



Search for third-generation vector-like leptons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

A search for vector-like leptons in multilepton (two, three, or four-or-more electrons plus muons) final states with zero or more hadronic τ -lepton decays is presented. The search is performed using a dataset of proton–proton collisions at a centre-of-mass energy of 13 TeV recorded by the ATLAS detector at the LHC, corresponding to an integrated luminosity of 139 fb^{-1} . To maximize the separation of signal and background, a machine-learning classifier is used. No excess of events is observed beyond the Standard Model expectation. Using a doublet vector-like lepton model, vector-like leptons coupling to third-generation Standard Model leptons are excluded in the mass range from 130 GeV to 900 GeV at the 95% confidence level, while the highest excluded mass is expected to be 970 GeV.

Contents

1	Introduction	2
2	ATLAS detector	3
3	Data and simulation samples	4
4	Object reconstruction	5
5	Analysis strategy	8
6	Background estimation	11
7	Systematic uncertainties	15
8	Results	16
9	Conclusions	21

1 Introduction

The predictions of the Standard Model (SM) of particle physics are in excellent agreement with measurements of proton–proton (pp) collisions at the Large Hadron Collider (LHC) [1]. Nonetheless, some aspects of our universe cannot be explained within the framework of the SM, such as the excess of matter over antimatter, the origin of the neutrino masses, the nature of dark matter and dark energy, and the hierarchy and fine-tuning problems. Many possible ways to find solutions have been proposed, including models based on supersymmetry, which help to explain why the Higgs boson’s mass is very far from the Planck scale. However, the measured Higgs boson branching fractions are in good agreement with the SM predictions, so it is not easy to accommodate new particles whose masses are generated via the Higgs mechanism. One class of particles that are motivated by a variety of phenomenological models based on string theory or large extra dimensions are vector-like fermions that transform as non-chiral representations of the unbroken SM gauge group [2]. They therefore have Dirac masses and decouple from the electroweak scale in the large-mass limit.

A large number of searches for vector-like quarks have been performed at the Tevatron [3, 4] and the LHC [5–22]. Following the suggestion in Refs. [23, 24], a search for vector-like leptons (VLL) in a doublet VLL model has been performed by ATLAS and is presented in this article. A similar search was performed by the CMS Collaboration using 138 fb^{-1} of pp collisions at $\sqrt{s} = 13 \text{ TeV}$ and excludes vector-like τ -lepton masses below 1045 GeV at the 95% confidence level (CL) [25]. The VLL doublet $L' = (\nu_{\tau}', \tau')$ comprises two fermions of approximately equal mass that couple only to the third-generation leptons. The VLL production cross section is dominated by $pp \rightarrow \nu_{\tau}' \tau'$, which is approximately 3.7 times greater than either the $pp \rightarrow \tau'^+ \tau'^-$ and $pp \rightarrow \nu_{\tau}' \bar{\nu}_{\tau}'$ modes, which have approximately equal cross sections. The ν_{τ}' decays exclusively into $W^+ \tau^-$, while the τ' decays are $\tau'^- \rightarrow Z \tau^-$ and $\tau'^- \rightarrow H \tau^-$ where the branching fraction of the former is larger but asymptotically approaches that of the latter with increasing τ' mass ($M_{\tau'}$) due to the Goldstone equivalence principle [26]. Examples of leading-order (LO) Feynman diagrams for VLL production and decay are displayed in Figure 1. Given the possible decays, the search is performed by

selecting events containing at least two charged light leptons, e^\pm or μ^\pm , zero or more τ -leptons decaying hadronically, and a momentum imbalance transverse to the beam. To achieve better background rejection than is possible with an event selection based on kinematic and topological variables, a boosted decision tree (BDT) algorithm is utilized as an event classifier [27, 28].

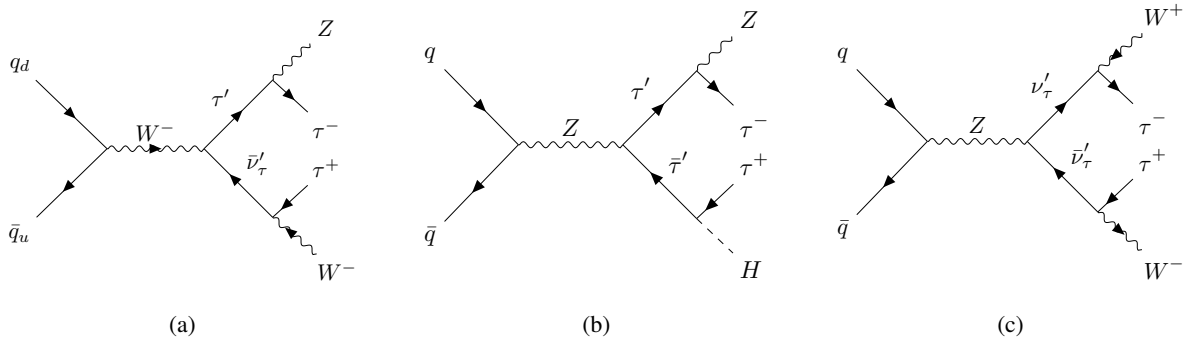


Figure 1: Examples of LO Feynman diagrams for VLL production and decay: (a) $\tau'v'_\tau$ production followed by τ' decay into $Z\tau$ and v'_τ decay into τ^+W^- , where q_u (q_d) represents a weak isospin $+\frac{1}{2}$ ($-\frac{1}{2}$) quark, (b) production of $\tau'\bar{\tau}'$ followed by the decay into $Z\tau^-$ and τ^+H , and (c) $v'_\tau\bar{v}'_\tau$ production followed by v'_τ decay into $W\tau$.

This article is organized as follows. A brief description of the ATLAS detector is given in Section 2. Section 3 presents the data and simulation samples used in this search. The reconstruction of objects used in the search for a VLL signal is delineated in Section 4. Section 5 describes techniques used to perform the event selection, while Section 6 outlines the method used to estimate the backgrounds. A discussion of the systematic uncertainties is given in Section 7. The statistical method used to arrive at the 95% CL upper limit on the VLL production cross section, and hence the mass exclusion region, is described in Section 8. Finally, the analysis and results are summarized in Section 9.

2 ATLAS detector

The ATLAS detector [29] at the LHC covers nearly the entire solid angle around the collision point.¹ It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadron calorimeters, and a muon spectrometer incorporating three large superconducting air-core toroidal magnets.

The inner-detector system (ID) is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the vertex region and provides up to four measurements per track, the first hit normally being in the insertable B-layer (IBL) installed before Run 2 [30, 31]. It is followed by the silicon microstrip tracker (SCT), which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$. The TRT also provides

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 center 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 $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadron calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadron endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimized for electromagnetic and hadronic energy measurements, respectively.

The muon spectrometer (MS) comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by the superconducting air-core toroidal magnets. The field integral of the toroids ranges between 2.0 and 6.0 Tm across most of the detector. A set of precision chambers covers the region $|\eta| < 2.7$ with three layers of monitored drift tubes, complemented by cathode-strip chambers in the forward region, where the particle flux is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

Interesting events are selected by the first-level trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger [32]. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the high-level trigger further reduces in order to record events to disk at about 1 kHz.

An extensive software suite [33] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

3 Data and simulation samples

The data sample used in this article corresponds to an integrated luminosity of 139 fb^{-1} of pp collisions at a centre-of-mass energy (\sqrt{s}) of 13 TeV collected by the ATLAS detector during the 2015–2018 data-taking periods after requiring stable LHC beams and that all detector subsystems were operational [34]. The primary luminosity measurement was performed using the LUCID-2 detector [35]. Candidate events are required to satisfy the dilepton triggers (ee , $\mu\mu$, $e\mu$) [32, 36, 37]. These triggers have looser identification and isolation requirements than the single-lepton triggers but have comparable signal efficiency for events satisfying the analysis selection. The lowest p_{T} threshold for the leading lepton ranged from 12 GeV to 22 GeV and for the sub-leading lepton it ranged from 8 GeV to 17 GeV.

To evaluate the effects of the detector resolution and acceptance on the signal and background, and to estimate the SM backgrounds, simulated event samples were produced using dedicated event generators. The detector response to the final-state particles was then modelled using a GEANT4-based Monte Carlo (MC) detector simulation [38, 39]. The simulated data must account for the fact that significantly more than one inelastic pp collision occurs per bunch crossing, with the average number ranging from 13 to 38 for the 2015–2018 data-taking periods, respectively. Inelastic collisions were simulated using PYTHIA 8.186 [40] with the A3 set of tuned parameters [41] and the NNPDF2.3LO [42] set of parton distribution functions (PDFs), and overlaid on the signal and background MC samples. These simulated events were reweighted to match the conditions of the collision data, specifically the number of additional pp interactions (pile-up).

In order to optimize the signal selection and to estimate its acceptance and efficiency, a simulated VLL sample was generated using the model described in Refs. [23, 24]. This model was implemented at LO, using MADGRAPH5_AMC@NLO 2.6.2 [43] with the NNPDF2.3_{LO} PDF to generate the parton-level process, and selecting events with two, three, or four-or-more light leptons and ≥ 0 hadronically decaying τ -leptons, with all leptons having $p_T > 18$ GeV and $|\eta| < 2.8$. The generated τ' mass points are 130 GeV and 200 GeV to 1300 GeV in steps of 100 GeV. These samples were then reweighted to next-to-leading order (NLO) in QCD. To perform the fragmentation and hadronization, these events were processed using PYTHIA 8.212 with the A14 set of tuned parameters [41] and the NNPDF2.3_{LO} PDFs.

Since each VLL can decay into a vector boson or a Higgs boson and a τ -lepton, the selection requirements are ≥ 2 light leptons and ≥ 0 τ -leptons decaying to hadrons (τ_{had}). The leptonic decays of τ -leptons are not explicitly considered in the analysis. Top quark and multi-vector-boson events can contribute to the background since they have similar final states. Therefore, to estimate the SM backgrounds, samples of simulated events were generated, containing multiple final-state leptons that result from decays of either directly produced vector bosons or vector bosons from top quark or H decays. Backgrounds from misidentified leptons are estimated with data-driven methods, as discussed in Section 6.

The vector-boson samples were produced using either SHERPA 2.2.1 or SHERPA 2.2.2 [44] with the NNPDF3.0_{NNLO} [45] set of PDFs, and the set of tuned parton-shower parameters developed by the authors of SHERPA [46]. The V +jets samples, where V is either a W or Z boson, and the semileptonic diboson samples were produced using SHERPA 2.2.1, while the fully leptonic diboson samples and the triboson samples were produced using SHERPA 2.2.2. The matrix element (ME) calculations were matched and merged with the SHERPA parton shower using the MEPS@NLO prescription [47–50] based on the Catani–Seymour dipole factorization procedure [51]. For the V +jets samples, the calculation was performed at NLO in QCD for up to two partons and at LO for up to four partons using the COMIX [52] and OPENLOOPS [53–55] libraries. For the diboson sample, this was performed for one parton at NLO and up to three partons at LO, while, for the triboson sample, the calculation was performed at NLO for the inclusive sample and at LO for up to two partons.

