Search for top squarks in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV $pp$ collisions with the ATLAS detector

M. Aaboud et al.*
(ATLAS Collaboration)
(Received 14 June 2016; published 19 September 2016)

The results of a search for the top squark, the supersymmetric partner of the top quark, in final states with one isolated electron or muon, jets, and missing transverse momentum are reported. The search uses the 2015 LHC $pp$ collision data at a center-of-mass energy of $\sqrt{s} = 13$ TeV recorded by the ATLAS detector and corresponding to an integrated luminosity of 3.2 fb$^{-1}$. The analysis targets two types of signal models: gluino-mediated pair production of top squarks with a nearly mass-degenerate top squark and neutralino and direct pair production of top squarks, decaying to the top quark and the lightest neutralino. The experimental signature in both signal scenarios is similar to that of a top quark pair produced in association with large missing transverse momentum. No significant excess over the Standard Model background prediction is observed, and level exclusion limits on gluino and top squark masses are set at 95% confidence level. The results extend the LHC run-1 exclusion limit on the gluino mass up to 1460 GeV in the gluino-mediated scenario in the high gluino and low top squark mass region and add an excluded top squark mass region from 745 to 780 GeV for the direct top squark model with a massless lightest neutralino. The results are also reinterpreted to set exclusion limits in a model of vectorlike top quarks.

DOI: 10.1103/PhysRevD.94.052009

I. INTRODUCTION

Supersymmetry (SUSY) [1–6] is a natural solution [7,8] to the hierarchy problem [9–12]. The top squark ($\tilde{t}$), which is the superpartner of the top quark, is expected to be relatively light due to its large contribution to the Higgs boson mass radiative corrections [13,14]. For reasons such as gauge unification [15] and the two-loop radiative corrections to the Higgs boson mass [16,17], one may also expect a TeV mass scale for the gluino ($\tilde{g}$), the superpartner of the gluon. A common theoretical strategy for avoiding strong constraints from the nonobservation of proton decay [18] is to introduce a multiplicative quantum number called $R$ parity. If $R$ parity is conserved [19], SUSY particles are produced in pairs and the lightest supersymmetric particle (LSP) is stable. This analysis follows the typical assumption that the lightest neutralino ($\tilde{\chi}_1^0$) is the LSP. Since the $\tilde{\chi}_1^0$ interacts only weakly, it can serve as a candidate for dark matter [20,21].

This paper presents a search targeting the lighter top squark$^2$ ($\tilde{t}_1$) in two scenarios: gluino-mediated pair production of the $\tilde{t}_1$ with a small $\tilde{t}_1$-LSP mass splitting and direct pair production of the $\tilde{t}_1$, both illustrated by the diagrams in Fig. 1. The former scenario refers to pair production of gluinos, each decaying to the top quark and the $\tilde{t}_1$. In this scenario, the mass difference between the gluino and the $\tilde{t}_1$ is assumed to be well above the top quark mass, while the mass difference between the $\tilde{t}_1$ and the LSP is assumed to be significantly smaller than the W boson mass. As a result, the visible $\tilde{t}_1$ decay products have low momentum, typically below the reconstruction and identification thresholds. This scenario is motivated by the dark matter relic density, which is generally too large in the Minimal Supersymmetric Standard Model [22,23] but can be regulated by coannihilation of the top squark and the neutralino [24]. In the second scenario, the two directly produced $\tilde{t}_1$ are each assumed to decay to the top quark and the LSP. This model is interesting as it is independent of the gluino mass, which is more weakly constrained by naturalness arguments than the top squark mass.

Experimentally, the final states of the two scenarios are similar [25], and the detector signature consists of the decay

$^2$The superpartners of the left- and right-handed top quarks, $\tilde{t}_L$ and $\tilde{t}_R$, mix to form the two mass eigenstates $\tilde{t}_1$ and $\tilde{t}_2$, where $\tilde{t}_1$ is the lighter one.
products of a pair of top quarks\(^3\) and large missing transverse momentum \(\vec{p}_{\text{miss}}\), where the magnitude is referred to as \(E_{\text{miss}}\). The main difference between the two scenarios is that the production cross section for gluino pairs is about a factor 50 higher than for \(\tilde{t}_1\) pairs of the same mass due to the additional spin and color states. The results are also reinterpreted in a model of strong-interaction direct pair production of vectorlike top quarks \(T\) (referred to as VLQ) [26–28], for which the decay mode \(T \rightarrow iZ\) with \(Z \rightarrow \nu\bar{\nu}\) has a signature similar to that of direct top squark pair production with \(\tilde{t}_1 \rightarrow iZ\).

The analysis presented here—which is based on previous ATLAS searches for the same signature [29,30]—targets the one-lepton final state where the W boson from one of the top quarks decays to an electron or muon (either directly or via a \(Z\) or \(\nu\bar{\nu}\)) and the W boson from the other top quark decays hadronically. The dominant Standard Model (SM) background processes are the production of \(t\bar{t}\); the associated production of a top quark and a W boson (single top \(Wt\)); \(t\bar{t} + Z\) (\(\nu\bar{\nu}\)); and the associated production of W bosons and jets (\(W + \text{jets}\)). The search uses the ATLAS data collected in proton-proton (\(pp\)) collisions in 2015 corresponding to an integrated luminosity of 3.2 \(fb^{-1}\) at a center-of-mass energy of \(\sqrt{s} = 13\) TeV. The ATLAS run-1 searches for gluino-mediated top squark production and direct top squark pair production are summarized in Refs. [31,32], respectively. Run-1 searches for VLQ production can be found in Refs. [33–35]. The CMS Collaboration has performed similar searches for gluino-mediated top squark production [36], direct top squark pair production [37–42], and VLQ production [43].

This document is organized as follows. The ATLAS detector, data set, and trigger are described in Sec. II, and the corresponding set of simulations are detailed in Sec. III. Section IV presents the reconstruction and selection of physics objects and the construction of discriminating variables. These variables are used in Sec. V to construct the signal event selections. The background estimation procedure (Sec. VI) and systematic uncertainties (Sec. VII) are described before the results are presented in Sec. VIII. Section IX contains concluding remarks.

### II. ATLAS DETECTOR AND DATA SET

The ATLAS detector [44] is a multipurpose particle physics detector with nearly \(4\pi\) coverage in solid angle around the collision point.\(^4\) It consists of an inner tracking detector (ID), surrounded by a superconducting solenoid providing a 2 T axial magnetic field, a system of calorimeters, and a muon spectrometer (MS) incorporating three large superconducting toroid magnets. The ID provides charged-particle tracking in the range \(|\eta| < 2.5\) using three technologies: silicon pixel and silicon microstrip tracking detectors and a transition radiation tracker. During the LHC shutdown between run 1 and run 2, a new innermost layer of silicon pixels was added, which improves the track impact parameter resolution and vertex position resolution [45]. High-granularity electromagnetic and hadronic calorimeters cover the region \(|\eta| < 4.9\). The central hadronic calorimeter is a sampling calorimeter with scintillator tiles as the active medium and steel absorbers. All the electromagnetic calorimeters, as well as the end cap and forward hadronic calorimeters, are sampling calorimeters with liquid argon as the active medium and lead, copper, or tungsten absorber. The MS consists of three layers of high-precision tracking chambers with coverage up to \(|\eta| = 2.7\) and dedicated chambers for triggering in the region \(|\eta| < 2.4\). Events are selected by a two-level trigger system: the first level is a hardware-based system and the second is a software-based system.

