ATLAS CONF Note

ATLAS-CONF-2019-016

20th May 2019

Search for top squarks in events with a Z boson using 139 fb^{-1} of pp collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector

The ATLAS Collaboration

A search for direct top squark pair production in events with a same-flavour opposite-sign dilepton pair with invariant mass consistent with a Z boson plus missing transverse momentum is presented. The analysis is performed using the proton-proton collision data at $\sqrt{s}$ = 13 TeV collected by the ATLAS experiment during the LHC Run-2, corresponding to an integrated luminosity of 139 fb^{-1}. No excess is observed in the data with respect to the SM predictions. The results are interpreted in simplified models featuring direct production of pairs of either the lighter top squark ($\tilde{t}_1$) or the heavier top squark ($\tilde{t}_2$), excluding at 95% confidence level $\tilde{t}_1$ and $\tilde{t}_2$ masses up to about 1140 and 875 GeV, respectively.

ATLAS-CONF-2019-016
28 May 2019

1 Introduction

Supersymmetry (SUSY) [1–6] is one of the most studied frameworks to extend the Standard Model (SM) beyond the electroweak scale. It predicts new bosonic (fermionic) partners for the known fermions (bosons). Assuming R-parity conservation [7], SUSY particles are produced in pairs and the lightest supersymmetric particle (LSP) is stable, providing a possible dark-matter candidate. The SUSY partners of the Higgs bosons and electroweak gauge bosons mix to form the mass eigenstates known as charginos ($\tilde{\chi}_k^{\pm}$, $k = 1, 2$) and neutralinos ($\tilde{\chi}_m^0$, $m = 1, \ldots, 4$), where the increasing index denotes increasing mass. The scalar partners of right-handed and left-handed quarks, $\tilde{q}_{\mathrm{R}}$ and $\tilde{q}_{\mathrm{L}}$ squarks, mix to form two mass eigenstates, $\tilde{q}_1$ and $\tilde{q}_2$, with $\tilde{q}_1$ defined to be the lighter of the two. To address the SM hierarchy problem [8–11], TeV-scale masses are required [12, 13] for the supersymmetric partners of the gluons and the top squarks [14, 15].

Top squark production with Z bosons in the decay chain can appear either in production of the lighter top squark mass eigenstate ($\tilde{t}_1$) decaying via $\tilde{t}_1 \to t\tilde{\chi}_2^0$ with $\tilde{\chi}_2^0 \to Z\tilde{\chi}_1^0$, or in production of the heavier top squark mass eigenstate ($\tilde{t}_2$) decaying via $\tilde{t}_2 \to Z\tilde{t}_1$ with $\tilde{t}_1 \to t^{(*)}\tilde{\chi}_1^0$, as illustrated in Figure 1. Unlike other top squark models, these signals can be efficiently discriminated from the SM top quark pair production ($t\bar{t}$) background by requiring a same-flavour opposite-sign (SF-OS) lepton pair originating from the $Z \to \ell^+\ell^-$ decay plus the presence of an additional lepton produced in the decay of the top quarks in the event.

Simplified models [16–18] are used for the analysis optimisation and interpretation of the results. In these models, direct top squark pair production is considered and all SUSY particles are decoupled except for the top squarks and the neutralinos involved in their decay. In all cases the $\tilde{\chi}_1^0$ is assumed to be the LSP. Simplified models featuring direct $\tilde{t}_1$ production with $\tilde{t}_1 \to t\tilde{\chi}_2^0$ and decays either via Z ($\tilde{\chi}_2^0 \to Z\tilde{\chi}_1^0$) or Higgs ($h$) ($\tilde{\chi}_2^0 \to h\tilde{\chi}_1^0$) bosons with a 50% branching ratio are considered. In these models, the $\tilde{\chi}_1^0$ is assumed to be very light and the $\tilde{\chi}_2^0 - \tilde{\chi}_1^0$ mass difference to be large enough to allow on-shell Z or Higgs boson decays. Additional simplified models featuring direct $\tilde{t}_2$ production with $\tilde{t}_2 \to Z\tilde{t}_1$ decays and $\tilde{t}_1 \to t^{(*)}\tilde{\chi}_1^0$ are also considered. The mass difference between the $\tilde{t}_1$ and $\tilde{\chi}_1^0$ is set to 40 GeV, and the four-body decay channel $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ is assumed to occur, where f and f' are two fermions from the W^* decay.

This note presents the results of a search for top squarks in final states with Z bosons at $\sqrt{s} = 13$ TeV using the complete data set collected by the ATLAS experiment [19] in proton–proton (pp) collisions during Run-2 (2015-2018), corresponding to 139 fb^{-1}. Searches for top squark production in events involving Z bosons have been previously performed by both ATLAS [20, 21] and CMS [22, 23].

2 ATLAS detector

The ATLAS experiment [19] at the LHC is a multi-purpose particle detector with a forward-backward symmetric cylindrical geometry and a near 4π coverage in solid angle.[1] It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer. The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors.

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

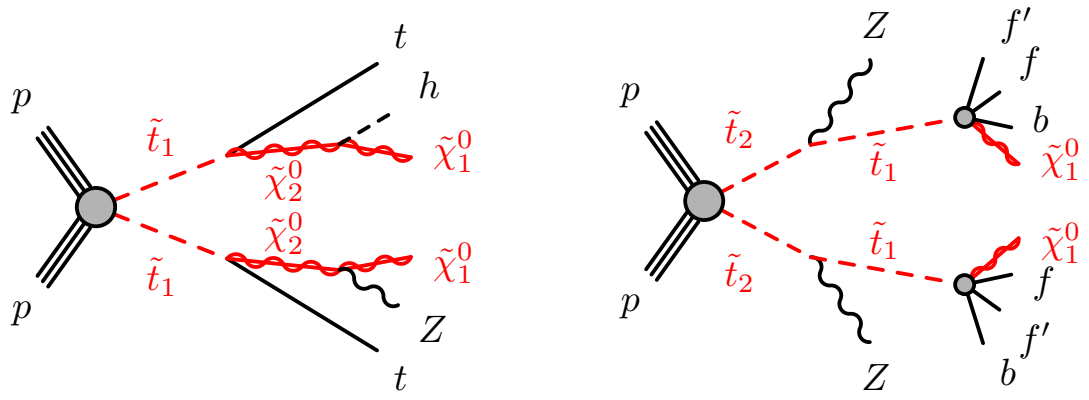

Figure 1: Diagrams for the top squark pair production processes considered in this analysis: (a) $\tilde{t}_1 \to t\tilde{\chi}_2^0$ with $\tilde{\chi}_2^0 \to h/Z\tilde{\chi}_1^0$ decays, and (b) $\tilde{t}_2 \to Z\tilde{t}_1$ with $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ decays.

Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A hadron (steel/scintillator-tile) calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The end-cap and forward regions are instrumented with LAr calorimeters for both EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large air-core toroidal superconducting magnets with eight coils each. The field integral of the toroids ranges between 2 and 6 T·m across most of the detector. The muon spectrometer includes a system of precision tracking chambers and fast detectors for triggering. A three-level trigger system is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most nearly 100 kHz. This is followed by two software-based trigger levels that together reduce the accepted event rate to 1 kHz on average depending on the data-taking conditions.

3 Data set and simulated event samples

The data were collected by the ATLAS detector during the LHC Run-2 (2015–2018) with a peak instantaneous luminosity of $\mathcal{L} = 2.1 \times 10^{34}$ cm^{-2}s^{-1}, resulting in a mean number of pp interactions per bunch crossing of $\langle\mu\rangle = 34$. Data quality requirements are applied to ensure that all subdetectors were operating at nominal conditions, and that LHC beams were in stable-collision mode. The integrated luminosity of the resulting data set is 139 fb^{-1}. The uncertainty in the combined 2015-2018 integrated luminosity is 1.7%. It is derived from the calibration of the luminosity scale using x-y beam-separation scans, following a methodology similar to that detailed in Ref. [24], and using the LUCID-2 detector for the baseline luminosity measurements [25].

Monte Carlo (MC) simulated event samples are used to aid in the estimation of the background from SM processes and to model the SUSY signal. The choices of MC event generator, parton shower and hadronisation, the cross-section normalisation, the parton distribution function (PDF) set and the set of tuned parameters (tune) for the underlying event of these samples are summarised in Table 1. More details of the event generator configurations can be found in Refs. [26–29]. For production of top quark pairs

Table 1: Simulated signal and background event samples: the corresponding event generator used for the hard-scatter process, the generator used to model the parton showering, the source of the cross-section used for normalisation, the PDF set and the underlying-event tune are shown.