The electroweak production of single-top events, and also $t\bar{t}$ and $t\bar{t}+H$ production, was simulated by POWHEG BOX v2 [56–58] at NLO with the h_{damp} parameter² set to $1.5 m_{\text{top}}$ [59], the NNPDF3.0_{NLO} [45] PDF set, and the top-quark mass set to 172.5 GeV. The diagram removal scheme [60] was employed to handle interference between single-top $t+W$ associated production and $t\bar{t}$ production [59]. Partons were hadronized and showered by PYTHIA 8.230 [61], using the A14 tune and the NNPDF2.3_{LO} PDF set.

The $t\bar{t}+V$, $t+Z$, $t+WZ$, ttt , $t\bar{t}\bar{t}$ and $t\bar{t}+WW$ events were generated by MADGRAPH5_AMC@NLO 2.3.3 [43] at NLO, using the NNPDF3.0_{NLO} PDF set for $t\bar{t}+V$, $t+Z$, $t+WZ$, ttt and $t\bar{t}+WW$, and the NNPDF3.1_{NLO} PDF set for $t\bar{t}\bar{t}$. Parton hadronization and showering was performed by PYTHIA 8.210 using the A14 tune and the NNPDF2.3_{LO} PDF set. A summary of simulated signal and background samples is provided in Table 1.

4 Object reconstruction

All events used in this analysis are required to contain a primary vertex [62]. It is selected as the pp collision vertex candidate with the highest sum of the squared transverse momenta of all associated tracks

² The h_{damp} parameter is a resummation damping factor and one of the parameters that controls the matching of POWHEG matrix elements to the parton shower. It effectively regulates the high- p_T radiation against which the $t\bar{t}$ system recoils.

Table 1: Simulated signal and background samples used in this analysis. The physics process is listed, along with the ME generator and order of the ME calculation, the parton hadronization and showering model, and the PDF used to simulate the process.

Process	ME Generator	ME Order	Showering Model	PDF
Signal	MADGRAPH5_AMC@NLO	LO	PYTHIA 8	NNPDF2.3LO
$Z \rightarrow \ell\ell$	SHERPA 2.2.1	NNLO	SHERPA	NNPDF3.0NNLO
$W \rightarrow \ell\ell$	SHERPA 2.2.1	NNLO	SHERPA	NNPDF3.0NNLO
VV (lep)	SHERPA 2.2.2	NLO	SHERPA	NNPDF3.0NNLO
VV (semi-lep)	SHERPA 2.2.1	NLO	SHERPA	NNPDF3.0NNLO
VVV	SHERPA 2.2.2	NLO	SHERPA	NNPDF3.0NNLO
t	POWHEG BOX v2	NLO	PYTHIA 8	NNPDF2.3LO
$t\bar{t}$	POWHEG BOX v2	NLO	PYTHIA 8	NNPDF2.3LO
$t\bar{t} + WW$	MADGRAPH5_AMC@NLO	NLO	PYTHIA 8	NNPDF2.3LO
$t\bar{t} + H$	POWHEG BOX v2	NLO	PYTHIA 8	NNPDF2.3LO
$t\bar{t} + V$	MADGRAPH5_AMC@NLO	NLO	PYTHIA 8	NNPDF2.3LO
$t\bar{t}t$	MADGRAPH5_AMC@NLO	NLO	PYTHIA 8	NNPDF2.3LO
$t\bar{t}\bar{t}$	MADGRAPH5_AMC@NLO	NLO	PYTHIA 8	NNPDF2.3LO
$t + Z$	MADGRAPH5_AMC@NLO	NLO	PYTHIA 8	NNPDF2.3LO
$t + WZ$	MADGRAPH5_AMC@NLO	NLO	PYTHIA 8	NNPDF2.3LO

with $p_T > 500$ MeV. In addition, there must be at least two tracks associated with that vertex.

Electron candidates are reconstructed from energy clusters in the EM calorimeter that match a reconstructed track [63]. They are required to have $p_T > 30$ GeV in order to suppress the number of fake electrons. In addition, they are required to be within $|\eta| < 2.47$ with the region $1.37 < |\eta| < 1.52$ being excluded since it contains a significant amount of non-sensitive material in front of the calorimeter. To match these objects to the primary vertex, the track's transverse and longitudinal impact parameters (d_0 and z_0 , respectively) are required to satisfy $|d_0|/\sigma(d_0) < 5$ and $|z_0 \sin \theta| < 0.5$ mm. Furthermore, these candidates must satisfy object identification criteria by applying one of several working points [63]. By applying different quality criteria, each working point offers a trade-off between identification efficiency and misidentification rate. A likelihood-based discriminant is constructed from a set of variables that enhance electron selection, while rejecting photon conversions and hadrons misidentified as electrons [63]. An η - and E_T -dependent selection is applied to the likelihood discriminant. For this search, the *Tight* likelihood working point is used, which has a 75% efficiency at $E_T = 30$ GeV, increasing to 88% at $E_T = 100$ GeV [63], when used to identify electrons from Z -boson decays. Electrons are also required to be isolated using criteria based on ID tracks and topological clusters in the calorimeter; the *Loose* isolation working point is applied and has an efficiency of approximately 99% [63]. Correction factors are applied to simulated electrons to take into account the small differences in reconstruction, identification and isolation efficiencies between data and MC simulation.

Muon candidates are reconstructed by combining a reconstructed track from the inner detector with one from the muon spectrometer [64], with the requirement that $p_T > 20$ GeV and $|\eta| < 2.5$. In addition, the transverse and longitudinal impact parameters of the track are required to satisfy $|d_0|/\sigma(d_0) < 3$ and $|z_0 \sin \theta| < 0.5$ mm. To reject misidentified muon candidates, primarily originating from pion and kaon decays, several quality requirements are imposed on the muon candidate. The *Medium* working

point is used to select muons with $p_T < 300$ GeV and the *HighPt* working point is used for those with $p_T > 300$ GeV [64]. This choice ensures a 97% efficiency at $p_T = 30$ GeV for the *Medium* working point, and 76% at $p_T = 500$ GeV for the *HighPt* working point [64]. An isolation requirement based on ID tracks and topological clusters in the calorimeter is imposed. The *TightTrackOnly* isolation working point is used, resulting in an efficiency between 94% and 99% for muons from W -boson decays in simulated $t\bar{t}$ events [64]. Similarly to electrons, correction factors are applied to simulated muons to account for the small differences between data and simulation.

Particle-flow (PFlow) jets within $|\eta| < 4.5$ are reconstructed using the anti- k_t algorithm [65, 66] with a radius parameter $R = 0.4$ [67], using neutral PFlow constituents and charged constituents associated with the primary vertex as input [68]. These jets are then calibrated to the particle level by applying a jet energy scale derived from simulation [69]. Furthermore, in situ corrections based on the collected data are applied [69]. A cleaning procedure is used to identify and remove jets arising from calorimeter noise or non-collision backgrounds. To suppress jets arising from pile-up, a discriminant called the ‘jet vertex tagger’ (JVT) is constructed using a two-dimensional likelihood method [70]. The *Medium* JVT working point is used, which has an average efficiency of 95%. Jets used in this analysis are required to have $p_T > 20$ GeV and $|\eta| < 2.5$. Jets originating from b -quarks are identified using the DL1r b -tagging algorithm [71, 72]. The working point used corresponds to a b -tagging efficiency of 77% [71, 72], measured in a sample of simulated $t\bar{t}$ events. The corresponding rejection factors are approximately 130, 4.9 and 14 for light-quark and gluon jets, c -jets, and hadronically decaying τ -leptons, respectively. Correction factors are applied to the simulated jets to take account of the small differences in reconstruction and identification efficiencies, and the energy scale and resolution differences, between data and MC simulation.

The event selection for this analysis considers only those τ -leptons that decay into final states containing a ν_τ -neutrino and hadrons, denoted by τ_{had} . Since the ν_τ escapes the detector volume undetected, only the hadronic decay products consisting of one or three charged hadrons and up to two neutral pions are visible and they are denoted by $\tau_{\text{had-vis}}$. Reconstruction of $\tau_{\text{had-vis}}$ [73, 74] is seeded by jets reconstructed from topological clusters by the anti- k_t algorithm with the radius parameter set to $R = 0.4$. The tracks associated with the jet are required to originate from the primary vertex by satisfying the impact parameter requirements $|d_0| < 1$ mm and $|z_0 \sin \theta| < 1.5$ mm. If these requirements are satisfied, the tracks are then required to be within a cone of size $\Delta R = 0.2$ around the jet axis, surrounded by a conical isolation region covering $0.2 < \Delta R < 0.4$, in order to be considered a $\tau_{\text{had-vis}}$ candidate. The direction of the $\tau_{\text{had-vis}}$ candidate in (η, ϕ) is calculated as the vector sum of the topological clusters within $\Delta R = 0.2$ of the jet axis, using the $\tau_{\text{had-vis}}$ vertex as the origin. A multivariate discriminant is used to select tracks that were produced by charged τ_{had} decay products [73, 74]. Reconstructed $\tau_{\text{had-vis}}$ objects are selected for the analysis if they have exactly one or three associated tracks (1- or 3-prong) with a total charge equal to ± 1 . The $\tau_{\text{had-vis}}$ objects must also satisfy the requirements $p_T > 20$ GeV and $|\eta| < 2.47$, excluding the region $1.37 < |\eta| < 1.52$. These requirements have an efficiency of about 85% for 1-prong and 70% for 3-prong $\tau_{\text{had-vis}}$ objects [73, 74] as estimated from simulated $Z/\gamma^* \rightarrow \tau^+\tau^-$ events. A multivariate regression technique trained on MC samples is used to determine the $\tau_{\text{had-vis}}$ energy scale using information from associated tracks, calorimeter energy clusters, and reconstructed neutral pions [73, 74].

A recurrent neural network (RNN) classifier [75] is employed to select τ_{had} -initiated jets and reject those initiated by quarks or gluons. The RNN is trained on simulated $Z/\gamma^* \rightarrow \tau^+\tau^-$ (for signal) and simulated dijet events (for background). The training variables are single-track variables, and reconstructed kinematic and topological variables. This analysis uses the *Medium* working point with an efficiency of 75% (60%) for 1-prong (3-prong) candidates and a background rejection factor of 35 (240). A boosted decision tree (eBDT) [73, 74] is used to reject electron backgrounds misidentified as 1-prong $\tau_{\text{had-vis}}$ objects.

