Search for heavy neutral leptons in decays of W bosons produced in 13 TeV pp collisions using prompt and displaced signatures with the ATLAS detector


Published in:
JHEP

DOI:
10.1007/JHEP10(2019)265

License
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Search for heavy neutral leptons in decays of $W$ bosons produced in 13 TeV $pp$ collisions using prompt and displaced signatures with the ATLAS detector

The ATLAS collaboration

$E$-mail: atlas.publications@cern.ch

ABSTRACT: The problems of neutrino masses, matter-antimatter asymmetry, and dark matter could be successfully addressed by postulating right-handed neutrinos with Majorana masses below the electroweak scale. In this work, leptonic decays of $W$ bosons extracted from 32.9 fb$^{-1}$ to 36.1 fb$^{-1}$ of 13 TeV proton–proton collisions at the LHC are used to search for heavy neutral leptons (HNLs) that are produced through mixing with muon or electron neutrinos. The search is conducted using the ATLAS detector in both prompt and displaced leptonic decay signatures. The prompt signature requires three leptons produced at the interaction point (either $\mu\mu\mu$ or $e_3e_2\mu$) with a veto on same-flavour opposite-charge topologies. The displaced signature comprises a prompt muon from the $W$ boson decay and the requirement of a dilepton vertex (either $\mu\mu$ or $\mu e$) displaced in the transverse plane by 4–300 mm from the interaction point. The search sets constraints on the HNL mixing to muon and electron neutrinos for HNL masses in the range 4.5–50 GeV.

KEYWORDS: Hadron-Hadron scattering (experiments)

ArXiv ePrint: 1905.09787
1 Introduction

The observation of neutrino oscillations implies that neutrinos are massive [1]. This requires the addition of a neutrino mass-generation mechanism to the Standard Model (SM), which can manifest itself in the form of right-handed neutrinos, Majorana neutrinos, or both. Adding a right-handed Majorana neutrino (denoted by heavy neutral lepton HNL, or simply $N$) gives rise to the so-called Type-1 Seesaw mechanism [2]. An SM neutrino then acquires a mass inversely proportional to the HNL Majorana mass, providing a natural explanation for neutrino masses and why they are so small compared with those of other fermions. Heavy neutral leptons could generate the observed amount of baryon asymmetry in the universe through a process known as leptogenesis [3], and an HNL with a mass of the order of keV would be a valid dark-matter candidate [4]. In fact, the addition of three HNLs with masses below the electroweak scale, two of which are potentially accessible by accelerator-based experiments in the range 0.1–90 GeV [5], could address the three
fundamental questions of the origins of neutrino masses, baryon asymmetry, and dark matter [6, 7]. Meeting these conditions requires small mixing angles between HNLs and neutrinos. Mixing requirements are relaxed if all three HNLs can participate in generating a baryon asymmetry [8, 9], which means however that none of the three HNLs is available as a dark-matter candidate. Depending on the mixing and mass parameters, the HNL may decay promptly or be long-lived. In this paper, searches exploiting both prompt-decay and displaced-decay signatures are reported.

Heavy neutral leptons with masses below 5 GeV can be produced in hadron decays. In this case, sensitivity to small coupling strengths has been achieved by fixed-target experiments with long decay volumes and by high-intensity collider experiments [10–21]. Higher HNL masses can only be directly accessed through the decays of $W$, $Z$ or $H$ bosons, and indirectly through precision tests of the SM. Within some assumptions about the relative HNL mixing angles to the different neutrino flavours, experiments sensitive for processes such as $\mu \rightarrow e\gamma$ or $\mu \rightarrow eee$ can provide indirect constraints which are competitive with direct searches for HNL masses above 30 GeV [22]. An analysis with the DELPHI experiment at LEP1 using $10^6$ neutrinos from $Z$ boson decays provided the best direct constraints prior to the LHC in the HNL mass range 2–75 GeV [23]. At hadron colliders, HNLs are better sought in $W$ rather than $Z$ boson decays due to trigger requirements and the higher production cross section. The CMS Collaboration presented results [24] which explore HNL masses in the range 1–1200 GeV and mixing to muon and electron neutrinos, using a signature of $W$ boson decays into three prompt leptons with lepton-number violation (LNV) similar to the prompt signature presented in this search. The high rate of $W$ boson production at the LHC, combined with the capability of the ATLAS and CMS experiments to reconstruct displaced vertices in their inner detectors, permits HNLs with longer lifetimes to be accessed in regions of the parameter space that are beyond the LEP constraints and allowed by cosmological constraints [25–29].

The present search relies on two distinct experimental signatures designed to probe both short and long HNL lifetimes. These select orthogonal sets of data taken in 2015–2016 with all ATLAS subsystems fully operational for proton–proton ($pp$) collisions at 13 TeV centre-of-mass energy. The dataset used in this analysis corresponds to an integrated luminosity of 36.1 fb$^{-1}$ for the prompt signature and 32.9 fb$^{-1}$ for the displaced signature. The uncertainty in the integrated luminosity is 2.1%, derived from the calibration of the luminosity scale using $x$-$y$ beam-separation scans, following a methodology similar to that detailed in ref. [30], and using the LUCID-2 detector for the baseline luminosity measurements [31].

The prompt signature features three leptons originating from the interaction point, either two muons and an electron or two electrons and a muon, with same-flavour leptons of the same charge. This latter requirement is important for rejecting the large backgrounds from prompt SM processes. The displaced signature features a prompt muon accompanied by a vertex significantly displaced from the interaction point, formed by either two muons or a muon and an electron. The prompt lepton (expected to originate from a $W$ boson decay) is essential for ensuring an efficient triggering of such events and the displaced dilepton decay is very characteristic, rendering this signature virtually background-free.
Figure 1. Feynman diagrams for $N$ production and decay in the channels which this search is sensitive to: (a) $\mu$ mixing, $\mu e$ decay, LNC (probed by displaced signature); (b) $\mu$ mixing, $\mu\mu$ decay, LNC (probed by displaced signature); (c) $\mu$ mixing, $\mu e$ decay, LNV (probed by prompt and displaced signatures); (d) $\mu$ mixing, $\mu\mu$ decay, LNV (probed by displaced signature); (e) $e$ mixing, $e\mu$ decay, LNV (probed by prompt signature). Analogous processes involving the decay of the charge-conjugate $W^-$ boson are also included in the search, but omitted in this figure.
The selection of a prompt muon and a displaced leptonic decay is chosen for the first search of this type because requiring lepton objects (favouring muons as they are less readily affected by misidentification of other objects than electrons) considerably reduces QCD backgrounds, even though the same signature with a prompt electron and/or a displaced semileptonic decay can also be exploited at the LHC [25–27].

The results of this search are interpreted in a model postulating a single right-handed Majorana neutrino $N$ produced in leptonic $W$ boson decays, with just two parameters: mass ($m_N$) and coupling strength ($|U|^2$). The heavy neutrino $N$ is allowed to mix exclusively with either $\nu_\mu$ or $\nu_e$ and to decay leptonically in a scenario with either lepton-number conservation (LNC) or LNV, as illustrated in figure 1. The Majorana nature of $N$ results in equal mixing to neutrinos and anti-neutrinos and the possibility for the process to violate the lepton number (although LNV is not guaranteed [32]), as shown in figures 1c, 1d and 1e.

The range $5 < m_N < 50 \text{ GeV}$ is explored using the prompt signature assuming LNV. The range $4 < m_N < 10 \text{ GeV}$, corresponding to decay lengths of the order mm–cm, is probed down to lower $|U|^2$ values using the displaced signature with HNL mixing to $\nu_\mu$ without any assumption regarding LNC or LNV, as depicted in figures 1a, 1b, 1c and 1d.

2 HNL modelling

This section details $N$ production via mixing with an electron or muon neutrino originating from an on-shell $W$ boson decay, as well as its leptonic decays via the same mixing, as illustrated in figure 1. The generation and simulation of Monte Carlo (MC) signal and background events is presented at the end of the section.

2.1 HNL production

The branching ratio of $W$ boson decays into an $N$ and a charged lepton, $\mathcal{B}(W \rightarrow \ell N)$, is proportional to the mixing matrix squared, or coupling strength, denoted by $|U|^2 = \sum_\ell |U_\ell|^2$, where the terms $U_\ell$ are the matrix elements for $N$ mixing to the different neutrino flavours. The signatures considered in this search are sensitive to mixing to either $\nu_\mu$ or $\nu_e$ and can thus only constrain $|U_\mu|^2$ or $|U_e|^2$ (or potentially $|U|^2$ in the case where either of them is dominant).

