Using displaced tracks to search for sterile neutrinos in the ATLAS detector

by

Robin Newhouse

B.Sc., The University of British Columbia, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2022

© Robin Newhouse 2022

CERN-THESIS-2022-277
28/11/2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Using displaced tracks to search for sterile neutrinos in the ATLAS detector

submitted by **Robin Newhouse** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Physics**.

Examining Committee:

Alison Lister, Associate Professor, Physics & Astronomy, UBC
Supervisor

Jeremy Heyl, Professor, Physics & Astronomy, UBC
Supervisory Committee Member

Michael Hasinoff, Professor, Physics & Astronomy, UBC
University Examiner

Michiel van de Panne, Professor, Computer Science, UBC
University Examiner

Todd Adams, Professor, Physics, Florida State University
External Examiner

Additional Supervisory Committee:

Colin Gay, Professor, Physics & Astronomy, UBC
Supervisory Committee Member

Ariel Zhitnitsky, Professor, Physics & Astronomy, UBC
Supervisory Committee Member

Abstract

The Standard Model of particle physics is a powerful theory of nature, yet it does not account for all physical observations. Notably, the nonzero masses of the three neutrino flavours and their transformations into one another suggest the need for an extension of the Standard Model. One such extension postulates the existence of Heavy Neutral Leptons (HNLs, $\mathcal{N}$) — right-handed neutrino states that do not interact with other particles except through mixing with Standard Model neutrinos. HNLs may generate light neutrino masses through the so-called "seesaw mechanism."

This dissertation presents a direct search for long-lived HNLs using 139 fb^{-1} of $\sqrt{s}$ = 13 TeV pp collision data collected by the ATLAS detector at the Large Hadron Collider. In this search, the $\mathcal{N}$ is produced via $W \to \mathcal{N}\mu$ or $W \to \mathcal{N}e$ and decays into a neutrino and two charged leptons, which form a displaced vertex in the inner detector. No signal is observed, and limits are set on the squared mixing angles of the $\mathcal{N}$ with the Standard Model neutrinos in the mass range 3 GeV $< m_{\mathcal{N}} <$ 15 GeV. For the first time in a collider search, results are presented in the context of realistic HNL mixing models consistent with neutrino oscillation data.

Such a displaced vertex search relies heavily upon nonstandard reconstruction algorithms. An additional pass of the track reconstruction algorithm with relaxed collision vertex pointing requirements is executed on a subset of the recorded data to improve sensitivity to long-lived particles in the inner detector. This Large-Radius Tracking (LRT) configuration is effective yet computationally expensive due to the reconstruction of many "fake" tracks not formed from individual charged particles. This algorithm was optimized in preparation for the Run 3 data-taking period, which began in July 2022. Fake LRT tracks are reduced by 95%, drastically improving the purity and reducing the computational load such that LRT is now run in the default reconstruction pipeline. This optimization is expected to substantially improve the sensitivity and reduce the complexity of long-lived particle analyses in ATLAS, including future HNL searches.

Lay Summary

The ATLAS detector at the Large Hadron Collider (LHC) tests the Standard Model of particle physics (SM) and searches for evidence of new phenomena by smashing protons together at record energies. A growing area of interest in particle physics involves searching for long-lived particles (LLPs), predicted by several theories, which travel a small but macroscopic distance in collider experiments. Detecting such hypothetical particles in ATLAS requires specialized algorithms. This dissertation explains how a “displaced tracking” algorithm was optimized to enable such searches for LLPs. The physics analysis presented here searches for one of these LLPs called a “Heavy Neutral Lepton” (HNL) using data collected by ATLAS. If discovered, the HNL would help explain the mysterious masses of neutrinos — elusive fundamental particles with small but nonzero masses. Unfortunately, no evidence of a new particle is found, but new limits are set, which will help guide future experiments.

Preface

The research presented in this dissertation is based on data collected by the ATLAS experiment. The ATLAS collaboration's ongoing success is made possible by more than 5500 members (nearly 3000 authors) from 245 institutes in 42 countries, including graduate students, postdoctoral researchers, and professors. These scientists develop simulation and reconstruction software, design and commission new detector technologies, monitor data-taking operations, and publish analyses of the recorded data. These activities are done with help from the broader scientific community, not the least of which is the successful operation of the Large Hadron Collider, without which none of this work would be possible.

I wrote this dissertation's text with some phrases and definitions taken from ATLAS publications and unpublished internal documentation, clearly attributed in each case. The captions in some figures are attributed to the source if they have minor or no alterations. Figures with no reference in the caption were produced by me or in collaboration with other ATLAS researchers.