Variables used for its training include information from the calorimeter, the tracking detector, and the visible momentum measured from the reconstructed tracks. The *Tight* working point with an efficiency of 75% is used. To assess the contribution from misidentified leptons (Section 6), less stringent (*Loose*) object identification requirements are applied. When selecting *Loose* $\tau_{\text{had-vis}}$ there is an additional requirement for the RNN score to be at least 0.01. In addition, a dedicated muon-veto criterion is used to reject muons reconstructed as $\tau_{\text{had-vis}}$. Correction factors are applied to simulated τ_{had} objects to take into account the small differences in reconstruction and identification efficiencies between data and MC simulation. The energy scale and resolution differences between data and MC simulation are also accounted for by applying scale factors. A summary of the lepton object definitions is presented in Table 2.

Table 2: Object definitions used in the analysis for leptons. *Tight* objects are those selected for the nominal analysis and pass the prescribed selection requirements. *Loose* objects are those used to assess the contribution from the fake-lepton background and fail one or more selection requirements.

	Electrons		Muons		$\tau_{\text{had-vis}}$
	<i>Tight</i> likelihood	<i>Loose</i> isolation	<i>Medium (HighPt)</i>	<i>TightTrackOnly</i> isolation	<i>Medium</i> RNN
<i>Tight</i>	Pass	Pass	Pass	Pass	Pass
<i>Loose</i>	Fail*	Fail*	Pass	Fail	Fail

**Loose* electrons are defined as failing either the *Tight* likelihood or the *Loose* isolation requirements, but not both.

The missing transverse momentum (with magnitude E_T^{miss}) [76] is reconstructed as the negative vector sum of the p_T of all the selected electrons, muons, jets, and $\tau_{\text{had-vis}}$. An extra track-based ‘soft term’ is built using additional tracks associated with the primary vertex, but not with any reconstructed object. The use of this track-based soft term is motivated by improved performance in E_T^{miss} in a high pile-up environment.

To avoid cases where the detector response to a single physical object is reconstructed as two separate final-state objects, an overlap removal procedure is used. If electron and muon candidates share a track, the electron candidate is removed. After that, if the $\Delta R_{y,\phi}$ distance³ between a jet and an electron candidate is less than 0.2, the jet is discarded. If multiple jets satisfy this requirement, only the closest jet is removed. For jet–electron distances between 0.2 and 0.4, the electron candidate is removed. If the distance between a jet and a muon candidate is less than 0.4, the muon candidate is removed if the jet has more than two associated tracks; otherwise the jet is removed. The τ_{had} candidates are seeded from jets, so this procedure removes any ambiguity in their selection.

5 Analysis strategy

Based on the production and decay modes of the τ' and ν_{τ}' given in Section 1, the multilepton final states are expected to maximize the signal sensitivity, and hence the multilepton final states are used to search for the VLL signal. To further optimize the signal sensitivity, the different decay modes are targeted by splitting the data into seven training regions (given in Table 3) based on the numbers of light leptons and

³ $\Delta R_{y,\phi}$ is the Lorentz-invariant distance in the rapidity–azimuthal-angle plane, defined as $\Delta R_{y,\phi} = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$ where the rapidity is $y = (1/2)[(E + p_z)/(E - p_z)]$.

τ_{had} , a $E_{\text{T}}^{\text{miss}}$ requirement. In all regions at least one light-flavour jet is required. Additional requirements are derived by maximizing the signal significance through the application of a BDT [27, 28]. This leads to the seven signal regions (SRs) defined in Table 4.

Table 3: The regions used in the training and optimization of the BDT. The training regions are split according to the number of light leptons (N_{ℓ}), and their charge (SS same charge, OS opposite charge) and flavour (SF same flavour, OF opposite flavour), the number of τ_{had} (N_{τ}), and the missing transverse momentum ($E_{\text{T}}^{\text{miss}}$). Every region has a requirement that it contain at least one jet ($N_{\text{jet}} > 0$).

Variables	BDT Training Regions						
BDT	2ℓ SSSF, 1τ	2ℓ SSOF, 1τ	2ℓ OSSF, 1τ	2ℓ OSOF, 1τ	$2\ell, \geq 2\tau$	$3\ell, \geq 1\tau$	$4\ell, \geq 0\tau$
N_{ℓ}	2	2	2	2	2	3	≥ 4
Charge/flavour	SSSF	SSOF	OSSF	OSOF	-	-	-
N_{τ}	1	1	1	1	≥ 2	≥ 1	≥ 0
$E_{\text{T}}^{\text{miss}}$ [GeV]	≥ 120	≥ 90	≥ 60	≥ 100	≥ 60	≥ 90	≥ 60

Table 4: The BDT score requirement imposed on the training regions which are used to define the signal regions. No other requirement differs between the two regions.

Variables	Signal Regions						
BDT	2ℓ SSSF, 1τ	2ℓ SSOF, 1τ	2ℓ OSSF, 1τ	2ℓ OSOF, 1τ	$2\ell, \geq 2\tau$	$3\ell, \geq 1\tau$	$4\ell, \geq 0\tau$
BDT Score	≥ 0.15	≥ 0.1	≥ 0.1	≥ 0.1	≥ -0.11	≥ 0.08	≥ 0.08

In addition to the seven SRs, three control regions (CRs) are used in order to normalize the dominant physics background ($t\bar{t} + Z$, WZ , and ZZ) estimates to data, and a fourth CR is used to assess fake τ_{had} objects originating from gluon-initiated jets and pile-up. These CRs are defined in Table 5. Since events in the WZ CR do not have a τ -lepton, its kinematic variables are set to zero when calculating the BDT score. To confirm that the CRs are modelled correctly and that the obtained background normalization factors are also valid in the regions with different numbers of light leptons and τ -leptons, three validation regions (VRs) are defined, as also shown in Table 5. An additional seven VRs are used to confirm that the BDT models the data correctly. These VRs are selected where the BDT distribution is expected to primarily contain background events, as shown in Figure 5. The SRs, CRs, and VRs are selected such that there is no overlap of events between the regions. The BDT distribution shape in the CRs is shown in Figure 2, which demonstrates good agreement between data and the background simulation.

To classify the events as signal or background, the AdaBoost BDT algorithm [27, 28] is used as implemented in the scikit-learn package [77]. The training and optimization of the BDT are performed in two steps. The first step is the optimization of the BDT hyperparameters and the second step is the optimization of the training through the selection of the variables used for the training.

To optimize the hyperparameters (Maximum Tree Depth, Maximum Features Per Split, Minimum Samples Per Leaf, Minimum Samples Per Split, Number of Estimators, and Learning Rate) the set of 34 kinematic and topological variables listed in Table 6 were used. These variables are chosen as they are expected to provide good separation between the background and signal topologies. To avoid biasing the search by selecting a specific mass, the simulated signal samples described in Section 3 are combined with equal weight for $M_{\tau'}$ set to 800, 900 and 1000 GeV to train the BDT. All of the background samples are considered in the training, including fake light leptons and τ -leptons. A scan over the hyperparameters is

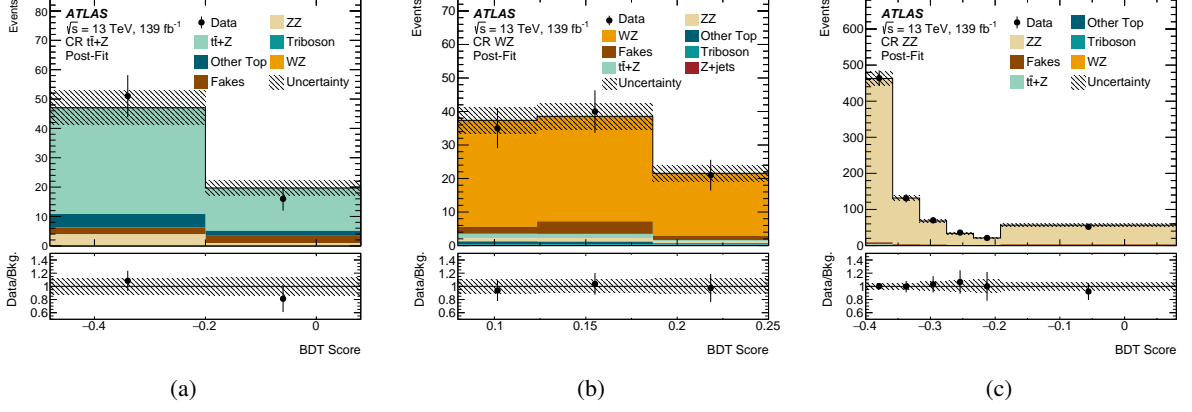


Figure 2: A comparison of the BDT scores for data and the predicted background after performing the fit described in Section 8 in the corresponding CRs used to estimate the three dominant backgrounds: (a) $t\bar{t}+Z$, (b) WZ , and (c) ZZ . The first bin of the $t\bar{t}+Z$ and ZZ distributions contains underflow events, while the last bin of the WZ distribution contains overflow events. The hatched band represents the combined statistical and systematic uncertainties.

Table 5: The definition of the CRs used to determine the normalization of the largest backgrounds, as well as the CR used to assess fake τ -leptons originating from gluon-initiated jets and pile-up. In addition, the VRs used to validate the CRs are also defined. Both the CRs and VRs are selected so that they do not overlap with the SRs or with each other but are similar enough to avoid problems when extrapolating between the three regions. Like the SRs, the CRs and VRs have the requirement that $N_{\text{jet}} > 0$.

	Control Regions				Validation Regions		
	$t\bar{t}+Z$	WZ	ZZ	Fake τ_{had}	$t\bar{t}+Z$	WZ	ZZ
BDT	$4\ell, \geq 0\tau$	$3\ell, \geq 1\tau$	$4\ell, \geq 0\tau$	2ℓ OSSF, 1τ	$3\ell, \geq 1\tau$	2ℓ SSOFF, 1τ	$3\ell, \geq 1\tau$
N_ℓ	≥ 4	3	≥ 4	2	3	2	3
N_τ	≥ 0	0	≥ 0	1	≥ 1	1	≥ 1
N_b	> 0	0	0	-	> 0	0	0
E_T^{miss} [GeV]	≥ 60	≥ 90	< 60	≥ 60	-	-	-
Charge/flavour	-	-	-	OSSF	-	SSOFF	-
BDT score	< 0.08	≥ 0.08	< 0.08	< -0.15	< 0.08	< 0.1	< 0.08

performed using the 5-fold cross-validation procedure to train the BDT, where the simulated data samples are split into five equal randomized samples, with four being used in the training and the fifth used as a testing sample. The training is performed five times so that each of the five samples is used as a testing sample. The set of hyperparameters with the highest receiver operating characteristic (ROC) score is selected.