Physics process	Generator	Parton shower	Cross-section normalisation	PDF set	Tune
SUSY Signals	MG5_aMC@NLO 2.6.2 [31]	Pythia 8.212 [32]	NNLO+NNLL [33–37]	NNPDF2.3LO [38]	A14 [39]
$t\bar{t}Z/\gamma^*$, $t\bar{t}W$	MG5_aMC@NLO 2.3.3	Pythia 8.210	NLO [31]	NNPDF2.3LO	A14
Diboson	Sherpa 2.2.2 [40]	Sherpa 2.2.2	Generator NLO	NNPDF3.0NNLO [41]	Sherpa default
$t\bar{t}h$	Powheg v2 [42]	Pythia 8.230	NLO [43]	NNPDF2.3LO	A14
Wh, Zh	Pythia 8.186 [44]	Pythia 8.186	NLO [43]	NNPDF2.3LO	A14
$t\bar{t}WW$, $t\bar{t}t\bar{t}$	MG5_aMC@NLO 2.2.2	Pythia 8.186	NLO [31]	NNPDF2.3LO	A14
$t\bar{t}t$	MG5_aMC@NLO 2.2.2	Pythia 8.186	LO	NNPDF2.3LO	A14
tZ	MG5_aMC@NLO 2.3.3	Pythia 8.186	LO	NNPDF2.3LO	A14
tWZ	MG5_aMC@NLO 2.3.3	Pythia 8.212	Generator NLO	NNPDF2.3LO	A14
Triboson	Sherpa 2.2.2	Sherpa 2.2.2	Generator NLO	NNPDF3.0NNLO	Sherpa default

in association with vector or Higgs bosons, cross-sections calculated at next-to-leading order (NLO) are used, and the event generator NLO cross-sections from Sherpa are used when normalising the multi-boson backgrounds. In all MC samples, except those produced by Sherpa, the EvtGen v1.2.0 program [30] is used to model the properties of the bottom and charm hadron decays.

SUSY signal samples are generated with MG5_aMC@NLO 2.6.2 [31] interfaced to Pythia 8.212 [32] for the parton showering (PS) and hadronization. The matrix element (ME) calculation was performed at tree level and includes the emission of up to two additional partons for all signal samples. The parton distribution function (PDF) set used for the generation of the signal samples is NNPDF2.3LO [38] with the A14 [39] set of tuned underlying-event and shower parameters (UE tune). The ME–PS matching was performed with the CKKW-L prescription [45], with a matching scale set to one quarter of the mass of the pair-produced superpartner mass. All signal cross-sections are calculated to approximate next-to-next-to-leading order in the strong coupling constant, adding the resummation of soft gluon emission at next-to-next-to-leading-logarithmic accuracy (approximate NNLO+NNLL) [46–49]. The nominal cross-section and its uncertainty are derived using the PDF4LHC15_mc PDF set, following the recommendations of Ref. [50].

To simulate the effects of additional pp collisions in the same and nearby bunch crossings (pile-up), additional interactions are generated using the soft QCD processes as provided by Pythia 8.186 with the A3 tune [51] and the MSTW2008LO PDF set [52], and overlaid onto each simulated hard-scatter event. The MC samples are reweighted so that the pile-up distribution matches the one observed in the data. The MC samples are processed through an ATLAS detector simulation [53] based on Geant4 [54] or, in the case of $t\bar{t}t$ and the SUSY signal samples, a fast simulation using a parameterisation of the calorimeter response and Geant4 for the other parts of the detector. All MC samples are reconstructed in the same manner as the data.

4 Event selection

Candidate events are required to have a reconstructed vertex [55] with at least two associated tracks with transverse momentum (p_{T}) larger than 500 MeV which are consistent with originating from the beam collision region in the x–y plane. The primary vertex in the event is the vertex with the highest sum of squared transverse momenta of associated tracks.

Two categories of leptons (electrons and muons) are defined: "candidate" and "signal" (the latter being a subset of the "candidate" leptons satisfying tighter selection criteria). Electron candidates are reconstructed from isolated electromagnetic calorimeter energy deposits matched to ID tracks and are required to have $|\eta| < 2.47$, a transverse momentum $p_{\mathrm{T}} > 4.5$ GeV, and to pass the "LooseAndBLayer" requirement defined in Ref. [56], which is based on a likelihood using measurements of shower shapes in the calorimeter and track properties in the ID as input variables.

Muon candidates are reconstructed in the region $|\eta| < 2.4$ from muon spectrometer tracks matching ID tracks. Candidate muons must have $p_{\mathrm{T}} > 4$ GeV and pass the medium identification requirements defined in Ref. [57], based on the number of hits in the different ID and muon spectrometer subsystems, and on the ratio of the charge and momentum (q/p) measured in the ID and MS divided by the sum in quadrature of the corresponding uncertainties.

The tracks associated to the lepton candidates must have a significance of the transverse impact parameter with respect to the reconstructed primary vertex, d_0, of $|d_0|/\sigma(d_0) < 5$ for electrons and $|d_0|/\sigma(d_0) < 3$ for muons, and a longitudinal impact parameter with respect to the reconstructed primary vertex, z_0, satisfying $|z_0 \sin\theta| < 0.5$ mm.

Jets are reconstructed from three-dimensional energy clusters in the calorimeter [58] using the anti-k_t jet clustering algorithm [59] with a radius parameter $R = 0.4$. Only jet candidates with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.8$ are considered in the analysis. Jets are calibrated using simulation with corrections obtained from in-situ techniques [60]. In order to reduce the effects of pile-up, jets with $p_{\mathrm{T}} < 120$ GeV and $|\eta| < 2.5$ must have a significant fraction of their associated tracks compatible with originating from the primary vertex, as defined by the jet vertex tagger [61]. This requirement reduces the fraction of jets from pile-up down to 1%, with an efficiency for pure hard scatter jets of about 90%. Events are discarded if they contain any jet with $p_{\mathrm{T}} > 20$ GeV not satisfying basic quality selection criteria designed to reject detector noise and non-collision backgrounds [62].

Identification of jets containing b-hadrons (b-tagging) is performed with a multivariate discriminant that makes use of track impact parameters and reconstructed secondary vertices [63, 64]. Jets are considered as b-tagged if fulfilling a requirement corresponding to a 77% average efficiency obtained for jets containing b-hadrons in simulated $t\bar{t}$ events. The rejection factors for light-quark and gluon jets, jets containing c-hadrons and jets containing hadronically decaying τ leptons in simulated $t\bar{t}$ events are approximately 113, 16 and 4, respectively.

Jet candidates with $p_{\mathrm{T}} < 200$ GeV within $\Delta R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2} = 0.2$ of an electron candidate are discarded, unless the jet has a value of the b-tagging discriminant larger than the value corresponding to approximately 85% b-tagging efficiency, in which case the lepton is discarded since it is likely to have originated from a semileptonic b-hadron decay. The same procedure is applied to jets within $\Delta R = 0.2$ of a muon candidate irrespective of the jet p_{T}. Any remaining electron candidate within $\Delta R = 0.4$ of a non-pile-up jet, and any muon candidate within $\Delta R = \min\{0.4, 0.04 + p_{\mathrm{T}}(\mu)/10\,\mathrm{GeV}\}$ of a non-pile-up jet is discarded. In the latter case, if the jet has fewer than three associated tracks, the muon is retained and the jet is discarded instead to avoid inefficiencies for high-energy muons undergoing significant energy loss in the calorimeter. Finally, any electron candidate sharing an ID track with a remaining muon candidate is also removed.

Tighter requirements on the lepton candidates are imposed, which are then referred to as "signal" electrons or muons. Signal electrons must satisfy the "medium" identification requirement as defined in Ref. [56] and signal muons must have $p_{\mathrm{T}} > 5$ GeV. Isolation requirements are applied to both the signal electrons and muons. The scalar sum of the p_{T} of tracks within a variable-size cone around the lepton, excluding its

own track, must be less than 6% of the lepton p_{T}; these tracks are required to be associated to the primary vertex to limit sensitivity to pile-up. The size of the track isolation cone for electrons (muons) is given by the smaller of $\Delta R = 10\,\mathrm{GeV}/p_{\mathrm{T}}$ and $\Delta R = 0.2\,(0.3)$, that is, a cone of size 0.2 (0.3) at low p_{T} but narrower for high-p_{T} leptons. In addition, in the case of electrons the energy of calorimeter energy clusters in a cone of $\Delta R_\eta = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.2$ around the electron (excluding the deposition from the electron itself) must be less than 6% of the electron p_{T}.

Simulated events are corrected for differences between data and MC simulation in jet vertex tagger and b-tagging efficiencies as well as b-tagging mis-tag rates [61, 65]. Corrections are also applied to account for minor differences between data and MC simulation in the signal lepton trigger, reconstruction, identification and isolation efficiencies.

The missing transverse momentum vector, whose magnitude is denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$, is defined as the negative vector sum of the transverse momenta of all identified electrons, muons and jets, and an additional soft term. The soft term is constructed from all tracks originating from the primary vertex which are not associated with any identified lepton or jet. In this way, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is adjusted for the best calibration of leptons and jets, while contributions from pile-up interactions are suppressed through the soft term [66, 67].

Events are accepted if they pass a trigger requiring either two electrons, two muons or an electron and a muon. The trigger-level requirements on the p_{T}, identification and isolation of the leptons involved in the trigger decision are looser than those applied offline to ensure that trigger efficiencies are constant in the relevant phase space [68]. A pre-selection of events of interest is performed by requiring the presence of at least three signal leptons (electrons or muons), with at least one SF-OS lepton pair whose invariant mass is compatible with the Z boson mass ($|m_{\ell\ell} - m_Z| < 15$ GeV, with $m_Z = 91.2$ GeV). In addition, the leading lepton is required to have $p_{\mathrm{T}} > 40$ GeV and the subleading to have $p_{\mathrm{T}} > 20$ GeV. These requirements are summarised in Table 2 and are applied all through the analysis unless otherwise stated.

