ATLAS CONF Note

ATLAS-CONF-2021-023

7th June 2021

Search for type-III seesaw heavy leptons in leptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS Collaboration

A search for the pair production of heavy leptons as predicted by the type-III seesaw mechanism is presented. The search uses proton–proton collision data at a centre-of-mass energy of 13 TeV, corresponding to 139 fb^{-1} of integrated luminosity recorded by the ATLAS detector during Run 2 of the Large Hadron Collider. The analysis focuses on final states with three or four electrons or muons from the possible decays of new heavy leptons via intermediate electroweak bosons. No significant deviations above the Standard Model expectation are observed and upper and lower limits on the production cross-section and on the heavy lepton masses are derived. These results are then combined for the first time with the ones already published by ATLAS using the channel with two leptons in the final state. The observed lower limit on the mass of the type-III seesaw heavy leptons combining two, three and four lepton channels together is 910 GeV at the 95% confidence level.

ATLAS-CONF-2021-023
18 June 2021

1 Introduction

Neutrino physics presents some of the biggest puzzles yet to be addressed in modern particle physics. The extremely small values of the neutrino masses compared to the masses of the other fermions appears unnatural in the Standard Model (SM) [1]. The seesaw mechanism [2–6] provides an elegant way to give a very small mass m_ν to each of the SM neutrinos by introducing a heavy Majorana neutrino with mass M. The Electroweak (EW) spontaneously broken symmetry explains the neutrino mass as a Yukawa coupling. Three types of seesaw mechanism have been proposed and their phenomenology can be tested at collider experiments. The type-III seesaw [7] introduces at least one extra fermionic $\mathrm{SU(2)_L}$ triplet field coupled to EW gauge bosons. These heavy charged and neutral leptons can in principle be produced by EW processes at the Large Hadron Collider (LHC).

Type-III seesaw heavy-lepton searches have already been performed in various decay channels by both the ATLAS and CMS Collaborations. In Run 1, ATLAS excluded heavy leptons with masses below 335 GeV [8] using final states containing two light leptons (electron or muons) and two jets. This limit was then improved to 470 GeV, still using Run-1 data, due to the addition of the three-lepton channel [9]. Using the full Run 2 data sample of proton–proton collisions at $\sqrt{s} = 13$ TeV, the CMS Collaboration has excluded heavy lepton masses up to 880 GeV [10] by analyzing three- and four-leptons final states, while ATLAS has excluded heavy lepton masses up to 810 GeV [11] using only the two-lepton plus jets final state. The analysis presented in this paper searches for type-III seesaw heavy lepton in three- and four-lepton final states. For the first time, a combination with the two-lepton plus jets final state is performed, giving a significant improvement in the sensitivity of the analysis.

The type-III seesaw model targeted in this search is described in Ref. [12], where the pair production of the neutral Majorana (N^0) and charged ($L^\pm$) heavy leptons proceeds via the s-channel production of virtual EW gauge bosons. The production cross-section depends only on the masses of the N^0 and $L^\pm$, which are assumed to be degenerate as the mass splitting due to electroweak radiative corrections is expected to be smaller than ~200 MeV [13]. The decays allowed in this model are $L^\pm \to H\ell^\pm, Z\ell^\pm, W^\pm\nu$ and $N^0 \to Z\nu, H\nu, W^\pm\ell^\mp$, where the SM leptons ($\ell = e, \mu, \tau$) can be of any flavour . The branching ratios $\mathcal{B}_\ell$ for the heavy lepton decays into ℓ plus one SM boson are determined by the mixing parameters V_ℓ between the new heavy leptons and the SM leptons. In this analysis it is assumed that all three mixing parameters are equal, namely $\mathcal{B}_e = \mathcal{B}_\mu = \mathcal{B}_\tau = 1/3$. The branching ratios $\mathcal{B}_Z, \mathcal{B}_W, \mathcal{B}_H$ for heavy lepton decays into any SM lepton plus Z, W or H, namely $L^\pm, N^0 \to Z, W, H$ are independent of the mixing parameters. For N^0 masses larger than few times the H mass, the decays into different SM bosons are independent from the heavy lepton mass, therefore $2\mathcal{B}_H \simeq 2\mathcal{B}_Z \simeq \mathcal{B}_W \simeq 1/2$. Examples of Feynman diagrams in three- and four-lepton final states are shown in Figure 1. These events are characterized by the production of two SM bosons (VV, VH or HH, where $V = W, Z$) and two charged leptons or neutrinos in the final state. This analysis focuses on events with high light lepton multiplicity, including light leptons from τ-lepton decays.

This paper is structured as follows: the ATLAS detector is described in Section 2; the data and simulated events used in the analysis are outlined in Section 3, and the event reconstruction procedure is detailed in Section 4. The analysis strategy and background estimation are presented in Section 5 and 6, respectively. Finally, results and their statistical interpretation are presented in Section 8, followed by the conclusions in Section 9.

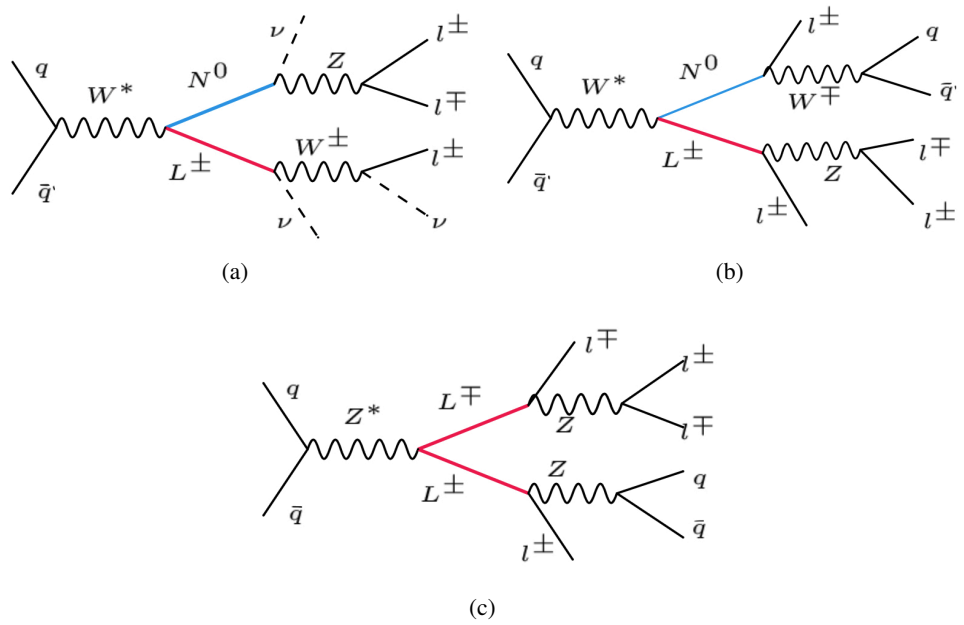


Figure 1: Examples of Feynman diagrams for the considered type-III seesaw model [12] producing three- and four-lepton final states.

2 ATLAS detector

The ATLAS detector [14] at the LHC is a multipurpose particle detector with a near-4π coverage in solid angle around the collision point and a cylindrical geometry[1] coaxial with the beam axis. It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2T magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroidal magnets with eight coils each.