The 2015 LHC collision data used in this analysis have a mean number of additional \(pp\) interactions per bunch crossing (pileup) of approximately 14 and a bunch spacing of 25 ns. Following requirements based on beam and detector conditions and data quality, the data set corresponds to an integrated luminosity of 3.2 \(fb^{-1}\) with an associated uncertainty of 5\%. The uncertainty is derived following the same methodology as that detailed in Ref. [46]. Events used for this search were recorded using a trigger logic that accepts events with \(E_{\text{miss}}\), calibrated to

\(^3\)Due to the Majorana nature of the gluino, in the gluino-mediated model, each of the two “visible” top quarks can independently be a top or an antitop quark. Hereafter, the term \(t\bar{t}\) can be taken to refer to any combination of \(t\) and \(\bar{t}\).

\(^4\)ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z axis along the beam pipe. The x axis points from the IP to the center of the LHC ring, and the y axis points upwards. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \(\phi\) being the azimuthal angle around the z axis. The pseudorapidity is defined in terms of the polar angle \(\theta\) as \(\eta = -\ln \tan(\theta/2)\). Angular distance is measured in units of \(\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}\).
the electromagnetic scale, above 70 GeV. The trigger is more than 95% efficient for events passing an offline-computed \(E_{T}^{\text{miss}} > 200\) GeV requirement and is > 99% efficient for events passing all signal selections. An additional data sample used to estimate one of the background processes was recorded with a trigger requiring a photon with transverse momentum \(p_T > 120\) GeV, which is > 99% efficient for the offline photon selection described in Sec. IV.

III. MONTE CARLO SIMULATIONS

Samples of Monte Carlo (MC) simulated events are used for the description of the background and to model the SUSY signals. Several matrix element (ME) generators are combined with parton shower (PS) and hadronization generators. Signal SUSY samples are generated at leading order (LO) with MG5\_AMC v2.2 [47] while VLQ signal samples are generated at LO with PROTOS v2.2 [48,49]. All signal samples are interfaced with PYTHIA 8.186 [50]. Background samples use one of three setups:

(i) MG5\_AMC v2.2 interfaced with PYTHIA 8 or HERWIG++ using the CKKW-L [51] or the MC@NLO method for matching a LO or next-to-leading-order (NLO) ME to the PS, respectively.

(ii) POWHEG-BOX [52–56] interfaced to PYTHIA 6 [57] or HERWIG++ using the POWHEG method [58,59] for matching the NLO ME to the PS.

(iii) SHERPA 2.1.1 [60] using Comix [61] and OpenLoops [62] ME generators interfaced with the SHERPA parton shower [63].

The CT10 [64] NLO parton distribution function (PDF) set is used for ME calculations with SHERPA and POWHEG-Box and the NNPDF2.3 [65] PDF set is used for samples generated with MG5\_AMC, except for the NLO samples, which use either CT10 or NNPDF3.0 [66]. The CTEQ6L1 [67] LO PDF set along with the P2012 [68] set of underlying-event tuned parameters (UE tune) is used for PYTHIA 6; the NNPDF2.3 LO PDF set and the A14 UE tune [69] is used for PYTHIA 8; and the CT10 PDF set with the default UE tune provided by the authors of SHERPA is used for the SHERPA samples. The samples produced with MG5\_AMC, POWHEG-Box, and PROTOS all use EvtGen v1.2.0 [70] for the modeling of \(b\)-hadron decays. The simulation setup is summarized in Table I and more details can be found in Refs. [71–74] for \(t\bar{t}\) and single top, W/Z + jets, dibosons, and \(t\bar{t} + W/Z\), respectively. Additional samples aside from those shown in Table I are used to assess theoretical modeling uncertainties and are discussed in Sec. VII.

In the gluino-mediated production the top squark is assumed to decay via \(t_1 \to c + \chi_{0}^{\pm}\) with a 100% branching ratio and with a default mass splitting \(m_{t_1} - m_{\chi_{0}^{\pm}} = 5\) GeV. Alternative samples with larger mass splitting and/or replacing the two-body top squark decay by a four-body top squark decay \(t_1 \to b f f' \chi_{1}^{0}\) where \(ff'\) is a fermion-antifermion pair, are produced for additional studies. The gluinos and top squarks are assumed to decay promptly. In the direct top squark pair production samples, the \(t_1\) is chosen to be mostly the partner of the right-handed top quark\(^5\) and the \(\chi_{1}^{0}\) to be a pure bino. This choice is consistent with a large branching ratio for the given \(t_1\) decay. Different hypotheses for the left-right mixing in the top squark sector and the nature of the neutralino lead to different acceptance values. The acceptance is affected because the polarization of the top quark changes as a function of the field content of the supersymmetric particles, which impacts the boost of the lepton in the top quark decay. Signal grids are generated for both the gluino and direct top squark pair production models. The spacing between grid points in the gluino-top squark and top squark-neutralino mass planes vary between 25 and 100 GeV.

All the MC samples are normalized to the highest-order (in \(\alpha_S\)) cross section available, as indicated in the last column of Table I. The cross sections for the pair and single

### Table I. Overview of the nominal simulated samples

<table>
<thead>
<tr>
<th>Process</th>
<th>ME generator</th>
<th>ME PDF</th>
<th>PS and hadronization</th>
<th>UE tune</th>
<th>Cross-section order</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t\bar{t})</td>
<td>POWHEG-BOX v2</td>
<td>CT10</td>
<td>PYTHIA 6</td>
<td>P2012</td>
<td>NNLO + NNLL [75–80]</td>
</tr>
<tr>
<td>Single top</td>
<td>POWHEG-BOX</td>
<td>CT10</td>
<td>PYTHIA 6</td>
<td>P2012</td>
<td>NNLO + NNLL [81–83]</td>
</tr>
<tr>
<td>W/Z + jets</td>
<td>SHERPA 2.1.1</td>
<td>CT10</td>
<td>SHERPA</td>
<td>Default</td>
<td>NNLO [84]</td>
</tr>
<tr>
<td>Diboson</td>
<td>SHERPA 2.1.1</td>
<td>CT10</td>
<td>SHERPA</td>
<td>Default</td>
<td>NLO</td>
</tr>
<tr>
<td>(t\bar{t} + W/Z)</td>
<td>MG5_AMC 2.2.2</td>
<td>NNPDF2.3</td>
<td>PYTHIA 8</td>
<td>A14</td>
<td>NLO [47]</td>
</tr>
<tr>
<td>(t\bar{t} + \gamma)</td>
<td>MG5_AMC 2.2.3</td>
<td>CTEQ6L1</td>
<td>PYTHIA 8</td>
<td>A14</td>
<td>NLO [47]</td>
</tr>
<tr>
<td>SUSY signal</td>
<td>MG5_AMC 2.2.2</td>
<td>NNPDF2.3</td>
<td>PYTHIA 8</td>
<td>A14</td>
<td>NLO + NLL [85]</td>
</tr>
<tr>
<td>VLQ signal</td>
<td>PROTOS v2.2</td>
<td>NNPDF2.3</td>
<td>PYTHIA 8</td>
<td>A14</td>
<td>NNLO + NNLL [75–80]</td>
</tr>
</tbody>
</table>

---

5The \(t_1\) component is given by the off-diagonal entry of the top squark mixing matrix. The \(t_1\) decays in the direct top squark pair production samples are performed by PYTHIA and produce unpolarized top quarks. The events are reweighted to obtain a top squark mixing equivalent to a matrix with on-diagonal entries of approximately 0.55 and off-diagonal entries of approximately ±0.83. The event weights depend on the angular distributions of the top decay products [86].