The cross section times branching ratio for $W$ boson production and decay into $N$ and a charged lepton $\ell$ can then be expressed as [33]:

$$\sigma(pp \rightarrow W) \cdot \mathcal{B}(W \rightarrow \ell N) = \sigma(pp \rightarrow W) \cdot \mathcal{B}(W \rightarrow \ell \nu) \cdot |U|^2 \left(1 - \frac{m_N^2}{m_W^2}\right)^2 \left(1 + \frac{m_N^2}{2m_W^2}\right).$$

The product of the cross section for $W$ boson production in 13 TeV $pp$ collisions $\sigma(pp \rightarrow W)$ and the branching ratio for leptonic $W$ boson decay into a single lepton flavor $\mathcal{B}(W \rightarrow \ell \nu)$ (for $\ell = \mu$ or $e$) is taken from the ATLAS measurement in ref. [34] to be $20.6 \pm 0.6 \text{ nb}$.

2.2 HNL decay

For this search, partial widths are calculated for all HNL decay channels including leptons and quarks. The calculations consider charged- and neutral-current-mediated interactions...
as well as QCD loop corrections, which are all described in ref. [35]. The HNL lifetime $\tau_N$ has a strong dependence on the coupling strength $|U|^2$ and also the mass $m_N$ due to phase-space effects. For a given $|U|^2$ and $m_N$, the total width $\Gamma = \sum_i \Gamma_i (m_N, |U|^2)$ is computed, and the mean lifetime is obtained as $\tau_N = \hbar/\Gamma$. In the relevant range $4.5 \leq m_N \leq 50 \text{ GeV}$, the result agrees within 2% with the following parameterisations given in ref. [36]: $\tau_{N_\mu} = (4.49 \cdot 10^{-12} \text{ s}) |U|^2 (m_N/1 \text{ GeV})^{-5.19}$ and $\tau_{N_e} = (4.15 \cdot 10^{-12} \text{ s}) |U|^2 (m_N/1 \text{ GeV})^{-5.17}$ for dominant mixing to $\nu_\mu$ and $\nu_e$, respectively. These relationships, however, assume no LNV decays. If LNV is allowed, twice as many decay channels are allowed, and $\tau_N$ is reduced by a factor of 2. More elaborate models do not necessarily allow for LNV [32] and thus may or may not contain this factor of 2. To account for this model dependence, both interpretations are considered in the case of the displaced signature, which is not limited to LNV processes.

Leptonic HNL decay branching ratios are determined from the partial decay widths relative to the total width. In the mass range $4.5-50 \text{ GeV}$, they have almost no mass dependence and yield $\mathcal{B}(N \rightarrow \ell \nu_\ell) = 0.060$ and $\mathcal{B}(N \rightarrow \ell' \nu_{\ell'}) = 0.106$ for dominant mixing to a given lepton species $\ell = \mu$ or $e$ ($\ell \neq \ell'$, including both charges). The difference between decays into leptons of the same flavour and different flavour is due to interference between decays through $W$ and $Z$ boson mediators, which is only present in the same-flavour case. This calculation and calculations found in the literature [5, 36, 37] can yield up to 5% relative differences, mainly due to different treatments of QCD corrections. This 5% difference is considered as a relative theoretical systematic uncertainty in the branching ratio.

### 2.3 Event generation and simulation

The signal MC samples were generated with a model of $W$ boson production in 13 TeV $pp$ collisions, with the $W$ exclusively decaying into a muon or electron and an HNL ($W \rightarrow \mu N$ or $W \rightarrow e N$). Separate signal samples are generated for multiple choices of the HNL mass, $m_N = 4.5 \text{ GeV}$, $5 \text{ GeV}$, $7.5 \text{ GeV}$, $10 \text{ GeV}$, $12.5 \text{ GeV}$, $15 \text{ GeV}$, $20 \text{ GeV}$, $30 \text{ GeV}$ or $50 \text{ GeV}$, and the mean proper decay length, $c\tau = 0.001 \text{ mm}$, $0.01 \text{ mm}$, $0.1 \text{ mm}$, $1 \text{ mm}$, $10 \text{ mm}$ or $100 \text{ mm}$, with $\sim 50000$ events per sample. The $N$ decay modes correspond to the diagrams shown in figure 1. This model is implemented in Pythia8 [38] (v8.210), using the NNPDF2.3 LO [39] parton distribution function (PDF) set and the A14 set of tuned parameters (tune) [40].

Background processes were generated using Powheg-Box [41–43] with the next-to-leading order (NLO) CT10 PDF set [44] for top-quark pair ($tt$) (using v2 in r3026) and single top-quark (using r2856) production, in combination with Pythia [45] (v6.428, for parton showering) using the CTEQ6L1 PDF set [46] and Perugia 2012 tune [47]. MadGraph5_AMC@NLO [48] (v2.2.2.p6) with the NNPDF2.3 LO PDF set and A14 tune was used for $ttW$ and $ttZ$, while Sherpa [49] (v2.2.1) with the NNPDF2.3 LO PDF set was used for $WW$, $ZZ$, $WZ$, $W+$jets, $Z+$jets and tribosons. The purely leptonic decays of dibosons were generated using Powheg-Box v2 (r2819) with NLO CT10 PDF set in combination with Pythia8 using the CTEQ6L1 PDF set with AZNLO tune [50]. Together with the Pythia8, Powheg-Box and MadGraph5 event generators, EvtGen [51] (v1.2.0) was used for $b$- and $c$-hadron decays. Also, together with Powheg-Box in processes involving
$t$-quarks, TAUOLA [52] (v27feb06) was used for $\tau$-lepton decays and PHOTOS [53] (v2.09) was used for QED corrections.

For the generation of $tt$ events, the top-quark mass was set to 172.5 GeV. The sample is normalised using the next-to-next-to-leading-order (NNLO) cross section, including soft-gluon resummation to next-to-next-to-leading-logarithm (NNLL) [54–60]. For events containing a $W$ or $Z$ boson with associated jets simulated using SHERPA, matrix elements were generated with up to two additional parton emissions at NLO accuracy and up to four additional parton emissions at LO accuracy using COMIX [61] and OPENLOOPS [62] and merged with the SHERPA parton shower [63] according to the ME+PS@NLO prescription [64]. Diboson processes with one of the bosons decaying hadronically and the other leptonically are calculated for up to one ($ZZ$) or zero ($WW$, $WZ$) additional partons at NLO and up to three additional partons at LO using COMIX and OPENLOOPS, and merged with the SHERPA parton shower according to the ME+PS@NLO prescription. Matrix elements of triboson processes containing two or more charged leptons were generated with SHERPA including off-shell contributions with up to one additional real emission at LO accuracy [65].

The generated events were processed through a full simulation of the ATLAS detector geometry and response [66] using the Geant4 [67] toolkit. The simulation includes multiple $pp$ interactions per bunch crossing (pile-up), as well as the effect on the detector response due to interactions from bunch crossings before or after the one containing the hard interaction. Pile-up was simulated with the soft processes of PYTHIA8 using the A2 tune [40] and the MSTW2008LO PDF [68]. Per-event weights were applied to the simulated events to reproduce the distribution of the average number of interactions per bunch crossing as measured in data.

3 The ATLAS detector

The ATLAS experiment [69–71] at the LHC is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and a nearly $4\pi$ coverage in solid angle. The detector consists of several layers of subdetectors. From the interaction point (IP) outwards there is an inner tracking detector (ID), electromagnetic and hadronic calorimeters, and a muon spectrometer (MS).

The ID extends from a cylindrical radius of about 33 to 1100 mm and to $|z|$ of about 3100 mm, and is immersed in a 2 T axial magnetic field. It provides tracking for charged particles within the pseudorapidity region $|\eta| < 2.5$. At small radii, silicon pixel layers and stereo pairs of silicon microstrip detectors provide high-resolution position measurements. The pixel system consists of four barrel layers, and three forward discs on either side of the IP. The barrel pixel layers are positioned at radii of 33.3, 50.5, 88.5 and 122.5 mm. The silicon microstrip tracker (SCT) comprises four double layers in the barrel and nine

\footnote{ATLAS uses a right-handed coordinate system with its origin at the nominal 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 upward. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$.}
forward discs on either side. The radial position of the innermost (outermost) SCT barrel layer is 299 mm (514 mm). The final component of the ID, the transition-radiation tracker (TRT), is positioned at larger radii, with coverage up to $|\eta| = 2.0$.

