misassignment</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Total (data-driven reducible background)</td>
<td>SN</td>
<td>83</td>
</tr>
<tr>
<td>$t\bar{t}H$ modeling</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td>Cross section</td>
<td>S</td>
<td>3</td>
</tr>
<tr>
<td>Renormalization and factorization scales</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Parton shower and hadronization model</td>
<td>N</td>
<td>4</td>
</tr>
<tr>
<td>Shower tune</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>$t\bar{t}W$ modeling</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td>Cross section</td>
<td>S</td>
<td>3</td>
</tr>
<tr>
<td>Renormalization and factorization scales</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Shower tune</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>$t\bar{t}Z$ modeling</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td>Cross section</td>
<td>S</td>
<td>3</td>
</tr>
<tr>
<td>Renormalization and factorization scales</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Shower tune</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Other background modeling</td>
<td>N</td>
<td>15</td>
</tr>
<tr>
<td>Cross section</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Shower tune</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Total (signal and background modeling)</td>
<td>SN</td>
<td>41</td>
</tr>
<tr>
<td>Total (overall)</td>
<td>SN</td>
<td>315</td>
</tr>
</tbody>
</table>
within uncertainties. Additional systematic uncertainties applied to the estimate include the statistical uncertainty from the data and the variation in the rates when the $Z$-peak range definition is varied. The total systematic uncertainty in the charge misassignment background estimate is about 30%, with the dominant contribution at low $p_T$ from the closure tests and at high $p_T$ from the statistical uncertainty.

**VII. SYSTEMATIC UNCERTAINTIES**

The sources of systematic uncertainty considered in this analysis are summarized in Table IX. They impact the estimated signal and background rates, the migration of events between categories and/or the shape of the BDT discriminants used in the final fit. Systematic uncertainties are implemented in the fit as normalization factors that affect the normalization of a process in a given analysis category or as a shape variation that only affects the distribution of a discriminant in a given category but not its normalization. The impact of all these systematic uncertainties on the measured signal strength is discussed quantitatively in Sec. VIII.

The uncertainty in the combined 2015 + 2016 integrated luminosity is 2.1%. It is derived, following a methodology similar to that detailed in Ref. [103], from a calibration of the luminosity scale using $x$-$y$ beam-separation scans performed in August 2015 and May 2016.

The experimental uncertainties are related to the reconstruction and identification of light leptons and hadronically decaying $\tau$ leptons, to the reconstruction and $b$ tagging of jets, and to the reconstruction of $E_T^{\text{miss}}$. The sources that contribute to the uncertainty in the jet energy scale [104,105] are decomposed into uncorrelated components and treated as independent sources in the analysis. The total jet uncertainty varies from 1.0% to 5.5% depending on the jet $p_T$. The largest impact of experimental uncertainties on the signal strength $\mu = \sigma_{\tilde{t}\tilde{t},0}\sigma_{\tilde{t}\tilde{t},\text{SM}}$ arises from the jet energy scale, in particular, contributions from the different responses to quark and gluon jets, pileup subtraction, and in situ calibration in data [86].

The uncertainties in the $b$-tagging efficiencies measured in dedicated calibration analyses [88] are also decomposed into uncorrelated components. The large number of components for $b$ tagging is due to the calibration of the distribution of the BDT discriminant. The approximate relative size of the $b$-tagging efficiency uncertainty is 2% for $b$ jets, 10% for $c$ jets and $\tau$s, and 30% for light jets. The impact of the tagging uncertainty for jets containing either $c$ hadrons or $\tau_{\text{had}}$ is significant and, due to the calibration procedure applied, is taken as fully correlated between the two jet flavors.

Uncertainties in light-lepton reconstruction, identification, isolation and trigger efficiencies have negligible impact. The uncertainty in the identification efficiency for $\tau_{\text{had}}$ is 6% [81].

The systematic uncertainties associated with the estimation of the fake and nonprompt lepton backgrounds, as well as electron charge misassignment, are discussed in Sec. VI. They have large effects on the background estimates in all channels.

The systematic uncertainties associated with the generation of signal and background processes are due to uncertainties in the assumed cross sections and acceptance modeling for each process, and they are assessed in each category. The former are evaluated by varying the cross section of each process within its uncertainty, as described in Sec. III. The latter are estimated by comparing the results with those obtained using alternative simulated samples detailed in Sec. III. The most important uncertainty arising from theoretical predictions is in the assumed SM cross sections and the modeling of the acceptance for $t\bar{t}H$, $t\bar{t}Z$ and $t\bar{t}W$ production. The uncertainty in the shape of the simulated $t\bar{t}W$ and $t\bar{t}Z$ backgrounds due to the choice of event generator varies by at most 10% between bins. The uncertainties for $t\bar{t}\gamma$, $tZ$, $twZ$, and $VV\to \ell\ell XX$ include extrapolation uncertainties into the analysis phase space.

**VIII. STATISTICAL MODEL AND RESULTS**

Table X (top part) shows a comparison of the predicted yields to data in the eight signal and four control regions defined in Sec. V.

A maximum-likelihood fit is performed on all these 12 categories simultaneously to extract the $t\bar{t}H$ signal cross section normalized to the prediction from the SM ($\mu$) with the signal acceptance in the different regions derived assuming the SM. The statistical analysis of the data uses a binned likelihood function $L(\mu, \tilde{\theta})$, which is constructed from a product of Poisson probability terms to estimate $\mu$. The Higgs boson branching fractions and the cross section normalizations for associated production of a Higgs boson and a single top quark, which is treated as background, are set to their SM values. The former are evaluated by varying the cross section of each process within its uncertainty, as described in Sec. III. The latter are estimated by comparing the results with those obtained using alternative simulated samples detailed in Sec. III. The most important uncertainty arising from theoretical predictions is in the assumed SM cross sections and the modeling of the acceptance for $t\bar{t}H$, $t\bar{t}Z$ and $t\bar{t}W$ production. The uncertainty in the shape of the simulated $t\bar{t}W$ and $t\bar{t}Z$ backgrounds due to the choice of event generator varies by at most 10% between bins. The uncertainties for $t\bar{t}\gamma$, $tZ$, $twZ$, and $VV\to \ell\ell XX$ include extrapolation uncertainties into the analysis phase space.