052009-3
production of top quarks as well as for the signal processes also include resummation of soft gluon emission to next-to-next-to-leading-logarithmic (NLLNLL) and next-to-leading-logarithmic (NLL) accuracy, respectively. As is described in Sec. VI A 3, it is important that the simulated $t\bar{t} + \gamma$ and $t\bar{t} + Z$ events are as similar as possible. Therefore, a small 4% correction is applied to the $t\bar{t} + \gamma$ cross section to account for a different PDF set, factorization and renormalization scale, and number of partons from the matrix element. The same NLO QCD $K$ factor is then applied to the $t\bar{t} + \gamma$ process as is used for the $t\bar{t} + Z(\rightarrow \nu\bar{\nu})$ process [47]. This choice is motivated by the similarity of QCD calculations for the two processes as well as empirical studies of the ratio of $K$ factors computed as a function of the boson $p_T$. Further information about the $K$ factor and its uncertainty is given in Sec. VII. The cross sections for the $t\bar{t}$, $W +$ jets, and $Wt$ processes are used for cross-checks and optimization studies, while for the final results these processes are normalized to data in control regions.

All background samples, except for the $t\bar{t} + \gamma$ sample, are processed with the full simulation of the ATLAS detector [88] based on GEANT 4 [89]. The signal samples and the $t\bar{t} + \gamma$ sample are processed with a fast simulation [90] of the ATLAS detector with parameterized showers in the calorimeters. All samples are produced with varying numbers of simulated minimum-bias interactions generated with PYTHIA 8 overlaid on the hard-scattering event to account for pileup from multiple $pp$ interactions in the same or nearby bunch crossings. The average number of interactions per bunch crossing is reweighted to match the distribution in data. Furthermore, the simulated samples are reweighted to account for small differences in the efficiencies of physics-object reconstruction and identification with respect to those measured in data.

IV. EVENT RECONSTRUCTION AND SELECTION

All events must satisfy a series of quality criteria before being considered for further use. The reconstructed primary vertex with the highest $\sum_{\text{tracks}} p_T^2$ must have at least two associated tracks. In this analysis, physics objects are labeled as either baseline or signal depending on various quality and kinematic requirements, where the latter label describes a tighter selection of the former. Baseline objects are used to distinguish between the physics objects in the event and to compute the missing transverse momentum. Baseline leptons (electrons and muons) are also used to apply a second-lepton veto to suppress dilepton $t\bar{t}$ and $Wt$ events.

Electron candidates are reconstructed from electromagnetic calorimeter cell clusters that are matched to ID tracks. Baseline electrons are required to have $p_T > 7$ GeV, $|\eta| < 2.47$, and satisfy “VeryLoose” likelihood identification criteria that are defined following the methodology described in Ref. [91]. Signal electrons must pass all baseline requirements and in addition have $p_T > 25$ GeV, satisfy the “Loose” likelihood identification criteria in Ref. [91], and have impact parameters with respect to the reconstructed primary vertex along the beam direction ($z_0$) and in the transverse plane ($d_0$) such that $|z_0 \sin \theta| < 0.5$ mm and $|d_0|/\sigma_{d_0} < 5$, where $\sigma_{d_0}$ is the uncertainty of $d_0$. Furthermore, signal electrons must be isolated, where the criteria use track-based information to obtain a 99% efficiency that is independent of $p_T$, as derived from $Z \rightarrow \ell\ell$ MC samples and confirmed in data.

Muons are reconstructed from combined tracks that are formed from ID and MS tracks, ID tracks matched to MS track segments, standalone MS tracks, or ID tracks matched to an energy deposit in the calorimeter compatible with a minimum-ionizing particle (referred to as calo-tagged muon) [92]. Baseline muons are required to have $p_T > 6$ GeV, $|\eta| < 2.7$, and satisfy the “Loose” identification criteria described in Ref. [92]. Signal muons must pass all baseline requirements and in addition have $p_T > 25$ GeV, and have impact parameters $|z_0| < 0.5$ mm and $|d_0|/\sigma_{d_0} < 3$. Furthermore, signal muons must be isolated according to isolation criteria similar to those used for signal electrons, yielding the same efficiency.

Photon identification is not used in the main event selection, and photons give rise to extra jet or electron candidates. Photons must be identified, however, for the $t\bar{t} + \gamma$ sample that is used in the data-driven estimation of the $t\bar{t} + Z$ background. In this case, photon candidates are reconstructed from calorimeter cell clusters and are required to satisfy the “Tight” identification criteria described in Ref. [93]. Furthermore, photons are required to have $p_T > 125$ GeV and $|\eta| < 2.37$, excluding the barrel-end cap calorimeter transition in the range $1.37 < |\eta| < 1.52$, so that the photon trigger is fully efficient. Photons must further satisfy isolation criteria based on both track and calorimeter information.

Jet candidates are built from topological clusters [94,95] in the calorimeters using the anti-$k_T$ algorithm with a jet radius parameter $R = 0.4$ [96]. Jets are corrected for contamination from pileup using the jet area method [97–99] and then calibrated to account for the detector response [100,101]. Jets in data are further calibrated based on in situ measurements of the jet energy scale. Baseline jets are required to have $p_T > 20$ GeV. Signal jets must have $p_T > 25$ GeV and $|\eta| < 2.5$. Furthermore, signal jets with $p_T < 50$ GeV are required to satisfy criteria, implemented in the jet vertex tagger algorithm [99], designed to reject jets originating from pileup. Events containing a jet that does not pass specific jet quality requirements are vetoed from the analysis in order to
SEARCH FOR TOP SQUARKS IN FINAL STATES WITH...

TABLE II. Overlap removal procedure. The first two rows list the types of overlapping objects: electrons (e), muons (μ), electron or muon (e/μ), jets (j), photons (γ), and hadronically decaying τ lepton (τ). All objects refer to the baseline definitions, except for γτ and τ where no distinction between baseline and signal definition is made. The third row specifies when an object pair is considered as overlapping, the fourth row describes an optional condition, and the last row lists which label is given to the ambiguous object. More information is given in the text.

<table>
<thead>
<tr>
<th>Object 1</th>
<th>e</th>
<th>e</th>
<th>μ</th>
<th>e/μ</th>
<th>e</th>
<th>μ</th>
<th>e</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object 2</td>
<td>μ</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>ΔR &lt;</td>
<td>0.01</td>
<td>0.2</td>
<td>0.2</td>
<td>min (0.4, 0.04 + 10/πT/GeV)</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Condition</td>
<td>Calo-tagged μ</td>
<td>j not b-tagged</td>
<td>j not b-tagged and</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Precedence</td>
<td>e</td>
<td>e</td>
<td>μ</td>
<td>j</td>
<td>e</td>
<td>γ</td>
<td>τ</td>
</tr>
</tbody>
</table>

If the “ΔR <” requirement in Table II is not met, then both objects under consideration are kept. The order of steps in the procedure is given by the columns in Table II, which are executed from left to right. The second (e/μ) and the third (μ/τ) steps of the procedure ensure that leptons and jets have a minimum ΔR separation of 0.2. Therefore, the fourth step (e/μ) only has an effect for ΔR > 0.2. The steps involving a photon are not applied in the main event selection but only for the event selection where photons are identified. For the remainder of the paper, all baseline and signal objects are those that have survived the overlap removal procedure.

Large-radius jets are clustered from all signal (small-radius R = 0.4) jets using the anti-kt algorithm with R = 1.0 or 1.2. To reduce the impact of soft radiation and pileup, the large-radius jets are groomed using reclustered jet trimming, where constituents with pT less than 5% of the unmerged jet pT are removed [111–114]. Electrons and muons are not included in the reclustering, since it was found that including them increases the background acceptance more than the signal efficiency. Large-radius jets are not used in the overlap removal procedure; however, the signal jets that enter the reclustering have passed the overlap removal procedure described above. The analysis uses a large-radius jet mass, where the squared mass is defined as the square of the four-vector sum of the constituent (small-radius) jets’ momenta.