Table 2: Definition of the event pre-selection used in the analysis (see text for details).

	Pre-selection
Number of signal leptons	≥ 3
Number of SF-OS pairs	≥ 1
Leading lepton p_{T} [GeV]	> 40
Subleading lepton p_{T} [GeV]	> 20
$\lvert m_{\ell\ell}^{\text{SF-OS}} - m_Z\rvert$ [GeV]	< 15

Four overlapping signal regions (SRs) are optimised for the best discovery sensitivity in the simplified models discussed in Section 1. The requirements in each SR are summarised in Table 3.

Signal region $\mathrm{SR_{1A}}$ is optimised for large $\tilde{\chi}_2^0 - \tilde{\chi}_1^0$ mass splittings in the $\tilde{t}_1 \to t\tilde{\chi}_2^0$ with $\tilde{\chi}_2^0 \to h/Z\tilde{\chi}_1^0$ model. It includes requirements on $m_{\mathrm{T2}}^{3\ell}$, a variation of the stransverse mass m_{T2} which is used to bound the masses of a pair of particles that are presumed to have each decayed semi-invisibly into one visible and one invisible particle [69, 70]. In the case of $m_{\mathrm{T2}}^{3\ell}$, the two visible legs of the two semi-invisible decays are set to be the system of the SF-OS lepton pair with an invariant mass closest to m_Z, and the third lepton.

Models with small mass differences between the $\tilde{\chi}_2^0$ and the $\tilde{\chi}_1^0$ are targeted with SR_{1B}, which instead features requirements on the transverse momentum of the SF-OS lepton pair ($p_T^{\ell\ell}$).

Two SRs are optimised for the $\tilde{t}_2 \to Z\tilde{t}_1$ with $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ model, SR_{2A} and SR_{2B}, targeting small and large mass splittings between the $\tilde{t}_2$ and the $\tilde{\chi}_1^0$, respectively. Due to overall soft kinematics of the particles in compressed $\tilde{t}_2$ signals, SR_{2A} features upper bounds on the p_T of the third leading lepton and on $p_T^{\ell\ell}$, as well as no requirement on the number of b-tagged jets since they are likely to be soft in this scenario.

Table 3: Definition of the signal regions used in the analysis (see text for details).

Requirement / Region	SR_{1A}	SR_{1B}	SR_{2A}	SR_{2B}
Third leading lepton p_T [GeV]	> 20	> 20	< 20	< 60
n_{jets} ($p_T > 30$ GeV)	≥ 4	≥ 5	≥ 3	≥ 3
$n_{b-\text{tagged jets}}$ ($p_T > 30$ GeV)	≥ 1	≥ 1	–	≥ 1
Leading jet p_T [GeV]	–	–	> 150	–
Leading b-tagged jet p_T [GeV]	–	> 100	–	–
E_T^{miss} [GeV]	> 250	> 150	> 200	> 350
$p_T^{\ell\ell}$ [GeV]	–	> 150	< 50	> 150
$m_{T2}^{3\ell}$[GeV]	> 100	–	–	–

5 Background estimation

The dominant SM background contribution to the SRs is expected to originate from $t\bar{t}Z$ production, with minor contributions from multi-boson production (mainly WZ) and backgrounds containing jets misidentified as leptons (hereafter referred to as "fake" leptons) or non-prompt leptons from decays of hadrons (mainly in $t\bar{t}$ events).

The normalisation of the $t\bar{t}Z$ and multi-boson backgrounds is obtained by fitting the yield from MC simulation to the observed data in dedicated control regions (CRs) enhanced in a particular background component, and then extrapolating this yield to the SRs. Backgrounds from other sources ($t\bar{t}W$, $t\bar{t}h$ and rare SM processes), which provide a subdominant contribution to the SRs, are determined from MC simulation only.

For each signal region, a simultaneous "background fit" is performed to the numbers of events found in the CRs, using a minimisation based on likelihoods with the HistFitter package [71]. In each fit, the normalisations of the background processes having dedicated CRs are allowed to float freely, while the other backgrounds are fixed to the estimate from simulation or data-driven techniques.

Systematic uncertainties related to the MC modelling affect the expected yields in the different regions and are taken into account to determine the uncertainty in the background prediction. Each source of uncertainty is described by a single nuisance parameter, and correlations between background processes and selections are taken into account. The fit does not significantly affect either the uncertainty or the

central value of these nuisance parameters. The systematic uncertainties considered in the fit are described in Section 6.

The level of agreement of the predictions for the main backgrounds is compared with data in dedicated validation regions (VRs), which are not used to constrain the background normalisation or nuisance parameters in the fit.

5.1 Fake/non-prompt lepton background

The background from fake/non-prompt (FNP) leptons is estimated from data with a method similar to that described in Refs. [72, 73]. Two types of lepton identification criteria are defined for this evaluation: "tight" and "loose", corresponding to the signal and candidate electrons and muons described in Section 4. The method relates the number of events containing prompt or FNP leptons to the number of observed events with tight or loose-not-tight leptons using the probability for loose prompt or FNP leptons to satisfy the tight criteria. The probability for prompt loose leptons to satisfy the tight selection is determined from $t\bar{t}Z$ MC simulation applying correction factors obtained by comparing $Z \to \ell^+\ell^-$ ($\ell = e, \mu$) events in data and MC simulation. The equivalent probability for loose FNP leptons to pass the tight selection is measured in a data sample enhanced in $t\bar{t}$ using events with one electron and one muon with the same charge plus at least one b-tagged jet.

The estimates for the FNP background are validated in dedicated regions with selection criteria similar to those defining the SRs but with reduced contributions from processes with three prompt leptons. Two VRs are defined as detailed in Table 4, with $\mathrm{VR_{1F}}$ probing the lepton p_{T} regime in $\mathrm{SR_{1A}}$ and $\mathrm{SR_{1B}}$, while $\mathrm{VR_{2F}}$ includes soft lepton requirements as in $\mathrm{SR_{2A}}$ and $\mathrm{SR_{2B}}$. These VRs veto events with SF-OS dilepton pairs and require to have at least one different-flavour opposite-sign (DF-OS) dilepton pair. The observed and expected yields in these VRs are shown in Table 5, with a purity of FNP leptons above 55% and with good agreement between data and the background estimates.

Table 4: Definition of the validation regions used for the FNP lepton estimation (see text for details).

Requirement / Region	$\mathrm{VR_{1F}}$	$\mathrm{VR_{2F}}$
	No SF-OS pair	
	At least one DF-OS pair	
Third leading lepton p_{T} [GeV]	> 20	< 60
n_{jets} ($p_{\mathrm{T}} > 30$ GeV)	≥ 3	≥ 3
$n_{b-\mathrm{tagged\,jets}}$ ($p_{\mathrm{T}} > 30$ GeV)	≥ 1	–
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 50	> 150

5.2 $t\bar{t}Z$ and multi-boson background

The two dedicated control regions used for the $t\bar{t}Z$ ($\mathrm{CR}_{t\bar{t}Z}$) and multi-boson (CR_{VV}) background estimation are defined in Table 6. To ensure orthogonality with the SRs, only events with $50 < E_{\mathrm{T}}^{\mathrm{miss}} < 100$ GeV

Table 5: Background fit results for the FNP validation regions. The "Others" category contains the contributions from $t\bar{t}h$, $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}t$, $t\bar{t}t\bar{t}$, Wh, and Zh production. Combined statistical and systematic uncertainties are given. The individual uncertainties can be correlated and do not necessarily add in quadrature to the total systematic uncertainty. The number of $t\bar{t}Z$ and multi-boson background events is estimated as described in Section 5.2.

	$\mathrm{VR}_{1\mathrm{F}}$	$\mathrm{VR}_{2\mathrm{F}}$
Observed events	84	98
Total (post-fit) SM events	98 ± 28	95 ± 33
Post-fit, multi-boson	0.7 ± 0.2	2.8 ± 0.7
Post-fit, $t\bar{t}Z$	8.7 ± 1.6	2.9 ± 0.6
Fake/non-prompt leptons	54 ± 27	76 ± 32
tZ, tWZ	0.9 ± 0.5	0.40 ± 0.21
Others	33 ± 5	13.4 ± 2.6

are included in $\mathrm{CR}_{t\bar{t}Z}$, while a b-tagged jet veto and $50 < E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV requirements are applied in CR_{VV}.

To validate the background estimates and provide a statistically independent cross-check of the extrapolation towards the SRs, three validation regions are defined, as show in Table 6. The $\mathrm{VR}_{t\bar{t}Z}$ region primarily validates the $t\bar{t}Z$ background estimate. The $\mathrm{VR}_{VV}^{\text{n-jet}}$ and $\mathrm{VR}_{VV}^{b\text{-tag}}$ regions validate the multi-boson background estimate, the former focusing on the extrapolation in jet multiplicity and the latter on the extrapolation from events without to events with b-tagged jets. The overlap between these multi-boson VRs is around 50%.

Table 7 shows the observed and expected yields in the CRs and VRs for each background source, and Figure 2 shows the jet multiplicity distribution after the background fit for these CRs and VRs, respectively. The normalisation factors for the $t\bar{t}Z$ and multi-boson backgrounds do not differ from unity by more than 20% and the post-fit MC-simulated jet multiplicity distributions agree well with the data.