The inner detector (ID) provides charged-particle tracking in the range $|\eta| < 2.5$ and is composed of, going from the innermost to the outermost layer, a high-granularity silicon pixel detector, a silicon microstrip tracker, and a transition radiation tracker that covers the region up to $|\eta| = 2.0$. The calorimeter system is composed of lead/liquid-argon calorimeters providing electromagnetic energy measurements, steel/scintillator-tile calorimeter in the $|\eta| < 1.7$ range for hadronic energy measurements, and endcap copper/liquid-argon calorimeters for both the electromagnetic and hadronic energy measurements up to $|\eta| < 4.9$. The muon spectrometer (MS) instruments the outer part of the detector and is composed of high-precision tracking chambers up to $|\eta| < 2.7$ and fast detectors for triggering up to $|\eta| < 2.4$. The MS is immersed in a magnetic field produced by three large superconducting air-core toroidal magnets with eight coils each.

A two-level trigger system is used to select events that are of interest for the ATLAS physics programme [15].

[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. Polar 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 $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The first-level trigger is implemented in hardware and reduces the event rate to below 100 kHz. A software-based trigger further reduces the recorded event rate to approximately 1 kHz.

3 Data and simulated events

This analysis uses data collected in proton–proton collisions at $\sqrt{s} = 13$ TeV with proton bunches colliding every 25 ns. After requiring that all ATLAS subdetectors collected high-quality data and were operating normally [16], the total integrated luminosity amounts to 139 fb^{-1}. The uncertainty in the combined 2015-2018 integrated luminosity is 1.7 % [17], obtained using the LUCID-2 detector [18] for the primary luminosity measurements. Events were collected using dilepton triggers selecting pairs of electrons [19] or muons [20]. The unprescaled trigger threshold on the transverse momentum (p_{T}) of the leading leptons evolved during the data taking due to the luminosity increasing, but was never larger than 24 GeV for electrons and 22 GeV for muons, respectively.

Simulated signal and background processes were generated using different Monte Carlo (MC) generators as listed in Table 1. The ATLAS detector was simulated using the GEANT 4 toolkit [21] and simulated events were reconstructed with the same algorithms as those applied to data [22]. The type-III seesaw signal model was implemented in the MADGRAPH5_aMC@NLO [23] generator at leading order (LO) using FEYNRULES [24] and the `NNPDF3.0lo` [25] parton distribution functions (PDF). All decays of $L^{\pm}$ and N^0 in the different leptonic flavor and subsequent decays of the W, Z and H are considered. Matrix element (ME) events were interfaced to PYTHIA 8.230 [26] for parton showering with the `A14` set of tuned parameters [27] and the `NNPDF2.3lo` PDF [28]. The signal cross-section and its uncertainty at next-to-leading-order (NLO) plus next-to-leading-logarithmic (NLL) accuracy were calculated from SU(2) triplet production in an electroweak chargino-neutralino model [29, 30]. The calculated cross-sections are compatible within uncertainties with the type-III seesaw NLO implementation [31, 32].

Simulated SM background samples include diboson processes, which are the dominant ones, followed by processes labelled *rare top-quark* that include multi-top quark production and top quark production associated with EW bosons processes ($t\bar{t}V, t\bar{t}H, tWZ$). Other SM simulated samples are multiboson (VVV), $t\bar{t}$, single top as well as Drell–Yan ($q\bar{q} \to Z/\gamma^* \to \ell^+\ell^-$ ($\ell = e, \mu, \tau$)) production processes and are mainly used for the reducible background estimation as described in Section 6. The MEPS@NLO prescription [33] was used in the generation of Drell–Yan processes to match the ME to the parton shower. The generators used in the MC sample production and the cross-section calculations used for MC sample normalisations are listed in Table 1. The normalisation of the dominant backgrounds, diboson and rare top-quark processes, are extracted from the final likelihood fit, as described in Section 8.

Multiboson processes (diboson VV and triboson VVV) were simulated with the SHERPA v2.2.1-2 generator [38]. Off-shell effects and Higgs boson contributions were included. ME calculations were matched and merged with the SHERPA parton shower based on the Catani–Seymour dipole factorisation [39, 40] using the MEPS@NLO prescription [33, 41–43]. The OPENLOOPS library [44, 45] provided QCD corrections. Diboson samples were produced with different lepton multiplicities using MEs at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. Loop-induced processes $gg \to VV$ were generated with LO MEs for emission of up to one additional parton for both the fully leptonic and semileptonic final states. Electroweak production of a diboson pair in association with two jets ($VVjj$) was also simulated at LO. The PDFs used for the nominal samples were `CT14` [46] and `MMHT2014` [47], respectively.

Table 1: Configurations used for event generation of signal and most relevant background processes. The cross section (XS) order is the order in the strong coupling constant of the perturbative calculation. If only one parton distribution function is shown, the same one is used for both the ME and parton shower generators; if two are shown, the first is used for the ME calculation and the second for the parton shower. Tune refers to the tune of the underlying-event parameters of the parton shower generator. The mass of the top quark and SM Higgs boson were set to 172.5 GeV and 125 GeV, respectively.The samples with negligible impact are only mentioned in the table but not discussed in the text.

Process	Generator	XS order	Parton shower	PDF Set	Tune
type-III seesaw					
L^+L^-, $L^\pm N^0$	MadGraph5_aMC@NLO	LO	Pythia 8.230	NNPDF3.0lo NNPDF2.3lo	A14
Top quark					
$t\bar{t}$	Powheg-Box v2 [34–37]	NNLO	Pythia 8.230	NNPDF3.0nnlo NNPDF3.0nlo	A14
Single t	Powheg-Box v2 [34–37]	NNLO	Pythia 8.230	NNPDF3.0nnlo NNPDF3.0nlo	A14
Rare Top quark					
$3t$, $4t$	MadGraph5_aMC@NLO	LO	Pythia 8.230	NNPDF3.0lo	A14
$t\bar{t}+W/Z/H$, tWZ	MadGraph5_aMC@NLO	NNLO	Pythia 8.230	MEN30NLO	A14
Diboson					
ZZ, WZ	Sherpa 2.2.1 & 2.2.2	NLO	Sherpa	NNPDF3.0nnlo	Sherpa default
Multiboson					
WWW, WWZ, WZZ, ZZZ	Sherpa 2.2.1 & 2.2.2	NNLO	Sherpa	NNPDF3.0nnlo	Sherpa default
Drell–Yan					
$Z/\gamma^* \to e^+e^-/\mu^+\mu^-/\tau^+\tau^-$	Sherpa 2.2.1	NLO	Sherpa	NNPDF3.0nnlo	Sherpa default

Samples for $t\bar{t}+V$ and tWZ processes were produced using the MadGraph5_aMC@NLO generator with NLO accuracy and were interfaced to the Pythia 8.230 parton shower with MEN30NLO PDFs. The h_{damp} parameter that controls the matching between the ME and the parton shower was set to $1.5\,m_{\text{top}}$ [48], using a top quark mass of $m_{\text{top}} = 172.5\,\text{GeV}$.