All events are required to have EmissT > 200 GeV, exactly one signal lepton, and no additional baseline leptons, as well as at least four signal jets. In addition, the events must have a transverse mass7 of the signal lepton and the missing transverse momentum satisfying mT > 30 GeV and have an azimuthal angle between each of the two leading jets and the missing transverse momentum of |Δϕ(jet, EmissT)| > 0.4 with i ∈ {1, 2}. The events must further pass an HTsig > 5 requirement, where HTsig = (HmissT − 100GeV)/σHmissT. The variable HmissT is the magnitude of the negative

7The transverse mass mT is defined as mT2 = 2lepT EmissT [1 − cos(Δϕ)], where Δϕ is the azimuthal angle between the lepton and the missing transverse momentum direction. The quantity pTlep is the transverse momentum of the charged lepton.
vector sum of the transverse momenta of signal jets and the signal jet; the resolution $\sigma_{\Delta p_T}$ is computed using the per-event jet energy resolution uncertainties (more details are given in Refs. [29,115]). The latter three event selection criteria for $m_{W}$, $[\Delta \phi(\text{jet}, \widetilde{p}_T^{\text{miss}})]$, and $H_{T,sig}^{\text{miss}}$ suppress multi-jet processes with misidentified or nonprompt leptons and mismeasured $E_T^{\text{miss}}$ to a negligible level. With the above event selection, the dominant backgrounds are $\bar{t}t$ events with at least one leptonically decaying $W$ boson and $W +$ jets production. A powerful technique for suppressing these background processes is to require $m_{T}$ to be greater than the $W$ boson mass. For example, an $m_{T} > 120$ GeV requirement removes more than 90% of $\bar{t}t$ and $W +$ jets events that pass the above event selection.

One of the dominant contributions to the residual background is from $\bar{t}t$ production where both $W$ bosons decay leptonically or one $W$ boson decays leptonically and the other via a hadronic $\tau$ decay. A series of additional variables, described in detail in Ref. [29], are used to discriminate between these backgrounds and the signal processes. The asymmetric $m_{T2}$ ($am_{T2}$) [116–119] and $m_{T2}^{\tau}$ are both variants of the variable $m_{T2}$ [120], a generalization of the transverse mass applied to signatures where two particles are not directly detected. Like the transverse mass, $m_{T2}$ is the minimum mass consistent with the observed transverse momenta and is bounded by the parent particle mass for particular topologies. The $am_{T2}$ variable targets dileptonic $\bar{t}t$ events where one lepton is not reconstructed. For these events, the $am_{T2}$ distribution has a kinematic endpoint near the top quark mass. The $m_{T2}^{\tau}$ variable targets $\bar{t}t$ events where one of the two $W$ bosons decays via a hadronically decaying $\tau$ lepton. In $\bar{t}t$ events where the hadronically decaying $\tau$ lepton is correctly identified, $m_{T2}^{\tau} < m_{W}$. Due to the extra missing momentum from neutralinos, signal events are characterized by large values of $am_{T2}$ and $m_{T2}^{\tau}$.

To further suppress dileptonic $\bar{t}t$ and other backgrounds, variables are constructed to tag the hadronic decays of top quarks. The $m_{top}^{\tau}$ variable is the invariant mass of the three jets in the event most compatible with the hadronic decay products of a top quark, where the three jets are selected by a $\chi^2$ fit to the lepton-jets $\bar{t}t$ hypothesis, with one term for the mass of the hadronic $W$ boson candidate and one term for the mass of the hadronic top quark candidate. The uncertainty on the multijet invariant masses is estimated from the jet-energy resolution. When the target top squark mass is large enough so that the resulting top quarks are significantly boosted in the lab frame, the mass of high $p_T$ large-radius jets is a powerful tool that outperforms $m_{top}^{\tau}$.

Another kinematic variable that targets the dileptonic $\bar{t}t$ background is topness [121]. As with $m_{top}^{\tau}$, topness is constructed by minimizing a $\chi^2$-type function. However, in contrast to $m_{top}^{\tau}$, topness quantifies the kinematic compatibility with the dileptonic $\bar{t}t$ event topology where one lepton is not reconstructed. Low values of topness are background like while high values of topness are kinematically less consistent with dilepton $\bar{t}t$ events.

An important change from the run-1 suite of tools is the treatment of hadronically decaying $\tau$ candidates in the $m_{T2}^{\tau}$ variable. To increase the efficiency and purity of selecting the $\tau$ lepton, a reconstructed hadronic $\tau$ candidate is used as one of the two visible objects in the $m_{T2}$ calculation. Events are removed if one of the selected jets is additionally identified as a hadronic $\tau$ candidate, with a corresponding $m_{T2}^{\tau} < 80$ GeV. For an event selection with a $E_T^{\text{miss}} > 200$ GeV requirement, this hadronic $\tau$ veto removes approximately 40% of simulated $\bar{t}t$ events where one $W$ boson decays leptonically and the other decays via a hadronically decaying $\tau$ lepton. For the considered signal models, the veto removes 1% of the events. The $\tau$ veto is applied in all following event selections except those defining the $\bar{t} + Z$ control region (since the veto would remove only about 1% of the events in this region).

V. SIGNAL REGIONS

Three signal event selections (called signal regions, or SR1–3) are constructed using the set of discriminating variables described in Sec. IV. The three signal regions are optimized, before looking at the data, to maximize the discovery sensitivity using three benchmark signal models from the gluino-mediated top squark models, each representing a distinct phenomenology. The benchmark models are defined by $(\tilde{g}, \tilde{\chi}_1^0)$ masses of (1100, 800), (1250, 750), and (1400, 400) GeV for SR1, SR2, and SR3, respectively. The benchmark model for SR1 has a production cross section and kinematic properties similar to those of a direct top squark model with $(\tilde{t}_1, \tilde{\chi}_1^0)$ masses of about (600, 260) GeV, while the benchmark models for SR2 and SR3 cannot be directly mapped to have both the same cross sections and similar kinematic properties. As a consequence, SR2 and SR3 have reduced sensitivity to direct top squark models.

The three signal regions are characterized by increasing $E_T^{\text{miss}}$ requirements. The SR1 benchmark has the softest $E_T^{\text{miss}}$ spectrum and the momentum of the hadronically decaying top quark is typically not sufficient to capture all of the decay products inside a single large-radius jet. As a result, the top quark mass computed using the $m_{top}^{\tau}$ variable which is based on small-radius jets is useful for rejecting dileptonic $\bar{t}t$ and other background events without a top quark that has hadronic decay products. In contrast, the boost of the hadronically decaying top quarks in the SR2 and SR3 benchmarks is often sufficient to capture all decay products inside a single large-radius jet. The angular separation between the decay products scales with the inverse of the momentum. Therefore, the optimal large-radius jet cone size is found to be larger for SR2 ($R = 1.2$) than for SR3 ($R = 1.0$). Additional requirements on topness and $am_{T2}$ further reduce the dileptonic $\bar{t}t$ background. Background events without a high-$p_T$ top quark that decays leptonically are suppressed by using a requirement on the
Table III. Overview of the event selections for all SRs and the associated $\tilde{t}$ (TCR), $W +$ jets (WCR), and $Wt$ (STCR) control regions. Round brackets are used to describe lists of values and square brackets denote intervals.