ATLAS submits final results to peer-reviewed journals and publishes preliminary results not subject to external peer review. In both cases, the figures are subject to formal internal review. Figures labelled "ATLAS", "ATLAS Preliminary", "ATLAS Simulation", "ATLAS Simulation Preliminary", or "ATLAS Online" are public results which use ATLAS simulation or data and are internally reviewed by the collaboration. "ATLAS Internal" or "ATLAS Simulation Internal" figures are not reviewed by the collaboration and do not indicate approval from all members.

Chapters 2 and 3 present the theoretical background knowledge necessary to understand and motivate the HNL analysis. The presentation of this subject matter is inspired by well-known textbooks cited in the bibliography. Chapters 4 and 5 summarize the experimental apparatus (the LHC and ATLAS) and the standard ATLAS reconstruction algorithms. Those chapters are drawn from several ATLAS design reports and performance studies. Chapters 6 and 7 contain my original work, with detailed contributions and acknowledgements listed below.

Chapter 6 describes my work in optimizing the large-radius tracking (LRT) algorithm for its inclusion in the default Run 3 reconstruction pipeline. The work was done with other members of the LRT optimization team: Jackson Burzynski, Matthias Danninger, Giuliano Gustavino, Bingxuan Liu, Amber Roepe, and John Stupak. Each of these individuals contributed significantly to the performance studies, algorithm tuning, software development, and presenting and gathering feedback from physics groups. Matthias Danninger led the project, coordinating with ATLAS management and other stakeholders to ensure the effort's success. Jackson Burzynski and Bingxuan Liu produced many of the Run 3 performance figures in this dissertation. We received support from the tracking performance group. The following summarizes my contributions to this optimization effort:

- I developed and computed performance metrics to determine the degree of impact for each optimization.
- I performed a study on pileup effects to demonstrate the robustness of the algorithm as a function of track multiplicity.
- I led the design, development, training, testing, and deployment of the DNN discriminant.
- I wrote the necessary code to execute an additional tracking pass and include the LRT collection in the ATLAS EDM.
- I presented several progress updates in internal ATLAS group meetings and workshops.
- I conducted a review of systematic uncertainties associated with LRT tracks, helping lay the foundation for ongoing uncertainty studies.

The campaign to include LRT in the default reconstruction was a major success. Physics and computing sub-groups were initially skeptical of the introduction of a new strategy which had the potential to negatively impact physics analyses. The perseverance of this small team and the continually demonstrated progress paid off. The LRT algorithm will now be applied to every event recorded by ATLAS in Run 3 and was also run in the reprocessing of existing Run 2 data. Our team was distinguished with an *ATLAS Outstanding Achievement Award*. This award, presented every two years to roughly eight groups or individuals, recognizes exceptional technical contributions to the collaboration. An ATLAS PUB Note (public technical document without external review) is currently being drafted, which will document the details and performance of the optimization. The LRT configuration and procedure described in section 6.2 are used in the analysis presented in chapter 7.

Chapter 7 presents a search for Heavy Neutral Leptons (HNLs) with a displaced signature. At the time of this writing, the analysis has been accepted by *Physical Review Letters*. Fifteen team members contributed to this analysis at various levels, and I was one of the three principal data analyzers, along with Dominique Trischuk and Christian Appelt. Dominique Trischuk led a significant portion of this analysis, including the development of the background estimation technique, the m_{HNL} discriminant, the custom vertexing configuration and its inclusion in a custom derivation. She also conducted studies on spin correlations, coordinated signal sample production, and developed control and validation regions. Christian Appelt carried out the statistical interpretation, including implementing the TRExFitter workspace, executing toy MC experiments, and interpolating the exclusion limit contours. Susumu Oda developed and studied the DRAW filters to preselect data. Alison Lister, Matthias Danninger, Abi Soffer, and Heiko Lacker defined the analysis strategy and provided support throughout, with Matthias Danninger and Margaret Lutz as the analysis contacts. Two *Short Term Associates* (external collaborators brought on for temporary authorship of an ATLAS publication) who contributed to this analysis are Oleg Ruchayskiy and Jean-Loup Tastet. Their substantial knowledge of HNL theory and phenomenology was crucial in producing the realistic model interpretations which set this analysis apart. The following summarizes my contributions to the analysis:

- I was the lead developer of the main analysis software, DHNLAlgorithm, with features implemented by Dominique Trischuk and Jackson Burzynski.