The impact of systematic uncertainties on the signal and background expectations is described by nuisance parameters (NPs), $\tilde{\theta}$, which are constrained by Gaussian or log-normal probability density functions. The latter are used for normalization factors to ensure that they are always positive. The expected numbers of signal and background events are functions of $\tilde{\theta}$. The prior for each NP is added as a penalty term to the likelihood, $L(\mu, \tilde{\theta})$, to decrease it...
TABLE X  Background, signal and observed yields in the 12 analysis categories in 36.1 fb\(^{-1}\) of data at \(\sqrt{s} = 13\) TeV. Uncertainties in the background estimates due to systematic effects and to limited simulation sample size are shown. “Nonprompt,” “Fake \(r_{\text{had}}\)” and “q mis-id” refer to the data-driven background estimates described in Sec. VI. Rare processes (\(t\bar{t}Z\), \(t\bar{t}W\), \(t\bar{t}Z\), \(t\bar{t}WW\), triboson production, \(t\bar{t}t\bar{t}\), \(t\bar{t}H\), rare top decay) are labeled as “Other.” In the top part, the prefit values are quoted, i.e. using the initial values of background systematic uncertainty nuisance parameters and the signal expected from the SM. In the bottom part, the corresponding postfit values are quoted. In the postfit case, the prediction and uncertainties for \(t\bar{t}H\) reflect the best-fit production rate of 1.6\(^{+0.5}_{-0.4}\) times the Standard Model prediction and the uncertainty in the total background estimate is smaller than for the prefit values due to anticorrelations between the nuisance parameters obtained in the fit.