To select the optimal set of training variables, a BDT is trained in each analysis region defined by lepton multiplicity using the full set of 34 variables. The ranking of the variables is evaluated using the procedure provided by the scikit-learn package. The lowest-ranked training variable is then removed and the BDT is retrained. This procedure is repeated until the ROC score decreases by more than 1%. This leads to each SR having its own unique set of training variables, which are listed in Table 7.

The final set of training variables is then compared with data to confirm that they are well modelled. To avoid any bias in the analysis, only events with a BDT score not satisfying the SR criteria (Table 4) are used in the comparison, since they are background dominated. Through the use of a χ^2 test, a probability of $> 5\%$ is found for agreement between data and simulation for all variables used in the BDT training. Figure 3 shows the distributions of some highly ranked variables in each of the BDT training regions.

6 Background estimation

There are two basic categories of backgrounds to the signal. One category, the irreducible backgrounds, is defined by those processes that yield the same final state as the signal. The other category, the reducible backgrounds, is defined by those that mimic the final state because of misidentified leptons or non-prompt leptons as well as misidentified lepton charge. The irreducible backgrounds are estimated from the simulated samples discussed in Section 3.

Similarly to Ref. [78] the simulated background in the WZ CR is found not to agree with data when examined as a function of the number of jets in the event. Since inverting the BDT score criteria for this CR yielded a similar mismodelling, the data in this region are used to calculate a scale factor to correct the MC simulation to the data in the WZ CR, which then agreed with data.

Charge misidentification for electrons arises from photon conversions or bremsstrahlung and is challenging to describe through detector simulation. Therefore, scale factors are derived and applied to simulated background events to match the charge misidentification probabilities observed in data. The scale factors are derived using a $Z \rightarrow ee$ data sample and are parameterized as a function of p_T and η [63].

The reducible backgrounds from misidentified leptons (electron, muons and τ_{had}) are estimated from data by using the fake-factor (FF) method to derive a transfer function from a background-dominated region to the SR [79]. The transfer function is the ratio of events passing the tight lepton selection to events passing a loose lepton selection in the background-dominated CR. Since the SRs require multiple leptons, the FF is used to estimate the probability of an object being either a misidentified lepton or a non-prompt lepton from an in-flight decay. The FF is calculated as a function of the lepton p_T .

Non-prompt or fake light leptons can originate from decays of bottom or charm hadrons, pion or kaon decays, jets misidentified as electrons, and electrons from photon conversions. To calculate the transfer function for electrons, the background CR requires exactly one electron with a loose selection and no other leptons, $E_T^{\text{miss}} < 40$ GeV, $N_{\text{jet}} \geq 2$, and no reconstructed b -jets. The background CR for muons requires

Table 6: List of the input variables used to train the BDT. The final set is reduced by assessing the impact of removing the lowest-ranked variables on the ROC score for each training region independently.

Variable	Description
E_T^{miss}	The missing transverse momentum in the event
$\mathcal{S}(E_T^{\text{miss}})$	The missing transverse momentum's significance in the event
L_T	The scalar sum of light lepton p_T in the event
$L_T + E_T^{\text{miss}}$	The scalar sum of light lepton p_T and the missing transverse momentum in the event
$L_T + p_T(\tau)$	The scalar sum of light lepton p_T and τ -lepton p_T in the event
$p_T(\ell_1)$	The leading light lepton's p_T in the event
$p_T(\ell_2)$	The sub-leading light lepton's p_T in the event
$p_T(j_1)$	The leading jet's p_T in the event
$p_T(\tau_1)$	The leading τ -lepton's p_T in the event
N_j	The number of jets in the event
N_b	The number of b -jets in the event
H_T	The scalar sum of jet p_T in the event
$L_T + H_T$	The scalar sum of light lepton p_T and jet p_T in the event
$M_{\ell\ell}$	The invariant mass of all light leptons in the event
$M_{\ell\tau}$	The invariant mass of all light leptons and τ -leptons in the event
$M_{\ell j}$	The invariant mass of all light leptons and jets in the event
M_{jj}	The invariant mass of all jets in the event
$M_{j\tau}$	The invariant mass of all jets and τ -leptons in the event
M_T	The transverse mass of the leading light lepton and E_T^{miss} in the event
M_{OSSF}	The invariant mass of the opposite-sign same-flavour light-lepton pair closest to the Z mass in the event
$\Delta\phi(j_1 E_T^{\text{miss}})$	$\Delta\phi$ between E_T^{miss} and the leading p_T jet in the event
$\Delta\phi(\ell_1 E_T^{\text{miss}})$	$\Delta\phi$ between E_T^{miss} and the leading p_T light lepton in the event
$\Delta\phi(\ell_1 \ell_2)$	$\Delta\phi$ between the leading and sub-leading p_T light leptons in the event
$\Delta\phi(\ell_1 j_1)$	$\Delta\phi$ between the leading p_T light lepton and jet in the event
$\Delta\phi(\tau_1 E_T^{\text{miss}})$	$\Delta\phi$ between E_T^{miss} and the leading p_T τ -lepton in the event
$\Delta\phi(\ell_1 \tau_1)$	$\Delta\phi$ between the leading p_T light lepton and τ -lepton in the event
$\Delta\phi(j_1 \tau_1)$	$\Delta\phi$ between the leading p_T jet and τ -lepton in the event
$\Delta R(j_1 E_T^{\text{miss}})$	ΔR between E_T^{miss} and the leading p_T jet in the event
$\Delta R(\ell_1 E_T^{\text{miss}})$	ΔR between E_T^{miss} and the leading p_T light lepton in the event
$\Delta R(\ell_1 \ell_2)$	ΔR between the leading and sub-leading p_T light leptons in the event
$\Delta R(\ell_1 j_1)$	ΔR between the leading p_T light lepton and jet in the event
$\Delta R(\tau_1 E_T^{\text{miss}})$	ΔR between E_T^{miss} and the leading p_T τ -lepton in the event
$\Delta R(\ell_1 \tau_1)$	ΔR between the leading p_T light lepton and τ -lepton in the event
$\Delta R(j_1 \tau_1)$	ΔR between the leading p_T jet and τ -lepton in the event

Table 7: The ranking of the variables used as inputs to BDT algorithm for each of the training regions. Variables are ranked relative to each other by counting the instances a particular variable is used by the BDT nodes in defining the signal–background separation. Variables that do not have a ranking are not included in the corresponding training region. The final variables in each training region are selected using an optimization procedure.

Variable	2ℓ SSSF, 1τ	2ℓ SSOF, 1τ	2ℓ OSSF, 1τ	2ℓ OSOF, 1τ	$2\ell, \geq 2\tau$	$3\ell, \geq 1\tau$	$4\ell, \geq 0\tau$
$p_T(\tau_1)$	1	1	1	1	2	2	
$M_{\ell\tau}$	2	2	5	3	1	1	
$L_T + E_T^{\text{miss}}$	3	3	2	2	23	4	1
E_T^{miss}	4	7	4	21	5	8	5
$\Delta\phi(\tau_1 E_T^{\text{miss}})$	5	6	6	13	3	3	
$\Delta R(\ell_1 \ell_2)$	6	24	7	7	15		17
M_{jj}	7	21	24	15	1	12	19
$M_{\ell j}$	8	11	26	11	27	14	2
$\Delta\phi(\ell_1 E_T^{\text{miss}})$	9	16	20	8	20	10	15
$\Delta R(\ell_1 \tau_1)$	10	8	12	6	16	15	
$\Delta R(j_1 \tau_1)$	11	9	17	25	25	23	
$\Delta R(\ell_1 E_T^{\text{miss}})$	12	29	11	19	17	11	10
$\Delta\phi(\ell_1 \ell_2)$	13	13	18	16	28	13	9
$\Delta R(\tau_1 E_T^{\text{miss}})$	14	27	9	5	12	9	
$p_T(j_1)$	15	19	10	12	22	19	11
M_T	16	23	16	18	8	17	7
$\Delta\phi(j_1 \tau_1)$	17	20	27	29	24		
$M_{\ell\ell}$	18	10	25	20	10	22	4
$p_T(\ell_1)$	19	4			30	5	16
$S(E_T^{\text{miss}})$	20	5	14	24	9	24	8
N_j	21	14	28	23	26		22
$L_T + p_T(\tau)$	22	22		26			
$p_T(\ell_2)$	23	15			18		
$\Delta R(j_1 E_T^{\text{miss}})$	24	18	23	10	31		21
$\Delta\phi(\ell_1 j_1)$	25	17	13	17	13	25	13
N_b	26	26	21	22	29	20	14
L_T	27	32			32		3
$M_{j\tau}$	28	31	15	9	6	18	
$\Delta R(\ell_1 j_1)$	29		8	4	11		18
$L_T + H_T$		12	3	14			
M_{OSSF}			22		7	6	12
$\Delta\phi(\ell_1 \tau_1)$		25	19		19	16	
$\Delta\phi(j_1 E_T^{\text{miss}})$				27	21		6
H_T		28		28	33	21	20