- I co-developed the event selection software, DHNLNtupleAnalysis, with Dominique Trischuk.
- I coordinated and monitored the relevant data processing on the LHC computing grid and our institute's local computing cluster.
- I implemented and validated the signal systematic uncertainties, both standard and nonstandard, including those related to displaced tracking and vertexing, displaced lepton identification, and filter emulation.
- I prepared the public electronic repository of our results on HEPdata along with Christian Appelt.
- I regularly presented updates at internal ATLAS meetings and provided guidance to other analyses.
- I wrote relevant portions of the internal documentation.

In addition to drawing from the published analysis [1], this chapter contains text and figures drawn from the analysis's unpublished internal documentation. Margaret Lutz, Christian Appelt, Dominique Trischuk, and I wrote a majority of the internal documentation. Abi Soffer and Margaret Lutz were the contact editors responsible for writing the paper and guiding it through publication.

To become an ATLAS author one must contribute at least 80 days of technical work to one of the collaboration's high-priority activities. My qualification task involved developing a machine-learning-based score to be used in the ambiguity solver of the standard track reconstruction algorithm. I created, trained, and implemented a DNN-based score which used various track trajectory and quality parameters as inputs. It was successful and robust, but failed to outperform the existing metric to a degree that could persuade the tracking group to adopt the new procedure. However, this work directly led to the insights and tools used to develop the DNN discriminant in the LRT optimization.

I was also part of a team that published an analysis comparing several algorithms used to distinguish between flavours of boosted jets [2]. In this analysis I conducted several background data studies with help from Oliver Majersky. In the ATLAS reconstruction software, I implemented a DNN-based top quark tagging algorithm developed by Jannicke Pearkes and Wojtek Fedorko. This "TopoDNN" score was compared to other identification algorithms in the publication and performed well.

I presented my work at several forums including international and national conferences, ATLAS-wide workshops, and internal collaboration meetings.

- R. Newhouse. *Proposed CMOS-based Sensors for the ATLAS Detector: CHESS-2*. 54th Winter Nuclear Particle Physics Conference. Banff, Canada, 2017.
- R. Newhouse, E. Khoda, S. Sen. *Machine Learning in Dense Tracking Environments*. ATLAS Tracking & Flavour Tagging Workshop. Hamburg, Germany. 2019.
- R. Newhouse, M. Danninger, G. Gustavino, J. Stupak, A. Roepe, B. Liu. *Re-optimization of large-radius tracking for Run 3*. ATLAS Exotics + SUSY Workshop 2020. Online. 2020.
- R. Newhouse. *Jet Substructure + Correlations in Hadronic Final States from ATLAS*. 9th Edition of the Large Hadron Collider Physics Conference. Online. 2021.

Table of Contents

List of Tables

List of Figures

Glossary

***ν*MSM** Neutrino Minimal Standard Model

1SFH One-HNL model with Single-Flavour mixing

2QDH Two Quasi-Degenerate HNL model

ALICE A Large Ion Collider Experiment

AOD Analysis Object Data

ATLAS A Toroidal LHC ApparatuS

BR Branching Ratio

BSM Beyond the Standard Model

CC Charged-current

CERN European Organization for Nuclear Research

CL Confidence Level

CMB Cosmic Microwave Background

CMS Compact Muon Solenoid

COM Centre of Mass

CP Charge-Parity

CR Control Region

CSC Cathode Strip Chamber

CTP Central Trigger Processor

DELPHI DEtector with Lepton, Photon, and Hadron Identification

DNN Deep Neural Network

DV Displaced Vertex

ECAL Electromagnetic Calorimeter

EDM Event Data Model

EM Electromagnetic

EW Electroweak

EWSB Electroweak Symmetry Breaking

FCC Future Circular Collider

GRL Good Run List

GSF Gaussian-sum filter

GUT Grand Unified Theory

h.c. Hermitian Conjugate (terms)

HCAL Hadronic Calorimeter

HL-LHC High-Luminosity Large Hadron Collider

HLT High-Level Trigger

HNL Heavy Neutral Lepton

IBL Insertable B-Layer

ID Inner Detector

IH Inverted Hierarchy (of neutrino masses)

IP Interaction Point

KS Kolmogorov-Smirnov test

L1 Level-1

LAr Liquid Argon

LEP Large Electron-Positron Collider

LH Likelihood

LHC Large Hadron Collider

LHCb Large Hadron Collider beauty

LLP Long-Lived Particle

LNC Lepton-Number Conserving

LNV Lepton-Number Violating

LO Leading Order

LRT Large-Radius Tracking

MC Monte Carlo

MDT Monitored Drift Tube

MLE Maximum Likelihood Estimate

MS Muon Spectrometer

NC Neutral-current

NH Normal Hierarchy (of neutrino masses)