<table>
<thead>
<tr>
<th>Category</th>
<th>Nonprompt</th>
<th>Fake (r_{\text{had}})</th>
<th>q mis-id</th>
<th>(t\bar{t}W)</th>
<th>(t\bar{t}Z)</th>
<th>Diboson</th>
<th>Other</th>
<th>Total Bkgd.</th>
<th>(t\bar{t}H)</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2\ell)SS</td>
<td>233 ± 39</td>
<td>...</td>
<td>33 ± 11</td>
<td>123 ± 18</td>
<td>41.4 ± 5.6</td>
<td>25 ± 15</td>
<td>28.4 ± 5.9</td>
<td>484 ± 38</td>
<td>42.6 ± 4.2</td>
<td>514</td>
</tr>
<tr>
<td>(3\ell)SR</td>
<td>14.5 ± 4.3</td>
<td>...</td>
<td>...</td>
<td>5.5 ± 1.2</td>
<td>12.0 ± 1.8</td>
<td>1.2 ± 1.2</td>
<td>5.8 ± 1.4</td>
<td>39.1 ± 5.2</td>
<td>11.2 ± 1.6</td>
<td>61</td>
</tr>
<tr>
<td>(3\ell)(t\bar{t}W) CR</td>
<td>13.3 ± 4.3</td>
<td>...</td>
<td>...</td>
<td>19.9 ± 3.1</td>
<td>8.7 ± 1.1</td>
<td>&lt; 0.2</td>
<td>4.53 ± 0.92</td>
<td>46.5 ± 5.4</td>
<td>4.18 ± 0.46</td>
<td>56</td>
</tr>
<tr>
<td>(3\ell)(t\bar{t}Z) CR</td>
<td>3.9 ± 2.5</td>
<td>...</td>
<td>...</td>
<td>2.71 ± 0.56</td>
<td>66.11</td>
<td>8.4 ± 5.3</td>
<td>12.9 ± 4.2</td>
<td>93 ± 13</td>
<td>3.17 ± 0.41</td>
<td>107</td>
</tr>
<tr>
<td>(3\ell)VV CR</td>
<td>27.7 ± 8.7</td>
<td>...</td>
<td>...</td>
<td>4.9 ± 1.0</td>
<td>21.3 ± 3.4</td>
<td>51 ± 30</td>
<td>17.9 ± 6.1</td>
<td>123 ± 32</td>
<td>1.67 ± 0.25</td>
<td>109</td>
</tr>
<tr>
<td>(3\ell)(t\bar{t}) CR</td>
<td>70 ± 17</td>
<td>...</td>
<td>...</td>
<td>10.5 ± 1.5</td>
<td>7.9 ± 1.1</td>
<td>7.2 ± 4.8</td>
<td>7.3 ± 1.9</td>
<td>103 ± 17</td>
<td>4.00 ± 0.49</td>
<td>85</td>
</tr>
<tr>
<td>(4\ell)Z-emr.</td>
<td>0.11 ± 0.07</td>
<td>...</td>
<td>...</td>
<td>&lt; 0.01</td>
<td>1.52 ± 0.23</td>
<td>0.43 ± 0.23</td>
<td>0.21 ± 0.09</td>
<td>2.26 ± 0.34</td>
<td>1.06 ± 0.14</td>
<td>2</td>
</tr>
<tr>
<td>(4\ell)Z-dep.</td>
<td>0.01 ± 0.01</td>
<td>...</td>
<td>...</td>
<td>&lt; 0.01</td>
<td>0.04 ± 0.02</td>
<td>&lt; 0.01</td>
<td>0.06 ± 0.03</td>
<td>0.11 ± 0.03</td>
<td>0.20 ± 0.03</td>
<td>0</td>
</tr>
<tr>
<td>(1\ell + 2r_{\text{had}})</td>
<td>...</td>
<td>65 ± 21</td>
<td>...</td>
<td>0.09 ± 0.09</td>
<td>3.3 ± 1.0</td>
<td>1.3 ± 1.0</td>
<td>0.98 ± 0.35</td>
<td>71 ± 21</td>
<td>4.3 ± 1.0</td>
<td>67</td>
</tr>
<tr>
<td>(2\ell)SS + 1(r_{\text{had}})</td>
<td>2.41 ± 1.4</td>
<td>1.80 ± 0.30</td>
<td>0.05 ± 0.02</td>
<td>0.88 ± 0.24</td>
<td>1.83 ± 0.37</td>
<td>0.12 ± 0.18</td>
<td>1.06 ± 0.24</td>
<td>8.2 ± 1.6</td>
<td>3.09 ± 0.46</td>
<td>18</td>
</tr>
<tr>
<td>(2\ell)OS + 1(r_{\text{had}})</td>
<td>...</td>
<td>756 ± 80</td>
<td>...</td>
<td>6.5 ± 1.3</td>
<td>11.4 ± 1.9</td>
<td>2.0 ± 1.3</td>
<td>5.8 ± 1.5</td>
<td>782 ± 81</td>
<td>14.2 ± 2.0</td>
<td>807</td>
</tr>
<tr>
<td>(3\ell + 1r_{\text{had}})</td>
<td>...</td>
<td>0.75 ± 0.15</td>
<td>...</td>
<td>0.04 ± 0.04</td>
<td>1.38 ± 0.24</td>
<td>0.002 ± 0.002</td>
<td>0.38 ± 0.10</td>
<td>2.55 ± 0.32</td>
<td>1.51 ± 0.23</td>
<td>5</td>
</tr>
<tr>
<td>Prefit yields</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2\ell)SS</td>
<td>211 ± 26</td>
<td>...</td>
<td>28.3 ± 9.4</td>
<td>127 ± 18</td>
<td>42.9 ± 5.4</td>
<td>20.0 ± 6.3</td>
<td>28.5 ± 5.7</td>
<td>459 ± 24</td>
<td>67 ± 18</td>
<td>514</td>
</tr>
<tr>
<td>(3\ell)SR</td>
<td>13.2 ± 3.1</td>
<td>...</td>
<td>...</td>
<td>5.8 ± 1.2</td>
<td>12.9 ± 1.6</td>
<td>1.2 ± 1.1</td>
<td>5.9 ± 1.3</td>
<td>39.0 ± 4.0</td>
<td>17.7 ± 4.9</td>
<td>61</td>
</tr>
<tr>
<td>(3\ell)(t\bar{t}W) CR</td>
<td>11.7 ± 3.0</td>
<td>...</td>
<td>...</td>
<td>20.4 ± 3.0</td>
<td>8.9 ± 1.0</td>
<td>&lt; 0.2</td>
<td>4.54 ± 0.88</td>
<td>45.6 ± 4.0</td>
<td>6.6 ± 1.9</td>
<td>56</td>
</tr>
<tr>
<td>(3\ell)(t\bar{t}Z) CR</td>
<td>3.5 ± 2.1</td>
<td>...</td>
<td>...</td>
<td>2.82 ± 0.56</td>
<td>70.4 ± 8.6</td>
<td>7.1 ± 3.0</td>
<td>13.6 ± 4.2</td>
<td>97.4 ± 8.6</td>
<td>5.1 ± 1.4</td>
<td>107</td>
</tr>
<tr>
<td>(3\ell)VV CR</td>
<td>22.4 ± 5.7</td>
<td>...</td>
<td>...</td>
<td>5.05 ± 0.94</td>
<td>22.0 ± 3.0</td>
<td>39 ± 11</td>
<td>18.1 ± 5.9</td>
<td>106.8 ± 9.4</td>
<td>2.61 ± 0.82</td>
<td>109</td>
</tr>
<tr>
<td>(3\ell)(t\bar{t}) CR</td>
<td>56.0 ± 8.1</td>
<td>...</td>
<td>...</td>
<td>10.7 ± 1.4</td>
<td>8.1 ± 1.0</td>
<td>5.9 ± 2.7</td>
<td>7.1 ± 1.8</td>
<td>87.8 ± 7.9</td>
<td>6.3 ± 1.8</td>
<td>85</td>
</tr>
<tr>
<td>(4\ell)Z-emr.</td>
<td>0.10 ± 0.07</td>
<td>...</td>
<td>...</td>
<td>&lt; 0.01</td>
<td>1.60 ± 0.22</td>
<td>0.37 ± 0.15</td>
<td>0.22 ± 0.10</td>
<td>2.29 ± 0.28</td>
<td>1.65 ± 0.47</td>
<td>2</td>
</tr>
<tr>
<td>(4\ell)Z-dep.</td>
<td>0.01 ± 0.01</td>
<td>...</td>
<td>...</td>
<td>&lt; 0.01</td>
<td>0.04 ± 0.02</td>
<td>&lt; 0.01</td>
<td>0.07 ± 0.03</td>
<td>0.11 ± 0.03</td>
<td>0.52 ± 0.09</td>
<td>0</td>
</tr>
<tr>
<td>(1\ell + 2r_{\text{had}})</td>
<td>...</td>
<td>58.0 ± 6.8</td>
<td>...</td>
<td>0.11 ± 0.11</td>
<td>3.31 ± 0.90</td>
<td>0.98 ± 0.75</td>
<td>0.98 ± 0.33</td>
<td>63.4 ± 6.7</td>
<td>6.5 ± 2.0</td>
<td>67</td>
</tr>
<tr>
<td>(2\ell)SS + 1(r_{\text{had}})</td>
<td>1.86 ± 0.91</td>
<td>1.86 ± 0.27</td>
<td>0.05 ± 0.02</td>
<td>0.97 ± 0.26</td>
<td>1.96 ± 0.37</td>
<td>0.15 ± 0.20</td>
<td>1.09 ± 0.24</td>
<td>7.9 ± 1.2</td>
<td>5.1 ± 1.3</td>
<td>18</td>
</tr>
<tr>
<td>(2\ell)OS + 1(r_{\text{had}})</td>
<td>...</td>
<td>756 ± 28</td>
<td>...</td>
<td>6.6 ± 1.3</td>
<td>11.5 ± 1.7</td>
<td>1.64 ± 0.92</td>
<td>6.1 ± 1.5</td>
<td>782 ± 27</td>
<td>21.7 ± 5.9</td>
<td>807</td>
</tr>
<tr>
<td>(3\ell + 1r_{\text{had}})</td>
<td>...</td>
<td>0.75 ± 0.14</td>
<td>...</td>
<td>0.04 ± 0.04</td>
<td>1.42 ± 0.22</td>
<td>0.002 ± 0.002</td>
<td>0.40 ± 0.10</td>
<td>2.61 ± 0.30</td>
<td>2.41 ± 0.68</td>
<td>5</td>
</tr>
</tbody>
</table>
when $\theta$ is shifted away from its nominal value. The statistical uncertainties in the simulated background predictions and the control regions used for the nonprompt and fake estimates are included as bin-by-bin NPs using the Beeston-Barlow technique [106].

The test statistic, $q_\mu$, is constructed from the profile log-likelihood ratio:

$$q_\mu = -2 \ln \Lambda_\mu = -2 \ln \left( \frac{L(\hat{\mu}, \hat{\theta})}{L(\hat{\mu}, \hat{\theta})} \right),$$

where $\hat{\mu}$ and $\hat{\theta}$ are the parameters that maximize the likelihood and $\hat{\theta}$ are the NPs that maximize the likelihood for a given $\mu$. The test statistic is used to quantify how well the observed data agrees with the background-only hypothesis.

The fitted $\hat{\mu}$ value is obtained by maximizing the likelihood function with respect to all parameters, and the total uncertainty $\sigma_\mu$ is obtained from the variation of $-2 \ln \Lambda_\mu$ by one unit from its minimum. Systematic uncertainties are found by subtracting in quadrature the statistical uncertainty, determined by fixing all NPs to their best-fit values, from the total 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 nonprompt estimates with all NPs set to their best-fit values obtained from the fit to data. The significance is obtained from the test statistic in the asymptotic limit [107]. As the $4\ell$ channel has few events,

### Table XI

Summary of the basic characteristics and analysis strategies of all channels. In the $4\ell$ channel, the two entries correspond to the $Z$-enriched and the $Z$-depleted categories; 1D and 5D refer to one- and five-dimensional BDTs, respectively.