<table>
<thead>
<tr>
<th>Variable</th>
<th>SR1</th>
<th>TCR1/WCR1</th>
<th>STCR1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_\text{miss}^T$ [GeV]</td>
<td>$&gt;260$</td>
<td>$&gt;200$</td>
<td>$&gt;200$</td>
</tr>
<tr>
<td>$E_\text{miss}^{T_{\text{tag}}}$</td>
<td>$&gt;14$</td>
<td>$&gt;5$</td>
<td>$&gt;5$</td>
</tr>
<tr>
<td>$m_T$ [GeV]</td>
<td>$&gt;170$</td>
<td>[30, 90]</td>
<td>[30, 120]</td>
</tr>
<tr>
<td>$</td>
<td>\Delta p_T^{\text{miss}}(\ell, p_T^\text{miss})</td>
<td>&gt; 0.4$ for $i \in {1, 2}$</td>
<td></td>
</tr>
<tr>
<td>Topness</td>
<td>$&gt;6.5$</td>
<td>$&gt;6.5$</td>
<td>$&gt;6.5$</td>
</tr>
<tr>
<td>$H^\text{top}$ [GeV]</td>
<td>$&lt;270$</td>
<td>$&lt;270$</td>
<td>$&lt;270$</td>
</tr>
<tr>
<td>$\Delta R(b, \ell)$</td>
<td>$&lt;3.0$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$\Delta R(b_1, b_2)$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$&gt;1.2$</td>
</tr>
<tr>
<td>Number of $b$ tags</td>
<td>$\geq 1$</td>
<td>$\geq 1/0$</td>
<td>$\geq 2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>SR2</th>
<th>TCR2/WCR2</th>
<th>STCR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_\text{miss}^T$ [GeV]</td>
<td>$&gt;350$</td>
<td>$&gt;250$</td>
<td>$&gt;200$</td>
</tr>
<tr>
<td>$E_\text{miss}^{T_{\text{tag}}}$</td>
<td>$&gt;20$</td>
<td>$&gt;15$</td>
<td>$&gt;5$</td>
</tr>
<tr>
<td>$m_T$ [GeV]</td>
<td>$&gt;200$</td>
<td>[30, 90]</td>
<td>[30, 120]</td>
</tr>
<tr>
<td>$</td>
<td>\Delta p_T^{\text{miss}}(\ell, p_T^\text{miss})</td>
<td>&gt; 0.4$ for $i \in {1, 2}$</td>
<td></td>
</tr>
<tr>
<td>Topness</td>
<td>$&gt;6.5$</td>
<td>$&gt;6.5$</td>
<td>$&gt;6.5$</td>
</tr>
<tr>
<td>$H^\text{top}$ [GeV]</td>
<td>$&lt;270$</td>
<td>$&lt;270$</td>
<td>$&lt;270$</td>
</tr>
<tr>
<td>$\Delta R(b, \ell)$</td>
<td>$&lt;3.0$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$\Delta R(b_1, b_2)$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$&gt;1.2$</td>
</tr>
<tr>
<td>Number of $b$ tags</td>
<td>$\geq 1$</td>
<td>$\geq 1/0$</td>
<td>$\geq 2$</td>
</tr>
<tr>
<td>Leading large-R jet $p_T$ [GeV]</td>
<td>$&gt;200$</td>
<td>$&gt;200$</td>
<td>$&gt;200$</td>
</tr>
<tr>
<td>Number of $b$ tags</td>
<td>$\geq 1$</td>
<td>$\geq 1/0$</td>
<td>$\geq 2$</td>
</tr>
<tr>
<td>Leading large-R jet mass [GeV]</td>
<td>$&gt;140$</td>
<td>$&gt;140$</td>
<td>$&gt;0$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>SR3</th>
<th>TCR3/WCR3</th>
<th>STCR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_\text{miss}^T$ [GeV]</td>
<td>$&gt;480$</td>
<td>$&gt;280$</td>
<td>$&gt;200$</td>
</tr>
<tr>
<td>$E_\text{miss}^{T_{\text{tag}}}$</td>
<td>$&gt;14$</td>
<td>$&gt;8$</td>
<td>$&gt;5$</td>
</tr>
<tr>
<td>$m_T$ [GeV]</td>
<td>$&gt;190$</td>
<td>[30, 90]</td>
<td>[30, 120]</td>
</tr>
<tr>
<td>$</td>
<td>\Delta p_T^{\text{miss}}(\ell, p_T^\text{miss})</td>
<td>&gt; 0.4$ for $i \in {1, 2}$</td>
<td></td>
</tr>
<tr>
<td>Topness</td>
<td>$&gt;9.5$</td>
<td>$&gt;0$</td>
<td>$&gt;9.5$</td>
</tr>
<tr>
<td>$\Delta R(b, \ell)$</td>
<td>$&lt;2.8$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$\Delta R(b_1, b_2)$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$&gt;1.2$</td>
</tr>
<tr>
<td>Number of $b$ tags</td>
<td>$\geq 1$</td>
<td>$\geq 1/0$</td>
<td>$\geq 2$</td>
</tr>
<tr>
<td>Leading large-R jet $p_T$ [GeV]</td>
<td>$&gt;280$</td>
<td>$&gt;200$</td>
<td>$&gt;200$</td>
</tr>
<tr>
<td>Number of $b$ tags</td>
<td>$\geq 1$</td>
<td>$\geq 1/0$</td>
<td>$\geq 2$</td>
</tr>
</tbody>
</table>

$\Delta R$ between the highest-$p_T$ $b$-tagged jet and the signal lepton. The signal regions have additional requirements on the $m_T$ and $H^{T_{\text{miss}}}_{\text{tag}}$ variables to further exploit the large genuine $E^{\text{miss}}_T$ from undetected neutralinos. A requirement of at least one $b$-tagged jet is used in each of SR1–3 in order to reduce the $W +$ jets and diboson backgrounds.

The signal region definitions are summarized in Table III. The signal regions are not mutually exclusive.
VI. BACKGROUND ESTIMATES

The dominant background processes are $t\bar{t}$, single top ($Wt$), $t\bar{t} + Z(\rightarrow \ell\nu\ell\nu)$, and $W + \text{jets}$. Most of the $t\bar{t}$ and $Wt$ events in the signal regions have both $W$ bosons decay leptonically (one of which is “lost” meaning it is either not reconstructed, not identified, or removed by the overlap removal procedure) or one $W$ boson decays leptonically and the other via a hadronically decaying $\tau$ lepton. Other background processes considered are semileptonic $t\bar{t}$, dibosons (denoted by $VV$ in figure legends), $t\bar{t} + W$, $Z + \text{jets}$, and multijet events. The $t\bar{t}$ background is shown separately in the three decay components discussed above, which are referred to as 2L, 1L$\tau$, and 1L respectively. The combined $t\bar{t} + W$ and $t\bar{t} + Z$ background is referred to as $t\bar{t} + V$.

The main background processes are estimated by isolating each of them in a dedicated control region (CR), described in Sec. VI A, normalizing simulation to match data in a simultaneous fit. The fit is performed separately for each SR with the associated CRs. The background modeling as predicted by the fits is tested in a series of validation regions (VRs), discussed in Sec. VI B. Figure 2 schematically illustrates the setup for one example SR and its associated CRs and VRs. The CRs for $Wt$ and $t\bar{t} + Z$ are new with respect to the run-1 analysis.

The multijet background is estimated from data using a fake-factor method [122]. The contribution is found to be negligible. All other (small) backgrounds are determined from simulation, normalized to the most accurate theoretical cross sections available. The $Z + \text{jets}$ background is found to be negligible.

A. Control regions

A series of control regions are defined as event selections that are kinematically close to the signal regions but with a few key variable requirements inverted to significantly reduce signal contamination and enhance the yield and purity of a particular background. These control regions are then used to constrain the background normalization. Each of the three signal regions has a dedicated control region for each of the following background processes: $t\bar{t}$ (TCR), $W + \text{jets}$ (WCR), single top (STCR), and $t\bar{t} + W/Z$ (TZCR). The general strategy in constructing the control regions is to invert the transverse mass requirement from a high threshold to a low window. The requirements on several variables are loosened to increase the statistical power of the CR. The details of the TCR and the WCR are described in Sec. VI A 1, while the STCR and TZCR are documented in Sec. VI A 2 and VI A 3, respectively. Table IIII presents an overview of the CR selections for the TCR, WCR, and STCR corresponding to SR1, SR2, and SR3.