All simulated events were overlayed with a simulation of multiple pp interactions occurring in the same or neighbouring bunch crossings. The hard-scattering event with simulated were generated simulating inelastic pp events with Pythia 8.186 using the NNPDF2.3lo set of PDFs and the A3 set of tuned parameters [49]. These effects are referred to as *pile-up*. The simulated events were reweighted such that the distribution of the average number of interactions per bunch crossing is compatible with that observed in the data.

4 Event reconstruction

Events are required to have at least one collision vertex reconstructed with at least two tracks with transverse momentum greater than 500 MeV. The primary vertex of the hard-scattering event is the one with the largest sum of the associated tracks' squared transverse momenta.

Electrons are reconstructed by matching a charged-particle track in the ID with an energy deposit in the electromagnetic calorimeter. Electron candidates are required to satisfy a *Loose* likelihood-based identification selection [50–52] and to be in the fiducial volume of the inner detector, $|\eta| < 2.47$. The transition region between the barrel and endcap calorimeters ($1.37 < |\eta| < 1.52$) is excluded because it is partially non-instrumented due to service infrastructures. The significance of the transverse impact parameter $|d_0|/\sigma(d_0)$ of the track associated to the candidate electron must be less than 5. This is required

to reduce the number of electrons originating from secondary decays. Similarly, the track longitudinal impact parameter relative to the primary vertex z_0 is required to satisfy $|z_0 \sin(\theta)| < 0.5$ mm. The minimum transverse energy is 10 GeV. After this preselection, to refine the electron quality, a *Tight* likelihood-based identification selection and a *Loose* isolation set of criteria based on both the calorimetric and the tracking information are applied to primarily select electrons coming from the decays of the heavy leptons or from the EW bosons.

Track segments in the MS are matched with ID tracks to reconstruct muons if the latter are within the η coverage of the ID. Muon candidates with p_{T} lower than 300 GeV are required to satisfy the *Medium* muon identification requirements, while for high-p_{T} muons, a specific identification working point is applied [53]. Muon candidates are required to have $|\eta| < 2.5$, a transverse impact parameter significance of $|d_0|/\sigma(d_0) < 3$ and longitudinal impact parameter value of $|z_0 \sin(\theta)| < 0.5$ mm. The minimum p_{T} is 10 GeV. An isolation based cut on tracking information only [53] is required.

Particle flow objects are combined using both the tracks measured in the ID not associated to identified leptons and the energy deposits in the calorimeters [54]. Particle flows objects are then clustered into jets using the anti-k_t algorithm [55] with radius parameter equal to 0.4. The measured jet p_{T} is corrected for detector effects to measure the particle energy before interacting with the material of the detector [56]. Energy deposits in the calorimeter from pile-up are subtracted considering a mean pile-up density multiplied by the η–ϕ area of the jet. Pile-up can also produce additional jets that are identified and rejected with the Jet Vertex Tagger (JVT) algorithm [57], which distinguishes them from jets originating from hard-scattering processes. Only jets with transverse energy larger than 20 GeV are considered. Jets originating from heavy flavour quarks are identified with a multivariate b-tagging algorithm MV2c10 using the 77% efficiency working point [58–60], with measured rejection factors of approximately 134, 6 and 22 for light-quark and gluon jets, c-jets, and hadronically-decaying τ-leptons, respectively.

Missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) is calculated as the magnitude of the missing transverse momentum vector $\vec{p}_T^{\,\mathrm{miss}}$, defined as the vectorial sum of the p_{T} of reconstructed jets and leptons in the event. A soft term taking into account tracks not associated with any hard object is then added to guarantee the best performance in a high pile-up environment [61]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance ($\mathcal{S}(E_{\mathrm{T}}^{\mathrm{miss}})$), calculated with a maximum-likelihood ratio method, which considers the direction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the calibrated objects as well as their respective resolutions, is used to define the various analysis regions defined in Section 5.

Different objects reconstructed closely in the η–ϕ plane could in principle have originated from the same primary object. Possible overlaps are resolved by an algorithm that appropriately removes one of the two close-by objects to avoid double-counting. If muon candidates are found to share an ID track with electron candidates, the electron candidate is rejected. If two electron candidates share an ID track, the one with the lowest p_{T} is rejected. Jets are rejected if they are within $\Delta R = 0.2$ of a lepton candidate, except if the candidate is a muon and three or more collinear tracks are found. Finaly, lepton candidates that are within $\Delta R = 0.4$ of any remaining jet are removed.

5 Analysis strategy

Once events have been classified based on the presence of three or four light leptons, the different lepton multiplicity categories are refined with dedicated selections. Signal regions (SRs) are defined maximising the significance of the signal event count predicted by the targeted model above the expected number of SM background events. SM backgrounds are normalized by performing a background-only fit simultaneously in

dedicated control regions (CRs) and in the SRs. The CRs are defined to be enriched in relevant background processes and depleted in events from signal processes. The fit is performed on a kinematic variable chosen to optimise sensitivity to small cross-sections. Validation regions (VRs), still depleted in signal events, are used to validate the extrapolation of the SM background expectations obtained from the background-only fit in independent regions kinematically close to the SRs.

Backgrounds are classified into two broad categories: reducible and irreducible backgrounds. Reducible backgrounds are caused by misreconstructed objects like jets, light- or heavy-quark decays or, for the prompt (originating in the hard process) electrons only, by photon conversions. These are called fake non-prompt (FNP) leptons and events containing at least one such lepton are referred to as the FNP background. Irreducible backgrounds are produced by SM processes with prompt leptons in the final states. The most important ones are diboson and rare top processes; the latter being due to $t\bar{t}$ pairs produced in association with an EW or Higgs boson. Kinematic distributions of these processes are obtained from MC simulation, while their normalisations are extracted from the fit. Due to the high value of the current mass exclusion limits on heavy leptons, signal events are characterized by objects having large momentum. Consequently final states with charged leptons with large momentum and large $\mathcal{S}(E_{\mathrm{T}}^{\mathrm{miss}})$ are expected due to the presence high-energy neutrinos.

Details of three- and four-lepton multiplicity analysis strategy are described below while those for the two-lepton analysis regions are given in Ref. [11].

5.1 Three-lepton channels

A summary of the selection criteria used to define three-lepton SRs, CRs and VRs is shown in Table 2. The ZL SR is characterized by a leptonically decaying Z boson, thus an opposite sign same flavour (OSSF) lepton pair compatible with the Z boson mass is required. Signal events are expected to have a large invariant mass of all three leptons and a large transverse mass of the two highest-p_{T} leptons. To further suppress FNP backgrounds, an additional requirement is made on the angular distance between the leading and subleading leptons.

A complementary ZLveto SR is defined vetoing events containing OSSF lepton pairs compatible with a leptonic decay of an on-shell Z boson. The H_{T} variable is defined as the scalar sum of p_{T} of all the selected objects in the event. If the scalar sum of p_{T} is restricted to only a subset of the objects they are specified. Signal events are characterized by large H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ values and by large scalar sum of the momenta of the same-sign leptons $H_T(SS)$. To account for possible hadronic channels of the electroweak bosons from multiboson background sources, an upper limit is placed on the dijet invariant mass.