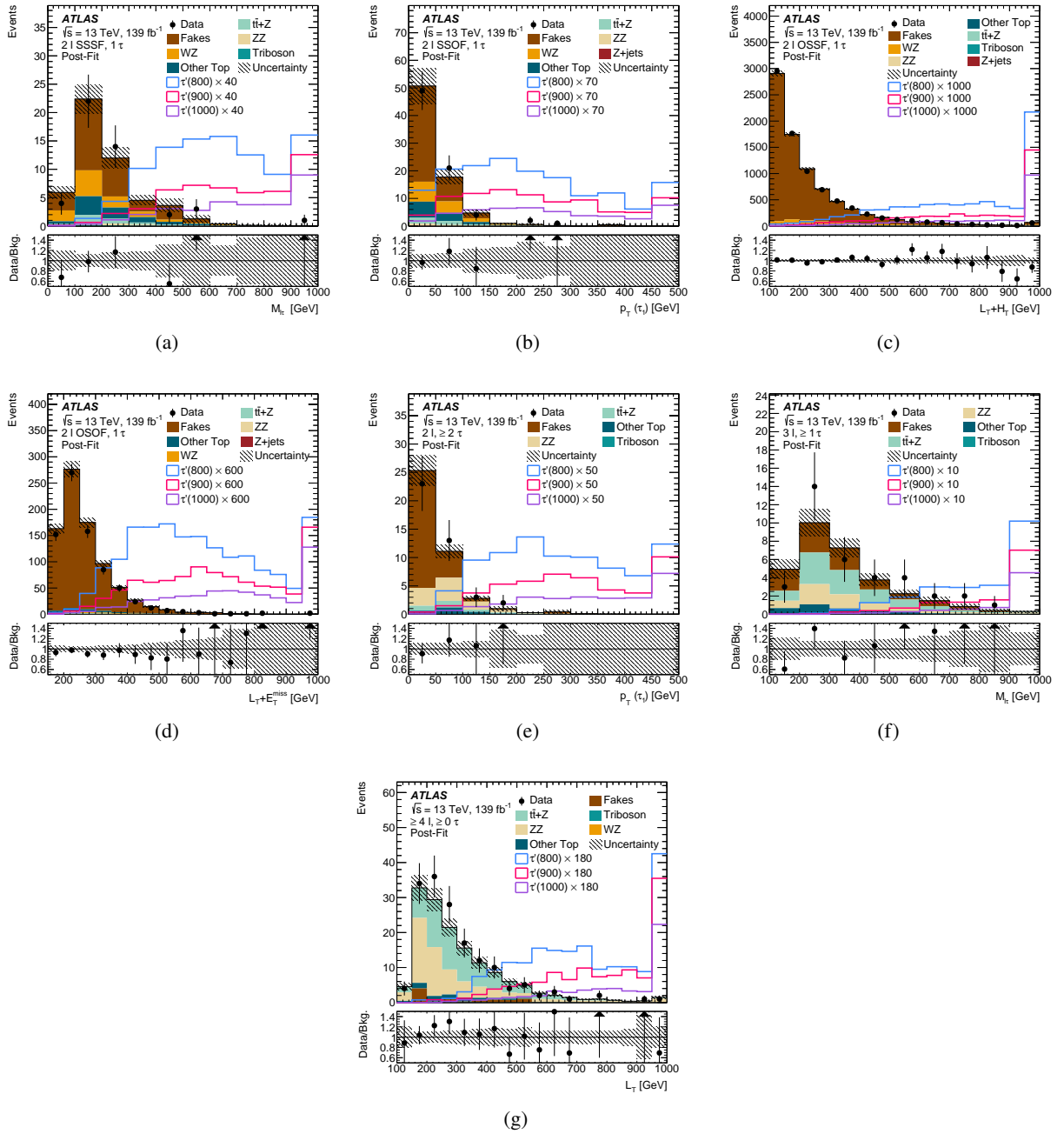


Figure 3: A comparison of the total background and signal distributions for $M_{\tau'} = 800, 900$ and 1000 GeV in variables that are highly ranked in the BDT training. The signal distribution is scaled by the value indicated in the legend. The background prediction is taken after performing the fit described in Section 8, while the signal prediction is taken before the fit. The hatched band represents the combined statistical and systematic uncertainties. The last bin contains overflow events. The arrows in the ratio plot are for points that are outside the range. (a) $M_{\ell\tau}$ for the 2ℓ SSSF, 1τ training region (b) $p_T(\tau_1)$ for the 2ℓ SSOF, 1τ training region (c) $L_T + H_T$ for the 2ℓ OSSF, 1τ training region (d) $L_T + E_T^{\text{miss}}$ for the 2ℓ OSOF, 1τ training region (e) $p_T(\tau_1)$ for the $2 \ell, \geq 2 \tau$ training region (f) $M_{\ell\tau}$ for the $3 \ell, \geq 1 \tau$ training region (g) L_T for the $\geq 4 \ell, \geq 0 \tau$ training region.

exactly one muon with a loose selection and zero electrons and τ -leptons, $E_T^{\text{miss}} < 40$ GeV, and at least two jets with the leading jet $p_T > 35$ GeV.

Since only τ_{had} decays are selected, fake τ -leptons originate from misidentified jets that can arise from several sources: light- and heavy-flavour jets, gluon radiation, and jets from pile-up. The SRs are dominated by light- and heavy-flavour jets misidentified as τ -leptons. Therefore, FFs are calculated independently in dedicated CRs corresponding to light- and heavy-flavour jets. The $Z + \text{jets}$ FF CR uses a $Z + \text{jets}$ sample, which is enriched in light-flavour fake τ -leptons with a requirement of $Z \rightarrow \mu^+ \mu^-$ decay and $|m_{\mu\mu} - m_Z| < 15$ GeV. In addition, the sample is required to have zero reconstructed b -jets and $E_T^{\text{miss}} < 60$ GeV. The $t\bar{t}$ FF CR is a dilepton $t\bar{t}$ sample, enriched in heavy-flavour fake τ -leptons with a requirement of exactly two light leptons that satisfy $|m_{\ell\ell} - m_Z| > 10$ GeV, exactly one τ -lepton that satisfies the loose criteria, and at least two jets with at least one satisfying the b -jet requirements. The FF is calculated separately for 1- and 3-prong τ -leptons. The FFs calculated in the $Z + \text{jets}$ and $t\bar{t}$ CRs are combined by performing a template fit of the BDT score in each training region. The fit uses MC ‘truth’-matched fake τ -leptons to compare the fake composition in the analysis region and the $Z + \text{jets}$ and $t\bar{t}$ CRs.

7 Systematic uncertainties

The systematic uncertainties considered in this analysis come from instrumental and theoretical sources, affecting both the overall event yield and the shape of the distribution. They are evaluated by varying each source around its nominal value as described below.

The uncertainties in the theoretical production cross sections used to simulate the background events are calculated following the same approach as in Ref. [80]. These uncertainties are assigned to every background process whose normalization is not determined by the fit. These backgrounds come from $Z + \text{jets}$, WW , triple vector boson, and top-quark processes.

Uncertainties from missing higher-order contributions are evaluated [81] by applying seven independent variations of the QCD factorization and renormalization scales by factors of one-half and two in the matrix elements after removing combinations that differ by a factor of four. The effect of the uncertainty in the strong coupling constant $\alpha_s(m_Z) = 0.118$ as well as the uncertainties in the nominal PDF set, used in the generation of simulated events, is evaluated by following the PDF4LHC recommendations [82]. In addition, the modelling uncertainty due to the choice of generator for $t\bar{t} + Z$ production is evaluated by comparing samples from the nominal generator MC@NLO + PYTHIA 8 and the alternative generator SHERPA 2.2.1. This gives an uncertainty of up to 7.9% in the $t\bar{t} + Z$ background estimate.

The uncertainty in the full integrated luminosity as obtained from the LUCID-2 detector is 1.7% [35] and is applied to the background and signal processes that are normalized to the theoretical predictions. Uncertainties associated with the pile-up reweighting procedure range from 0.2% to 3.5% [83].

Instrumental uncertainties are evaluated for each object that is considered in the analysis. For the selected leptons, they originate from the reconstruction and identification efficiency, the energy (momentum) scale and resolution, and the isolation efficiency [63, 64, 84]. The uncertainty related to trigger efficiencies is also included. For jets, the uncertainties originate from the jet energy scale and resolution [68, 85], the matching of jets to the primary vertex [83], and the identification of b -jets [71, 72]. Furthermore, the

impact of a possible miscalibration of the soft-track component of the E_T^{miss} is derived from a data–MC comparison of the p_T balance of the hard and soft E_T^{miss} components [86].

Systematic uncertainties associated with the light-lepton fake factors [79] account for the jet composition differences between the region where the fake factors are derived and the analysis regions where they are applied, as well as the differences between the values of fake factors calculated in regions with different selection requirements. Additionally, the uncertainty arising from the modelling and limited sample sizes is taken into account. The uncertainties are treated independently for each of the different sources of misidentified objects.

Uncertainties from several sources are evaluated for the FF derived for τ -leptons. Most notably, to account for the limited numbers of simulated events in the regions used to determine the composition of fake τ -lepton samples, the uncertainty of the fitted fractions from the template fit to the $Z + \text{jets}$ and $t\bar{t}$ FF CRs, is evaluated as the difference between the two values 0% and 100%, for each SR. Furthermore, a systematic uncertainty to account for the gluon-initiated and pile-up fake τ -lepton contribution is applied. To assess this uncertainty a new CR is defined with a less restrictive RNN score requirement on the loose $\tau_{\text{had-vis}}$ in the $Z + \text{jets}$ FF CR, which increases this contribution from roughly 40% to 60% in the CR. The systematic uncertainty in the estimated number of fake τ -leptons is then taken as the difference between applying the nominal FF in each SR and the FF calculated with the less restrictive RNN requirement. The nominal FF is calculated by applying a minimum RNN requirement of 0.01, while for the assessment of its systematic uncertainty a minimum value of 0.005 is used. Since gluon-initiated and pile-up fake τ -leptons typically have low p_T , the τ -lepton FF systematic uncertainties are split into <40 GeV and >40 GeV regions and assessed independently. The gluon-initiated and pile-up contribution is only taken for fake τ -leptons with p_T less than 40 GeV, and the large uncertainty of 28% is constrained in the Fake τ_{had} CR (Table 5), which is a highly populated region dominated by gluon- and pile-up-initiated fake τ -leptons. However, the gluon-initiated and pile-up fake τ -lepton fraction is less than 20% in the SRs, where the BDTs tend to select τ -leptons with p_T greater than 40 GeV, so this uncertainty does not have an impact on them. Additionally, uncertainties are derived for the τ -lepton fake factors by accounting for the numbers of events in the $Z + \text{jets}$ and $t\bar{t}$ FF CRs.

8 Results

In order to test for the presence of a VLL signal, the BDT score templates for signal and background events are fitted to the data using a binned maximum-likelihood (ML) approach in the RooFit and RooStats frameworks [87, 88]. The normalizations of the WZ , ZZ , $t\bar{t} + Z$, and VLL templates are allowed to vary, while the other backgrounds' normalizations are assigned Gaussian constraints based on their respective normalization uncertainties. In addition, the systematic uncertainties are included in the fit as nuisance parameters with correlations across regions and processes taken into account. To quantify the statistical significance of the fit and its resulting power to reject the background-only hypothesis, a test statistic is constructed using the profile likelihood ratio [89].

After performing the simultaneous fit of the seven SRs and four CRs to the data, the fitted normalization factors relative to the theoretical expectations for the main backgrounds are 1.06 ± 0.14 for WZ , 1.02 ± 0.07 for ZZ , and 1.18 ± 0.18 for $t\bar{t} + Z$. Figure 4 shows a comparison between the data and the signal and background yields for the SRs, CRs, and corresponding VRs. Figure 5 shows the templates and data versus BDT score for each of the analysis regions after applying the selection requirements in Table 4 but before the SR's BDT requirement is applied. The regions with BDT score values less than the SR requirement are

used as VRs. In all SRs the number of observed events is compatible with the background hypothesis. Tables 8 and 9 show the total background and signal yields in all SRs and CRs after fitting to data.

Table 8: Total observed yields as computed by the fit for signal regions. The uncertainty contains both the systematic and statistical uncertainties. The prediction for each background sample is taken after a likelihood fit is performed to measure the VLL production cross section. Background normalization factors are also applied. The ‘Other Top’ sample includes contributions from single top, $t\bar{t}$, $t\bar{t}+W$, $t+Z$, $t+WZ$, $t\bar{t}+H$, $t\bar{t}+WW$, $t\bar{t}t\bar{t}$, and ttt . The prediction from the signal samples is taken before the likelihood fit is performed. The background contributions may not add up to equal the total background due to rounding.