The electromagnetic calorimeter is based on lead absorbers and liquid argon and provides coverage for $|\eta| < 3.2$. Hadronic calorimetry uses steel absorbers and scintillator tiles in the region $|\eta| < 1.7$, and copper absorbers with liquid argon in the endcaps ($1.5 < |\eta| < 3.2$). A forward hadronic calorimeter using copper and tungsten absorbers with liquid argon completes the calorimeter coverage up to $|\eta| = 4.9$.

The MS is the outermost ATLAS subdetector. It is designed to detect muons in the pseudorapidity region up to $|\eta| = 2.7$, and to provide momentum measurements with a relative resolution better than 3% over a wide $p_T$ range and up to 10% at $p_T \approx 1$ TeV. The MS consists of one barrel ($|\eta| < 1.05$) and two endcap sections ($1.05 < |\eta| < 2.7$). A system of three large superconducting air-core toroidal magnets, each with eight coils, provides a magnetic field with a bending integral of about 2.5 Tm in the barrel and up to 6 Tm in the endcaps. Resistive plate chambers (three doublet layers for $|\eta| < 1.05$) and thin gap chambers (one triplet layer followed by two doublets for $1.0 < |\eta| < 2.4$) provide triggering capability to the detector as well as ($\eta, \phi$) position measurements with typical spatial resolution of 5 – 10 mm. A precise momentum measurement for muons with pseudorapidity up to $|\eta| = 2.7$ is provided by three layers of monitored drift tube chambers (MDT). For $|\eta| > 2$, the inner layer is instrumented with a quadruplet of cathode strip chambers (CSC) instead of MDTs. The single-hit resolution in the bending plane for the MDT and the CSC is about 80 $\mu$m and 60 $\mu$m, respectively. The muon chambers are aligned with a precision between 30 $\mu$m and 60 $\mu$m.

A two-level trigger system is used to select events [72]. The first-level trigger is implemented in custom electronics and uses information from the muon trigger chambers and the calorimeters. This is followed by a software-based high-level trigger system, which runs reconstruction algorithms similar to those used in offline reconstruction. Combined, the two levels reduce the 40 MHz bunch-crossing rate to approximately 1 kHz of events saved for further analysis.

## 4 Prompt-trilepton signature

The prompt-lepton search for HNLs is conducted in two channels: $W^\pm \rightarrow \mu^\pm \mu^\pm e^\mp \nu_e$ (muon channel) and $W^\pm \rightarrow e^\pm e^\mp \mu^\mp \nu_\mu$ (electron channel). It considers the case where the vertex displacement is small enough that an ID track can be reconstructed with standard ATLAS tracking algorithms. The standard reconstruction of tracks in the ID is optimised for charged particles that originate from the beam spot, the region where the proton beams intersect. This set-up restricts the detection for decay products of particles whose decay vertex is significantly displaced from the beam spot, especially for transverse displacements greater than approximately 20 mm [73]. By requiring the final state to have three isolated leptons and no opposite-charge same-flavour lepton pairs, background events from Drell-Yan pair production, $W$+jets and $t\bar{t}$ could be rejected.
4.1 Trigger and preselection (prompt signature)

Events are required to have a primary vertex defined as the reconstructed vertex with the largest sum of squared track transverse momenta. For the muon channel, events are selected with a dimuon trigger with transverse momentum $p_T$ thresholds of 18 and 8 GeV for events recorded in 2015 and of 22 and 8 GeV for events recorded in 2016. For the electron channel, events are selected with a logical OR between single-electron triggers with different transverse energy thresholds and different electron identification criteria: either 24 GeV (“medium”), 60 GeV (“medium”) or 120 GeV (“loose”) for the 2015 dataset and either 26 GeV (“tight”), 60 GeV (“medium”) or 140 GeV (“loose”) for the 2016 dataset. The different identification criteria “loose”, “medium” and “tight” are defined in ref. [74].

A further trigger match requirement is imposed between the reconstructed leptons and the corresponding triggers.

Muon candidates are reconstructed from tracks in the MS, matched with tracks found in the ID within $|\eta| < 2.5$. Electron candidates are reconstructed from energy deposits (clusters) in the electromagnetic calorimeter which are associated with a reconstructed track in the ID, within the fiducial region of $|\eta| < 2.47$, where $\eta$ is the pseudorapidity of the calorimeter energy deposit associated with the electron candidate. Electron candidates within the transition region between the barrel and endcap electromagnetic calorimeters, $1.37 < |\eta| < 1.52$, are excluded. Muons are required to have a $p_T$ of at least 4 GeV, while the lowest $p_T$ threshold for electrons is 4.5 GeV for 2016 data and 7 GeV for 2015 data. Both the muon and electron candidates are required to satisfy “loose” sets of identification criteria [74, 75] while for the electron channel the electron that passed the single-electron trigger needs to satisfy “tight” identification criteria for events in the dataset collected in 2016. The highest-$p_T$ (leading) lepton is required to pass a transverse impact parameter significance requirement, $d_0/\sigma(d_0) < 3$ for muons and $< 5$ for electrons, where the impact parameter $d_0$ is the transverse distance between the primary vertex and the point of closest approach of the lepton trajectory. For other leptons, no such requirement is imposed to allow for a slight displacement of leptons from HNLs. The leading lepton is also required to have $|\Delta z_0 \sin(\theta)| < 0.5$ mm while $< 1$ mm is required for the remaining leptons, where $\Delta z_0$ is the distance along the beam direction between the primary vertex and the point of closest approach of the lepton trajectory. To reduce the contribution from non-prompt leptons (e.g. from semileptonic $b$- or $c$-hadron decays), photon conversions and hadrons, “loose” calorimeter and track isolation criteria as defined in refs. [74, 75] are applied to lepton candidates, with a 99% efficiency. Scale factors are applied to simulated lepton distributions to take into account the small differences in reconstruction, identification, isolation, and trigger efficiencies between MC simulation and data. Energy/momentum scale and resolution corrections are also applied to leptons [76].

Jets are reconstructed from energy deposits in the calorimeter using the anti-$k_T$ algorithm [77] with a radius parameter value of $R = 0.4$. A multivariate technique (Jet Vertex Tagger) [78] is applied in order to identify and select jets originating from the hard-scatter interaction, at a working point corresponding to a 92% efficiency for identifying such jets and allowing an observed fake rate of 2% from pile-up jets. Jets are also calibrated
using energy- and $\eta$-dependent corrections [79]. Only jets with a transverse momentum $p_T > 20$ GeV and $|\eta| < 4.5$ are considered in the following. In order to identify jets containing $b$-hadrons (referred to as $b$-jets in the following), the MV2c10 algorithm is used, which combines impact parameter information with the explicit identification of secondary and tertiary vertices within the jet into a multivariate discriminant [80]. Operating points are defined by a single threshold in the range of the discriminant output values and are chosen to provide a specific $b$-jet efficiency in simulated $t\bar{t}$ events. The working point used in this analysis has an efficiency of 77% for $b$-jet tagging. It has rejection factors of 6, 22, and 134 against $c$-jets, hadronic decays of $\tau$-leptons, and jets from light quarks or gluons, respectively. A scale factor is applied to account for $b$-jet tagging efficiency differences between MC simulation and data [81].

The magnitude of the missing transverse momentum, $E_T^{\text{miss}}$ [82], is reconstructed from the negative vector sum of transverse momenta of reconstructed and calibrated particles and jets. The reconstructed particles are electrons, photons, $\tau$-leptons and muons. Additionally, there is a second contribution calculated from ID tracks that are matched to the primary vertex and not associated with any of the selected objects (soft objects).

To avoid assigning a single detector response to more than one reconstructed object, a sequential overlap-removal procedure is adopted. Jets are removed if found to be within $\Delta R = 0.2$ of an electron or muon track, unless they satisfy the $b$-tagging requirements. In the electron channel, jets are not removed if their $p_T$ is at least 20% higher than that of the electron. In the muon channel, a jet is not removed if it has at least three tracks originating from the primary vertex. Electrons or muons within a sliding-size cone around the remaining jets, defined as $\Delta R = \min\{0.04 + 10 \text{ GeV}/p_T(\ell), 0.4\}$, are rejected. Muons that can be matched to an energy deposit in the calorimeter (calorimeter-tagged muons) are removed if they share tracks with an electron. Electrons are then removed if they share tracks with one of the remaining muons.