Finally, the JNLow SR targets decays where the electroweak bosons decay leptonically, therefore events with a small jet multiplicity are selected. A lower bound on the invariant mass of the OSSF lepton pair is imposed and a large value of the scalar sum of the p_{T} of the three leptons,$H_T(\ell\ell\ell)$, is required. Two transverse mass values, $m_T(\ell_1)$ and $m_T(\ell_2)$, are computed: one using the leading-p_{T} lepton in the event, and the other one using the subleading one. Both values are used to further reduce fake-lepton background by requiring a minimal value of $m_T(\ell_{1,2})$. A minimal angular separation between the two leptons is required to reduce the FNP contribution.

The dominant SM backgrounds in the three-lepton SRs are diboson events, which amount to $\sim 60\%$, $\sim 80\%$ and $\sim 40\%$ in ZL, JNLow and ZLVeto regions, respectively. The second most important background in ZL and ZLVeto regions are rare top processes involving one or more top quarks, accounting for a $\sim 40\%$

and ∼ 50% contribution, respectively. Therefore a CR targeting the normalization of diboson background is defined using the same requirements as the ZL SR except for the requirement on the transverse mass constructed from the subleading lepton, which is inverted, $m_T(\ell_2) \leq 200\,\text{GeV}$.

Two VRs are defined to validate background estimations for events containing a Z boson decaying into leptons, both obtained by inverting the $\Delta R(\ell_1, \ell_2)$ selection of the ZL SR, and applying additional cuts. The DB-VR also requires a b-tag veto, while in RT-VR the presence of at least one b-tagged jet is required. These VRs validate the predictions and normalisation of diboson and rare top processes respectively. An additional JNLow-VR is obtained from the JNLow SR by inverting the transverse mass cut on the subleading lepton, $m_T(\ell_2) \leq 240\,\text{GeV}$. Moreover, the Fake-VR is defined starting from the leptons p_T cuts and inverting the $\mathcal{S}(E_\text{T}^\text{miss})$ selection common to all the other regions without applying any additional requirement. This enriches the region in contributions from FNP backgrounds.

In the three-lepton channel SRs, the kinematic variable used as the final discriminant is the magnitude of the vector sum of the three-lepton p_T and the missing p_T, $m_{T,3\ell} = |\sum_i^{3\ell} \vec{p}_{T,i} + \vec{p}_T^\text{miss}|$.

Table 2: Summary of the selection criteria used to define relevant regions in the three-lepton analysis.

		ZL				**ZLveto**	**JNLow**	
	Fake-VR	CR	DB-VR	RT-VR	SR	SR	VR	SR
	$p_\text{T}(l_1) > 40$ GeV							
	$p_\text{T}(l_2) > 40$ GeV							
	$p_\text{T}(l_3) > 15$ GeV							
$\mathcal{S}(E_\text{T}^\text{miss})$	< 5	≥ 5						
$N(jet)$		≥ 2				≥ 2	≤ 1	
$N(bjet)$		-	0	≥ 1	-			
$m_{ll}(OSSF)$ [GeV]		$80 - 100$				≥ 115	≥ 80	
$H_T + E_\text{T}^\text{miss}$ [GeV]						≥ 600		
m_{lll} [GeV]		-	≥ 300			≥ 300		
m_{jj} [GeV]						< 300		
$H_T(SS)$ [GeV]						≥ 300		
$H_T(lll)$ [GeV]							≥ 230	
$m_T(l_1)$ [GeV]			≥ 200				< 240	≥ 240
$m_T(l_2)$ [GeV]		< 200	≥ 200				≥ 150	
$\Delta R(l_1, l_2)$			< 1.2		$1.2 - 3.5$		≥ 1.3	

5.2 Four-lepton channel

In the four-lepton channel, the p_T of the third and fourth leptons must be greater than 10 GeV. Events are classified using the sum $\sum q_\ell$ of the charges of the four leptons in the final state. The conditions $\sum q_\ell = 0$ and $|\sum q_\ell| = 2$ identify the *Zero Charge* (Q0) and *Double Charge* (Q2) regions, respectively. The Q2 region is characterized by a lower background than the Q0 region since it is very rare for a SM process to produce a doubly charged final state. A summary of the selection criteria defining the four-lepton regions is shown in Table 3.

To reduce the ZZ contribution in the Q0 region, no more than one OSSF lepton pair is allowed in the event to be compatible with a leptonic Z decay. Signal events are characterized by large H_{T}, large invariant mass of the four-lepton system, $m_{\ell\ell\ell\ell}$, and large $\mathcal{S}(E_{\mathrm{T}}^{\mathrm{miss}})$ from the high momentum neutrino in the final state. A DB-CR targeting diboson backgrounds is built by requiring a b-jet veto and defining an invariant mass window of the four-lepton system (170 GeV $\leq m_{\ell\ell\ell\ell} \leq$ 300 GeV). A RT-CR targeting rare top background is obtained by requiring at least two b-tagged jets and $m_{\ell\ell\ell\ell} \leq$ 500 GeV. To ensure orthogonality with the CRs, VRs are defined by considering events with exactly one b-jet. To increase the contribution of diboson and rare top background in the VRs, $m_{\ell\ell\ell\ell}$ is required to satisfy 170 GeV $\leq m_{\ell\ell\ell\ell} \leq$ 300 GeV in DB-VR and 300 GeV $< m_{\ell\ell\ell\ell} \leq$ 500 GeV in RT-VR. The main sources of background are diboson or rare top events where the electric charge of one of the electrons is mismeasured. The only additional selections used to define the Q2 SR are that both H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ and the invariant mass of the four-lepton system should be greater than 300 GeV. A dedicated VR for the Q2 charge combination is obtained by requiring $m_{\ell\ell\ell\ell}$ to be less then 200 GeV or $H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$ less then 300 GeV.

In the four-lepton channel SRs, the kinematic variable used in the fit that maximises the sensitivity with respect to the expected signal is $H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$.

Table 3: Summary of the selection criteria used to define relevant regions in the four-lepton analysis. N_Z is the number of leptonically reconstructed Z, using opposite sign same flavour leptons.

	Q0					**Q2**	
	DB-VR	RT-VR	DB-CR	RT-CR	SR	VR	SR
$\lvert\sum q_\ell\rvert$	0					2	
$N_{b\text{-jet}}$	1	1	0	≥ 2	0		
m_{llll} [GeV]	170 – 300	300 – 500	170 – 300	< 500	≥ 300	< 200	≥ 300
						OR	
$H_T + E_{\mathrm{T}}^{\mathrm{miss}}$[GeV]		≥ 400			≥ 300	< 300	≥ 300
N_Z					≤ 1		
$\mathcal{S}(E_{\mathrm{T}}^{\mathrm{miss}})$		≥ 5			≥ 5		

6 Background composition and estimation

The background estimation techniques used in the analysis, combining simulations and data-driven methods common to all channels, are discussed in this section.

Irreducible-background predictions are obtained directly from simulations, but normalization of diboson and rare top quark processes are obtained from the fit. To avoid double counting between background estimates derived from MC simulation and the data-driven reducible background predictions, a specific check is performed. Events from irreducible background MC samples are considered only if generator-level prompt leptons can be associated with their reconstructed counterpart. These events are not used for the data-driven estimation.

There are two reducible background sources: the FNP leptons and events in which at least one lepton charge is misidentified. This latter source of background is relevant only in the Q2 four-lepton signal region

where an event classified under the Q0 category can migrate to the Q2 classification if the charge of one of the leptons is mismeasured.