Signal Regions	2ℓ SSSF, 1τ	2ℓ SSOF, 1τ	2ℓ OSSF, 1τ	2ℓ OSOF, 1τ	$2\ell, \geq 2\tau$	$3\ell, \geq 1\tau$	$\geq 4\ell, \geq 0\tau$
Observed Events	6	3	37	7	4	2	8
Total Background	4.60 ± 0.60	3.70 ± 0.40	28.00 ± 1.80	3.80 ± 0.50	5.70 ± 0.40	3.90 ± 0.30	7.20 ± 0.70
Other Top	0.81 ± 0.19	0.85 ± 0.24	1.60 ± 0.40	0.63 ± 0.19	0.76 ± 0.09	0.35 ± 0.07	0.44 ± 0.13
$t\bar{t}+Z$	0.14 ± 0.04	0.28 ± 0.07	2.10 ± 0.40	0.08 ± 0.05	0.82 ± 0.16	1.11 ± 0.20	3.10 ± 0.60
ZZ	0.09 ± 0.02	0.13 ± 0.02	1.83 ± 0.14	0.05 ± 0.01	2.06 ± 0.28	1.10 ± 0.12	2.00 ± 0.25
WZ	0.95 ± 0.16	1.28 ± 0.19	12.40 ± 1.60	0.51 ± 0.09	< 0.01	< 0.01	0.06 ± 0.04
Triboson	0.14 ± 0.02	< 0.01	0.47 ± 0.04	< 0.01	0.05 ± 0.01	0.10 ± 0.01	0.19 ± 0.02
Fakes	2.40 ± 0.50	1.18 ± 0.17	9.60 ± 1.50	2.50 ± 0.40	2.03 ± 0.17	1.20 ± 0.13	1.39 ± 0.34
$\tau'(130\text{ GeV})^*$	3.40 ± 2.40	8.00 ± 4.00	11.00 ± 5.00	6.00 ± 2.40	25.00 ± 6.00	13.00 ± 4.00	5.40 ± 3.50
$\tau'(500\text{ GeV})^*$	12.60 ± 0.80	13.60 ± 0.80	29.60 ± 1.40	10.70 ± 0.70	13.80 ± 1.00	13.70 ± 0.70	4.70 ± 0.40
$\tau'(800\text{ GeV})^*$	2.35 ± 0.13	2.20 ± 0.12	6.16 ± 0.30	2.48 ± 0.14	1.72 ± 0.13	2.25 ± 0.13	0.94 ± 0.06
$\tau'(1000\text{ GeV})^*$	0.73 ± 0.04	0.67 ± 0.04	1.90 ± 0.09	0.82 ± 0.04	0.51 ± 0.04	0.68 ± 0.04	0.29 ± 0.02

*Pre-fit

Table 10 shows the impact of each source of systematic uncertainty on the signal strength μ in the signal-plus-background fit. Signal strength is defined as the ratio of the signal cross section estimated using the data to the predicted signal cross section. The nuisance parameters are grouped according to their origin. To evaluate the impact of each source of systematic uncertainty, the source is removed from the full fit and the signal strength and its uncertainty are recalculated. The square of the impact is defined as the decrease in the squared signal-strength uncertainty. The nuisance parameters associated with the background normalization have the highest impact on μ , while the systematic uncertainties associated with the fake-lepton estimation have the second-highest impact.

Using the VLL doublet model in Refs. [23, 24], the predicted significance of the signal is expected to be greater than 5 standard deviations for $M_{\tau'}$ in the range from 130 GeV to 600 GeV and greater than 3 standard deviations for values of $M_{\tau'}$ up to 800 GeV. Above this mass the expected significance decreases to ≈ 1 standard deviation at 1 TeV. The observed significance is found to be $\lesssim 1$ standard deviation over the entire $M_{\tau'}$ range probed, as shown in Table 11.

No significant deviation from the SM prediction is observed. Therefore, the 95% CL exclusion limit on the VLL production cross section as a function of $M_{\tau'}$ is calculated and shown in Figure 6. In order to estimate the 95% CL upper limit on the VLL cross section, a simultaneous binned likelihood fit of the seven SRs and four CRs is performed, using the CL_s method [90] with the asymptotic approximation. The expected limit is shown with the black dashed line, and the shaded regions correspond to its one and two standard-deviation uncertainty bands. The observed 95% CL exclusion limit is shown with the solid black line. The expected lower limit on $M_{\tau'}$ is found to be 970 GeV, and the observed limit to be 900 GeV, by comparing the NLO theory prediction with the expected and observed 95% CL cross-section limits.

Table 9: Total observed yields as computed by the fit for control regions. The uncertainty contains both the systematic and statistical uncertainties. The prediction for each background sample is taken after a likelihood fit is performed to measure the VLL production cross section. Background normalization factors are also applied. The ‘Other Top’ sample includes contributions from single top, $t\bar{t}$, $t\bar{t}+W$, $t+Z$, $t+WZ$, $t\bar{t}+H$, $t\bar{t}+WW$, $t\bar{t}t\bar{t}$, and ttt . The prediction from the signal samples is taken before the likelihood fit is performed. The background contributions may not add up to equal the total background due to rounding.

Control Regions	$t\bar{t}+Z$	WZ	ZZ	Fake τ
Observed events	67	96	774	7743
Total background	67 ± 8	97 ± 9	774 ± 28	7760 ± 90
Z+jets	< 0.03	0.1 ± 0.1	< 0.03	2.55 ± 0.32
Other Top	6.3 ± 0.9	2.0 ± 0.6	0.73 ± 0.14	66 ± 7
$t\bar{t}+Z$	51 ± 8	3.1 ± 0.6	6.5 ± 1.2	52 ± 9
ZZ	4.8 ± 0.5	2.5 ± 0.2	753 ± 28	95 ± 5
WZ	10.0 ± 0.1	82 ± 9	0.17 ± 0.03	370 ± 50
Triboson	< 0.03	0.9 ± 0.1	0.44 ± 0.04	1.71 ± 0.15
Fakes	4.8 ± 0.8	7.0 ± 0.7	12.4 ± 1.2	7170 ± 100
$\tau'(130 \text{ GeV})^*$	11 ± 4	12 ± 4	285 ± 21	1380 ± 100
$\tau'(500 \text{ GeV})^*$	0.7 ± 0.1	7.4 ± 0.5	0.64 ± 0.13	3.71 ± 0.31
$\tau'(800 \text{ GeV})^*$	< 0.03	3.1 ± 0.2	< 0.03	0.06 ± 0.02
$\tau'(1000 \text{ GeV})^*$	< 0.03	1.2 ± 0.1	< 0.03	< 0.03

*Pre-fit

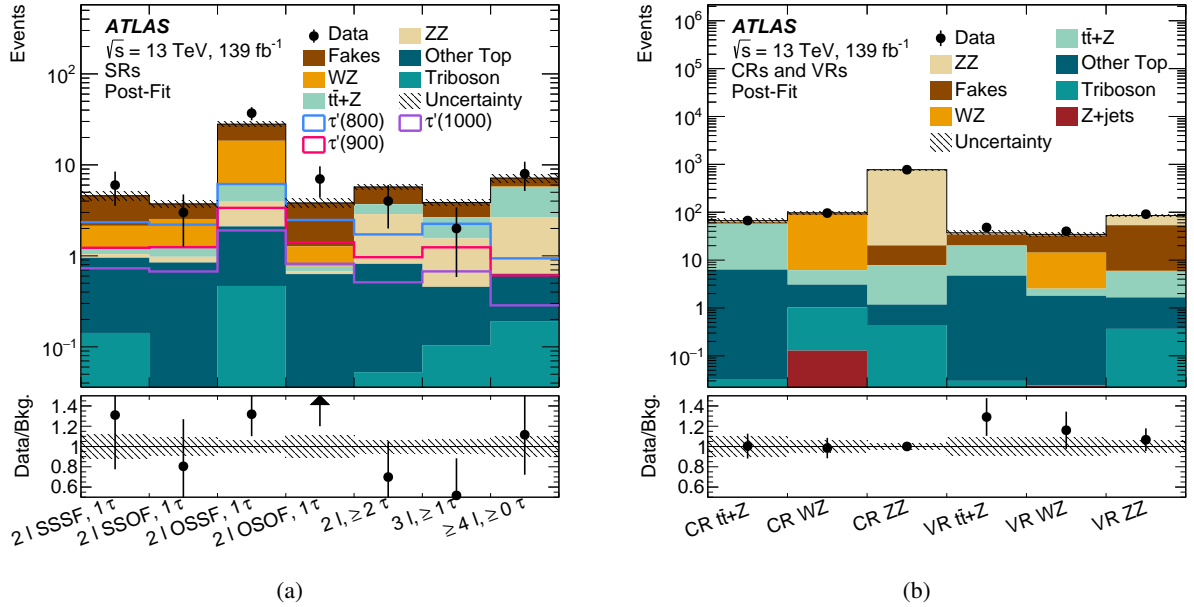


Figure 4: Summary of post-fit yields for data and background, and pre-fit yields for signal modelling in the (a) SRs and (b) CRs and corresponding VRs. Uncertainty bands contain both the systematic and statistical uncertainties. Background normalization factors are also applied. The ‘Other Top’ sample includes contributions from single top, $t\bar{t}$, $t\bar{t}+W$, $t+Z$, $t+WZ$, $t\bar{t}+H$, $t\bar{t}+WW$, $t\bar{t}t\bar{t}$, and ttt . The arrow in the ratio plot is for the point that is outside the range.