A likelihood fit is performed for each SR and involves the SR and the associated CRs [123]. The expected number of events in each region is given by the sum over all background processes and optionally a signal model. The normalizations of the $t\bar{t}$, $t\bar{t} + W/Z$, single top, and $W + \text{jets}$ backgrounds are controlled by four free parameters [normalization factors (NFs)] in the fit. Furthermore, a signal strength parameter to normalize the cross section of a given signal model can be included in the fit. Each fit is based on up to five observables: the total yields in four control regions and the total yield in one signal region. The electron and muon channels are always added together. To obtain a set of background predictions that are independent of the observations in the SRs, the fit can be configured to use only the CRs to constrain the fit parameters: the SR bins are removed from the likelihood and any potential signal contribution is neglected everywhere. This fit configuration is referred to as the background-only fit.

1. Top and $W$ CRs

The TCRs and WCRs are constructed by modifying the $m_T$ selection in the SRs to be a window whose upper edge is near the $W$ boson mass. An additional upper bound on $am_T^2$ is applied to the TCRs in order to make them orthogonal to the STCRs, described in the next section. Furthermore, some other kinematic requirements are relaxed or removed to increase the event yields in the CRs. The resulting selections yield 238, 102, and 121 events in TCR1, TCR2, and TCR3, respectively, which are enriched in semileptonic $t\bar{t}$ events with purities that vary between 75% and 85%. The WCRs are built from the TCRs by changing the $b$-jet requirement to a $b$-jet veto and

---

8The letter L is used to denote an electron or muon, including those from a leptonic $\tau$ decay; the $\tau$ symbol is used to denote a hadronic $\tau$ decay.
relaxing the $am_{T2}$ requirement. The $b$-jet veto suppresses $\bar{t}t$ events and results in a $W + \text{jets}$ purity of approximately 75% in all three regions. The selections yield 558, 135, and 352 events in WCR1, WCR2, and WCR3, respectively.

2. Single-top CR

All of the expected single-top contributions in the three SRs are in the $Wt$ channel. This process can evade kinematic bounds from selections targeting the suppression of $\bar{t}t$.

Nonetheless, isolating a pure sample of $Wt$ events kinematically close to the SRs is challenging due to the similarity of $Wt$ and $\bar{t}t$. The $Wt$ events that pass event selections similar to those for the SRs often have a second $b$-jet within the acceptance. The $am_{T2}$ variable is useful for discriminating between $\bar{t}t$ and $Wt$ because the mass of the $Wb$ system not from the resonant top quark is typically higher than for an on-shell top quark in the phase space selected by this analysis. Therefore, the STCRs are characterized by $am_{T2} > 200 \text{ GeV}$. Furthermore, to increase the purity of $Wt$ and reduce the $W + \text{jets}$ contamination,
events are required to have two \( b \)-tagged jets. Top quark pair events often exceed the \( am_{T_2} \) kinematic bound when one of the two \( b \) tags used in the \( am_{T_2} \) calculation is a jet produced from a charm quark from the \( W \) decay. When this jet is from the same top quark as the other \( b \)-tagged jet, the \( \Delta R \) between them tends to be smaller than for \( Wt \) events that have two \( b \) jets from \( b \) quarks that are naturally well separated. Therefore, to further increase the \( Wt \) purity, events in the STCRs are required to have \( \Delta R(b_1, b_2) > 1.2 \), where \( b_1 \) and \( b_2 \) are the two highest-\( p_T \) \( b \)-tagged jets. Figure 3 shows distributions of the key variables for STCR1 with all requirements applied except for that on the quantity plotted. The expected purity for \( Wt \) events is approximately 40% in all three STCRs, and the selections yield 62, 71, and 45 events in STCR1, STCR2, and STCR3, respectively.

3. \( \bar{t}t + Z \) CR

Top quark pair production in association with a \( Z \) boson that decays into neutrinos is an irreducible background. The expected contributions of \( \bar{t}t + W \) in the three SRs are less than 10% with respect to the expected \( \bar{t}t + Z \) yields, and the two processes are combined in the analysis. A CR using \( Z \) boson decays to charged leptons is not feasible given the small branching ratio to leptons and the limited data set available. However, a data-driven approach is still possible using a similar process: \( \bar{t}t + \gamma \). Similar techniques have been used for estimating \( Z(\to \nu\bar{\nu}) + \text{jets from }\gamma + \text{jets} \) [124] and the method was studied as a VR in the direct top quark search with one lepton with run-1 data [29]. An event selection is constructed requiring a high-\( p_T \) photon that is then treated as \( \slashed{E}_T \) in direct analogy to \( Z \to \nu\bar{\nu} \).

The CR is designed to minimize the differences between the two processes, in order to reduce the theoretical uncertainties in the extrapolation. The Feynman diagrams for the production of \( \bar{t}t + Z \) and \( \bar{t}t + \gamma \) are identical, except for a negligible production contribution where the \( Z \) boson is radiated from a neutrino (the coupling is absent for photons). The main differences arise from the \( Z \) boson mass, which reduces the available phase space, causing differences in kinematic distributions. In addition, the bremsstrahlung rate for \( Z \) bosons is highly suppressed at LHC energies, while there is a large contribution to the \( \bar{t}t + \gamma \) cross section from photons radiated from the top quark or its decay products. Both of these differences are mitigated if the boson \( p_T \) is larger than the \( Z \) boson mass. In this limit, the impact of the mass difference on the available phase space is reduced and the rate of photon radiation from bremsstrahlung is suppressed [87]. This small fraction of photons is fully accounted for in the simulation and any uncertainty in their modeling is subdominant compared to the uncertainties described in Sec. VII. In high-\( E_T \) \( \bar{t}t + Z(\to \nu\bar{\nu}) \) events, the \( Z \) boson \( p_T \) is the dominant source of \( \slashed{E}_T \) and so most \( \bar{t}t + Z \) events in the SRs have large \( Z \) boson \( p_T \).

Two \( \bar{t}t + \gamma \) CRs are designed to be kinematically close to SR1 and SR2/SR3. The event selection for TZCR2 is the same as for TZCR3. The regions require at least one signal photon, exactly one signal lepton and no additional baseline leptons, and at least four signal jets, of which at least one must be \( b \) tagged. In addition, the regions have the same jet \( p_T \) thresholds as the corresponding signal regions. To mimic the \( Z \to \nu\bar{\nu} \) decay, the highest-\( p_T \) photon is vectorially added to \( \slashed{p}_T \) and this sum is used to construct

---

FIG. 4. Comparison of data with estimated backgrounds in the \( \slashed{E}_T \) and \( \slashed{m}_T \) distributions with the TZCR1 event selection except for the requirement (indicated by an arrow) on the shown variable. The variables \( \slashed{E}_T \) and \( \slashed{m}_T \) are constructed in the same way as \( \slashed{E}_T \) and \( m_T \) but treating the leading photon transverse momentum as invisible. The predicted backgrounds are scaled with the NFs documented in Table IV. The uncertainty band includes statistical and all experimental systematic uncertainties. The last bin includes overflow.
TABLE IV. The numbers of observed events in the three SRs together with the expected numbers of background events and their uncertainties as predicted by the background-only fits, the scaling factors for the background predictions in the fit (NF), the probabilities (represented by the $p_0$ values) that the observed numbers of events are compatible with the background-only hypothesis, as well as the expected and observed 95% C.L. upper limits on the number of non-SM events.