Charge misidentification for muons is well described by the simulation and occurs only for high momentum muons where detector misalignments degrade the muon momentum resolution. On the contrary, electrons are more susceptible to charge misidentification, as a combination of effects from bremsstrahlung and photon conversions might not be adequately described by the detector simulation. Correction factors (scale factors) accounting for charge misreconstruction are applied to the simulated background events to compensate for inadequate descriptions of bremsstrahlung and photon conversion processes in the simulation. They are derived by comparing the charge misidentification probability measured in data with the one in simulations and are parameterised as function of p_{T} and η. The charge misidentification probability is extracted by performing a likelihood fit in a dedicated $Z \rightarrow ee$ data sample, as described in Ref. [62].

FNP leptons are produced by secondary decays of light- or heavy-flavour mesons into light leptons embedded within jets. Although the b-jet veto and lepton isolation significantly reduce the number of FNP leptons, a fraction still satisfy the selection requirements. For electrons, another significant component of FNP leptons arises from photon conversions, as well as from jets that are misreconstructed as electrons. MC samples are not used to estimate these background sources because the simulation of jet production and hadronisation has large intrinsic uncertainties.

Instead, the FNP background is estimated with a data-driven approach, known as the *fake factor* (FF) method, as described in Ref. [63]. The FF is measured in dedicated FNP-enriched regions with single-lepton triggers without isolation requirements, low $E_{\mathrm{T}}^{\mathrm{miss}}$ contribution, containing no b-jets and having only one reconstructed lepton that satisfies the lepton identification preselection described in Section 5. *Loose* leptons satisfy a looser object selection criteria than the ones used to identify the leptons considered in the analysis regions, which are named *tight* leptons. Electron and muon FFs are then defined as the ratio of the number of tight leptons to the number of loose leptons, and are parameterised as functions of p_{T} and η. The FNP background is then estimated in the SRs by applying FF as an event weight in a template region defined with the same selection criteria as the corresponding SRs, except that at least one of the leptons must satisfy the loose selection but fail to satisfy the tight one. The prompt-lepton contribution is subtracted from the template region using the irreducible-background MC samples to account for the prompt-lepton contamination in the adjacent regions [62]. The Fake-VR, as defined in Table 2, is used to verify the data-driven FNP-lepton estimate.

The final normalisations of the diboson samples are derived separately for the three- and four-lepton channels with a likelihood fit to the data in the ZL CR and Q0 DB-CR, respectively, accounting for the different neutral and charged currents contributions in the two channels. The three- and four-lepton channel diboson normalization factors amount to 0.80 ± 0.09 and 1.08 ± 0.03, respectively. The normalization and shape distributions are validated in the ZL DB-VR and Q0 DB-VR by comparing data and MC distributions after the fit. The rare top quark contribution normalization is estimated to be 1.3 ± 0.2 in the four-lepton channel Q0 RT-CR and is then extrapolated to the rest of the SRs. The background modelling is validated in the ZL RT-VR and Q0 RT-VRs for the three- and four-lepton channels, respectively.

7 Systematic uncertainties

Uncertainties affect this search in different aspects. Experimental uncertainties related to the trigger selection, lepton reconstruction, identification, momentum measurement and isolation selection affect both the global efficiency selection and the shape of the kinematical distributions used in the fit. Uncertainties are extimated mainly from $Z \to \ell\ell$ and $J/\psi \to \ell\ell$ processes [56, 58–61]. Uncertainties on jet energy scales and resolution are evaluated from MC simulations and from data using multijet, Z + jets and γ+jets events [56]; they are estimated to be smaller than 2% in the range of jet momentum of interest and affect both the selection efficiency measurement and kinematical distributions used in the fit. Uncertainties in the b-tagging efficiencies are evaluated from data on dileptonic $t\bar{t}$ events [58–60]. They are estimated to range from 8% at low momentum to 1% at high momentum. These uncertainties affect the analysis region selection efficiencies. $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement uncertainties are based on data to MC comparison in $Z \to \mu\mu$ final states without jets as described in Ref. [61]. $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainties affect both selection efficiencies and kinematical distributions used in the fit. The charge misidentification uncertainty affects only the Q2 analysis regions. The uncertainty in the charge misidentification SF is estimated to be less than 10% from the comparison between data and MC in a dielectron same-sign sample with the same electron selection as used in this analysis, and with $|m_{ee} - m(Z)| < 10\,\mathrm{GeV}$ as described in Ref. [62]. The uncertainty in the pile-up simulation, derived from comparison of data with simulation, is also taken into account [64].

FNP background uncertainty comes from the modelling and normalization of the prompt leptons subtracted in the FF estimation. The origin of the FNP is then varied by selecting slightly modified fake enriched regions, where the FF is measured. The fake enriched regions variation are, for example, obtained varying the jet multiplicity requirement and the $E_{\mathrm{T}}^{\mathrm{miss}}$ selection. The resulting uncertainty of the FF is dependent on lepton momentum and pseudorapity and goes from 5 –40% and 10 –30% for electrons and muons respectively.

The Theoretical uncertainties affect both the signal and background predictions. For both, the uncertainties from missing higher orders are evaluated by varying independently the QCD factorisation and renormalisation scales in the ME by up to a factor of two [65]. The uncertainties on PDFs are evaluated using the `LHAPDF` toolkit [66]. An additional uncertainty of 10% is added to the diboson cross section, to take in account the difference observed in the data to MC agreement of VV processes with different jet multiplicity regions. For rare top quark backgrounds, uncertainties in the $t\bar{t}W$ ($t\bar{t}H$) cross-section include $\pm 12\%$ ($^{+5.8}_{-9.2}\%$) measured by varying the QCD factorisation and renormalisation scales, and $\pm 4\%$ ($\pm 3.6\%$) from PDF+α_{S} variations. Since the yield of the rare top quark and diboson backgrounds is derived from the likelihood fit to the data in the CRs, the systematic variations have little impact on the final yields of the background predictions in the CR and SR.

8 Statistical analysis and results

The HistFitter [68] statistical package is used to fit data to predictions in the CRs and SRs. The fit considers the $m_{T,3\ell}$ and $H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the three- and four-lepton channels, respectively. The binning is chosen to optimize the signal sensitivity. The various components of background predictions are assessed in the corresponding VRs. Background and signal contributions are modelled by a product of independent Poisson probability density functions representing the likelihood of the fit. Systematic uncertainties are modelled as a Gaussian distribution probability density function centred on the pre-fit prediction of the nuisance parameters and widths that correspond to the magnitudes of these uncertainties.

Four different fitting procedures are performed: the three-lepton channel on its own, the four-lepton channel on its own, the three- and four-lepton channel combined together, and finally the two-, three- and four-lepton channel combined, where results for the two-lepton channel are taken from Ref. [11]. All the contributions from the experimental uncertainties in the lepton, jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ selections and reconstructions, pile-up simulation, background simulation, theoretical calculations and irreducible background estimations are considered correlated among the different multiplicity channels in multi-channel fits.

Event yields after the likelihood fit for the analysis regions in the three- and four-lepton channels are shown in Figure 2. Good agreement within statistical and systematic uncertainties between data and SM predictions is observed in all regions, demonstrating the validity of the background estimation procedure.