4.2 Reconstruction and selection (prompt signature)

All selected events are required to contain leptons which must satisfy flavour and charge requirements, such that the event consists of exactly $\mu^+\mu^+e^+$ in the muon channel and $e^+e^+\mu^+$ for the electron channel. Furthermore, a requirement is imposed on the three-lepton invariant mass constructed from the three final-state leptons: $40 < m(\ell, \ell', \ell'') < 90$ GeV. Its distribution for signal events is centred below the mass of the $W$ boson, as the neutrino escapes undetected.

If the mass of the HNL is smaller than half the mass of the $W$ boson, the leading lepton will generally originate from the $W$ boson decay and the other leptons from the $N$ decay. Accordingly, the dilepton invariant mass $m(\ell, \ell')$ is determined as the invariant mass of the $e\mu$ combination which excludes the leading $e$ or $\mu$ in the electron and muon channels, respectively. Its distribution is centred below the mass of the HNL.

In the muon channel, both the higher-$p_T$ and lower-$p_T$ (subleading) muons are required to have $p_T > 23$ GeV and $p_T > 14$ GeV, respectively. In the electron channel, the leading and subleading electrons are required to have a transverse momentum $p_T > 27$ GeV and $p_T > 10$ GeV, respectively. Additionally, to reject the large number of $Z$+jets background
Muon channel | Electron channel
---|---
exactly $\mu^\pm \mu^\mp e^\mp$ signature | exactly $e^\pm e^\mp \mu^\mp$ signature

- $p_T(\mu) > 4$ GeV
- $p_T(e) > 7$ GeV (2015), 4.5 GeV (2016)

leading muon $p_T > 23$ GeV | leading electron $p_T > 27$ GeV
subleading muon $p_T > 14$ GeV | subleading electron $p_T > 10$ GeV
$m(e,e) < 78$ GeV

- $40 < m(\ell,\ell,\ell') < 90$ GeV
- $b$-jet veto
- $E_T^{\text{miss}} < 60$ GeV

Table 1. Signal region selection criteria for the prompt trilepton analysis.

<table>
<thead>
<tr>
<th>Channel</th>
<th>$m_N = 5$ GeV</th>
<th>$m_N = 10$ GeV</th>
<th>$m_N = 20$ GeV</th>
<th>$m_N = 30$ GeV</th>
<th>$m_N = 50$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muon</td>
<td>$(0.6\pm0.1)%$</td>
<td>$(1.8\pm0.2)%$</td>
<td>$(6.8\pm0.4)%$</td>
<td>$(8.8\pm0.5)%$</td>
<td>$(9.9\pm0.5)%$</td>
</tr>
<tr>
<td>Electron</td>
<td>$(0.3\pm0.1)%$</td>
<td>$(1.8\pm0.2)%$</td>
<td>$(6.9\pm0.4)%$</td>
<td>$(7.9\pm0.5)%$</td>
<td>$(5.1\pm0.3)%$</td>
</tr>
</tbody>
</table>

Table 2. Prompt-trilepton signal efficiencies for muon and electron channels after applying all selection criteria. The uncertainties are statistical only.

events that contain electrons with a misidentified charge, the invariant mass of the electrons must satisfy $m(e,e) < 78$ GeV. In both channels, the events must not have $b$-tagged jets and the $E_T^{\text{miss}}$ value must be less than 60 GeV. A summary of the signal region selection criteria is given in table 1.

The impact of the final event selection on the efficiency for the signal samples is shown in table 2. The selection efficiency is defined as the fraction of generated events that were reconstructed and satisfied the selection criteria for the muon and electron channel $(pp \to W(\mu)N(\to \mu e\nu))$ or $(pp \to W(e)N(\to e\mu\nu))$. The black triangles in figure 2 show the typical dependence of the final selection efficiency in the muon channel on the mean proper decay length of the HNL. The efficiency is constant for mean decay length up to about 0.1 mm, beyond which it decreases due to a sharp reduction in the efficiency of the standard tracking algorithm. This illustrates the complementarity with the displaced-vertex search described in section 5.

4.3 Backgrounds and signal extraction (prompt signature)

The SM backgrounds that can lead to the same signature as the one from the prompt heavy neutral lepton are a mixture of prompt real leptons and non-prompt leptons and leptons from pile-up. These backgrounds can be split into two broad categories, irreducible and reducible types. The irreducible background is composed of exactly three leptons, where
Figure 2. HNL search event selection efficiency as a function of mean proper decay length evaluated from simulation. The efficiency for the prompt signature in the muon channel is shown for an HNL mass of 10 GeV, while the efficiencies for the displaced signature are shown for four different HNL masses. Error bars represent MC statistical uncertainties.

the only SM sources are diboson and triboson events as well as $t\bar{t}V$ ($V = W, Z$). These backgrounds are negligible due to the small cross section of these processes in this selection.

Reducible background events contain what is referred to as fake leptons. In the case of electrons, these include contributions from semileptonic decays of $b$- and $c$-hadrons, photon conversions and jets with large electromagnetic energy (from hadronisation into $\pi^0$’s or from early showering in the calorimeter). In the case of muons, they can originate from semileptonic decays of $b$- and $c$-hadrons, from charged hadron decays in the tracking volume, or from punch-through particles emerging from high-energy hadronic showers. A large fraction of such events originates from $W+\text{jets}$ and multijets events, in the following referred to as multi-fake events. This background is determined from data.

Other subdominant contributions with two or more real leptons such as $Z+\text{jets}$, single-top-quark, diboson ($WW$, $WZ$ and $ZZ$) and triboson events with hadronic decays are evaluated with simulation. The dominant $t\bar{t}$ background is derived from CRs in data, with shape extracted from MC simulation as described below.

To properly describe the backgrounds, a simultaneous binned maximum-likelihood fit is performed in three control regions (CR) and the signal region (SR). Normalisation factors are obtained for the most dominant MC background $t\bar{t}$ ($t\bar{t}$) and the multi-fake background ($\mu_{\text{mf}}$).

As described in section 4.2, the SR is defined by selecting events with four criteria: 0) Same-charge and same-flavour (SCSF) lepton pairs, 1) $40 < m(\ell, \ell') < 90$ GeV, 2) A $b$-jet veto and 3) $E_{T}^{\text{miss}} < 60$ GeV. The three CRs used are obtained by inverting one of the requirements of the SR definition (CR-1, CR-2 and CR-3). The maximum-likelihood fit uses the $p_T$ distribution of the third lepton for the three CRs, which provides good separation of the different background sources, while the $m(\ell, \ell')$ distribution is used for the SR, as it provides good separation of the background from the various signal models.

An overview of the criteria is given in table 3.
<table>
<thead>
<tr>
<th>Criterion</th>
<th>Signal region</th>
<th>Control region</th>
<th>Fit distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>exactly one SCSF lepton pair</td>
<td>(m(\ell, \ell'))</td>
<td>(m(\ell, \ell'))</td>
</tr>
<tr>
<td>1</td>
<td>(40 &lt; m(\ell, \ell') &lt; 90) GeV</td>
<td>(m(\ell, \ell') \leq 40) GeV</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>b-jet veto</td>
<td>at least one b-jet</td>
<td>(p_T(\ell'))</td>
</tr>
<tr>
<td>3</td>
<td>(E_T^{\text{miss}} &lt; 60) GeV</td>
<td>(E_T^{\text{miss}} \geq 60) GeV</td>
<td>(p_T(\ell'))</td>
</tr>
</tbody>
</table>

Table 3. Signal and control region selection criteria for the prompt HNL analysis and the corresponding distribution used in the binned maximum-likelihood fit in the SR (criterion 0) and the three CRs (criteria 1–3). In addition, estimation regions corresponding to the SR and the three CRs are defined by requiring all leptons to have the same charge. Criteria 0–3 are all used for the SR. Only one of them is inverted to define the corresponding CR.

Three estimation regions (ER) are constructed the same way as the CRs except that all three leptons should have the same charge. In addition, a fourth estimation region is defined with only the requirement of all leptons having the same charge (and the rest of the SR requirements). Additionally, isolation requirements on the leptons are loosened in all ERs to increase the number of events in data and also the number of fake leptons. The shape and initial normalisation of the multi-fake background are measured in these estimation regions as the difference between the data and the simulated backgrounds that contain two or more prompt real leptons. The third-lepton \(p_T\) distribution is used for the three estimation regions corresponding to the three CRs, while the \(m(\ell, \ell')\) distribution is used for the estimation region corresponding to the SR. The resulting statistical uncertainty dominates the multi-fake background uncertainties and is propagated as a bin-by-bin systematic uncertainty in the final maximum-likelihood fit.