<table>
<thead>
<tr>
<th>BDT trained against</th>
<th>$2\ell$SS</th>
<th>$3\ell$</th>
<th>$4\ell$</th>
<th>$1\ell + 2\tau_{\text{had}}$</th>
<th>$2\ell + 1\tau_{\text{had}}$</th>
<th>$2\ell + 1\tau_{\text{had}}$</th>
<th>$3\ell + 1\tau_{\text{had}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discriminant</td>
<td>Fakes and $t\bar{t}V$, $t\bar{t}W$, $t\bar{t}Z$, $VV$</td>
<td>$t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, $VV$</td>
<td>$t\bar{t}$, $t\bar{t}Z$</td>
<td>Event count</td>
<td>BDT</td>
<td>BDT</td>
<td>Event count</td>
</tr>
<tr>
<td>Number of bins</td>
<td>6</td>
<td>5</td>
<td>1/1</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Control regions</td>
<td>...</td>
<td>4</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

![FIG. 9](image.png)

The fitted $\hat{\mu}$ value is obtained by maximizing the likelihood function with respect to all parameters, and the total uncertainty $\sigma_\mu$ is obtained from the variation of $-2 \ln \Lambda_\mu$ by one unit from its minimum. Systematic uncertainties are found by subtracting in quadrature the statistical uncertainty, determined by fixing all NPs to their best-fit values, from the total 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 nonprompt estimates with all NPs set to their best-fit values obtained from the fit to data. The significance is obtained from the test statistic in the asymptotic limit [107]. As the $4\ell$ channel has few events,
the validity of this assumption was verified using pseudoexperiments.

As described in Sec. VII, a large number of systematic uncertainties, whose effects are accounted for using NPs, affect the final results. In total, 315 NPs are considered, most having experimental origin. The experimental uncertainties are fully correlated across categories, with the exception of those related to the quark or gluon jet composition and some uncertainties associated with the fake and nonprompt lepton background determinations, which are specific to the different categories, as detailed in Sec. VI. As the residual prompt (mainly $t\bar{t}W$ and $VV$) background contribution is subtracted from the control regions to extract the fake and nonprompt leptons, the associated nuisance parameters are taken as fully correlated with the theoretical cross-section systematic uncertainties. The same treatment is used for the uncertainty associated to the measurement of the background from charge misassignment, which is also subtracted from the control regions.

The fit uses templates constructed from the predicted yields for the signal and the various backgrounds in the bins of the input distribution in each region. The systematic uncertainties are encoded in templates of variations relative to the nominal template for each upward or downward ($\pm \sigma$) variation. A smoothing procedure is applied to remove large local fluctuations in the templates for some background processes in certain regions. Systematic uncertainties that have a negligible impact on the final results are removed to improve the speed of the fit: a normalization or a shape uncertainty is not applied if the associated variation is below 1% in all bins; this reduces the number of nuisance parameters to 230. Most of the neglected nuisance parameters are those related to flavor tagging.

![Graph](image-url)

**Fig. 10.** Comparison of prediction to data after the fit in the eight signal and four control regions. The background contributions after the global fit are shown as filled histograms. The total background before the fit is shown as a dashed blue histogram. The Higgs boson signal ($m_{H} = 125$ GeV), scaled according to the results of the fit, is shown as a filled red histogram superimposed on the fitted backgrounds. The size of the combined statistical and systematic uncertainty in the sum of the signal and fitted background is indicated by the blue hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel. The yields in each region are shown in Table X.
The behavior of the global fit is studied by performing a number of checks including evaluating how much each NP is pulled from its nominal value, how much its uncertainty decreases from the nominal uncertainty and which correlations develop between initially uncorrelated systematic uncertainties. The stability of the results was tested by performing fits for each channel independently and in combination.

The impact of each systematic uncertainty on the final result is assessed by performing the fit with the parameter fixed to its fitted value varied up or down by its fitted uncertainty, with all the other parameters allowed to vary.

**FIG. 11.** The distribution of the discriminating variables observed in data (points with bars indicating the statistical errors) and expected (histograms) in the (a) 2eSS, (b) 3ℓ, (c) 4ℓ (Z-enriched) and (d) 4ℓ (Z-depleted) signal regions. The background contributions after the global fit are shown as filled histograms. The total background before the fit is shown as a dashed blue histogram. The Higgs boson signal ($m_H = 125$ GeV), scaled according to the results of the fit, is shown as a filled red histogram superimposed on the fitted backgrounds. The size of the combined statistical and systematic uncertainty in the sum of the signal and fitted background is indicated by the blue hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.
and calculating the $\Delta \mu$ to the baseline fit. The ranking obtained for those nuisance parameters with the largest contribution to the uncertainty in the signal strength is shown in Fig. 9. The NP with the largest pull from its nominal value is the uncertainty in the nonprompt lepton estimate due to the nonclosure in the $3\ell$ channel. This is mainly due to the slight deficit observed in the $3\ell$ control region relative to the background prediction. As the fit includes bins with high purity of nonprompt light leptons and fake $\tau$ had backgrounds, the precision of these estimates is increased, as is shown in Table X. The correlations between the nuisance parameters were checked and no

FIG. 12. The distribution of the discriminating variables observed in data (points with bars indicating the statistical errors) and expected (histograms) in the (a) $2\ell SS + 1\tau_{\text{had}}$, (b) $2\ell OS + 1\tau_{\text{had}}$, (c) $1\ell + 2\tau_{\text{had}}$ and (d) $3\ell + 1\tau_{\text{had}}$ signal regions. The background contributions after the global fit are shown as filled histograms. The total background before the fit is shown as a dashed blue histogram. The Higgs boson signal ($m_H = 125$ GeV), scaled according to the results of the fit, is shown as a filled red histogram superimposed on the fitted backgrounds. The size of the combined statistical and systematic uncertainty in the sum of the signal and fitted background is indicated by the blue hatched band. The ratio of the data to the sum of the signal and fitted background is shown in the lower panel.
unexpected correlations were observed. The impact of the most important groups of systematic uncertainties on the measured value of $\mu$ is shown in Table XII. The uncertainties with the largest impact are those associated with the signal modeling, the jet energy scale and the nonprompt light-lepton estimate. The signal uncertainty is separated into two components to show the uncertainty due to the acceptance and the one due to the cross section. The uncertainties in the nonprompt light-lepton estimates, the fake $\tau_{\text{had}}$ estimates and the charge misassignment have large statistical components due to the small data sample size. The large impact of the luminosity uncertainty is due to its effect on both the signal and simulated background predictions. Although the individual groups are initially largely uncorrelated, a small correlation is introduced by the fit to data.