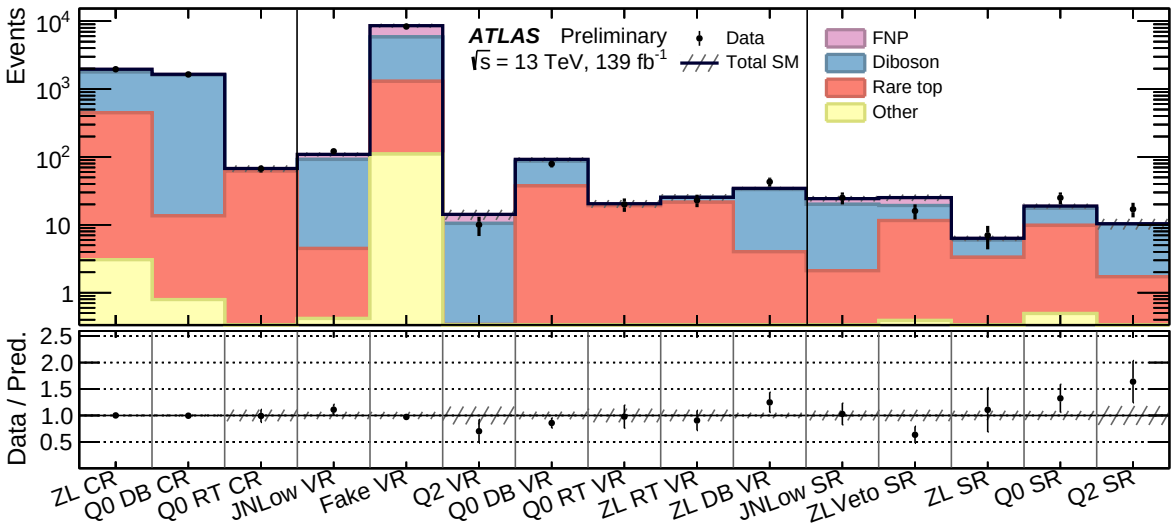
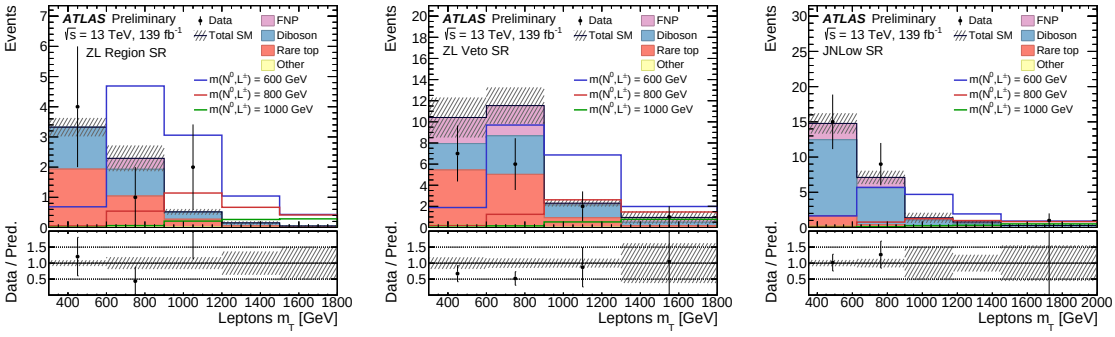


Figure 2: Observed and expected event yields the CRs, VRs and SRs for the three- and four-lepton channels after the fit procedure described in the text. *Diboson* indicates background from diboson processes. *Rare top* indicates background from $t\bar{t}+V$ and tWZ processes. *FNP* includes fake non-prompt leptons contribution. *Other* indicates all other background that contribute less than 2%. The hatched bands include systematic uncertainties with the correlations between various sources taken into account. The lower panel shows the ratio of the observed data to the predicted SM background after the likelihood fit.

The post-fit kinematical distributions of the variables used for the likelihood fit are shown in Figures 3 and 4 for the signal regions in the three- and four-lepton channels, respectively, with the binning used in the fit. After the fit, the compatibility between data and expected background is assessed. Good agreement is observed. The p-value [2] is 0.38 and 0.089 for the three- and four-lepton channels respectively.

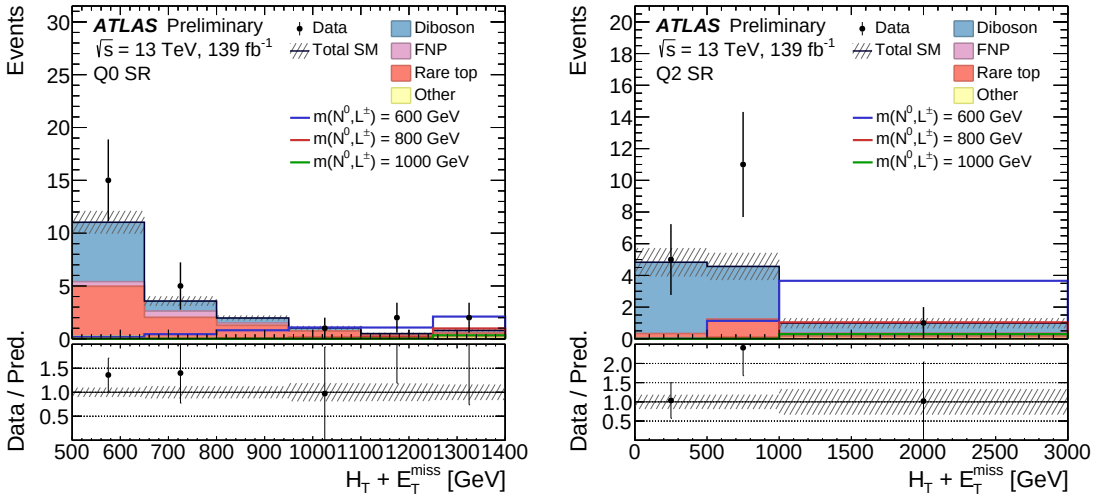
The relative uncertainties in the background yield estimates are shown in Figure 5 for all analysis regions in the three- and four-lepton channels. The dominant uncertainty in the SRs, and in most of the other regions, is the statistical uncertainty in data that, depending on the signal region, varies from 20% to 37%. The MC statistical uncertainty varies from 2% up to 7% instead. In the Q2-SR a contribution close to the data statistical uncertainty comes from the charge misidentified background, considered in the *Experimental* category. Predicted background yields in the SRs are listed together with the observed number of events in data in Table 4.

[2] The p-value is defined as the probability of observing an excess at least as large as the one observed in data, in the absence of signal.



(a) (b) (c)

Figure 3: Distributions of $m_{T,3\ell}$ in the three lepton signal regions, namely (a) the ZL signal region, (b) the ZLVeto signal region and (c) the JNLow signal region, after the background-only fit. The coloured lines correspond to signal samples with the N^0 and $L^\pm$ mass values stated in the legend. The hatched bands include all statistical and systematic post-fit uncertainties with the correlations between various sources taken into account. The lower panel shows the ratio of the observed data to the predicted SM background. The highest bin in the distributions contains the overflow.