Table 6: Definition of the control and validation regions used for the $t\bar{t}Z$ and multi-boson background estimation.

Requirement / Region	$\mathrm{CR}_{t\bar{t}Z}$	$\mathrm{VR}_{t\bar{t}Z}$	CR_{VV}	$\mathrm{VR}_{VV}^{\text{n-jet}}$	$\mathrm{VR}_{VV}^{b\text{-tag}}$
Second lepton p_{T} [GeV]	> 20	> 20	> 20	> 40	> 40
Third lepton p_{T} [GeV]	> 20	> 20	> 20	> 40	> 40
$n_{\mathrm{jets}}(p_{\mathrm{T}} > 30$ GeV)	≥ 4	≥ 3	≥ 3	≥ 3	3
$n_{b-\text{tagged jets}}$ ($p_{\mathrm{T}} > 30$ GeV)	≥ 1	≥ 1	0	0	≥ 0
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50 – 100	100 – 150	50 – 200	200 – 300	200 – 300

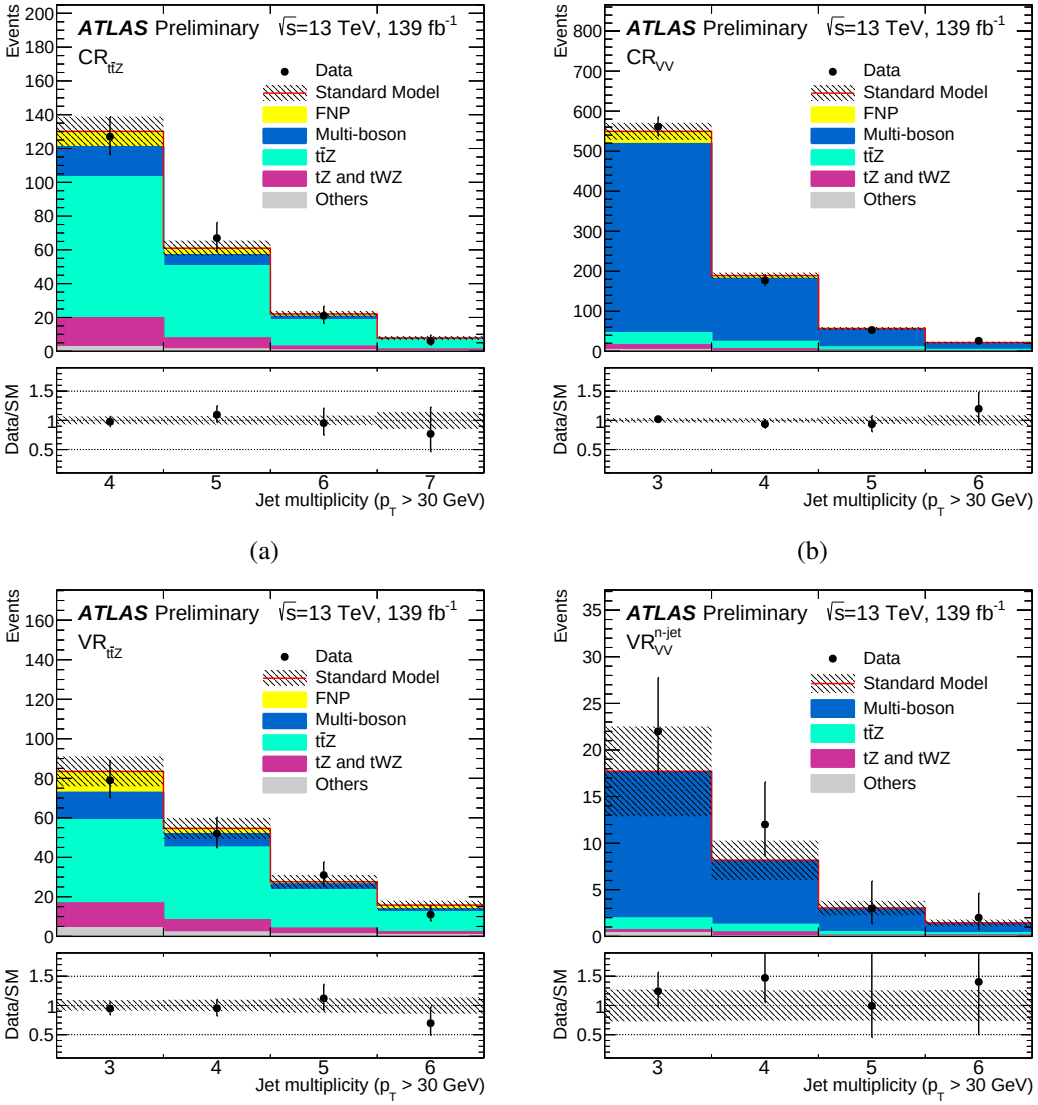

Figure 2: Jet multiplicity distributions in control regions (a) $\mathrm{CR}_{t\bar{t}Z}$, (b) CR_{VV}, (c) $\mathrm{VR}_{t\bar{t}Z}$ and (d) $\mathrm{VR}_{VV}^{\text{n-jet}}$ after normalising the $t\bar{t}Z$ and multi-boson background processes via the simultaneous fit described in Section 5. The contributions from all SM backgrounds are shown as a histogram stack; the bands represent the total uncertainty in the background prediction. The "Others" category contains the contributions from $t\bar{t}h$, $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}t$, $t\bar{t}t\bar{t}$, Wh, and Zh production. The "FNP" category represents the background from fake or non-prompt leptons. The last bin in each figure contains the overflow. The lower panels show the ratio of the observed data to the total SM background prediction, with bands representing the total uncertainty in the background prediction.

Table 7: Background fit results for the control and validation regions for the $t\bar{t}Z$ and multi-boson backgrounds. The nominal predictions from MC simulation are given for comparison for those backgrounds ($t\bar{t}Z$, multi-boson) that are normalised to data. The "Others" category contains the contributions from $t\bar{t}h$, $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}t$, $t\bar{t}t\bar{t}$, Wh, and Zh production. Combined statistical and systematic uncertainties are given. The individual uncertainties can be correlated and do not necessarily add in quadrature to the total systematic uncertainty. The number of events with fake/non-prompt leptons is estimated with the data-driven technique described in Section 5.1.

	$CR_{t\bar{t}Z}$	$VR_{t\bar{t}Z}$	CR_{VV}	$VR_{VV}^{\text{n-jet}}$	$VR_{VV}^{b\text{-tag}}$
Observed events	221	173	817	39	34
Total (post-fit) SM events	221 ± 15	182 ± 16	817 ± 29	30 ± 8	27 ± 6
Post-fit, multi-boson	27 ± 7	25 ± 8	693 ± 34	26 ± 8	18 ± 6
Post-fit, $t\bar{t}Z$	148 ± 22	109 ± 19	59 ± 11	2.7 ± 0.5	6.1 ± 1.2
Fake/non-prompt leptons	15.1 ± 1.7	16.7 ± 1.9	41 ± 15	<1.5	$0.9^{+1.1}_{-0.9}$
tZ, tWZ	27 ± 14	23 ± 12	19 ± 10	0.9 ± 0.5	1.7 ± 0.9
Others	4.6 ± 0.8	8.0 ± 1.5	5.5 ± 2.0	0.42 ± 0.08	0.59 ± 0.11
Pre-fit, multi-boson	32.1 ± 3.1	30 ± 7	830 ± 190	31 ± 6	21 ± 4
Pre-fit, $t\bar{t}Z$	153 ± 13	112.2 ± 2.5	61.3 ± 2.7	2.82 ± 0.20	6.3 ± 0.4

6 Systematic uncertainties

The main sources of systematic uncertainty affecting the analysis SRs are related to the background theoretical and modelling uncertainties, the limited statistics in the CRs and MC simulated samples, the uncertainties on the FNP probabilities, as well as the jet energy scale and resolution. The effects of the systematic uncertainties are evaluated for all signal samples and background processes. Since the normalisation of the dominant background processes is extracted in dedicated CRs, the systematic uncertainties only affect the extrapolation to the SRs in these cases. Table 8 summarises the contributions from the different sources of systematic uncertainty to the total SM background predictions in the signal regions.

The jet energy scale and resolution uncertainties are derived as a function of the p_{T} and η of the jet, as well as of the pile-up conditions and the jet flavour composition (more quark-like or gluon-like) of the selected jet sample. They are determined using a combination of data and simulated event samples, through measurements of the jet response asymmetry in dijet, Z+jet and γ+jet events [60, 74].

Systematic uncertainties on the b-tagging efficiency are estimated by varying the η-, p_{T}- and flavour-dependent scale factors applied to each jet in the simulation within a range that reflects the systematic uncertainty in the measured tagging efficiency and mis-tag rates in data [63, 65].

Other detector-related systematic uncertainties, such as those related to the $E_{\mathrm{T}}^{\mathrm{miss}}$ modelling, as well as lepton reconstruction efficiency, energy scale and energy resolution are found to have a small impact on the results.

Table 8: Summary of the main systematic uncertainties and their impact (in %) on the total SM background prediction in each of the signal regions studied. The total systematic uncertainty can be different from the sum in quadrature of individual sources due to the correlations between them resulting from the fit to the data. The quoted theoretical uncertainties include modelling and cross-section uncertainties.