![FIG. 13. The observed best-fit values of the $t\bar{t}H$ signal strength $\mu$ and their uncertainties by final-state category and combined. The individual $\mu$ values for the channels are obtained from a simultaneous fit with the signal-strength parameter for each channel floating independently. The SM prediction is $\mu = 1$.](image)

![FIG. 14. 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 signal regions are combined into bins of $\log_{10}(S/B)$, where $S$ is the expected signal yield and $B$ the background yield from the unconditional fit. The background yields are shown as the fitted values, while the signal yields are shown for the fitted value ($\mu = 1.6$) and the SM prediction ($\mu = 1$). The total background before the fit is shown as a dashed blue histogram. 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 $\mu = 1.6$ ($\mu = 1$) and background relative to the background-only prediction. The background is also shown after the fit to data assuming zero signal contribution as well as its pull (dotted black line) relative to the background from the nominal fit.](image)

Figure 10 and Table X (bottom part) compare the data to the yields after the predictions were adjusted by the fit in the 12 signal and control regions. Figures 11 and 12 show the

![TABLE XIII. Observed and expected best-fit values of the signal strength $\mu$ and associated significance under the SM background-only hypothesis. The expected values are shown for the prefit background estimates. The observed significance is indicated with $\sigma$—for the channels where $\mu$ is negative.](image)
assuming that the observed signal is due to the SM Higgs strengths in the seven channels are compatible is 34%. When for each channel. The probability that the fitted signal from the full fit but with a separate parameter of interest

Table XIII. The individual channel results are extracted and the combination of all channels are shown in Fig. 13 and Table XIII. The individual channel results are extracted from the full fit but with a separate parameter of interest for each channel. The probability that the fitted signal strengths in the seven channels are compatible is 34%. When assuming that the observed signal is due to the SM Higgs boson, the excess over the SM signal-plus-background

Distributions of the discriminating variables used by the fit in the eight signal regions. Distributions are shown both before and after the fit to the data. An excess of events over the expected Standard Model background is found with an observed (expected) significance of 4.1 (2.8) standard deviations. The observed (expected) best-fit value of \( \mu \) is \( 1.6^{+0.3}_{-0.3} \) (stat) \( +0.4 \) (syst) = \( 1.6^{+0.5}_{-0.3} \) (stat) \( +0.3 \) (syst) = \( 1.00^{+0.4}_{-0.3} \). The best-fit value of \( \mu \) for each individual channel and the combination of all channels are shown in Fig. 13 and Table XIII. The individual channel results are extracted from the full fit but with a separate parameter of interest for each channel. The probability that the fitted signal strengths in the seven channels are compatible is 34%. When assuming that the observed signal is due to the SM Higgs boson, the excess over the SM signal-plus-background

Table XIV. Summary of the observed and expected \( \mu \) measurements and \( t\bar{t}H \) production significance from individual analyses and the combination. As no events are observed in the \( H \rightarrow 4\ell \) analysis, a 68% confidence level (C.L.) upper limit on \( \mu \), computed using the C.L., method [108], is reported.

![Figure 15](image_url)

**FIG. 15.** Summary of the measurements of \( \mu \) from individual analyses and the combined result. “ML” refers to the multi-lepton decay channels discussed in Sec. VIII. The best-fit values of \( \mu \) for the individual analyses are extracted independently, and systematic uncertainty nuisance parameters are only correlated for the combination. As no events are observed in the \( H \rightarrow 4\ell \) analysis, a 68% confidence level (C.L.) upper limit on \( \mu \), computed using the C.L., method [108], is reported.

**Table XV.** Summary of the uncertainties affecting the combined value of \( \mu \).

<table>
<thead>
<tr>
<th>Uncertainty source</th>
<th>( \Delta \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t\bar{t} ) modeling in ( H \rightarrow b\bar{b} ) analysis</td>
<td>+0.15 –0.14</td>
</tr>
<tr>
<td>( t\bar{t}H ) modeling (cross section)</td>
<td>+0.13 –0.06</td>
</tr>
<tr>
<td>Nonprompt light-lepton and ( \tau ) had estimates</td>
<td>+0.09 –0.09</td>
</tr>
<tr>
<td>Simulation statistics</td>
<td>+0.08 –0.08</td>
</tr>
<tr>
<td>Jet energy scale and resolution</td>
<td>+0.08 –0.07</td>
</tr>
<tr>
<td>( iV ) modeling</td>
<td>+0.07 –0.07</td>
</tr>
<tr>
<td>( iH ) modeling (acceptance)</td>
<td>+0.07 –0.04</td>
</tr>
<tr>
<td>Other non-Higgs boson backgrounds</td>
<td>+0.06 –0.05</td>
</tr>
<tr>
<td>Other experimental uncertainties</td>
<td>+0.05 –0.05</td>
</tr>
<tr>
<td>Luminosity</td>
<td>+0.05 –0.04</td>
</tr>
<tr>
<td>Jet flavor tagging</td>
<td>+0.03 –0.02</td>
</tr>
<tr>
<td>Modeling of other Higgs boson production modes</td>
<td>+0.01 –0.01</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>+0.27 –0.23</td>
</tr>
<tr>
<td>Statistical uncertainty</td>
<td>+0.19 –0.19</td>
</tr>
<tr>
<td>Total uncertainty</td>
<td>+0.34 –0.30</td>
</tr>
</tbody>
</table>

Hypothesis has a significance of 1.4\( \sigma \). A model-dependent extrapolation is made to the inclusive phase space, and the measured \( t\bar{t}H \) production cross section is \( \sigma (t\bar{t}H) = 790^{+150}_{-150} (stat) +170_{-150} (syst) \) fb = \( 790^{+230}_{-210} \) fb. The predicted cross section is \( \sigma (iH) = 507^{+35}_{-50} \) fb.