A simultaneous binned maximum-likelihood fit is performed to data in order to obtain normalisation factors for the dominant \(t\bar{t}\) background (\(\mu_{t\bar{t}}\)) and the derived multi-fake background (\(\mu_{\text{mf}}\)). A value \(\mu_{t\bar{t}} = 1.0\) would imply no normalisation change relative to the MC prediction, while \(\mu_{\text{mf}} = 1.0\) would imply the same normalisation of the multi-fake background as was determined in the estimation regions. Table 4 shows the resulting normalisation factors from a fit under the background-only hypothesis using only the control regions CR-1 to CR-3, as well as using these same three control regions together with the signal region. The statistical uncertainty associated with multi-fakes in the control regions for the electron channel is large, giving enough freedom for the fit to estimate the normalisation factor. After adding the signal region, the total statistical uncertainty is smaller because more data are added, constraining the normalisation factor. In the muon channel, the statistical uncertainty is lower, reducing the impact of adding the signal region.

The total uncertainty in the expected signal yield is presented for the different signal points in table 5 using the post-fit systematic uncertainties. The relative statistical uncertainty arising from the limited number of events in the MC samples and in the CRs for the multi-fake background estimation is shown in parentheses. Systematic uncertainties do not have a large impact on the signal yield except for the lowest masses. The most prominent systematic uncertainties are related to the variation of the \(p_T\) resolution correction of
Table 4. Normalisation factors obtained for multi-fake $\mu_{nf}$ and $t\ell$ $\mu_{t\ell}$ in different fit regions after the background-only fit. Shown are full post-fit uncertainties of the normalisation factors taking into account all statistical and systematic uncertainties.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Fit configuration</th>
<th>$\mu_{nf}$</th>
<th>$\mu_{t\ell}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muon</td>
<td>only CRs</td>
<td>0.97(38)</td>
<td>0.90(14)</td>
</tr>
<tr>
<td></td>
<td>CRs + SR</td>
<td>1.48(34)</td>
<td>0.88(13)</td>
</tr>
<tr>
<td>Electron</td>
<td>only CRs</td>
<td>0.42(92)</td>
<td>1.02(16)</td>
</tr>
<tr>
<td></td>
<td>CRs + SR</td>
<td>0.81(28)</td>
<td>0.99(15)</td>
</tr>
</tbody>
</table>

Table 5. Prompt-trilepton relative uncertainty of signal yields for muon and electron channels after applying the selection criteria. Uncertainties correspond to those obtained after the fit. Uncertainties associated to the limited number of events in the MC samples and in the CRs for the multi-fake background estimation are shown in parentheses.

<table>
<thead>
<tr>
<th>Channel</th>
<th>$m_N = 5\text{GeV}$</th>
<th>$m_N = 10\text{GeV}$</th>
<th>$m_N = 20\text{GeV}$</th>
<th>$m_N = 30\text{GeV}$</th>
<th>$m_N = 50\text{GeV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$c\tau = 1\text{mm}$</td>
<td>$c\tau = 1\text{mm}$</td>
<td>$c\tau = 0.1\text{mm}$</td>
<td>$c\tau = 10\mu\text{m}$</td>
<td>$c\tau = 1\mu\text{m}$</td>
</tr>
<tr>
<td>Muon</td>
<td>19% (15%)</td>
<td>8.1% (6.7%)</td>
<td>4.2% (4.1%)</td>
<td>4.3% (3.9%)</td>
<td>3.5% (3.3%)</td>
</tr>
<tr>
<td>Electron</td>
<td>14% (14%)</td>
<td>7.8% (7.6%)</td>
<td>4.8% (4.1%)</td>
<td>4.6% (3.6%)</td>
<td>4.7% (3.6%)</td>
</tr>
</tbody>
</table>

For all simulated backgrounds, systematic uncertainties related to the reconstruction and identification of leptons and jets as well as the missing transverse momentum are applied. Uncertainties associated with charged leptons arise from the trigger selection, the object reconstruction, the isolation criteria, as well as the lepton momentum scale and resolution [74, 75]. Uncertainties associated with jets arise from the jet reconstruction and identification efficiencies related to the jet energy scale (JES) and jet energy resolution, and from the Jet Vertex Tagger efficiency [79]. The JES-related uncertainties contain 23 components which are treated as statistically independent and uncorrelated. Some of these components are related to jet flavour, pile-up corrections, $\eta$ dependence and high-$p_T$ jets. The uncertainties associated with $E_T^{\text{miss}}$ are propagated from the uncertainties in the reconstruction of leptons and jets since they are used for the $E_T^{\text{miss}}$ calculation. Uncertainties due to soft objects are also considered [82]. Additional cross-section uncertainties of the various backgrounds estimated in MC simulation are considered. The systematic uncertainties of the backgrounds are smaller than the statistical uncertainty in the phase space selected by this prompt HNL analysis. The most prominent systematic uncertainties are related to the identification of the analysis objects, namely (relative uncertainty in parentheses) calibration variations of the jet energy resolution and scale (6%), jet pseudorapidity inter-calibration (5%) and efficiency scale factor of $b$-jet tagging (2%) [81]. Other systematic uncertainties such as the variation of the $p_T$ resolution correction of muon tracks in the MS and ID and the variation of the energy scale for
the calibration of electrons [76] account for systematic effects smaller than $\sim 1\%$. A 50\% systematic uncertainty is also applied to the yield of $Z+$jets events, accounting for possible mismodelling of non-prompt leptons in the simulation.

Systematic uncertainties are parameterised by nuisance parameters with Gaussian constraints in the likelihood fit. These nuisance parameters are shared between all samples, while statistical uncertainties are modelled using Gamma distributions. The statistical uncertainty of the total MC-based background is treated with a single nuisance parameter for each bin in the likelihood [83, 84].

The third-lepton $p_T$ distributions for each control region using the post-fit normalisation factors as measured in control and signal regions combined are shown in figures 3 and 4 for the muon and electron channels, respectively. The number of multi-fake events in the electron channel is about a third of the ones in the muon channel. When performing the CR-only fit, this is reduced further by a factor of about two, yielding a large uncertainty in the electron channel $\mu_{\text{MC}}$ CR-only result. The invariant mass distributions of the second and third leptons in the signal region for the signal-plus-background post-fit are shown in figure 5. This variable provides a good signal and background separation as most signal mass points concentrate in the first three bins, also allowing for distinction between different signal mass values.

The difference between data and the multi-fake estimate in the lowest bin of CR-3 (see figure 3, right) is responsible for a change of $\mu_{\text{data}}$ by $1.3\sigma$ when including the SR in the fit for the muon channel (see table 4). This was found to be anti-correlated with the statistical uncertainty parameterisations of MC and multi-fake backgrounds in the lowest bin of CR-3, as a consequence of a statistical fluctuation of data in the low-$p_T$ spectrum of this region. Further checks were conducted, including a $p$-value test in the signal region using normalisation factors as measured in the CRs. The $p$-values are determined without taking into account the fit parameters associated with statistical uncertainty in the SR and found to be 7.2\% for the muon channel and 13.2\% for the electron channel. As this lies above the usual rejection level of 5\%, the test is considered satisfactory. Upon using the background-only hypothesis normalised in both CRs and SR very good compatibility was found.

Figure 3. Third-lepton $p_T$ distributions in the control regions CR-1 (left), CR-2 (middle) and CR-3 (right) for the prompt HNL analysis in the muon channel, showing post-fit background-only hypothesis including all the uncertainties and normalised in both CRs+SR. The total uncertainty in the background is shown as dashed regions.
Figure 4. Third-lepton $p_T$ distributions in the control regions CR-1 (left), CR-2 (middle) and CR-3 (right) for the prompt HNL analysis in the electron channel, showing post-fit background-only hypothesis including all the uncertainties and normalised in both CRs+SR. The total uncertainty in the background is shown as dashed regions.