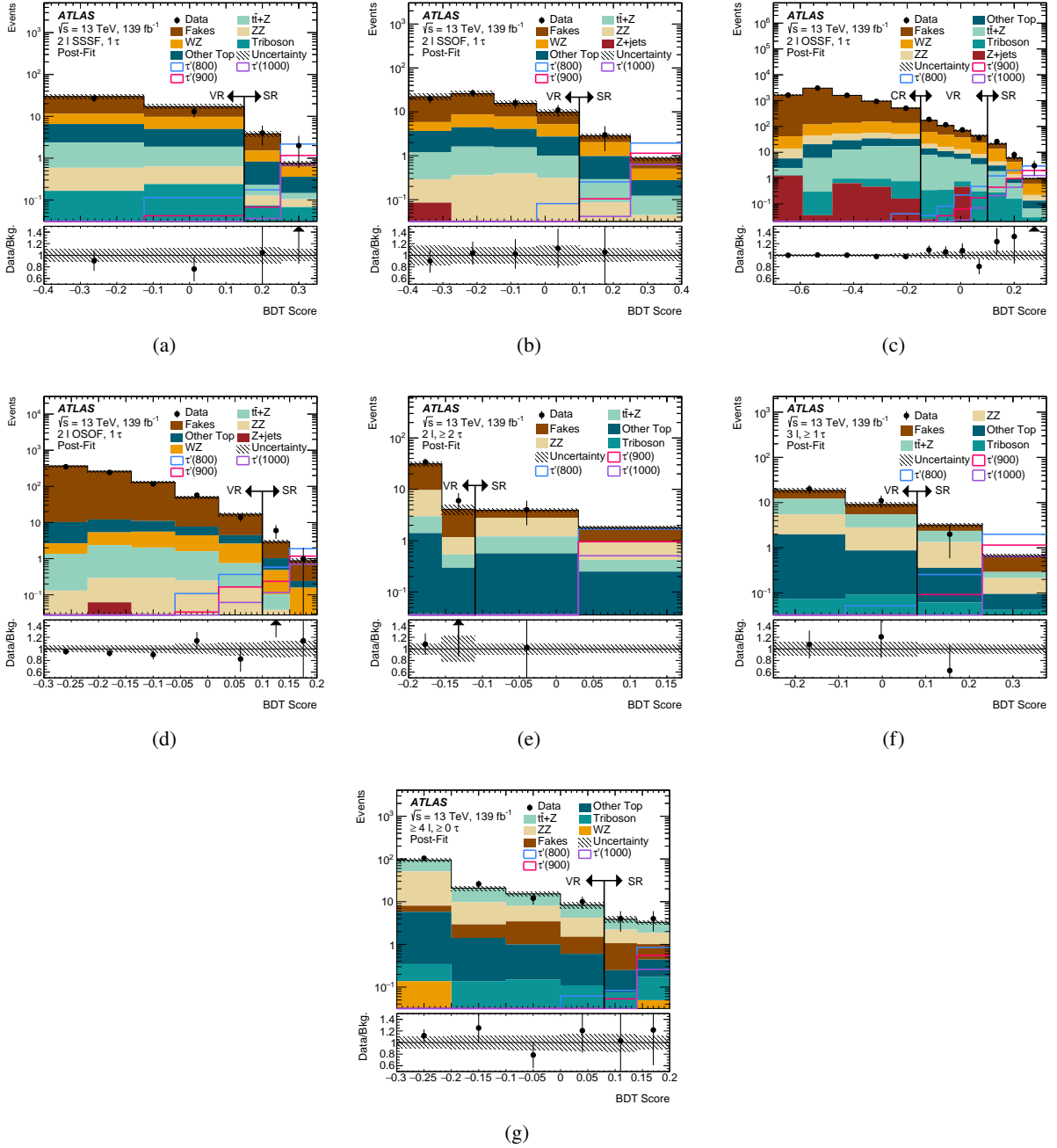


Figure 5: The post-fit BDT score distributions of data, background, and pre-fit signal modelling. The arrows indicate the point where there is a break between the regions and indicate the SRs used in the likelihood fit. The remaining distribution is treated as a VR and is not used in the fit (with the exception of (c) 2ℓ OSSF, where the low BDT score region is used as a CR and is included in the fit); however, fitted nuisance parameters are propagated to these regions. The uncertainty bands contain both the systematic and statistical components. The first bin contains underflow events and the last bin contains overflow events. The arrows in the ratio plot are for points that are outside the range. (a) 2ℓ SSSF, 1τ ; (b) 2ℓ SSOFF, 1τ ; (c) 2ℓ OSSF, 1τ ; (d) 2ℓ OSOF, 1τ ; (e) 2ℓ , $\geq 2 \tau$; (f) 3ℓ , $\geq 1 \tau$; (g) $\geq 4 \ell$, $\geq 0 \tau$.

Table 10: Impact of each source of systematic uncertainty on the signal strength μ in the signal-plus-background fit. $M_{\tau'} = 900$ GeV is considered. Both the positive and negative impacts on μ are considered. Individual sources of uncertainty are grouped for brevity but are treated as independent and uncorrelated in the fitting procedure. Impact of the total statistical uncertainty is also presented.

Uncertainty	$+\Delta\mu$	$-\Delta\mu$
Normalization factors	0.046	0.059
Fakes	0.033	0.050
Other theory uncertainties	0.017	0.012
τ -lepton ID and reconstruction	0.016	0.045
Light-lepton ID, reco., energy scale and momentum resolution	0.016	0.033
Jet energy scale and resolution, JVT, pile-up reweighting	0.016	0.012
τ -lepton energy scale	0.011	0.010
E_T^{miss}	0.008	0.003
Luminosity	0.015	0.006
Flavour tagging	0.003	0.003
MC and fake-background statistical uncertainty	0.066	0.120
Total systematic uncertainty	0.098	0.160
Total statistical uncertainty	0.450	0.350

Table 11: Results from the fitting procedure for the VLL mass points studied. The expected and observed signal significances, and the 95% CL exclusion limits are shown.

τ' Mass [GeV]	Significance		Exclusion Limit [fb]	
	Expected	Observed	Expected	Observed
130	5.9	-0.4	1110_{-310}^{+520}	953
200	12	-0.4	100_{-28}^{+44}	90
300	15	0.1	$17.0_{-4.8}^{+7.5}$	18
400	12	0.3	$7.4_{-2.1}^{+3.3}$	8.7
500	10	0.7	$3.6_{-1.0}^{+1.6}$	4.7
600	7.6	0.9	$2.3_{-0.6}^{+1.0}$	3.1
700	5.1	1.0	$1.8_{-0.5}^{+0.8}$	2.4
800	3.4	1.1	$1.5_{-0.4}^{+0.7}$	2.1
900	2.1	1.1	$1.3_{-0.4}^{+0.6}$	1.8
1000	1.3	1.2	$1.2_{-0.3}^{+0.6}$	1.7
1100	0.8	1.1	$1.2_{-0.3}^{+0.6}$	1.7
1200	0.5	1.1	$1.2_{-0.3}^{+0.6}$	1.7
1300	0.3	1.1	$1.2_{-0.3}^{+0.6}$	1.7

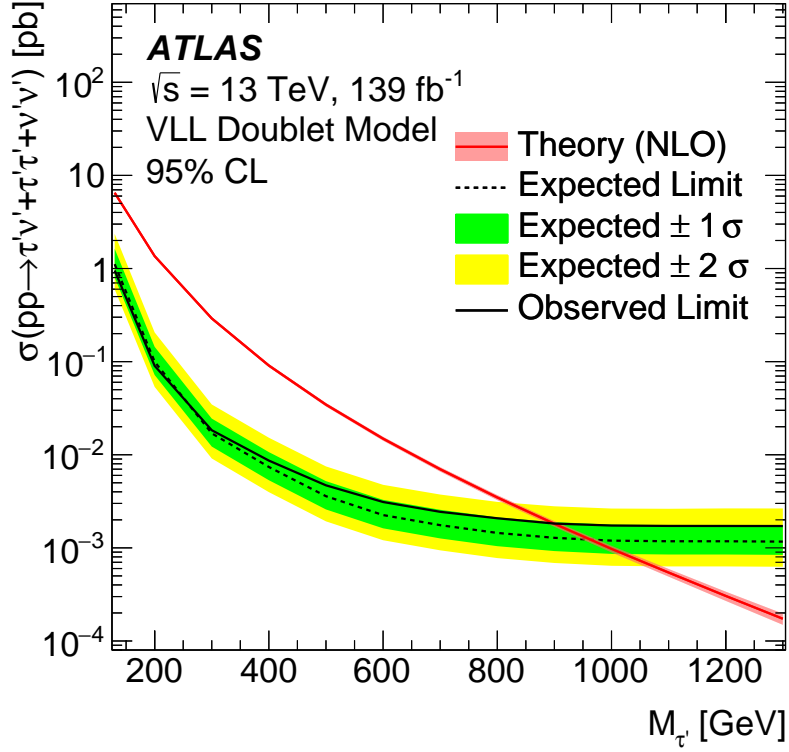


Figure 6: The 95% CL exclusion limit on the VLL production cross section as a function of VLL mass. The black dashed line represents the expected limit while the shaded regions are its one and two standard-deviation uncertainty bands. The solid black line is the observed limit as a function of VLL mass. The red curve is the NLO theory prediction along with its uncertainty.

9 Conclusions

A search for vector-like leptons in a doublet model is performed using 139 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13 \text{ TeV}$ by the ATLAS detector at the LHC. The search is performed using events with final states containing multiple light leptons and τ_{had} . Observing no excess of events above the SM expectation, a 95% CL upper limit on the cross section is calculated using the CL_s method. Using a doublet model where the vector-like leptons couple to the third-generation SM leptons, the observed mass range from 130 GeV to 900 GeV is excluded at the 95% CL, while the highest excluded mass is expected to be 970 GeV.

References

- [1] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.
- [2] P. H. Frampton, P. Q. Hung and M. Sher, *Quarks and leptons beyond the third generation*, *Phys. Rept.* **330** (2000) 263, arXiv: [hep-ph/9903387](#).
- [3] D0 Collaboration, *Search for Single Vector-Like Quarks in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96 \text{ TeV}$* , *Phys. Rev. Lett.* **106** (2011) 081801, arXiv: [1010.1466 \[hep-ex\]](#).

- [4] CDF Collaboration, *Search for New Bottomlike Quark Pair Decays $Q\bar{Q} \rightarrow (tW^\mp)(\bar{t}W^\pm)$ in Same-Charge Dilepton Events*, *Phys. Rev. Lett.* **104** (2010) 091801, arXiv: [0912.1057 \[hep-ex\]](#).
- [5] CMS Collaboration, *Search for a Vectorlike Quark with Charge 2/3 in $t + Z$ Events from pp Collisions at $\sqrt{s} = 7$ TeV*, *Phys. Rev. Lett.* **107** (2011) 271802, arXiv: [1109.4985 \[hep-ex\]](#).
- [6] CMS Collaboration, *Search for pair produced fourth-generation up-type quarks in pp collisions at $\sqrt{s} = 7$ TeV with a lepton in the final state*, *Phys. Lett. B* **718** (2012) 307, arXiv: [1209.0471 \[hep-ex\]](#).
- [7] ATLAS Collaboration, *Search for Pair Production of a New b' Quark that Decays into a Z Boson and a Bottom Quark with the ATLAS Detector*, *Phys. Rev. Lett.* **109** (2012) 071801, arXiv: [1204.1265 \[hep-ex\]](#).
- [8] CMS Collaboration, *Search for vectorlike charge 2/3 T quarks in proton–proton collisions at $\sqrt{s} = 8$ TeV*, *Phys. Rev. D* **93** (2016) 012003, arXiv: [1509.04177 \[hep-ex\]](#).
- [9] CMS Collaboration, *Search for pair-produced vectorlike B quarks in proton–proton collisions at $\sqrt{s} = 8$ TeV*, *Phys. Rev. D* **93** (2016) 112009, arXiv: [1507.07129 \[hep-ex\]](#).
- [10] CMS Collaboration, *Inclusive search for a vector-like T quark with charge $\frac{2}{3}$ in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett. B* **729** (2014) 149, arXiv: [1311.7667 \[hep-ex\]](#).
- [11] CMS Collaboration, *Search for vector-like T quarks decaying to top quarks and Higgs bosons in the all-hadronic channel using jet substructure*, *JHEP* **06** (2015) 080, arXiv: [1503.01952 \[hep-ex\]](#).
- [12] ATLAS Collaboration, *Search for pair production of a new heavy quark that decays into a W boson and a light quark in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Phys. Rev. D* **92** (2015) 112007, arXiv: [1509.04261 \[hep-ex\]](#).
- [13] ATLAS Collaboration, *Search for production of vector-like quark pairs and of four top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **08** (2015) 105, arXiv: [1505.04306 \[hep-ex\]](#).
- [14] ATLAS Collaboration, *Search for pair production of vector-like top quarks in events with one lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *JHEP* **08** (2017) 052, arXiv: [1705.10751 \[hep-ex\]](#).
- [15] CMS Collaboration, *Search for pair production of vector-like T and B quarks in single-lepton final states using boosted jet substructure techniques at $\sqrt{s} = 13$ TeV*, *JHEP* **11** (2017) 085, arXiv: [1706.03408 \[hep-ex\]](#).
- [16] CMS Collaboration, *Search for pair production of vector-like quarks in the $bW\bar{b}W$ channel from proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **779** (2018) 82, arXiv: [1710.01539 \[hep-ex\]](#).
- [17] CMS Collaboration, *Search for vector-like T and B quark pairs in final states with leptons at $\sqrt{s} = 13$ TeV*, *JHEP* **08** (2018) 177, arXiv: [1805.04758 \[hep-ex\]](#).