For the 4\( \ell \), 2\( \ell OS + 1\tau \) had, and 3\( \ell \) + 1\( \tau \) had channels, the uncertainties in \( \mu \) are mainly statistical, while the statistical and systematic uncertainties are of comparable size for the 2\( \ell SS \), 3\( \ell \), 2\( \ell SS + 1\tau \) had, and 1\( \ell \) + 2\( \tau \) had channels. Figure 14 shows the data, background and signal

![Figure 16](image_url)

**FIG. 16.** Summary of the best-fit values of \( \mu \) broken down by Higgs boson decay mode. The decays \( H \rightarrow WW^* \) and \( H \rightarrow ZZ^* \) are assumed to have the same signal-strength modification factor and are shown together as \( VV \). All systematic uncertainties are correlated as in the nominal result.
yields, where the final-discriminant bins in all signal regions are combined into bins of log(S/B), S being the expected signal yield and B the fitted background yield. The most sensitive $2\ell\ell$SS, $3\ell$ and $2\ell\ell$ + 1$\ell_{\text{had}}$ analyses were cross-checked with simpler cut-and-count analyses with reduced sensitivity. The observed significance relative to the background-only hypothesis is 1.2$\sigma$, 2.3$\sigma$ and 2.3$\sigma$, respectively. The observed signal strengths in the cross-check analyses are found to be statistically compatible with those from the nominal analyses.

An alternative fit where $t\bar{t}W$ and $t\bar{t}Z$ normalizations were left free together with $\mu$ was performed as a cross-check. The expected sensitivity to $\mu$ is 15% worse than with the nominal fit. The observed best-fit value of $\mu$ is $1.6^{+0.6}_{-0.5}$, in agreement with the result obtained with the nominal fit. The fitted $t\bar{t}W$ and $t\bar{t}Z$ cross-section modifiers are $0.92 \pm 0.32$ and $1.17^{+0.25}_{-0.22}$, respectively, in agreement with the SM predictions.

**IX. COMBINATION OF ATLAS $t\bar{t}H$ SEARCHES**

In addition to the results reported in Sec. VIII (referred to hereafter as the multilepton analysis), the ATLAS Collaboration has carried out searches for $t\bar{t}H$ production at $\sqrt{s} = 13$ TeV using other Higgs boson decay modes:

(i) $H \rightarrow b\bar{b}$, in the lepton + jets and dileptonic $t\bar{t}$ final states [36],

(ii) $H \rightarrow \gamma\gamma$, in lepton+jets/dileptonic and all-hadronic $t\bar{t}$ decay channels [37]. In addition, specialized categories sensitive to $tHqb/WtH$ production also have significant $t\bar{t}H$ acceptance and are included.

(iii) $H \rightarrow ZZ^{*} \rightarrow 4\ell$ (hereafter $H \rightarrow 4\ell$), in a single category including all $t\bar{t}$ decay channels [38].

All analyses use the same 36.1 fb$^{-1}$ of data. The overlap between the signal and control regions of all the analyses was checked and found to be negligible. All analyses use the same Monte Carlo event generators for $t\bar{t}H$ production and use nominal Higgs boson decay branching fractions from Ref. [31] assuming $m_H = 125$ GeV.

For the extraction of the $t\bar{t}H$ signal strength $\mu$, the single top quark and Higgs boson associated production processes $tHqb$ and $WtH$ are considered backgrounds and fixed to their SM predictions with appropriate theoretical uncertainties. All other Higgs boson production mechanisms contribute negligibly to the multilepton and $H \rightarrow b\bar{b}$ analyses and are ignored. The searches for $t\bar{t}H$ production in $H \rightarrow \gamma\gamma$ and $H \rightarrow 4\ell$ reported in Refs. [37,38] both utilize categories targeting $t\bar{t}H$ production in global analyses of all Higgs boson production; in the following result, only the $t\bar{t}H$ -enhanced categories from those results are considered. These categories have non-negligible contamination from other production mechanisms (4%–21% in the $t\bar{t}H$ categories with $H \rightarrow \gamma\gamma$, 21%–64% in the $tHqb/WtH$ categories with $H \rightarrow \gamma\gamma$, 23% in $H \rightarrow 4\ell$). The best-fit values for $\mu$ obtained in those analyses result from multiple-parameter-of-interest fits that allow other Higgs boson production mode signal strengths to take on non-SM values. In the following discussion, non-$t\bar{t}H$ Higgs boson production mechanism cross sections and all Higgs boson branching fractions are set to SM expectations with theoretical errors considered as systematic uncertainties [31]. This results in slightly different $\mu$ values than reported in the stand-alone analyses. Details of the modeling and simulation of non-$t\bar{t}H$ production modes can be found in Refs. [37,38].

The combined likelihood function $L(\mu, \vec{\theta})$ is obtained from the product of likelihood functions of the individual analyses. The nuisance parameters associated with the same sources in the different analyses are treated as follows:

(i) **Higgs boson production and decay.**—All analyses use the same nominal production cross sections and decay branching fractions. All theoretical uncertainties associated with these parameters are fully correlated between analyses.

(ii) **Background uncertainties.**—The cross-section and modeling uncertainties for MC-estimated $t\bar{t}Z$, $t\bar{t}W$, $tZqb/WtZ$, $WZ/ZZ$, $tW$, and $tWt$ production are correlated between the $H \rightarrow b\bar{b}$ and multilepton analyses. The modeling systematic uncertainties of the dominant background of $t\bar{t}$ in the $H \rightarrow b\bar{b}$ analyses are not applied to any other channels, as the relevant regions of phase space are not similar and other channels have independent methods of estimating the relevant $t\bar{t}$ background.

(iii) **Experimental uncertainties.**—The dominant experimental systematic uncertainties are associated with the jet energy scale, jet energy resolution, and flavor tagging. Nuisance parameters related to the jet energy scale are correlated between the analyses with the exception of the uncertainty in the fractions of jets initiated by quarks and by gluons, which differs between the channels. The jet energy resolution is correlated between all channels except for the control regions of the $H \rightarrow b\bar{b}$ analysis, to avoid constraining this systematic uncertainty in the signal regions; this gives a conservative estimate of the impact. The $H \rightarrow \gamma\gamma$ and $H \rightarrow 4\ell$ analyses use a different calibration for the flavor-tagging efficiencies and mistag rates compared to the $H \rightarrow b\bar{b}$ and multilepton analyses. Due to this, the flavor-tagging uncertainties are correlated between $H \rightarrow \gamma\gamma$ and $H \rightarrow 4\ell$ and between $H \rightarrow b\bar{b}$ and multilepton analyses, but are uncorrelated between the two pairs. The flavor-tagging uncertainties are constrained significantly by the $H \rightarrow b\bar{b}$ analysis, due to its large samples of $b$ and $c$ jets, which carries over to the multilepton analysis.

Other experimental systematic uncertainties such as luminosity, pileup effects, lepton identification,
isolation, and trigger efficiencies are treated as correlated, except for statistical uncertainties associated with efficiency measurements for different working points.

None of the NPs in the fit are strongly constrained by more than one analysis, and the value of \( \mu \) obtained from the combined fit does not depend on the choice of the correlation scheme.