Figure 5. Data-to-prediction comparison of dilepton invariant mass distributions in the signal regions for the prompt signature in the muon (left) and electron (right) channels. The prediction including its uncertainty is post-fit. The total uncertainty in the background is shown as dashed regions. Signal yields, overlaid as lines, correspond to observed 95% confidence-level exclusion limits for different mass scenarios: for the muon channel, the signals have $(m_N, c_T, |U|^2)$ values of (5 GeV, 4.5 mm, $7.1 \times 10^{-5}$), (10 GeV, $2.7 \times 10^{-1}$ mm, $3.2 \times 10^{-5}$), (20 GeV, $2.1 \times 10^{-2}$ mm, $1.1 \times 10^{-5}$), (30 GeV, $2.2 \times 10^{-3}$ mm, $1.3 \times 10^{-5}$) and (50 GeV, $8.2 \times 10^{-5}$ mm, $2.5 \times 10^{-5}$); for the electron channel, the signals have $(m_N, c_T, |U|^2)$ values of (5 GeV, 1.6 mm, $1.9 \times 10^{-4}$), (10 GeV, $2.1 \times 10^{-1}$ mm, $4.1 \times 10^{-5}$), (20 GeV, $1.6 \times 10^{-2}$ mm, $1.4 \times 10^{-5}$), (30 GeV, $1.9 \times 10^{-3}$ mm, $1.5 \times 10^{-5}$) and (50 GeV, $5.4 \times 10^{-5}$ mm, $3.8 \times 10^{-5}$).
For this search, two models are compared with the data: the background-only model describing the SM processes and a signal-plus-background model in which the signal under investigation is added to the SM background. First, a discovery hypothesis test is performed in which a rejection of the background-only model given the observed data is examined. Afterwards, the signal strength in the signal-plus-background model is fixed to different values and the CL$_S$ [85] method is used to exclude various signal strengths. Exclusion limits on the signal-strength fit parameter translate directly into limits on $|U|^2$. A signal with a given mixing angle is considered excluded once the $p$-value for the CL$_S$ = CL$_{s+b}$/CL$_b$ value is below 0.05 and therefore corresponds to a 95% confidence-level exclusion limit. An overview of different signal shapes in the SR is shown in figure 5 for which the signal yield corresponds to a 95% confidence-level exclusion limit for each mass point. The exclusion limits in $|U_\mu|^2$ and $|U_e|^2$ for the different signal hypotheses are shown in figure 6 together with those obtained from the displaced HNL search. The expected limits are determined using the results of a CR-only fit as input to the limit extraction procedure. The offset by almost 1 $\sigma$ from the expected limit for the muon channel shown in figure 6 originates from the use of central values for the statistical uncertainty Gamma distributions in the SR for the Asimov dataset. In figure 5, scaled (post-fit) Gamma nuisance parameters are used for the statistical uncertainty parameters, which largely dominate over systematic uncertainties. For the electron channel, the post-fit statistical uncertainty parameters deviate less from their central value as $\mu_{\text{ml}}$ has more freedom in the CRs, and consequently a general offset is not observed. The deviation for the muon channel is a consequence of the aforementioned $\mu_{\text{ml}}$ constraint in the lowest bin of CR-3.

5 Displaced-vertex signature

For $m_N \lesssim 20$ GeV, as the HNL lifetime gets longer for lower masses and coupling strengths, the searches relying on standard prompt objects, as defined above, become highly inefficient. The signature of a displaced vertex (DV) is needed to explore these complementary regions of the parameter space. Another advantage of this approach is that the requirement of a DV detached from the primary $pp$ interaction by 4 mm or more in the transverse plane eliminates the vast majority of SM backgrounds. In addition to probing lower masses and coupling strengths, this allows the requirement of same-charge same-flavour leptons to be removed and thus the search to be performed without relying on LNV.

Searches for long-lived neutral particles using DV signatures conducted so far by the ATLAS [86–93] and CMS [94–100] Collaborations considered the new particles to be high-$p_T$ decay products of other massive particles, with relatively large branching ratios. None of these provided any relevant sensitivity to HNLs due to high-$p_T$ requirements or to the requirement that two displaced vertices must be reconstructed in the same event. It was noted [25–27] that HNLs from $W$ boson decays produce a distinct signature which had not yet been probed at the LHC: a prompt lepton from the $W$ boson and a DV formed with tracks of relatively low $p_T$, among which there is at least one belonging to a lepton.
5.1 Trigger and preselection (displaced signature)

The displaced signature explored in this search comprises a prompt and isolated muon accompanied by a DV formed by either two muons or a muon and an electron. A prompt lepton from the $W$ boson decay is essential for triggering the event. The single-muon trigger with a $p_T$ threshold of 26 GeV is used.

Displaced vertices in the ID can be reconstructed at radial distances up to the first SCT layer at $\sim$300 mm due to the application of a large-radius tracking (LRT) algorithm optimised for tracks with large unsigned transverse impact parameters relative to the primary vertex ($d_0$) [101]. Large-radius tracking is computationally intensive and is therefore not performed on the entirety of the dataset but rather on a subset of events preselected according to a specific set of criteria which must rely on the standard track-reconstruction algorithm. In this analysis, this preselection requires the presence of at least two muons with $|\eta| < 2.5$ and is applied to the 2016 dataset, corresponding to an integrated luminosity of 32.9 fb$^{-1}$. One muon, meant to originate from an on-shell $W$ boson decay, is required to have an ID track matched to an MS track segment, have $p_T > 28$ GeV, and satisfy “loose” isolation and identification criteria as defined in ref. [75]. Another muon, which targets a displaced muon from an HNL decay, is required to satisfy a “loose” isolation criterion and have $p_T > 5$ GeV; it must have an MS track which either has no matched track in the ID or, if it has a matched track, must have either $d_0 > 0.1$ mm, or $d_0 < 0.1$ mm and $\chi^2$/dof > 5 for the MS-ID track matching, as defined in ref. [75]. The $\chi^2$/dof selection is added to recover tracks which have relatively low $d_0$ values but still correspond to a displaced muon. For instance, it could be that, during standard reconstruction, the MS track from a muon is incorrectly matched to an unrelated ID track.

5.2 Reconstruction and selection (displaced signature)

Large-radius tracking is performed on the dataset which satisfies the preselection, producing an extended collection of tracks. Using these new tracks, muon and electron candidates are reconstructed as described in section 4.1. Displaced-vertex candidates are also formed using this extended track collection. The following additional requirements are imposed on the tracks to be considered as seeds of the vertex reconstruction algorithm: they must have $p_T > 1$ GeV, $d_0 > 2$ mm, at least two SCT hits, and at least one TRT hit or two pixel hits. The vertex reconstruction algorithm described in ref. [90] is used, based on the incompatibility graph method and iteratively merging all two-track vertices which are found within 1 mm distance from each other. For an HNL decay length of the order of a cm, the DV reconstruction efficiency (including the track reconstruction efficiency) is about 20%, as studied in ref. [102].

The event selection requires the presence of at least one DV which satisfies the properties described below. The DV must be within the fiducial volume defined as $4 < r_{DV} < 300$ mm, where $r_{DV}$ is the distance to the beam axis. Since TRT hits do not have a well defined $z$ coordinate, the upper bound on $r_{DV}$ ensures that tracks have a sufficient number of SCT hits for high-quality track and DV reconstruction. Then the DV is required to be formed by exactly two tracks with opposite charges in order to be identified as decay prod-
ucts of a neutral particle. A cosmic-ray veto, \( \sqrt{\sum \eta^2 + (\pi - \Delta \phi)^2} > 0.04 \) (in which the sum \( \sum \eta \) and difference \( \Delta \phi \) refer to the two tracks forming the DV), is applied to eliminate high-mass vertices from a single cosmic-ray muon which is reconstructed as two back-to-back muons. The DV must be formed by at least one tight muon and an additional tight lepton (either muon or electron), with a “tight” identification identical to the standard one \([74, 75]\) except that it does not impose a minimum number of pixel hits such as to be efficient for DVs originating beyond the first pixel layer. Given a reconstructed track originating from the HNL decay, tight-lepton reconstruction efficiencies are around 70% and 50% for muons and electrons, respectively. Finally, a requirement is applied on the DV invariant mass \( m_{DV} \) as defined by \( m_{DV}^2 = (\sum E_i)^2 - (\sum \vec{p}_i)^2 \), where the sum runs over the two tracks forming the vertex. The requirement is chosen to be \( m_{DV} > 4 \) GeV as a compromise between keeping good signal efficiencies for HNL masses of 5 GeV and above, and rejecting low-mass backgrounds from material interactions and decays of metastable SM particles.