<table>
<thead>
<tr>
<th>Signal region</th>
<th>Observed</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total background</td>
<td>12</td>
<td>$5.50 \pm 0.72$</td>
<td>$1.25 \pm 0.26$</td>
<td>$1.03 \pm 0.18$</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>2.21 $\pm 0.60$</td>
<td>$0.29 \pm 0.10$</td>
<td>$0.20 \pm 0.07$</td>
<td></td>
</tr>
<tr>
<td>(1L, 1L1t, 2L) in %</td>
<td>(6, 48, 46)</td>
<td>(0, 58, 42)</td>
<td>(0, 36, 64)</td>
<td></td>
</tr>
<tr>
<td>Single top</td>
<td>0.46 $\pm 0.39$</td>
<td>$0.09 \pm 0.08$</td>
<td>$0.10 \pm 0.09$</td>
<td></td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>0.71 $\pm 0.43$</td>
<td>$0.15 \pm 0.09$</td>
<td>$0.20 \pm 0.09$</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + V$</td>
<td>1.90 $\pm 0.42$</td>
<td>$0.61 \pm 0.14$</td>
<td>$0.41 \pm 0.10$</td>
<td></td>
</tr>
<tr>
<td>Diboson</td>
<td>0.23 $\pm 0.15$</td>
<td>$0.11 \pm 0.07$</td>
<td>$0.12 \pm 0.07$</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ NF</td>
<td>1.10 $\pm 0.14$</td>
<td>$0.10 \pm 0.14$</td>
<td>$0.80 \pm 0.13$</td>
<td></td>
</tr>
<tr>
<td>Single top NF</td>
<td>0.62 $\pm 0.46$</td>
<td>$0.65 \pm 0.49$</td>
<td>$0.71 \pm 0.42$</td>
<td></td>
</tr>
<tr>
<td>$W +$ jets NF</td>
<td>0.75 $\pm 0.12$</td>
<td>$0.78 \pm 0.15$</td>
<td>$0.93 \pm 0.12$</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + W/Z$ NF</td>
<td>1.42 $\pm 0.24$</td>
<td>$1.45 \pm 0.24$</td>
<td>$1.46 \pm 0.24$</td>
<td></td>
</tr>
<tr>
<td>$p_0$</td>
<td>0.012 (2.3 $\sigma$)</td>
<td>0.50 (0.0 $\sigma$)</td>
<td>0.50 (0.0 $\sigma$)</td>
<td></td>
</tr>
<tr>
<td>$\lambda_{\text{non-SM}}^\text{exp.}$</td>
<td>6.4 $^{+3.2}_{-2.0}$</td>
<td>3.6 $^{+2.3}_{-1.3}$</td>
<td>3.5 $^{+2.2}_{-1.2}$</td>
<td></td>
</tr>
<tr>
<td>$\lambda_{\text{non-SM}}^\text{obs.}$</td>
<td>13.3</td>
<td>3.4</td>
<td>3.4</td>
<td></td>
</tr>
</tbody>
</table>

SR1 before the requirement on the plotted variable is applied. The contribution from events not involving top quarks is negligible. The predicted backgrounds in the figure are scaled with the NFs documented in Table IV. Without scaling, the total number of events in data is about 40% higher than in simulation, but there is no significant evidence of mismodeling of the shapes of the various distributions within uncertainties.

B. Validation regions

The background estimates are tested using validation regions, which are disjoint to both the control and signal regions. Background normalizations determined in the control regions are extrapolated to the VRs and compared with the observed data. Each signal region has associated validation regions for the $t\bar{t}$ (TVR) and $W +$ jets (WVR) processes, and these are constructed with the same selection as the TCR/WCR except that $m_T$ is between 90 and 120 GeV.\textsuperscript{9} The validation regions are not used to constrain parameters in the fit but provide a statistically independent test of the background estimates made using the CRs. In Fig. 5, background estimates in all the associated VRs are compared to the observed data. The potential signal contamination in the VRs is studied for all considered signal models (and SUSY mass ranges) and found to be negligible.

A second set of validation regions, not associated with any of the three signal regions, is used for general monitoring purposes. Two of the more significant backgrounds are dileptonic $t\bar{t}$ and lepton+hadronic $\tau$ $t\bar{t}$ events.

\textsuperscript{9}A $Wt$ VR is not defined since the $m_T$ range in the STCR is extended upward to 120 GeV to accept more events.
To pass the four-jet requirement, such events must have at least one hard jet that does not originate from the \( \bar{t}t \) decay (two hard jets for dileptonic \( \bar{t}t \)). The modeling of these extra jets is validated in dedicated VRs that require either two signal leptons (electron or muon) or one signal lepton and one hadronic \( \tau \) candidate. In Fig. 6 the jet multiplicity distributions are shown for event selections requiring an electron-muon pair (left) and one lepton plus one \( \tau \) candidate (right). Additional validation regions are constructed by considering (i) events with high \( E_{\text{miss}}^{\text{miss}} \), high \( m_T \), and low am\( m_T \) for dilepton \( \bar{t}t \) events with a lost lepton or (ii) high \( m_T \) and a \( b \)-jet veto to probe the modeling of the resolution-induced \( m_T \) tail in \( W + \) jets events (using the WVR-tail region in Fig. 2). There are no significant indications of mismodeling in any of the validation regions.

VII. SYSTEMATIC UNCERTAINTIES

The systematic uncertainties in the signal and background estimates arise both from experimental sources and from the uncertainties in the theoretical predictions and modeling. Since the yields from the dominant background sources, \( \bar{t}t \), single top, \( t\bar{t}V \), and \( W + \) jets, are all obtained in dedicated control regions, the modeling uncertainties for these processes affect only the extrapolation from the CRs into the signal regions (and between the various control regions) but not the overall normalization. The systematic uncertainties are included as nuisance parameters with Gaussian constraints and profiled in the likelihood fits.

The dominant experimental uncertainties arise from imperfect knowledge of the jet energy scale (JES) and jet energy resolution (JER) [101], the modeling of the \( b \)-tagging efficiencies for \( b \), \( c \) and light-flavor jets [125,126] as well as the contribution to the \( E_T^{\text{miss}} \) soft term, i.e., from tracks neither associated with any reconstructed objects nor identified as originating from pileup. From these sources, the resulting uncertainties in the extrapolation factors for going from the four CRs to the SRs are 4%–15% for JES, 0%–9% for JER, 0%–6% for \( b \) tagging, and 0%–3% for the \( E_T^{\text{miss}} \) soft term. Other sources of experimental uncertainty are the modeling of lepton- and photon-related quantities (energy scales, resolutions, reconstruction and identification efficiencies, isolation, hadronic-\( \tau \) identification) and the uncertainty in the integrated luminosity. These uncertainties have a small impact on the final results.

The uncertainties in the modeling of the single-top and \( \bar{t}t \) backgrounds include effects related to the MC event generator, the hadronization and fragmentation modeling, and the amount of initial- and final-state radiation [71]. The MC generator uncertainty is estimated by comparing events produced with POWHEG-BOX+HERWIG++ and with MG5_AMC+HERWIG++. Events generated with POWHEG-BOX are hadronized with either PYTHIA or HERWIG++ to estimate the effect from the modeling of the fragmentation and hadronization. The impact of altering the amount of initial- and final-state radiation is estimated from comparisons of POWHEG-BOX+PYTHIA samples with different parton shower radiation, NLO radiation, and modified factorization and renormalization scales. One additional uncertainty stems from the modeling of the interference between the \( \bar{t}t \) and \( Wt \) processes at NLO. The uncertainty is estimated using inclusive \( WWbb \) events, generated using MG5_AMC, which are compared with the sum of the \( \bar{t}t \) and \( Wt \) processes [71]. The resulting theoretical uncertainties in the extrapolation factors for going from the \( \bar{t}t \) and \( Wt \) CRs to the SRs are 19%–26% for \( \bar{t}t \) and 38%–57% for \( Wt \) events, where the latter is dominated by the interference term.
extrapolated to higher jet multiplicities, are applied in all regions requiring at least one $b$-tagged jet. The resulting theoretical uncertainties in the extrapolation from the $W + \text{jets}$ CR to the SR amount to about 40%.