(a) (b)

Figure 4: Distributions of $H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$ in the four lepton signal regions, namely (a) the Q0 signal region where the sum of lepton charge is zero and (b) the Q2 signal region where the sum of lepton charge is ±2, after the background-only fit. The coloured lines correspond to signal samples with the N^0 and $L^\pm$ mass values stated in the legend. The hatched bands include all statistical and systematic post-fit uncertainties with the correlations between various sources taken into account. The lower panel shows the ratio of the observed data to the predicted SM background. The last bin in the distributions contains the overflow.

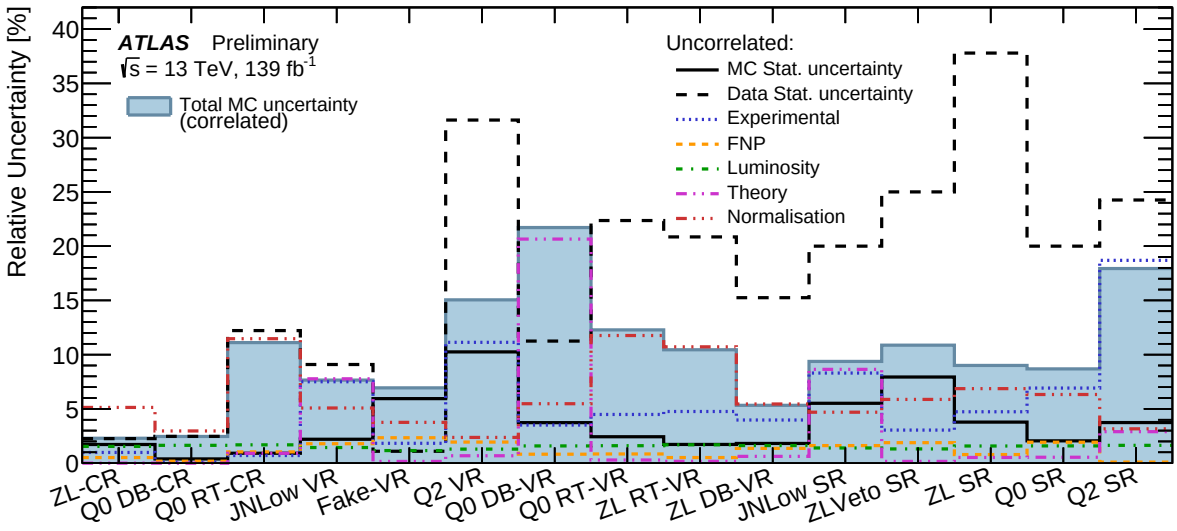


Figure 5: Relative contributions of different sources of statistical and systematic uncertainty in the total background yield estimates after the fit. *Experimental* uncertainties are related to the lepton, jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ selection and reconstruction additionally to lepton charge misidentification. *FNP* includes fake non-prompt leptons contribution. *Luminosity* is related to the luminosity uncertainty that affect the background simulation yields. *Theory* affects the theoretical uncertainties connected to the theory uncertainties like PDF, α_{S}, normalisation and factorization scale. *Normalisation* is related to the diboson and rare top quark normalisation factor extracted by the likelihood fit. Systematic uncertainties are calculated in an uncorrelated way by changing each nuisance parameter from its fit value by one standard deviation, keeping all the other parameters at their central values, and comparing the resulting event yield with the nominal yield. Individual uncertainties can be correlated, and do not necessarily add in quadrature to the total background uncertainty, which is shown as *Total MC uncertainty(correlated)*. *Data Stat. uncertainty* refers to the statistical uncertainty on collected data.

Table 4: Observed data and background yields in the three- and four-lepton signal regions after the background only fit in the combined three- and four-lepton regions.

	Three lepton channels			Four lepton channels	
	ZL SR	ZLVeto SR	JNLow SR	Q0-SR	Q2-SR
Data	7	16	25	25	17
Total background	6.25 ± 0.52	25.2 ± 2.8	24.4 ± 2.3	19.0 ± 1.6	10.3 ± 1.9
Diboson	2.62 ± 0.27	7.64 ± 0.95	18.0 ± 2.1	7.70 ± 0.78	8.5 ± 1.6
Rare Top	3.23 ± 0.46	11.2 ± 1.7	1.82 ± 0.32	9.4 ± 1.4	1.63 ± 0.35
Fakes	0.29 ± 0.05	5.98 ± 0.85	4.3 ± 0.5	1.37 ± 0.36	0.07 ± 0.37
Other	0.11 ± 0.01	0.36 ± 0.12	0.33 ± 0.03	0.49 ± 0.04	0.10 ± 0.01

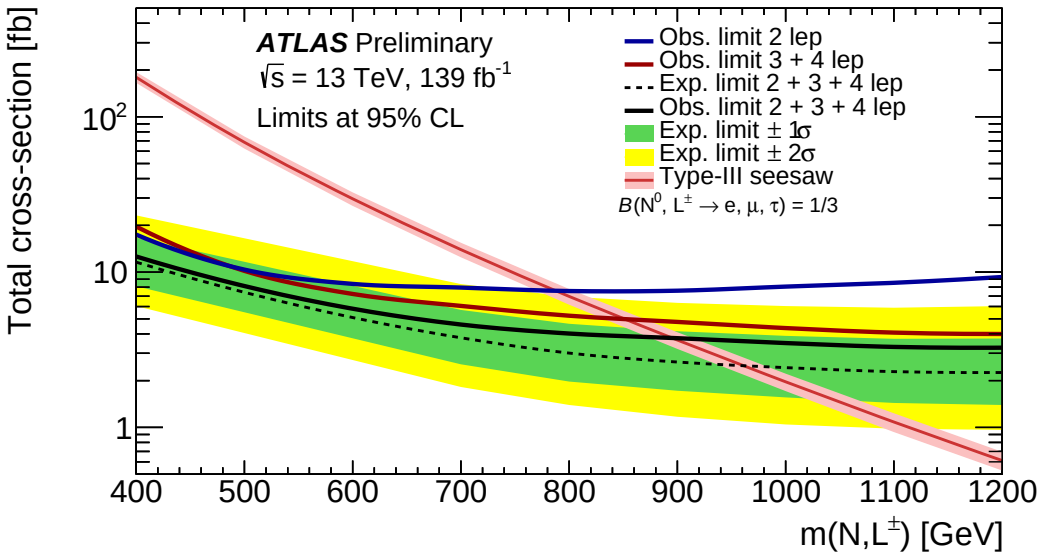


Figure 6: Expected and observed 95% CL_s exclusion limits in the two-lepton channel (from [11]), three- and four-lepton channels, and two-, three- and four-lepton channels for the type-III seesaw process with the corresponding one- and two-standard-deviation uncertainty bands, showing the 95% CL upper limit on the cross-section. The theoretical signal cross-section prediction, given by the NLO calculation [29, 30], with its corresponding uncertainty band is also shown.

In the absence of a significant deviation from SM expectations, 95% confidence level (CL) upper limits on the signal production cross-section are derived using the CL_s method [69]. The upper limits on the production cross-sections of the $pp \to W^* \to N^0 L^\pm$ and $pp \to Z^* \to L^\pm L^\mp$ processes at the 95% CL are evaluated as a function of the heavy lepton mass using the three- and four-lepton channels. By comparing the upper limits on the cross-section with the theoretical cross-section calculation as a function of the heavy-lepton mass, the lower mass limit on the mass of the type-III seesaw heavy leptons N^0 and $L^\pm$ is derived. The observed (expected) exclusion limits are 870 GeV (900^{+80}_{-80} GeV).