The best-fit value of the \( t\bar{t}H \) signal strength, as determined from the combined likelihood function, is

\[
\mu = 1.17 \pm 0.19 \text{(stat)} \pm 0.27 \pm 0.23 \text{ (syst).}
\]

The background-only hypothesis (\( \mu = 0 \)) is excluded at 4.2\( \sigma \), with an expectation of 3.8\( \sigma \) in the case of a SM signal. This constitutes evidence for \( t\bar{t}H \) production.

The values of \( \mu \) obtained in each analysis, and the result of the combination, are shown in Fig. 15 and Table XIV. The probability that the signal strengths from the SM, from the individual analyses are compatible with the combined value of \( \mu \) is 38\%. The impact of various uncertainties on the combination is shown in Table XV. The leading systematic uncertainties are those associated with the \( t\bar{t}H \) signal modeling and cross section and the \( t\bar{t} \) background modeling in the H \( \rightarrow \) \( b\bar{b} \) analysis. The cross section for \( t\bar{t}H \) production corresponding to the best-fit value of \( \mu \) is 590^{+160}_{-150} \text{ fb}, as compared to the SM prediction of \( \sigma(t\bar{t}H) = 507^{+35}_{-30} \text{ fb}. \)

Due to the different acceptances for the different analysis categories for different Higgs boson decay modes, it is possible to independently determine \( \mu \) in different Higgs boson decay modes. In particular, the multilepton analysis has categories with zero and \( \geq 1 \) \( \tau_{\text{had}} \) candidates, which are enriched in \( H \rightarrow WW^* \) and \( H \rightarrow \tau \tau \), respectively (see Fig. 4). The result of a fit for four signal strengths is shown in Fig. 16. Due to very weak sensitivity for \( H \rightarrow ZZ^* \), the ratio of branching fractions of \( H \rightarrow ZZ^* \) and \( H \rightarrow WW^* \) are assumed to be as in the SM and a single combined signal strength for \( H \rightarrow VV \) is computed. For \( H \rightarrow b\bar{b} \) and \( H \rightarrow \gamma\gamma \) the result is essentially the same as for the individual analyses, due to the high purity of those signal regions for the respective Higgs boson decays. The \( H \rightarrow WW^* \) and \( H \rightarrow \tau\tau \) decays are distinguished only by their different contributions to the various multilepton signal regions, resulting in a significant anticorrelation. Two-dimensional scans of the signal strengths are shown in Fig. 17 for \( H \rightarrow b\bar{b} \) versus \( H \rightarrow VV \) and for \( H \rightarrow \tau\tau \) versus \( H \rightarrow VV \); in these plots the two signal strengths not shown are profiled in the scan.

The \( t\bar{t}H \) analyses are sensitive to the \( Htt, Hbb, \) and \( H\tau\tau \) fermion couplings, the \( HWW \) and \( HZZ \) gauge boson couplings, and the effective \( H\gamma\gamma \) coupling. Accordingly, constraints can be placed on deviations of these couplings from the SM. An interpretation is made using the \( \kappa \) parameterization, in which Higgs boson couplings to particle species \( i \) are linearly scaled by factors \( \kappa_i \). Here, all fermion couplings are assumed to scale by a common factor \( \kappa_F \) and the \( WW/ZZ \) couplings by a common factor \( \kappa_V \). As only the relative sign of the \( \kappa \) factors is meaningful, the convention that \( \kappa_V \geq 0 \) is chosen. Modifications to loop-induced processes are determined by multiplying the contributing SM amplitudes by the relevant \( \kappa \) factors; no contributions from non-SM particles are considered and no non-SM Higgs boson decay modes are allowed. The relevant parameterizations are given in Ref. [18]. In particular, the factor \( \kappa_\gamma \) modifying the effective \( H\gamma\gamma \) coupling is expressed in terms of \( \kappa_X \) and \( \kappa_F \), and \( \kappa_L \) is set equal to \( \kappa_F \). The total width of the Higgs boson is modified appropriately.

The \( t\bar{t}H \) analyses, especially the \( H \rightarrow \gamma\gamma \), multilepton, and \( H \rightarrow 4\ell \) channels, have acceptance for \( tHqb \) and \( WhH \) production. The amplitudes for the \( H \rightarrow \gamma\gamma \) decay and the production of \( tHqb \) and \( WhH \) involve interference between the \( Htt \) and \( HWW \) couplings. In the SM, the interference is destructive, almost completely in the case of \( tHqb \) and...
are in good agreement with the Standard Model values

\( \kappa \) terms of the two couplings.

In this parameterization, is able to resolve the relative sign of \( WtH \) mechanisms and decays are assumed to be constant as the analysis acceptances for all Higgs boson production mechanisms are varied over the scanned region, with only parameters being modified. The results are shown in Fig. 18 and are in good agreement with the Standard Model values \( \kappa_F \), \( \kappa_V = 1 \). The possibility that \( \kappa_F < 0 \) is excluded at 95% C.L. in this parameterization.

**X. CONCLUSIONS**

A search for \( t\bar{t}H \) production in multilepton final states using a data set corresponding to an integrated luminosity of 36.1 fb\(^{-1}\) of proton-proton collision at \( \sqrt{s} = 13 \) TeV recorded by the ATLAS experiment at the LHC is presented. Seven final states, targeting Higgs boson decays to \( WW^*, \tau\tau \), and \( ZZ^* \), categorized by the number and flavor of charged-lepton candidates, are analyzed. An excess of events over the expected background from SM processes is found, which is interpreted as an observed significance of 4.1 standard deviations for a SM Higgs boson of mass 125 GeV. The expected significance for a SM Higgs boson is 2.8 standard deviations. The best-fit result of the observed production cross section is \( \sigma(t\bar{t}H) = 590^{+180}_{-150} \) fb, in agreement with the SM prediction. This provides evidence for the \( t\bar{t}H \) production mode.

**ACKNOWLEDGMENTS**

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; STHC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DLR and DMSN, Denmark; IN2P3-CNRS, CEA-DRF/IFRF, France; SRNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNI\( \kappa S \) and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Sklodowska-Curie Actions, European Union; Investissements d'Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partagé le Savoir, France; DFG and AvH Foundation, Germany; Herakleitus, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSI, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (United Kingdom) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [109].
EVIDENCE FOR THE ASSOCIATED PRODUCTION OF THE …

PHYS. REV. D 97, 072003 (2018)

[38] 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, arXiv:1712.02304.


EVIDENCE FOR THE ASSOCIATED PRODUCTION OF THE …

PHYS. REV. D 97, 072003 (2018)
EVIDENCE FOR THE ASSOCIATED PRODUCTION OF THE...