	SR_{1A}	SR_{1B}	SR_{2A}	SR_{2B}
Total systematic uncertainty (%)	13	13	29	15
Diboson theoretical uncertainties (%)	2	3	11	5
$t\bar{t}Z$ theoretical uncertainties (%)	3	6	4	5
Other theoretical uncertainties (%)	6	9	2	9
MC and FNP statistical uncertainties (%)	6	<1	14	7
Diboson fitted normalisation (%)	2	3	11	6
$t\bar{t}Z$ fitted normalisation (%)	5	9	2	7
Fake/non-prompt leptons efficiency (%)	4	<1	14	2
Jet energy resolution (%)	4	3	2	2
Jet energy scale (%)	1	4	<1	1
b-tagging (%)	3	5	1	5

The diboson background MC modelling uncertainties are estimated by varying the renormalisation, factorisation and resummation scales used to generate the samples [75]. For $t\bar{t}Z$, the predictions from the MG5_aMC@NLO and Sherpa event generators are compared, and the uncertainties related to the choice

of renormalisation and factorisation scales are assessed by varying the corresponding event generator parameters up and down by a factor of two around their nominal values [76].

The cross-sections used to normalise the MC samples are varied according to the uncertainty in the cross-section calculation, i.e. 6% for diboson, 12% for $t\bar{t}Z$ and 13% for $t\bar{t}W$ production [31]. For $t\bar{t}WW$, tZ, tWZ, $t\bar{t}h$, Wh, Zh, $t\bar{t}t$, $t\bar{t}t\bar{t}$, and triboson production processes, which constitute a small background, a 50% uncertainty in the event yields is assumed.

Systematic uncertainties are assigned to the estimated background from fake/non-prompt leptons to account for potentially different compositions (heavy flavour, light flavour or conversions) between the signal and control regions, as well as the contamination from prompt leptons in the regions used to measure the probabilities for loose fake/non-prompt leptons to pass the tight criteria.

7 Results

The observed number of events and expected yields are shown in Table 9 for each of the four inclusive SRs, and Figure 3 show kinematic distributions after applying all the SR selection requirements except those on $E_{\mathrm{T}}^{\mathrm{miss}}$ or $p_{\mathrm{T}}^{\ell\ell}$. Data agree with the SM background prediction and these results are interpreted as exclusion limits for several beyond-the-SM (BSM) scenarios.

The HistFitter framework, which utilises a profile-likelihood-ratio-test statistic [77], is used to estimate 95% confidence intervals using the $\mathrm{CL_s}$ prescription [78]. The likelihood is built from the product of a probability density function describing the observed number of events in the SR and the associated CRs. The statistical uncertainties in the CRs and SRs are modelled using Poisson distributions. Systematic uncertainties enter the likelihood as nuisance parameters which are constrained by Gaussian distributions whose widths correspond to the sizes of these uncertainties. Table 9 also shows upper limits (at the 95% CL) on the visible BSM cross-section $\sigma_{\mathrm{vis}} = S_{\mathrm{obs}}^{95}/\mathcal{L}dt$, defined as the product of the production cross-section, acceptance and efficiency.

Model-dependent limits are also set in specific classes of SUSY models. For each signal hypothesis, the background fit is redone taking into account the signal contamination in the CRs, which is found to be below 12% for signal models close to the existing exclusion limits [20]. Correlations of the uncertainties between the SM backgrounds and the signals are taken into account. To enhance the exclusion power in the considered SUSY models, some of the SRs are split into several bins in different kinematic variables, as detailed below. The observed number of events and expected yields in all these bins is shown in Figure 4.

Figure 5 shows the exclusion limits in the $\tilde{t}_1 \to t\tilde{\chi}_2^0$ with $\tilde{\chi}_2^0 \to h/Z\tilde{\chi}_1^0$ simplified model. These results are obtained from a relaxed version of $\mathrm{SR_{1A}}$ which is split in four $E_{\mathrm{T}}^{\mathrm{miss}}$ bins: $200 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$ GeV, $250 < E_{\mathrm{T}}^{\mathrm{miss}} < 300$ GeV, $300 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$ GeV, and $E_{\mathrm{T}}^{\mathrm{miss}} > 350$ GeVand a version of $\mathrm{SR_{1B}}$ which is split across four bins in $p_{\mathrm{T}}^{\ell\ell}$: $150 < p_{\mathrm{T}}^{\ell\ell} < 300$ GeV, $300 < p_{\mathrm{T}}^{\ell\ell} < 450$ GeV, $450 < p_{\mathrm{T}}^{\ell\ell} < 600$ GeV, and $p_{\mathrm{T}}^{\ell\ell} > 600$ GeV. For each combination of sparticle masses, only the SR (taken as the statistical combination of the composing bins) with best expected sensitivity is considered for the final limit setting. For $\tilde{\chi}_2^0$ masses above 250 GeV, $\tilde{t}_1$ masses up to about 1050 GeV are excluded at 95% CL, while $\tilde{t}_1$ masses below 1140 GeV are excluded for a $\tilde{\chi}_2^0$ mass of 850 GeV. These results improve the existing limits on the $\tilde{t}_1$ mass in this model by approximately 200 GeV [20].

Figure 6 shows the exclusion limits in the $\tilde{t}_2 \to Z\tilde{t}_1$ with $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ simplified model. These results are obtained from the statistical combination of $\mathrm{SR_{2A}}$ and a relaxed version of $\mathrm{SR_{2B}}$ that is split across

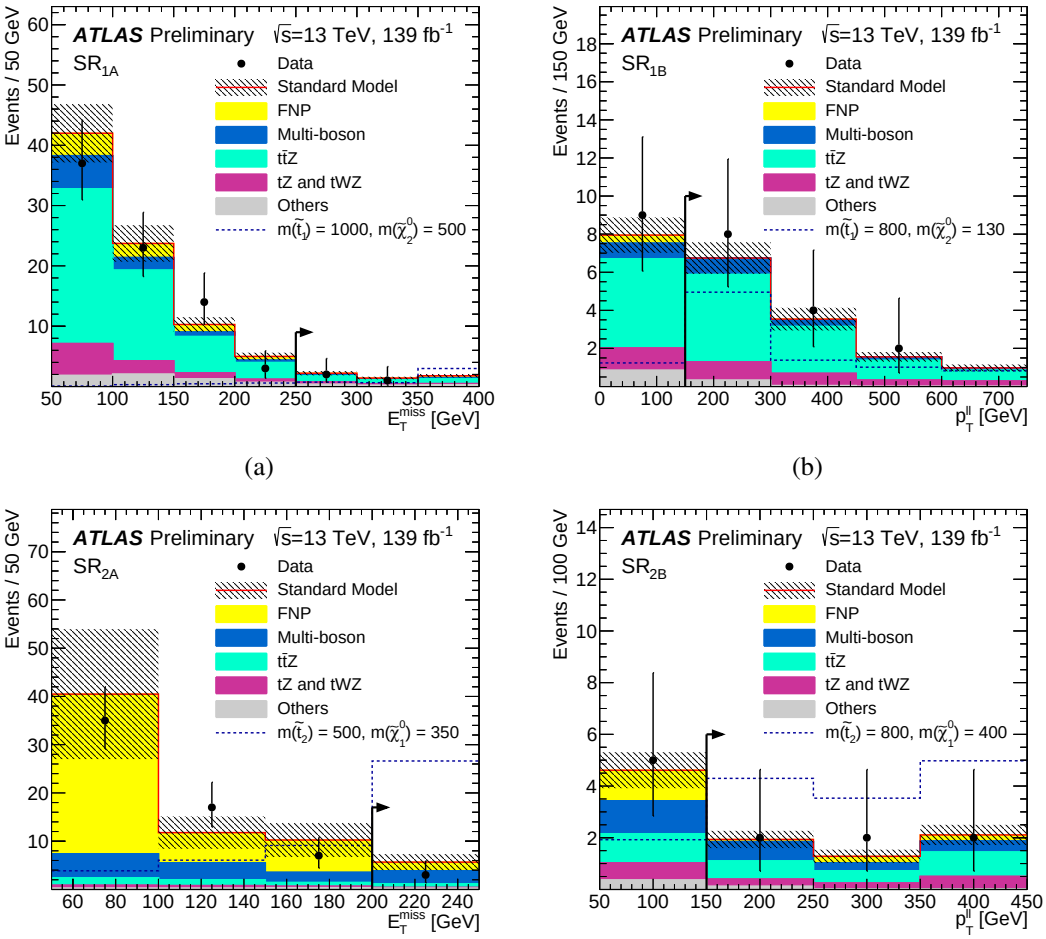

Figure 3: Distributions of (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ in $\mathrm{SR_{1A}}$, (b) $p_{\mathrm{T}}^{\ell\ell}$ in $\mathrm{SR_{1B}}$, (c) $E_{\mathrm{T}}^{\mathrm{miss}}$ in $\mathrm{SR_{2A}}$, and (d) $p_{\mathrm{T}}^{\ell\ell}$ in $\mathrm{SR_{2B}}$ for events passing all the SR requirements except those on the variable being plotted (which are indicated by the arrows). The contributions from all SM backgrounds are shown after the background fit described in Section 5; the hashed bands represent the total uncertainty. The "Others" category contains the contributions from $t\bar{t}h$, $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}t$, $t\bar{t}t\bar{t}$, Wh, and Zh production. The "FNP" category represents the background from fake or non-prompt leptons. The expected distributions for selected signal models are also shown as dashed lines. The last bin in each figure contains the overflow.

bins in both $E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\ell\ell}$: $300 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$ GeV, and $E_{\mathrm{T}}^{\mathrm{miss}} > 350$ GeV; $50 < p_{\mathrm{T}}^{\ell\ell} < 150$ GeV, and $p_{\mathrm{T}}^{\ell\ell} > 150$ GeV. The shape of the contour is driven by $\mathrm{SR_{2A}}$ and $\mathrm{SR_{2B}}$ being most sensitive to respectively small and large mass splittings between the $\tilde{t}_2$ and the $\tilde{\chi}_1^0$. Masses of the $\tilde{t}_2$ up to 875 GeV are excluded at 95% CL for a $\tilde{\chi}_1^0$ mass of about 350 GeV and $\tilde{\chi}_1^0$ masses of approximately 520 (450) GeV are excluded for $\tilde{t}_2$ masses of 650 (800) GeV, extending the previous limits on the $\tilde{\chi}_1^0$ mass from Ref. [80] by up to 160 GeV.