The overall signal efficiency, defined as the fraction of generated signal events that were reconstructed and satisfied the selection criteria, depends on the HNL mass and lifetime and is typically (1–2)% in the regions probed by the displaced HNL search. Signal losses are largely dominated by inefficiencies for displaced tracks and DV reconstruction. For a given HNL mass, the efficiency for any value of the mean proper decay length \( \tau \) is obtained from the simulated samples generated with \( \tau = 1, 10, 100 \) mm. To evaluate the efficiency for a given value of \( \tau \), each reconstructed event is weighted so that the generated event sample is distributed according to \( \exp(-t_{true}/\tau) \), where \( t_{true} \) is the true proper decay time. The resulting efficiencies are shown in figure 2. The efficiencies evaluated with this method agree with those from the fully simulated MC samples with \( \tau = 1, 10 \) and 100 mm within statistical uncertainties. Efficiencies increase with increasing \( m_N \) due to the requirement on the reconstructed vertex mass \( m_{DV} > 4 \) GeV and due to the fact that the decay leptons have larger momenta. The search is not sensitive to values of \( \tau \) lower than 0.1 mm or higher than 1000 mm due to the requirement of a DV in the fiducial volume \( 4 < r_{DV} < 300 \) mm.

Relevant systematic uncertainties that can affect the signal efficiencies (with relative magnitude in between parentheses) include uncertainties in the efficiencies for the prompt-lepton reconstruction and identification (1%), displaced track and DV reconstruction (15%), displaced-lepton identification (5%), as well as uncertainties in the modelling of lepton kinematic distributions and individual decay branching ratios (10%), in the modelling of multiple \( pp \) interactions in the bunch crossing (10%), and due to MC statistical uncertainties (10%). To evaluate the dominant 15% uncertainty due to the modelling of displaced tracks and DV reconstruction, a sample of \( K_S^0 \) mesons is selected from two-pion vertices in the invariant mass window 488–508 MeV. The rate of \( K_S^0 \) reconstruction is parameterised as a function of the sum of the \( p_T \) of the two tracks \( (\sum p_T) \) and the radial distance \( (r_{DV}) \). In each window of \( \sum p_T \) and \( r_{DV} \), the efficiency is obtained by dividing the event yield by the expected exponentially falling distance distribution, and a weight is computed as the ratio between measured and simulated efficiencies. These weights are normalised to be equal to one at small radii and then used to reweight the DV efficiencies in the signal samples, and a relative difference of 15% is found as the maximum effect in the final selection efficiency. In addition, uncertainties in the \( W \) boson production cross section
(3%) and the integrated luminosity (2.2%) are taken into account for the interpretation. The total systematic uncertainty, with all contributions added in quadrature, is 24%.

5.3 Backgrounds (displaced signature)

Possible background sources which can result in two-track DVs include hadronic interactions in material, decays of metastable particles such as bottom, charm, and strange hadrons, accidental crossings of charged particles produced in the collisions, and cosmic-ray muons which either cross a charged particle from the collision or are reconstructed as two back-to-back muons. All of these are reduced by over an order of magnitude when requiring a prompt muon in the same event, and all except cosmic-ray muons are significantly reduced when requiring both tracks to be matched to objects satisfying tight lepton identification. No backgrounds from single cosmic-ray muons reconstructed as two back-to-back muons remain after applying the cosmic-ray veto described in section 5.2. Other backgrounds can arise in processes such as dijets and $W$+jets. These are processes with large cross sections combined with a very low probability to produce a DV satisfying the selection criteria, making these backgrounds extremely difficult to reproduce in simulations due to the large numbers of events which need to be simulated. Therefore, they are studied and evaluated in a fully data-driven manner using suitable control regions.

A study of the different types of background sources was performed using a control sample of events which fail the requirement of a prompt muon with a matched track in the ID, which is part of the preselection (section 5.1). This control sample has undergone LRT and thus consists of events collected by a variety of different triggers and preselections. With the requirement of at least one DV, this sample contains 12 times more events than the sample of events passing the preselection. Sub-categories are defined for DVs containing no reconstructed lepton (0-lepton) and only one reconstructed lepton (1-lepton), as requiring two leptons (2-lepton) would lead to an insufficient number of events for this study. For $m_{\text{DV}} < 2.5$ GeV, more events with 0-lepton and 1-lepton DVs are observed in the high-density material region as compared to the low-density region (as expected from hadronic interactions in material). Likewise, more opposite-charge events are observed than same-charge events (as expected from metastable particle decays). This shows that these two types of backgrounds contribute at low mass. Above $m_{\text{DV}} = 2.5$ GeV, the $m_{\text{DV}}$ distributions fall quickly and coincide within 5% regardless of the DV being in a high-density or low-density material region, regardless of the tracks being same-charge or opposite-charge, and regardless of the number of tight leptons identified in the DV. The conclusion of this study is that, in the signal region ($m_{\text{DV}} > 4$ GeV), backgrounds from hadronic interactions and metastable particle decays provide a contribution of less than 5% relative to other sources of background.

Using the same control sample as above with two muons identified in the DV, peaks in the $m_{\text{DV}}$ distribution are observed at the mass value of the $J/\psi$ and $\psi(2S)$ mesons (3.1 GeV and 3.7 GeV, respectively). These peaks correspond to decay distances in the range 4–40 mm, consistent with the decays of $b$-hadrons into $J/\psi$ and $\psi(2S)$. From integrating fits to these two peaks in the region $m_{\text{DV}} > 4$ GeV, taking into account the fact that this sample is statistically enhanced relative to the samples used in the search as well as fit uncertainties, a contribution of less than 0.005 background events from $J/\psi$ and $\psi(2S)$ decays is estimated...
in the signal region. No contributions from the very short-lived \( \Upsilon \) meson are expected due to the fact that the probability for them to result in displaced decays is negligible, and indeed no two-muon DVs are found for \( m_{DM} > 3.85 \text{ GeV} \) in the control sample.

The number of background events in the signal region, which requires opposite-charge DVs, is estimated using a control region of same-charge DVs, all other requirements being the same. This is done by applying a transfer factor from 0-lepton DVs to 2-lepton DVs obtained in the same-charge control region. This provides an unbiased estimate of all backgrounds remaining after the selection for which the ratio of 2-lepton background DVs to 0-lepton background DVs does not depend on the DV charge configuration. While this is the case for accidental crossings including those involving cosmic-ray muons, it does not include single cosmic-ray muons reconstructed as back-to-back muons nor decays of neutral hadrons which are either metastable or produced in material interactions, which can be neglected in the \( m_{DM} > 4 \text{ GeV} \) region as discussed above. The validity of the method is verified by performing the estimate on a validation sample of 1-lepton DVs. The numbers of events observed in the control regions are reported in table 6. The estimated numbers of events in the 1-lepton DV validation region for electron and muon are compatible with the observed numbers within statistical uncertainties. From a 90\% confidence-level limit of 2.3 background events with 2-lepton DVs in the same-charge DV control region (where 0 are observed), an upper limit of \( 2.3 \times 168037/169254 \sim 2.3 \) is obtained for the 2-lepton DVs in the signal region (where 0 are observed).