The signal hypothesis in the three- and four-lepton channel result is also tested in a combined fit with the similar type-III seesaw search regions in the two lepton channel [11]. All the CRs, VRs and SRs in the various multiplicity regions are statistically independent. The reconstruction algorithms and working points are the same in all cases, and the FNP and lepton charge misidentification backgrounds are estimated by using the same method. The parameter of interest, namely the number of signal events, as well as common systematic uncertainties are treated as correlated. Background normalization of diboson, $t\bar{t}$ (for the two lepton multiplicity region) and rare top quark (for the three- and four-lepton multiplicity regions) are treated as uncorrelated since they account for different physics processes and different acceptances in each final state. The combined exclusion limit is shown in Figure 6. The observed (expected) excluded limits are 910 GeV (960^{+90}_{-80} GeV).

9 Conclusion

ATLAS has searched for pair produced heavy leptons predicted by the type-III seesaw model in the 139 fb^{-1} of data from proton–proton collisions at $\sqrt{s} = 13$ TeV, recorded during the 2015–2018 data-taking period.

The lower mass limit on the mass of the type-III seesaw heavy leptons N^0 and $L^{\pm}$ is derived for final states with three- and four-light leptons. No significant deviation from SM expectations are observed. The observed (expected) exclusion limit on the heavy lepton masses is 870 GeV (900^{+80}_{-80} GeV). The results are combined with those from the two lepton analysis, which used very similar experimental methodologies and treatment of statistics. This combination provides the most stringent limit of the type-III seesaw model at LHC. In the full combination, heavy leptons with masses below 910 GeV are excluded at the 95% CL while the expected lower mass limit is 960^{+90}_{-80} GeV.

References

[1] I. Esteban, M. C. Gonzalez-Garcia, A. Hernandez-Cabezudo, M. Maltoni and T. Schwetz, *Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of θ_{23}, δ_{CP}, and the mass ordering*, JHEP **01** (2019) 106, arXiv: 1811.05487 [hep-ph].

[2] P. Minkowski, *$\mu \to e\gamma$ at a rate of one out of 10^9 muon decays?*, Phys. Lett. B **67** (1977) 421.

[3] T. Yanagida, *Proceedings of the Workshop on Unified Theory and the Baryon Number in the Universe*, KEK Report No. 79-18 **95** (1979), ed. O. Sawada, A. Sugamoto.

[4] S. Glashow, *Proceedings of the 1979 Cargèse Summer Institute on Quarks and Leptons*, 1980.

[5] M. Gell-Mann, P. Ramond and R. Slansky, *Supergravity: Proceedings of the Supergravity Workshop at Stony Brook, New York, 1979*, ed. P. Van Nieuwenhuizen, D. Z. Freedman, 1980.

[6] R. N. Mohapatra and G. Senjanović, *Neutrino Mass and Spontaneous Parity Nonconservation*, Phys. Rev. Lett. **44** (1980) 912.

[7] R. Foot, H. Lew, X. G. He and G. C. Joshi, *See-saw neutrino masses induced by a triplet of leptons*, Z. Phys. C **44** (1989) 441.

[8] ATLAS Collaboration, *Search for type-III seesaw heavy leptons in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, Phys. Rev. D **92** (2015) 032001, arXiv: 1506.01839 [hep-ex].

[9] ATLAS Collaboration, *Search for heavy lepton resonances decaying to a Z boson and a lepton in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **09** (2015) 108, arXiv: 1506.01291 [hep-ex].

[10] CMS Collaboration, *Search for physics beyond the standard model in multilepton final states in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2020) 051, arXiv: 1911.04968 [hep-ex].

[11] ATLAS Collaboration, *Search for type-III seesaw heavy leptons in dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2020) 218, arXiv: 2008.07949 [hep-ex].

[12] C. Biggio and F. Bonnet, *Implementation of the type III seesaw model in FeynRules/MadGraph and prospects for discovery with early LHC data*, Eur. Phys. J. C **72** (2012) 1899, arXiv: 1107.3463 [hep-ph].

[13] A. Arhrib et al., *Collider Signatures for Heavy Lepton Triplet in Type I+III Seesaw*, Phys. Rev. D **82** (2010) 053004, arXiv: 0904.2390 [hep-ph].

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

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

[16] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det].

[17] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: https://cds.cern.ch/record/2677054.

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

[19] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex].

[20] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv: 2004.13447 [hep-ex].

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

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

[23] 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].

[24] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, Comput. Phys. Commun. **185** (2014) 2250, arXiv: 1310.1921 [hep-ph].

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

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

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

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

[29] B. Fuks, M. Klasen, D. R. Lamprea and M. Rothering, *Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV*, JHEP **10** (2012) 081, arXiv: 1207.2159 [hep-ph].

[30] B. Fuks, M. Klasen, D. R. Lamprea and M. Rothering, *Precision predictions for electroweak superpartner production at hadron colliders with RESUMMINO*, Eur. Phys. J. C **73** (2013) 2480, arXiv: 1304.0790 [hep-ph].

[31] R. Ruiz, *QCD Corrections to Pair Production of Type III Seesaw Leptons at Hadron Colliders*, JHEP **12** (2015) 165, arXiv: 1509.05416 [hep-ph].

[32] Y. Cai, T. Han, T. Li and R. Ruiz, *Lepton Number Violation: Seesaw Models and Their Collider Tests*, Front. in Phys. **6** (2018) 40, arXiv: 1711.02180 [hep-ph].

[33] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph].

[34] S. Frixione, P. Nason and G. Ridolfi, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph].

[35] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[36] 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].

[37] 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].

[38] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph].

[39] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph].

[40] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph].

[41] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: 1111.1220 [hep-ph].

[42] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231.

[43] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph].

[44] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph].

[45] 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].

[46] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (2016) 033006, arXiv: 1506.07443 [hep-ph].

[47] L. Harland-Lang, A. Martin, P. Motylinski and R. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, arXiv: 1412.3989 [hep-ph].

[48] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168.

[49] 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.

[50] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv: 1407.5063 [hep-ex].

[51] 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*, Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [hep-ex].

[52] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016, URL: https://cds.cern.ch/record/2157687.

[53] 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].

[54] 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].

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

[56] 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].

[57] 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].

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

[59] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016, URL: https://cds.cern.ch/record/2160731.

[60] 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.

[61] 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].

[62] 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].

[63] ATLAS Collaboration, *Search for anomalous production of prompt same-sign lepton pairs and pair-produced doubly charged Higgs bosons with $\sqrt{s} = 8$ TeV pp collisions using the ATLAS detector*, JHEP **03** (2015) 041, arXiv: 1412.0237 [hep-ex].

[64] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: https://cds.cern.ch/record/1700870.

[65] 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].

[66] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, Eur. Phys. J. C **75** (2015) 132, arXiv: 1412.7420 [hep-ph].

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

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

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

[70] ATLAS Collaboration, *ATLAS Computing Acknowledgements*, ATL-SOFT-PUB-2020-001, URL: https://cds.cern.ch/record/2717821.