Table 9: Observed and expected numbers of events in the four signal regions. The pre-fit predictions from MC simulation are given for comparison for those backgrounds ($t\bar{t}Z$, multi-boson) that are normalised to data in dedicated control regions. The "Others" category contains the contributions from $t\bar{t}h$, $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}t$, $t\bar{t}t\bar{t}$, Wh, and Zh production. Combined statistical and systematic uncertainties are given. The table also includes model-independent 95% CL upper limits on the visible number of BSM events (S^{95}_{obs}), the number of BSM events given the expected number of background events (S^{95}_{exp}) and the visible BSM cross-section (σ_{vis}), as well as the discovery p-value (p_0) for the background-only hypothesis, all calculated with pseudo-experiments. The value of p_0 is capped at 0.5 if the number of observed events is below the number of expected events.

	$\text{SR}_{1\text{A}}$	$\text{SR}_{1\text{B}}$	$\text{SR}_{2\text{A}}$	$\text{SR}_{2\text{B}}$
Observed events	3	14	3	6
Total (post-fit) SM events	5.4 ± 0.7	12.8 ± 1.6	5.7 ± 1.7	5.4 ± 0.8
Post-fit, multi-boson	0.50 ± 0.22	1.5 ± 0.5	2.7 ± 1.0	1.5 ± 0.6
Post-fit, $t\bar{t}Z$	2.5 ± 0.5	8.7 ± 1.6	0.73 ± 0.29	2.1 ± 0.5
Fake or non-prompt leptons	0.74 ± 0.24	0.04 ± 0.02	1.8 ± 1.1	0.65 ± 0.11
tZ, tWZ	0.9 ± 0.4	2.2 ± 1.2	0.19 ± 0.11	1.0 ± 0.5
Others	0.78 ± 0.17	0.37 ± 0.08	0.21 ± 0.06	0.16 ± 0.03
Pre-fit, multi-boson	0.59 ± 0.18	1.8 ± 0.5	3.2 ± 0.9	1.8 ± 0.4
Pre-fit, $t\bar{t}Z$	2.6 ± 0.4	8.9 ± 1.0	0.76 ± 0.27	2.2 ± 0.4
S^{95}_{obs}	4.6	10.9	4.9	7.0
S^{95}_{exp}	$6.1^{+2.5}_{-1.5}$	$9.4^{+3.3}_{-1.9}$	$6.2^{+2.5}_{-1.7}$	$6.5^{+2.5}_{-1.8}$
σ_{vis} [fb]	0.03	0.08	0.03	0.05
p_0	0.5	0.37	0.50	0.38

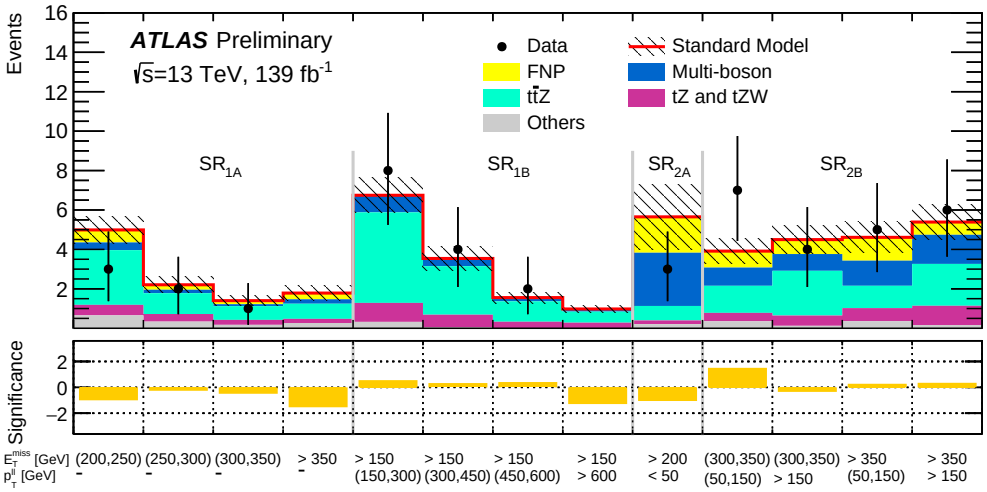

Figure 4: Comparison of the observed and expected event yields in all the SRs and bins used for the model-dependent exclusion limits. The "Others" category contains the contributions from $t\bar{t}h$, $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}t$, $t\bar{t}t\bar{t}$, Wh, and Zh production. The "FNP" category represents the background from fake or non-prompt leptons. The lower panel shows the significance in each SR bin, computed as described in Ref. [79].

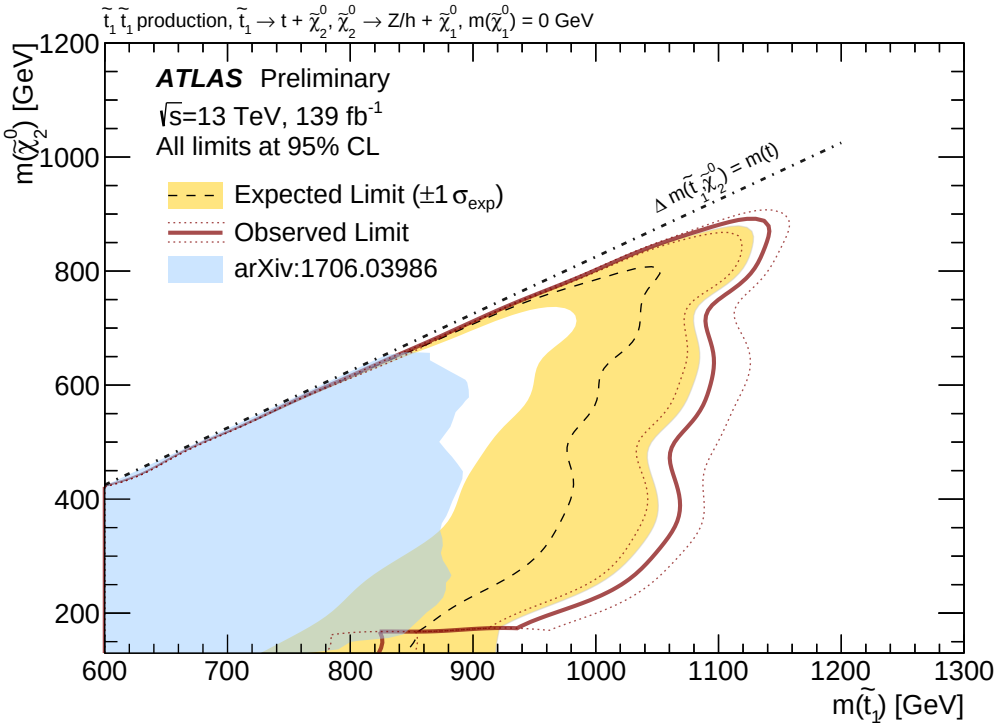

Figure 5: Exclusion limits at 95% CL on the masses of the $\tilde{t}_1$ and $\tilde{\chi}_2^0$, for a fixed $m(\tilde{\chi}_1^0) = 0$ GeV, assuming $\mathcal{B}(\tilde{\chi}_2^0 \to Z\tilde{\chi}_1^0) = 0.5$ and $\mathcal{B}(\tilde{\chi}_2^0 \to h\tilde{\chi}_1^0) = 0.5$. The dashed line and the shaded band are the expected limit and its $\pm 1\sigma$ uncertainty, respectively. The thick solid line is the observed limit for the central value of the signal cross-section. The expected and observed limits do not include the effect of the theoretical uncertainties in the signal cross-section. The dotted lines show the effect on the observed limit when varying the signal cross-section by $\pm 1\sigma$ of the theoretical uncertainty. Results are compared with the observed limits obtained by the previous ATLAS search in Ref. [20].