- [18] CMS Collaboration, *Search for vector-like quarks in events with two oppositely charged leptons and jets in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 364, arXiv: [1812.09768 \[hep-ex\]](#).
- [19] ATLAS Collaboration, *Search for pair production of heavy vector-like quarks decaying to high- p_T W bosons and b quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **10** (2017) 141, arXiv: [1707.03347 \[hep-ex\]](#).
- [20] ATLAS Collaboration, *Search for pair production of up-type vector-like quarks and for four-top-quark events in final states with multiple b -jets with the ATLAS detector*, *JHEP* **07** (2018) 089, arXiv: [1803.09678 \[hep-ex\]](#).
- [21] ATLAS Collaboration, *Search for pair production of heavy vector-like quarks decaying into high- p_T W bosons and top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **08** (2018) 048, arXiv: [1806.01762 \[hep-ex\]](#).
- [22] CMS Collaboration, *Search for pair production of vector-like quarks in the fully hadronic final state*, *Phys. Rev. D* **100** (2019) 072001, arXiv: [1906.11903 \[hep-ex\]](#).
- [23] N. Kumar and S. P. Martin, *Vectorlike Leptons at the Large Hadron Collider*, *Phys. Rev. D* **92** (2015) 115018, arXiv: [1510.03456 \[hep-ph\]](#).
- [24] P. N. Bhattiprolu and S. P. Martin, *Prospects for vectorlike leptons at future proton-proton colliders*, *Phys. Rev. D* **100** (2019) 015033, arXiv: [1905.00498 \[hep-ph\]](#).
- [25] CMS Collaboration, *Inclusive nonresonant multilepton probes of new phenomena at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **105** (2022) 112007, arXiv: [2202.08676 \[hep-ex\]](#).
- [26] J. M. Cornwall, D. N. Levin and G. Tiktopoulos, *Derivation of Gauge Invariance from High-Energy Unitarity Bounds on the S Matrix*, *Phys. Rev. D* **10** (1974) 1145, Erratum: *Derivation of gauge invariance from high-energy unitarity bounds on the S matrix*, *Phys. Rev. D* **11** (4 1975) 972.
- [27] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, *Journal of Computer and System Sciences* **55** (1997) 119, ISSN: 0022-0000.
- [28] L. Breiman, J. H. Friedman, R. A. Olshen and C. J. Stone, *Classification and regression trees*, Routledge, 2017.
- [29] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [30] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*, ATLAS-TDR-19; CERN-LHCC-2010-013, 2010, URL: <https://cds.cern.ch/record/1291633>, Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009, 2012, URL: <https://cds.cern.ch/record/1451888>.
- [31] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, *JINST* **13** (2018) T05008, arXiv: [1803.00844 \[physics.ins-det\]](#).
- [32] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, *Eur. Phys. J. C* **77** (2017) 317, arXiv: [1611.09661 \[hep-ex\]](#).

- [33] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATL-SOFT-PUB-2021-001, 2021, URL: <https://cds.cern.ch/record/2767187>.
- [34] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, *JINST* **15** (2020) P04003, arXiv: [1911.04632](https://arxiv.org/abs/1911.04632) [[physics.ins-det](#)].
- [35] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, *JINST* **13** (2018) P07017.
- [36] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, *Eur. Phys. J. C* **80** (2020) 47, arXiv: [1909.00761](https://arxiv.org/abs/1909.00761) [[hep-ex](#)].
- [37] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, *JINST* **15** (2020) P09015, arXiv: [2004.13447](https://arxiv.org/abs/2004.13447) [[hep-ex](#)].
- [38] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 – a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [39] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: [1005.4568](https://arxiv.org/abs/1005.4568) [[physics.ins-det](#)].
- [40] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852, arXiv: [0710.3820](https://arxiv.org/abs/0710.3820) [[hep-ph](#)].
- [41] 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>.
- [42] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244, arXiv: [1207.1303](https://arxiv.org/abs/1207.1303) [[hep-ph](#)].
- [43] 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](https://arxiv.org/abs/1405.0301) [[hep-ph](#)].
- [44] E. Bothmann et al., *Event generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034, arXiv: [1905.09127](https://arxiv.org/abs/1905.09127) [[hep-ph](#)].
- [45] R. D. Ball et al., *Parton distributions for the LHC run II*, *JHEP* **04** (2015) 040, arXiv: [1410.8849](https://arxiv.org/abs/1410.8849) [[hep-ph](#)].
- [46] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, *JHEP* **03** (2008) 038, arXiv: [0709.1027](https://arxiv.org/abs/0709.1027) [[hep-ph](#)].
- [47] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049, arXiv: [1111.1220](https://arxiv.org/abs/1111.1220) [[hep-ph](#)].
- [48] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, *JHEP* **04** (2013) 027, arXiv: [1207.5030](https://arxiv.org/abs/1207.5030) [[hep-ph](#)].
- [49] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) 063, arXiv: [hep-ph/0109231](https://arxiv.org/abs/hep-ph/0109231).
- [50] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053, arXiv: [0903.1219](https://arxiv.org/abs/0903.1219) [[hep-ph](#)].

- [51] J.-C. Winter, F. Krauss and G. Soff, *A modified cluster-hadronization model*, *Eur. Phys. J. C* **36** (2004) 381, arXiv: [hep-ph/0311085](#).
- [52] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039, arXiv: [0808.3674 \[hep-ph\]](#).
- [53] F. Buccioni et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866, arXiv: [1907.13071 \[hep-ph\]](#).
- [54] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, arXiv: [1111.5206 \[hep-ph\]](#).
- [55] A. Denner, S. Dittmaier and L. Hofer, *COLLIER: A fortran-based complex one-loop library in extended regularizations*, *Comput. Phys. Commun.* **212** (2017) 220, arXiv: [1604.06792 \[hep-ph\]](#).
- [56] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, arXiv: [1002.2581 \[hep-ph\]](#).
- [57] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: [hep-ph/0409146](#).
- [58] 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 \[hep-ph\]](#).
- [59] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: <https://cds.cern.ch/record/2216168>.
- [60] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029, arXiv: [0805.3067 \[hep-ph\]](#).
- [61] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159, arXiv: [1410.3012 \[hep-ph\]](#).
- [62] 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>.
- [63] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, *JINST* **14** (2019) P12006, arXiv: [1908.00005 \[hep-ex\]](#).
- [64] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **81** (2021) 578, arXiv: [2012.00578 \[hep-ex\]](#).
- [65] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: [0802.1189 \[hep-ph\]](#).
- [66] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett. B* **641** (2006) 57, arXiv: [hep-ph/0512210](#).
- [67] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, arXiv: [1111.6097 \[hep-ph\]](#).
- [68] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466, arXiv: [1703.10485 \[hep-ex\]](#).

- [69] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2020) 689, arXiv: [2007.02645 \[hep-ex\]](#).
- [70] 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 \[hep-ex\]](#).
- [71] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 970, arXiv: [1907.05120 \[hep-ex\]](#).
- [72] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, (2022), arXiv: [2211.16345 \[physics.data-an\]](#).
- [73] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2017-029, 2017, URL: <https://cds.cern.ch/record/2261772>.
- [74] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, ATL-PHYS-PUB-2015-045, 2015, URL: <https://cds.cern.ch/record/2064383>.
- [75] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, ATL-PHYS-PUB-2019-033, 2019, URL: <https://cds.cern.ch/record/2688062>.
- [76] 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>.
- [77] F. Pedregosa et al., *Scikit-learn: Machine Learning in Python*, *Journal of Machine Learning Research* **12** (2011) 2825.
- [78] ATLAS Collaboration, *Search for doubly and singly charged Higgs bosons decaying into vector bosons in multi-lepton final states with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **06** (2021) 146, arXiv: [2101.11961 \[hep-ex\]](#).
- [79] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, (2022), arXiv: [2211.16178 \[hep-ex\]](#).
- [80] ATLAS Collaboration, *Search for new phenomena in three- or four-lepton events in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **824** (2021) 136832, arXiv: [2107.00404 \[hep-ex\]](#).
- [81] E. Bothmann, M. Schönherr and S. Schumann, *Reweighting QCD matrix-element and parton-shower calculations*, *Eur. Phys. J. C* **76** (2016) 590, arXiv: [1606.08753 \[hep-ph\]](#).
- [82] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys. G* **43** (2016) 023001, arXiv: [1510.03865 \[hep-ph\]](#).
- [83] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: <https://cds.cern.ch/record/1700870>.

- [84] 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 \[hep-ex\]](#).
- [85] 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 \[hep-ex\]](#).
- [86] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **78** (2018) 903, arXiv: [1802.08168 \[hep-ex\]](#).
- [87] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: [physics/0306116 \[physics.data-an\]](#).
- [88] L. Moneta et al., *The RooStats Project*, 2011, arXiv: [1009.1003 \[physics.data-an\]](#).
- [89] 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, arXiv: [1007.1727 \[physics.data-an\]](#), Erratum: *Eur. Phys. J. C* **73** (2013) 2501.
- [90] A. L. Read, *Presentation of search results: the CL_S technique*, *J. Phys. G* **28** (2002) 2693.