### 6 Results

Observations in the signal regions are consistent with background expectations in both the prompt and displaced signatures described in sections 4 and 5, respectively. For a given HNL mass and lifetime, the selection efficiency is obtained from MC simulations, and the uncertainty in the efficiency is evaluated. Combining this information with the event yield obtained from the integrated luminosity of 36.1 fb\(^{-1}\) and 32.9 fb\(^{-1}\) for the prompt and displaced analyses, respectively, and the HNL production cross section and branching ratio (eq. (2.1)), a set of choices of HNL coupling strengths (\( |U|^2 \)) and masses (\( m_N \)) are

<table>
<thead>
<tr>
<th>Leptons in DV</th>
<th>Same-charge DV</th>
<th>Opposite-charge DV</th>
<th>Opposite-charge DV estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>0 (signal region)</td>
<td>&lt; 2.3 at 90% CL</td>
</tr>
<tr>
<td>1 ((\mu))</td>
<td>83</td>
<td>89</td>
<td>82.4±9.0</td>
</tr>
<tr>
<td>1 ((e))</td>
<td>28</td>
<td>35</td>
<td>27.8±5.3</td>
</tr>
<tr>
<td>0</td>
<td>169254</td>
<td>168037</td>
<td></td>
</tr>
</tbody>
</table>

**Table 6.** Observed number of events in the control and validation regions used for evaluating the backgrounds in the displaced HNL signature, including all sources of background except for single cosmic-ray muons reconstructed as back-to-back muons and metastable particles. Transfer factors to translate the number of 0-lepton DVs to 2-lepton DVs are obtained using a same-charge DV control sample. Validation is performed using 1-lepton DVs. The number of events in the signal and validation regions is also indicated.
Figure 6. Observed 95% confidence-level exclusion in $|U_\mu|^2$ (top) and $|U_e|^2$ (bottom) versus the HNL mass for the prompt signature (the region above the black line is excluded) and the displaced signature (the region enclosed by the red line is excluded). The solid lines show limits assuming lepton-number violation (LNV) for 50% of the decays and the long-dashed line shows the limit in the case of lepton-number conservation (LNC). The dotted lines show expected limits and the bands indicate the ranges of expected limits obtained within 1σ and 2σ of the median limit, reflecting uncertainties in signal and background yields.

excluded at the 95% confidence level. Calculations of confidence intervals and hypothesis testing are performed using a frequentist method with the CLs formalism as implemented in RooStats [103]. The exclusion limits are shown in figure 6, in the cases of dominant mixing to $\nu_\mu$ (top) and $\nu_e$ (bottom), for the cases of LNV (both signatures, solid lines) and LNC (displaced signature, long-dashed line).

Limits from the prompt signature cover the mass range 5–50 GeV. In the mass range 20–30 GeV, the regions in $|U_\mu|^2$ and $|U_e|^2$ above $1.4 \times 10^{-5}$ are excluded, a reach which is limited by the integrated luminosity of the analysed data, as well as the selection efficiency and the signal-to-background ratio. At higher masses, the sensitivity decreases due to a kinematic suppression of HNL production from the $W$ boson decay. For masses below
20 GeV, the long decay path causes large efficiency losses. This is the region where
the displaced signature becomes more sensitive.

Limits from the displaced signature cover the mass range 4.5–10 GeV, in which they
exclude coupling strengths down to $|U_\mu|^2 \sim 2 \times 10^{-6} (1.5 \times 10^{-6})$ assuming LNV (LNC). For
comparison, the best previous constraints on $|U_\mu|^2$ in this mass range were obtained
with the DELPHI experiment at LEP1 [23], excluding values down to $\sim 1.5 \times 10^{-5}$. The limit
contour of the displaced signature takes the shape of an oblique ellipse which approximately
corresponds to HNL proper decay lengths in the range 1–30 mm. It is also limited from
below by the product of integrated luminosity and efficiency. The interpretation with LNV
provides weaker limits because the search is sensitive to long lifetimes and, for a given
coupling strength, the lifetime is reduced by a factor of two when LNV decays are allowed.

7 Conclusions

A search for heavy neutral leptons (HNLs) produced in leptonic decays of on-shell $W$
bosons has been performed using data recorded by the ATLAS detector at the LHC in proton–
proton collisions at a centre-of-mass energy of 13 TeV corresponding to an integrated lumi-
nosity of up to 36.1 fb$^{-1}$, using two distinct signatures. The prompt signature requires three
prompt leptons (either muons or electrons) with no same-flavour opposite-charge configura-
tion. It probes mean HNL proper decay lengths of 1 mm or less, with the assumption of
lepton-number violation. The displaced signature, explored for the first time at the LHC,
features a prompt muon accompanied by a vertex displaced in the radial direction by 4–
300 mm from the beam line containing two opposite-charge leptons (either two muons or a
muon and an electron) with a reconstructed vertex mass $m_{DV} > 4$ GeV. It does not require
lepton-number violation and probes longer lifetimes, corresponding to lower masses.

Observations are consistent with background expectations and results of the search
are presented as exclusion contours in the HNL coupling strength versus mass plane in a
model postulating a single HNL mixing either to muon or electron neutrinos. The prompt
signature excludes coupling strengths above $4 \times 10^{-5}$ in the mass range 10–50 GeV, with
a most stringent limit of $1.1 \times 10^{-5}$ for a mass of 20 GeV. In the case of mixing to muon
neutrinos, the displaced signature excludes coupling strengths down to $2 \times 10^{-6} (1.5 \times 10^{-6})$
at best in the mass range 4.5–10 GeV assuming lepton-number violation (conservation),
surpassing the best previous constraints using on-shell $Z$ boson decays at LEP1 by one
order of magnitude.

A notable characteristic of the displaced signature is that backgrounds fall off sharply
with the vertex mass. Heavy neutral leptons with smaller coupling strengths (longer life-
times) at higher masses can be probed by increasing the $W$ boson yield at higher luminosi-
ties. Also, even though this was not done in this work, it is possible to probe long-lived
HNL mixing to the electron flavour, and also to consider their semileptonic decays.
Acknowledgments

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, CRC and Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Skłodowska-Curie Actions, European Union; Investissements d’ Avenir Labex and Idex, ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya, Spain; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [104].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


[94] CMS collaboration, Search for long-lived particles that decay into final states containing two electrons or two muons in proton-proton collisions at $\sqrt{s} = 8$ TeV, Phys. Rev. D 91 (2015) 052012 [arXiv:1411.6977] [insPIRE].


[98] CMS collaboration, Search for new particles decaying to a jet and an emerging jet, JHEP 02 (2019) 179 [arXiv:1810.10069] [insPIRE].


Physics Department, Southern Methodist University, Dallas TX; United States of America
Physics Department, University of Texas at Dallas, Richardson TX; United States of America
National Centre for Scientific Research “Demokritos”, Agia Paraskevi; Greece
(a) Department of Physics, Stockholm University; (b) Oskar Klein Centre, Stockholm; Sweden
Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen; Germany
Lehrstuhl für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund; Germany
Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden; Germany
Department of Physics, Duke University, Durham NC; United States of America
SUPA — School of Physics and Astronomy, University of Edinburgh, Edinburgh; United Kingdom
INFN e Laboratori Nazionali di Frascati, Frascati; Italy
Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg; Germany
II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen; Germany
Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland
(a) Department of Physics, University of Genova, Genova; (b) INFN Sezione di Genova; Italy
II. Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen; Germany
SUPA — School of Physics and Astronomy, University of Glasgow, Glasgow; United Kingdom
LPSC, Université Grenoble Alpes, CNRS/IN2P3, Grenoble INP; Grenoble; France
Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge MA; United States of America
(a) Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei; (b) Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao; (c) School of Physics and Astronomy, Shanghai Jiao Tong University, KLPPAC-MoE, SKLPPC, Shanghai; (d) Tsung-Dao Lee Institute, Shanghai; China
Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg; Germany
Faculty of Applied Information Science, Hiroshima Institute of Technology, Hiroshima; Japan
(a) Department of Physics, Chinese University of Hong Kong, Shatin, N.T., Hong Kong; (b) Department of Physics and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong; China
Department of Physics, National Tsing Hua University, Hsinchu; Taiwan
Department of Physics, Indiana University, Bloomington IN; United States of America
(a) INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine; (b) ICTP, Trieste; (c) Dipartimento Politecnico di Ingegneria e Architectura, Università di Udine, Udine; Italy
(a) INFN Sezione di Lecco; (b) Dipartimento di Matematica e Fisica, Università del Salento, Lecce; Italy
(a) INFN Sezione di Milano; (b) Dipartimento di Fisica, Università di Milano, Milano; Italy
(a) INFN Sezione di Napoli; (b) Dipartimento di Fisica, Università di Napoli, Napoli; Italy
(a) INFN Sezione di Pavia; (b) Dipartimento di Fisica, Università di Pavia, Pavia; Italy
(a) INFN Sezione di Pisa; (b) Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa; Italy
(a) INFN Sezione di Roma; (b) Dipartimento di Fisica, Sapienza Università di Roma, Roma; Italy
(a) INFN Sezione di Roma Tor Vergata; (b) Dipartimento di Fisica, Università di Roma Tor Vergata, Roma; Italy
(a) INFN Sezione di Roma Tre; (b) Dipartimento di Matematica e Fisica, Università Roma Tre, Roma; Italy
(a) INFN-INFN TIFPA; (b) Università degli Studi di Trento, Trento; Italy
Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck; Austria
University of Iowa, Iowa City IA; United States of America
Department of Physics and Astronomy, Iowa State University, Ames IA; United States of America
Joint Institute for Nuclear Research, Dubna; Russia