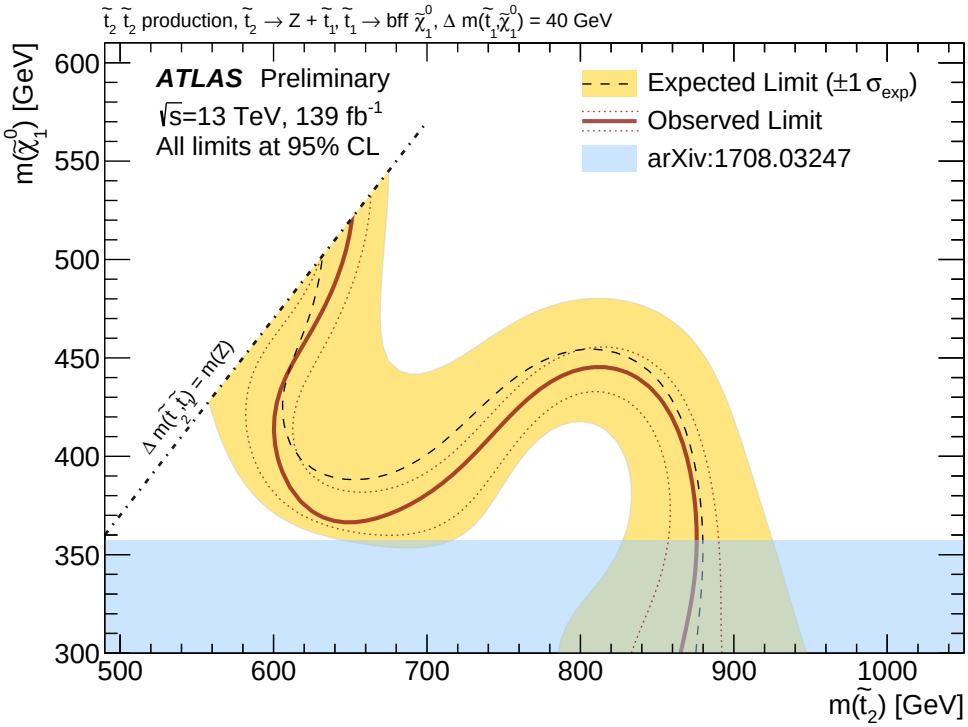

Figure 6: Exclusion limits at 95% CL on the masses of the $\tilde{t}_2$ and $\tilde{\chi}_1^0$, for a fixed $m(\tilde{t}_1) - m(\tilde{\chi}_1^0) = 40$ GeV and assuming $\mathcal{B}(\tilde{t}_2 \to Z\tilde{t}_1) = 1$. The dashed line and the shaded band are the expected limit and its $\pm 1\sigma$ uncertainty, respectively. The thick solid line is the observed limit for the central value of the signal cross-section. The expected and observed limits do not include the effect of the theoretical uncertainties in the signal cross-section. The dotted lines show the effect on the observed limit when varying the signal cross-section by $\pm 1\sigma$ of the theoretical uncertainty. Results are compared with the observed limits obtained by the previous ATLAS search in Ref. [80].

8 Conclusion

A search for direct top squark pair production in events with a leptonically decaying Z boson is presented in this note. The analysis uses 139 fb^{-1} of proton–proton collision data at $\sqrt{s}$ = 13 TeV recorded by ATLAS. No excess over the SM background predictions is observed, and exclusion limits are presented on a selection of simplified models. Limits exclude at 95% confidence level $\tilde{t}_1$ masses up to 1140 GeV in models featuring $\tilde{t}_1$ production and $\tilde{t}_1 \to t\tilde{\chi}_2^0$ with $\tilde{t}_2 \to Z/h\tilde{\chi}_1^0$ decays, and $\tilde{t}_2$ masses up to 875 GeV in models featuring $\tilde{t}_2$ production and $\tilde{t}_2 \to Z\tilde{t}_1$ with $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ decays. Compared to previous limits, these results extend the mass parameter space exclusion by up to 200 GeV in $\tilde{t}_1$ mass and by up to 160 GeV in $\tilde{\chi}_1^0$ mass in the considered $\tilde{t}_2$ model.

References

[1] Yu. A. Golfand and E. P. Likhtman, *Extension of the Algebra of Poincare Group Generators and Violation of p Invariance*, JETP Lett. **13** (1971) 323, [Pisma Zh. Eksp. Teor. Fiz. **13** (1971) 452].

[2] D. V. Volkov and V. P. Akulov, *Is the Neutrino a Goldstone Particle?*, Phys. Lett. B **46** (1973) 109.

[3] J. Wess and B. Zumino, *Supergauge Transformations in Four-Dimensions*, Nucl. Phys. B **70** (1974) 39.

[4] J. Wess and B. Zumino, *Supergauge Invariant Extension of Quantum Electrodynamics*, Nucl. Phys. B **78** (1974) 1.

[5] S. Ferrara and B. Zumino, *Supergauge Invariant Yang-Mills Theories*, Nucl. Phys. B **79** (1974) 413.

[6] A. Salam and J. A. Strathdee, *Supersymmetry and Nonabelian Gauges*, Phys. Lett. B **51** (1974) 353.

[7] G. R. Farrar and P. Fayet, *Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry*, Phys. Lett. B **76** (1978) 575.

[8] N. Sakai, *Naturalness in Supersymmetric Guts*, Z. Phys. C **11** (1981) 153.

[9] S. Dimopoulos, S. Raby and F. Wilczek, *Supersymmetry and the Scale of Unification*, Phys. Rev. D **24** (1981) 1681.

[10] L. E. Ibanez and G. G. Ross, *Low-Energy Predictions in Supersymmetric Grand Unified Theories*, Phys. Lett. B **105** (1981) 439.

[11] S. Dimopoulos and H. Georgi, *Softly Broken Supersymmetry and SU(5)*, Nucl. Phys. B **193** (1981) 150.

[12] R. Barbieri and G. F. Giudice, *Upper Bounds on Supersymmetric Particle Masses*, Nucl. Phys. B **306** (1988) 63.

[13] B. de Carlos and J. A. Casas, *One loop analysis of the electroweak breaking in supersymmetric models and the fine tuning problem*, Phys. Lett. B **309** (1993) 320, arXiv: hep-ph/9303291.

[14] K. Inoue, A. Kakuto, H. Komatsu and S. Takeshita, *Aspects of Grand Unified Models with Softly Broken Supersymmetry*, Prog. Theor. Phys. **68** (1982) 927, Erratum: Prog. Theor. Phys. **70** (1983) 330.

[15] J. R. Ellis and S. Rudaz, *Search for Supersymmetry in Toponium Decays*, Phys. Lett. B **128** (1983) 248.

[16] J. Alwall, M.-P. Le, M. Lisanti and J. G. Wacker, *Searching for Directly Decaying Gluinos at the Tevatron*, Phys. Lett. B **666** (2008) 34, arXiv: 0803.0019 [hep-ph].

[17] J. Alwall, P. Schuster and N. Toro, *Simplified Models for a First Characterization of New Physics at the LHC*, Phys. Rev. D **79** (2009) 075020, arXiv: 0810.3921 [hep-ph].

[18] D. Alves et al., *Simplified Models for LHC New Physics Searches*, J. Phys. G **39** (2012) 105005, arXiv: 1105.2838 [hep-ph].

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

[20] ATLAS Collaboration, *Search for direct top squark pair production in events with a Higgs or Z boson, and missing transverse momentum in* $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, JHEP **08** (2017) 006, arXiv: 1706.03986 [hep-ex].

[21] ATLAS Collaboration, *Search for direct top squark pair production in events with a Z boson, b-jets and missing transverse momentum in $\sqrt{s}$ = 8 TeV pp collisions with the ATLAS detector*, Eur. Phys. J. C **74** (2014) 2883, arXiv: 1403.5222 [hep-ex].

[22] CMS Collaboration, *Search for top-squark pairs decaying into Higgs or Z bosons in pp collisions at $\sqrt{s}$ = 8 TeV*, Phys. Lett. B **736** (2014) 371, arXiv: 1405.3886 [hep-ex].

[23] CMS Collaboration, *Search for top-squark pairs decaying into Higgs or Z bosons in pp collisions at $\sqrt{s}$=8 TeV*, Phys. Lett. B **736** (2014) 371, arXiv: 1405.3886 [hep-ex].

[24] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653, arXiv: 1608.03953 [hep-ex].

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

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

[27] ATLAS Collaboration, *Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2*, ATL-PHYS-PUB-2016-003, 2016, URL: https://cds.cern.ch/record/2120133.

[28] ATLAS Collaboration, *Multi-boson simulation for 13 TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: https://cds.cern.ch/record/2119986.

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

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

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

[32] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph].

[33] W. Beenakker, R. Hopker, M. Spira and P. Zerwas, *Squark and gluino production at hadron colliders*, Nucl. Phys. B **492** (1997) 51, arXiv: hep-ph/9610490.

[34] A. Kulesza and L. Motyka, *Threshold resummation for squark-antisquark and gluino-pair production at the LHC*, Phys. Rev. Lett. **102** (2009) 111802, arXiv: 0807.2405 [hep-ph].

[35] A. Kulesza and L. Motyka, *Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC*, Phys. Rev. D **80** (2009) 095004, arXiv: 0905.4749 [hep-ph].

[36] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen et al., *Soft-gluon resummation for squark and gluino hadroproduction*, JHEP **12** (2009) 041, arXiv: 0909.4418 [hep-ph].

[37] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen et al., *Squark and gluino hadroproduction*, Int. J. Mod. Phys. A **26** (2011) 2637, arXiv: 1105.1110 [hep-ph].

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

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

[40] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph].