NLO Next-to-Leading Order

NN Neural Network

NNLO Next-to-Next-to-Leading Order

NP Nuisance Parameters

OS Opposite Sign (charge)

PDF Parton Distribution Function

PLR Profile Likelihood Ratio

PMNS Pontecorvo-Maki-Nakagawa-Sakata (matrix)

POI Parameter of Interest

PS Proton Synchrotron

PSB Proton Synchrotron Booster

PV Primary Vertex

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

QFT Quantum Field Theory

RF Radio Frequency

ROI Region of Interest

RPC Resistive Plate Chamber

RPVLL R-Parity Violating Long-Lived

RTC Random Track Crossing (background)

SCT Semiconductor Tracker

SM Standard Model (of Particle Physics)

SNO Sudbury Neutrino Observatory

SPS Super Proton Synchrotron

SR Signal Region

SS Same Sign (charge)

SUSY Supersymmetry

TDAQ Trigger and Data Acquisition system

TGC Thin Gap Chamber

TRT Transition Radiation Tracker

VEV Vacuum Expectation Value

VH4b Vector boson to Higgs to Four b quarks (process)

VR Validation Region

VSI Vertexing Secondary Inclusive (algorithm)

WP Working Point

Acknowledgements

First, I thank my professor, Alison Lister, for making this journey possible. For over a decade, she has opened doors, instilled knowledge, and supported my education in the world of physics. Alison has been an outstanding leader in the CERN community and in our lab and exemplifies the influence of an advisor who cares deeply about her students. Most of all, she's been a wonderful friend.

I thank my supervisory committee, Colin Gay, Ariel Zhitnitsky, and Jeremy Heyl, for their enlighting feedback and for their instruction in my undergraduate courses many years ago. It has been a pleasure to have your consistent expertise throughout my education.

I gratefully acknowledge the financial support provided by UBC through the International Doctoral Fellowship award. I also appreciate the support provided by ATLAS Canada, which enabled me to work at CERN for a large portion of my degree and attend several international workshops and conferences. I thank CERN for the successful operation of the LHC that made this work possible; it is a truly monumental achievement.

To those in the ATLAS UBC lab, rarely have I had the pleasure of working among such wonderful people. You are outstanding both as colleagues and friends. My fellow students, Sebastien, Steffen, Felix, Dominique, Robin H, Elham, Chris, and Vincent, and my mentors Matthias, Wojtek, and Colin, thank you for many years of guidance and support. May thanks also to Nick Zacchia whose infectious love of science is a joy to witness.

Thanks to the CERN community for the warm welcome during my year abroad. I will never forget the hikes, runs, and parties that brought so much levity to a challenging time. I would also like to thank the displaced HNL team for the brilliance and determination that produced an excellent publication. Christian, Margaret, Dominique, and others on the team, collaborating with such remarkable scientists was a unique opportunity.

To my community, thank you for everything. You made sure that life outside of my academic bubble was outrageously fun! Your rejuvenating presence kept me afloat in challenging times and for that, I am immensely grateful. Thank you to my family, and especially to my incredible parents. You have encouraged me from my earliest years to value inquiry and education and have endlessly supported me in accomplishing my academic goals.

I would also like to express my gratitude to the land. This precious corner of our world has been a reservoir of tranquility for me. The northwest coast, with its sublime forests, resolute mountains, and pensive seas, is a place of deep beauty from which we can learn so much.

And finally, saving the best for last, I express my infinite gratitude to my beloved partner Breanna. You held me up when I was down, celebrated my successes, and taught me, by example, how to best care for myself. Having a partner in life with such grace and capacity is what has made any of this possible. Thank you for being with me, full of love, on this adventure.

Chapter 1

Introduction

The Standard Model of Particle Physics (SM) is a powerful theory of matter. It describes the known elementary particles and their interactions, and its predictions match experimental data with astounding precision. Nevertheless, our understanding of the universe remains full of questions that have yet to be satisfactorily answered.

The neutrino is an elusive and enigmatic particle that may play a meaningful role in answering many of these open questions. This nearly undetectable particle, which comes in three types, was thought to be massless until only twenty years ago. However, the discovery of its peculiar tendency to change its variety has shown physicists that neutrinos do have mass. Extensions to the SM, which account for these masses, may provide hints to the universe's mysteries of dark matter and matter-antimatter asymmetry. One such extension predicts the existence of a new type of massive neutrino, or heavy neutral lepton (HNL), which interacts even more rarely with other SM particles.