A. Ventura, M. Venturi, N. Venturi, V. Vercesi, M. Verducci, W. Verkerke, A. T. Vermeulen,
J. C. Vermeulen, M. C. Vetterli, N. Viaux Maia, O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
V. B. Vinogradov, A. Vishwakarma, C. Vittori, I. Vivarelli, S. M. W. Wang, R. Wang, S. M. Weber,
J. P. Wells, T. Wenaus, J. Wallinger, A. J. Wall, H. Wallberg, S. Wahrmund, K. Wakamiya, J. Walder,
R. Walker, J. H. Vossebeld, N. Vranjes, M. Vranjes Milosavljevic, V. Vrba, M. C. Vetterli, N. Viaux Maira,
O. Viazlo, I. Vichou, T. Vickey, O. E. Vickey Boeri,
G. H. A. Viehhauser, S. Viel, L. Viganò, M. Villa, M. Villaplana Perez, E. Vilucchi, M. G. Vincter,
16 Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, California, USA
17 Department of Physics, Humboldt University, Berlin, Germany
18 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
19 School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
20 a) Department of Physics, Bogazici University, Istanbul, Turkey
20 b) Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey
20 c) Istanbul Bilgi University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey
20 d) Bahcesehir University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey
21 Centro de Investigaciones, Universidad Antonio Narino, Bogota, Colombia
22 a) INFN Sezione di Bologna, Italy
22 b) Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy
23 Physikalisches Institut, University of Bonn, Bonn, Germany
24 a) Department of Physics, Boston University, Boston, Massachusetts, USA
24 b) Department of Physics, Brandeis University, Waltham, Massachusetts, USA
25 Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil
26 a) Electrical Circuits Department, Federal University of Juiz de Fora (UFJF), Juiz de Fora, Brazil
26 b) Federal University of Sao Joao del Rei (UFSJ), Sao Joao del Rei, Brazil
26 c) Instituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil
27 Physics Department, Brookhaven National Laboratory, Upton, New York, USA
28 a) Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania
28 b) Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi, Romania
28 c) National Institute for Research and Development of Isotopic and Molecular Technologies, Physics Department, Cluj Napoca, Romania
28 d) West University in Timisoara, Timisoara, Romania
29 Departamento de Fisica, Universidad de Buenos Aires, Buenos Aires, Argentina
30 Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
31 Department of Physics, Carleton University, Ottawa, Ontario, Canada
32 CERN, Geneva, Switzerland
33 Enrico Fermi Institute, University of Chicago, Chicago, Illinois, USA
34 a) Departamento de Fisica, Pontificia Universidad Católica de Chile, Santiago, Chile
34 b) Departamento de Fisica, Universidad Técnica Federico Santa María, Valparaíso, Chile
35 a) Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China
35 b) Department of Physics, Nanjing University, Jiangsu, China
35 c) Physics Department, Tsinghua University, Beijing 100084, China
36 a) University of Chinese Academy of Science (UCAS), Beijing, China
36 b) School of Physics, Shandong University, Shandong, China
36 c) School of Physics and Astronomy, Key Laboratory for Particle Physics, Astrophysics and Cosmology, Ministry of Education; Shanghai Key Laboratory for Particle Physics and Cosmology, Tsung-Dao Lee Institute, Shanghai Jiao Tong University, China
36 d) Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Anhui, China
37 Université Clermont Auvergne, CNRS/IN2P3, LPC, Clermont-Ferrand, France
38 Nevis Laboratory, Columbia University, Irvington, New York, USA
39 Niels Bohr Institute, University of Copenhagen, Kobenhavn, Denmark
40 a) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Italy
40 b) Dipartimento di Fisica, Università della Calabria, Rende, Italy
41 a) AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Krakow, Poland
41 b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow, Poland
42 Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland
43 Physics Department, Southern Methodist University, Dallas, Texas, USA
44 Physics Department, University of Texas at Dallas, Richardson, Texas, USA
45 DESY, Hamburg and Zeuthen, Germany
46 Lehrstuhl für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany
47 Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany
48 Department of Physics, Duke University, Durham, North Carolina, USA
96 Research Institute for Nuclear Problems of Byelorussian State University, Minsk, Republic of Belarus
97 Group of Particle Physics, University of Montreal, Montreal, Quebec, Canada
98 P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, Russia
99 Institute for Theoretical and Experimental Physics (ITEP), Moscow, Russia
100 National Research Nuclear University MEPhI, Moscow, Russia
101 D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
102 Fakultät für Physik, Ludwig-Maximilians-Universität München, München, Germany
103 Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany
104 Nagasaki Institute of Applied Science, Nagasaki, Japan
105 Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan
106 INFN Sezione di Napoli, Italy
107 Dipartimento di Fisica, Università di Napoli, Napoli, Italy
108 Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico, USA
109 Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, Netherlands
110 Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, Netherlands
111 Department of Physics, Northern Illinois University, DeKalb, Illinois, USA
112 Budker Institute of Nuclear Physics, SB RAS, Novosibirsk, Russia
113 Department of Physics, New York University, New York, New York, USA
114 Ohio State University, Columbus, Ohio, USA
115 Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman, Oklahoma, USA
116 Department of Physics, Oklahoma State University, Stillwater, Oklahoma, USA
117 Palacký University, RCPTM, Olomouc, Czech Republic
118 Center for High Energy Physics, University of Oregon, Eugene, Oregon, USA
119 LAL, Université Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France
120 Graduate School of Science, Osaka University, Osaka, Japan
121 Department of Physics, University of Oslo, Oslo, Norway
122 Department of Physics, Oxford University, Oxford, United Kingdom
123 INFN Sezione di Pavia, Italy
124 Dipartimento di Fisica, Università di Pavia, Pavia, Italy
125 Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania, USA
126 INFN Sezione di Pisa, Italy
127 Departamento de Física Teórica y del Cosmos, Universidad de Granada, Granada, Portugal
128 Dep Física and CEFITEC of Faculdade de Ciencias e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
129 Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
130 Czech Technical University in Prague, Praha, Czech Republic
131 Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic
132 State Research Center Institute for High Energy Physics (Protvino), NRC KI, Russia
133 Particle Physics Department, Rutherford Appleton Laboratory, Didcot, United Kingdom
134 INFN Sezione di Roma, Italy
135 INFN Sezione di Roma Tor Vergata, Italy
136 INFN Sezione di Roma Tre, Italy
137 Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies - Université Hassan II, Casablanca, Morocco
EVIDENCE FOR THE ASSOCIATED PRODUCTION OF THE …

PHYS. REV. D 97, 072003 (2018)

072003-43