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

[42] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[43] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, (2012), arXiv: 1201.3084 [hep-ph].

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

[45] L. Lönnblad and S. Prestel, *Matching tree-level matrix elements with interleaved showers*, JHEP **03** (2012) 019, arXiv: 1109.4829 [hep-ph].

[46] W. Beenakker, C. Borschensky, M. Krämer, A. Kulesza and E. Laenen, *NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation*, JHEP **12** (2016) 133, arXiv: 1607.07741 [hep-ph].

[47] W. Beenakker, M. Kramer, T. Plehn, M. Spira and P. Zerwas, *Stop production at hadron colliders*, Nucl. Phys. B **515** (1998) 3, arXiv: hep-ph/9710451 [hep-ph].

[48] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen et al., *Supersymmetric top and bottom squark production at hadron colliders*, JHEP **08** (2010) 098, arXiv: 1006.4771 [hep-ph].

[49] W. Beenakker et al., *NNLL resummation for stop pair-production at the LHC*, JHEP **05** (2016) 153, arXiv: 1601.02954 [hep-ph].

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

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

[52] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, arXiv: 0901.0002 [hep-ph].

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

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

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

[56] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, CERN-EP-2018-273 (2019), arXiv: 1902.04655 [hep-ex].

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

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

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

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

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

[62] ATLAS Collaboration, *Selection of jets produced in* 13 *TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015, URL: https://cds.cern.ch/record/2037702.

[63] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex].

[64] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: https://cds.cern.ch/record/2273281.

[65] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018, URL: https://cds.cern.ch/record/2306649.

[66] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector in the first proton–proton collisions at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-027, 2015, URL: https://cds.cern.ch/record/2037904.

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

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

[69] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, Phys. Lett. B **463** (1999) 99, arXiv: hep-ph/9906349 [hep-ph].

[70] A. Barr, C. G Lester and P. Stephens, *A variable for measuring masses at hadron colliders when missing energy is expected; m T2: The truth behind the glamour*, J. Phys. G **29** (2003) 2343.

[71] M. Baak et al., *HistFitter software framework for statistical data analysis*, Eur. Phys. J. C **75** (2015) 153, arXiv: 1410.1280 [hep-ex].

[72] ATLAS Collaboration, *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s}$ = 7 TeV*, Eur. Phys. J. C **71** (2011) 1577, arXiv: 1012.1792 [hep-ex].

[73] ATLAS Collaboration, *Measurement of the top quark pair production cross section in pp collisions at $\sqrt{s}$ = 7 TeV in dilepton final states with ATLAS*, Phys. Lett. B **707** (2012) 459, arXiv: 1108.3699 [hep-ex].

[74] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-015, 2015, URL: https://cds.cern.ch/record/2037613.

[75] ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: https://cds.cern.ch/record/2119986.

[76] ATLAS Collaboration, *Modelling of the $t\bar{t}H$ and $t\bar{t}V$ ($V = W, Z$) processes for $\sqrt{s}$ = 13 TeV ATLAS analyses*, ATL-PHYS-PUB-2015-022, 2016, URL: https://cds.cern.ch/record/2120826.

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

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

[79] R. D. Cousins, J. T. Linnemann, J. Tucker, *Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process*, Nucl. Instrum. Meth. A **595** (2008) 480, arXiv: physics/0702156 [physics].

[80] ATLAS Collaboration, *Search for direct top squark pair production in final states with two leptons in $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 898, arXiv: 1708.03247 [hep-ex].

Appendix

Tables 10–13 show cut flow tables for the four SRs defined in the text. Signal points from the $\tilde{t}_1 \to t\tilde{\chi}_2^0$ with $\tilde{\chi}_2^0 \to h/Z\tilde{\chi}_1^0$ simplified model are used for SR_{1A} (Table 10) and SR_{1B} (Table 11), while signal points from the $\tilde{t}_2 \to Z\tilde{t}_1$ with $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ simplified model are used for SR_{2A} (Table 12) and SR_{2B} (Table 13).

Table 10: Cut flow for the $\tilde{t}_1 \to t\tilde{\chi}_2^0$ with $\tilde{\chi}_2^0 \to h/Z\tilde{\chi}_1^0$ simplified model signal point with $m(\tilde{t}_1) = 1000$ GeV and $m(\tilde{\chi}_2^0) = 500$ GeV in SR_{1A}. The number of events is normalised to the cross-section and to an integrated luminosity of 139 fb^{-1}.

Selection name	Events ($\mathcal{L}$=139 fb^{-1})
Total	950
Trigger ≥ 3 signal leptons n_{jets} ($p_{\text{T}} > 30$ GeV) ≥ 3 $E_{\text{T}}^{\text{miss}} > 50$ GeV Leading lepton $p_{\text{T}} > 40$ GeV Subleading lepton $p_{\text{T}} > 20$ GeV	15.1
Third leading lepton $p_{\text{T}} > 20$ GeV	12.3
$\lvert m_{\ell\ell}^{\text{SF-OS}} - m_Z \rvert < 15$ GeV	8.88
$n_{b-\text{tagged jets}}$ ($p_{\text{T}} > 30$ GeV) ≥ 1	8.43
n_{jets} ($p_{\text{T}} > 30$ GeV) ≥ 4	7.86
$E_{\text{T}}^{\text{miss}} > 250$ GeV	5.40
$m_{\text{T2}}^{3\ell} > 100$ GeV	4.29

Table 11: Cut flow for the $\tilde{t}_1 \to t\tilde{\chi}_2^0$ with $\tilde{\chi}_2^0 \to h/Z\tilde{\chi}_1^0$ simplified model signal point with $m(\tilde{t}_1)$ = 850 GeV and $m(\tilde{\chi}_2^0)$ = 130 GeV in $\mathrm{SR_{1B}}$. The number of events is normalised to the cross-section and to an integrated luminosity of 139 fb^{-1}.

Selection name	Events ($\mathcal{L}$=139 fb^{-1})
Total	3000
Trigger ≥ 3 signal leptons n_{jets} ($p_{\text{T}} > 30$ GeV) ≥ 3 $E_{\text{T}}^{\text{miss}} > 50$ GeV Leading lepton $p_{\text{T}} > 40$ GeV Subleading lepton $p_{\text{T}} > 20$ GeV	45.1
Third leading lepton $p_{\text{T}} > 20$ GeV	38.1
$\lvert m_{\ell\ell}^{\text{SF-OS}} - m_Z \rvert < 15$ GeV	27.1
$n_{b-\text{tagged jets}}$ ($p_{\text{T}} > 30$ GeV) ≥ 1	25.1
n_{jets} ($p_{\text{T}} > 30$ GeV) ≥ 5	18.1
$E_{\text{T}}^{\text{miss}} > 150$ GeV	11.7
$p_{\text{T}}^{\ell\ell} > 150$ GeV	9.64
Leading b-tagged jet $p_{\text{T}} > 100$ GeV	8.09

Table 12: Cut flow for the $\tilde{t}_2 \to Z\tilde{t}_1$ with $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ simplified model signal point with $m(\tilde{t}_1)$ = 600 GeV and $m(\tilde{\chi}_1^0)$ = 450 GeV in $\mathrm{SR_{2A}}$. The number of events is normalised to the cross-section and to an integrated luminosity of 139 fb^{-1}.

Selection name	Events ($\mathcal{L}$=139 fb^{-1})
Total	28500
Trigger ≥ 3 signal leptons n_{jets} ($p_{\text{T}} > 30$ GeV) ≥ 3 $E_{\text{T}}^{\text{miss}} > 50$ GeV Leading lepton $p_{\text{T}} > 40$ GeV Subleading lepton $p_{\text{T}} > 20$ GeV	130
Third leading lepton $p_{\text{T}} < 20$ GeV	85.0
$\lvert m_{\ell\ell}^{\text{SF-OS}} - m_Z \rvert < 15$ GeV	80.3
Leading jet $p_{\text{T}} > 150$ GeV	34.4
$E_{\text{T}}^{\text{miss}} > 200$ GeV	21.4
$p_{\text{T}}^{\ell\ell} < 50$ GeV	5.56

Table 13: Cut flow for the $\tilde{t}_2 \to Z\tilde{t}_1$ with $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ simplified model signal point with $m(\tilde{t}_1)$ = 800 GeV and $m(\tilde{\chi}_1^0)$ = 400 GeV in SR_{2B}. The number of events is normalised to the cross-section and to an integrated luminosity of 139 fb^{-1}.

Selection name	Events ($\mathcal{L}$=139 fb^{-1})
Total	4531
Trigger ≥ 3 signal leptons n_{jets} ($p_{\text{T}} > 30$ GeV) ≥ 3 $E_{\text{T}}^{\text{miss}} > 50$ GeV Leading lepton $p_{\text{T}} > 40$ GeV Subleading lepton $p_{\text{T}} > 20$ GeV	90.7
Third leading lepton $p_{\text{T}} < 60$ GeV	82.7
$\lvert m_{\ell\ell}^{\text{SF-OS}} - m_Z \rvert < 15$ GeV	78.5
$n_{b-\text{tagged jets}}$ ($p_{\text{T}} > 30$ GeV) ≥ 1	34.2
$E_{\text{T}}^{\text{miss}} > 350$ GeV	14.7
$p_{\text{T}}^{\ell\ell} > 150$ GeV	12.7