Since the diboson backgrounds are not normalized in a CR, the analysis is sensitive to the uncertainty in the total cross section, estimated to be 6%. In addition, the estimate from the nominal SHERPA sample is compared to that from a POWHEG-BOX+PYTHIA sample to account for differences related to the MC event generator modeling. The resulting theoretical uncertainties for the diboson yields in the three SRs are about 50%.

The SUSY signal cross-section uncertainty is taken from an envelope of cross-section predictions using different PDF sets and factorization and renormalization scales, as described in Ref. [129], and the resulting uncertainties range from 13% to 23%. The uncertainty in the VLQ signal cross section is 10% [80].

**VIII. RESULTS**

Table IV (top part) and Fig. 5 (right part) show the number of observed events together with the predicted number of background events in the three SRs. The prediction is obtained using the background-only fit configuration described in Sec. VI. The SR2 and SR3 predicted yields agree well with the observed data in those regions. Table IV (middle part) also lists the results for the four free fit parameters that control the normalization of the four main backgrounds (NFs), together with the associated fit uncertainties. To quantify the compatibility of the SM background-only hypothesis with the observations in the SRs, a profile likelihood ratio test is performed. These fits are configured to include the SR bin in the likelihood. Table IV reports the resulting $p$ values ($p_0$), which are set to 0.5 for SR2 and SR3 since the observation lies below the prediction. The data exceeds the background prediction in

![Image](image_url)
Exclusion limits are also derived for the gluino-mediated top squark and direct top squark pair production models. The signal uncertainties and potential signal contributions to all regions are taken into account. All uncertainties except those in the theoretical signal cross section are included in the fit. Combined exclusion limits are obtained by selecting a priori the signal region with the lowest expected CL$_{s}$ value for each signal model.

Figure 8 shows the expected and observed exclusion contours for both gluino-mediated and direct pair production of top squarks. The ±1σ$_{\text{exp}}$ (yellow) uncertainty band indicates the impact on the expected limit of all uncertainties included in the fit. The ±1σ$_{\text{th}}$ (dotted red) uncertainty lines around the observed limit illustrate the change in the observed limit as the nominal signal cross section is scaled up and down by the theoretical cross-section uncertainty. The gap in the observed exclusion between about 600 and 750 GeV in the direct top squark model is due to a transition between signal regions and the excess observed in SR1. For any model point, the single signal region used for the observed exclusion is chosen to be the one with the best expected CL$_{s}$ value.
pair-produced \( T \) quarks are used to reinterpret the results. The \( T \) quark is assumed to decay in three possible ways: \( T \rightarrow Z, T \rightarrow tH, \) and \( T \rightarrow bW \). The search described in this paper has sensitivity mostly to the \( T \rightarrow iZ \) decay mode with \( Z \rightarrow \ell \bar{\nu} \) due to the large \( E_{T}^{miss} \) requirements in the analysis. The direct \( T \) pair production cross section is higher for top squarks due to additional spin states, but after accounting for the \( Z \rightarrow \ell \bar{\nu} \) branching ratio, the models have a similar predicted yield. For a \( T \) quark with mass 800 GeV (just beyond the run-1 limit \([34,140]\)), a branching ratio \( B(T \rightarrow iZ) \) above about 90\% is excluded. Figure 9 shows the exclusion limit as a function of the \( T \) quark mass. Assuming a branching ratio for \( T \rightarrow iZ \) of 100\%, \( T \) masses up to about 850 GeV are excluded.

**IX. CONCLUSION**

This paper presents a search for pair production of gluino-mediated top squarks with a small mass splitting between the top squark and the LSP and direct pair production of top squarks decaying to two top quarks and two lightest neutralinos in final states with one isolated lepton, jets, and missing transverse momentum. Three signal region selections are optimized for discovery in benchmark models just beyond the exclusion limits from LHC run-1 searches with the same \( \bar{t} + E_{T}^{miss} \) signature. The search uses 3.2 fb\(^{-1} \) of LHC \( pp \) collision data collected by the ATLAS experiment at a center-of-mass energy of \( \sqrt{s} = 13 \) TeV. The observed data are consistent with data-driven background estimates in all three regions. The largest difference between data and the corresponding prediction is in the most inclusive signal region (SR1) and corresponds to 2.3 standard deviations above the estimated background. In the absence of a significant excess, exclusion limits at 95\% C.L. are derived in the gluino and top squark pair production models. These extend the LHC run-1 exclusion limits on the gluino mass upward to 1460 GeV in the gluino-mediated top squark pair production model for low top squark masses. For the direct top squark pair production models the results expand the LHC run-1 exclusion limits by excluding the top squark mass region from 745 to 780 GeV for a massless lightest neutralino. The analysis results are also reinterpreted to set exclusion limits in a model of vectorlike top quarks (\( T \)). Assuming a branching ratio for \( T \rightarrow iZ \) of 100\%, \( T \) masses up to about 850 GeV are excluded.

**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; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; CMSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, 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; FOM and NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC Ki, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, 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, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; 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), INP-NCNF (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. \[141\].


[34] ATLAS Collaboration, Search for pair and single production of new heavy quarks that decay to a Z boson and a third-generation quark in $pp$ collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, J. High Energy Phys. 11 (2014) 104.


M. Czakon, P. Fiedler, and A. Mitov, Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(a_s^2)$, Phys. Rev. Lett. 110, 252004 (2013).


P. Bärnreuther, M. Czakon, and A. Mitov, Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$, Phys. Rev. Lett. 109, 132001 (2012).


model $Z(\rightarrow \nu \bar{\nu}) +$ jets background to new physics processes at the LHC, J. High Energy Phys. 10 (2011) 058.


SEARCH FOR TOP SQUARKS IN FINAL STATES WITH

PHYSICAL REVIEW D 94, 052009 (2016)
SEARCH FOR TOP SQUARKS IN FINAL STATES WITH...
Also at Louisiana Tech University, Ruston, LA, USA.
2 Also at Institut Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain.
3 Also at Graduate School of Science, Osaka University, Osaka, Japan.
4 Also at Department of Physics, National Tsing Hua University, Taiwan.
5 Also at Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, Netherlands.
6 Also at Department of Physics, The University of Texas at Austin, Austin, TX, USA.
7 Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia.
8 Also at CERN, Geneva, Switzerland.
9 Also at Georgian Technical University (GTU), Tbilisi, Georgia.
a Also at Ochadai Academic Production, Ochanomizu University, Tokyo, Japan.
b Also at Manhattan College, New York, NY, USA.
c Also at Hellenic Open University, Patras, Greece.
d Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan.
e Also at School of Physics, Shandong University, Shandong, China.
f Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia.
g Also at Section de Physique, Université de Genève, Geneva, Switzerland.
h Also at Eotvos Lorand University, Budapest, Hungary.
i Also at International School for Advanced Studies (SISSA), Trieste, Italy.
j Also at Department of Physics and Astronomy, University of South Carolina, Columbia, SC, USA.
k Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China.
l Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria.
m Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia.
n Also at Institute of Physics, Academia Sinica, Taipei, Taiwan.
o Also at National Research Nuclear University MEPhI, Moscow, Russia.
p Also at Department of Physics, Stanford University, Stanford, CA, USA.
q Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary.
r Also at Flensburg University of Applied Sciences, Flensburg, Germany.
s Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia.
t Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France.