Colliding subatomic particles into one another at outrageous speeds, aside from being quite fun, helps probe these neutrino masses and many other puzzles in physics. The ATLAS detector at The Large Hadron Collider (LHC) studies the rare particles created when the kinetic energy from accelerated protons is converted into exotic states of matter.

Most searches for new physics concentrate on very short-lived phenomena, focusing the reconstruction algorithms on particles which are formed and decay almost instantaneously. However, many theories predict particles with long enough lifetimes that they would decay at some macroscopic distance from the proton-proton (pp) collision point. A restructuring of ATLAS's charged particle reconstruction algorithms is necessary to be sensitive to such long-lived particle (LLP) signatures. By reconstructing displaced tracks (charged particle trajectories which originate some distance from the detector's centre), ATLAS can extend its physics analyses into uninvestigated regions of phase space.

Certain theories predict HNLs which decay within the inner detector volume of the ATLAS detector and whose detection would rely on such modified algorithms. The following dissertation conducts a search for HNLs in $\sqrt{s} = 13$ TeV pp collisions, which uses displaced track reconstruction to identify potential signatures. Recent improvements to the displaced tracking algorithm will substantially increase the sensitivity of this analysis. Such refinements and their implications are also discussed.

This dissertation is organized as follows: Chapter 2 introduces the Standard Model and some of its known limitations. Chapter 3 presents a short history of the discovery of neutrino masses and the theoretical background which motivates the existence of HNLs. It also summarizes the results of HNL searches in several experiments. Chapter 4 introduces the experimental facilities, focusing on key systems in the LHC and the ATLAS detector. Chapter 5 describes the reconstruction algorithms used to compute representations of particles interacting with the ATLAS detector. Particular attention is paid to the reconstruction of charged particles, so-called tracks, in the inner detector. Chapter 6 discusses the large-radius tracking (LRT) algorithm and its use in the reconstruction pipeline. The algorithm used in this dissertation's HNL analysis is presented, as well as the optimizations which will allow for even more sensitive analyses in the future. Chapter 7 details the search for HNLs performed with pp collisions recorded from 2015–2018. No excess signal was found, and exclusions limits are set for a range of HNL masses and coupling strengths.

Chapter 2

The Standard Model of Particle Physics

At the heart of the rich field of high-energy physics lies the question: "what is the universe made of, and how do its constituents interact?" Over several centuries of investigation, a robust theory has emerged that has revolutionized our understanding of the fundamental nature of matter and given us extraordinarily powerful tools to predict how this matter behaves. This theory, called the Standard Model of particle physics (SM), provides an elegant description of the fundamental particles contained in the universe and the forces through which they interact.

Although the SM provides the most accurate picture of our universe to date, it is clear that the story is incomplete. Carefully designed experiments have revealed phenomena that the current theory does not readily explain. Extensions to the SM have been proposed and are the active subjects of many experiments. What is described below is a summary of the particles and forces of the SM as currently understood, and in particular, those most relevant to this thesis. The approach is guided by references [3], [4], and [5], which the reader is encouraged to consult for a more rigorous description.

2.1 Formalism

The Standard Model is written in the formalism of quantum field theory (QFT) which describes quantum mechanical interactions in a way consistent with special relativity. It describes the particle content of the universe as well as the dynamics of interaction between the various particles. In this theory, each field corresponds to a type of particle. A Lagrangian describes the interactions between these fields as well as their kinematics. The SM obeys certain constraints that dictate the allowable properties and interactions of these fields.

2.2 Particle content

In QFT, quantum-mechanical fields represent physical particles. In the SM, matter is composed of spin-half particles called fermions that interact by exchanging spin-one particles called bosons. These are so named for the statistical mechanics each type of particle follows: spin-half particles obey Fermi-Dirac statistics, while spin-one particles obey Bose-Einstein statistics. Every field in the SM has been experimentally observed, and the properties of each have been measured to often very high degrees of precision.

As a quantum-mechanical state, each field can be fully described by a set of *quantum numbers* — conserved quantities which correspond to eigenvalues of operators that commute with the Hamiltonian. Quantum numbers are also referred to as *charges*. Colour charge is related to the strong force. Isospin, weak hypercharge, and electric charge are related to the unified electroweak (EW) force (see section 2.6). The particles in the SM and their properties are depicted in fig. 2.1.

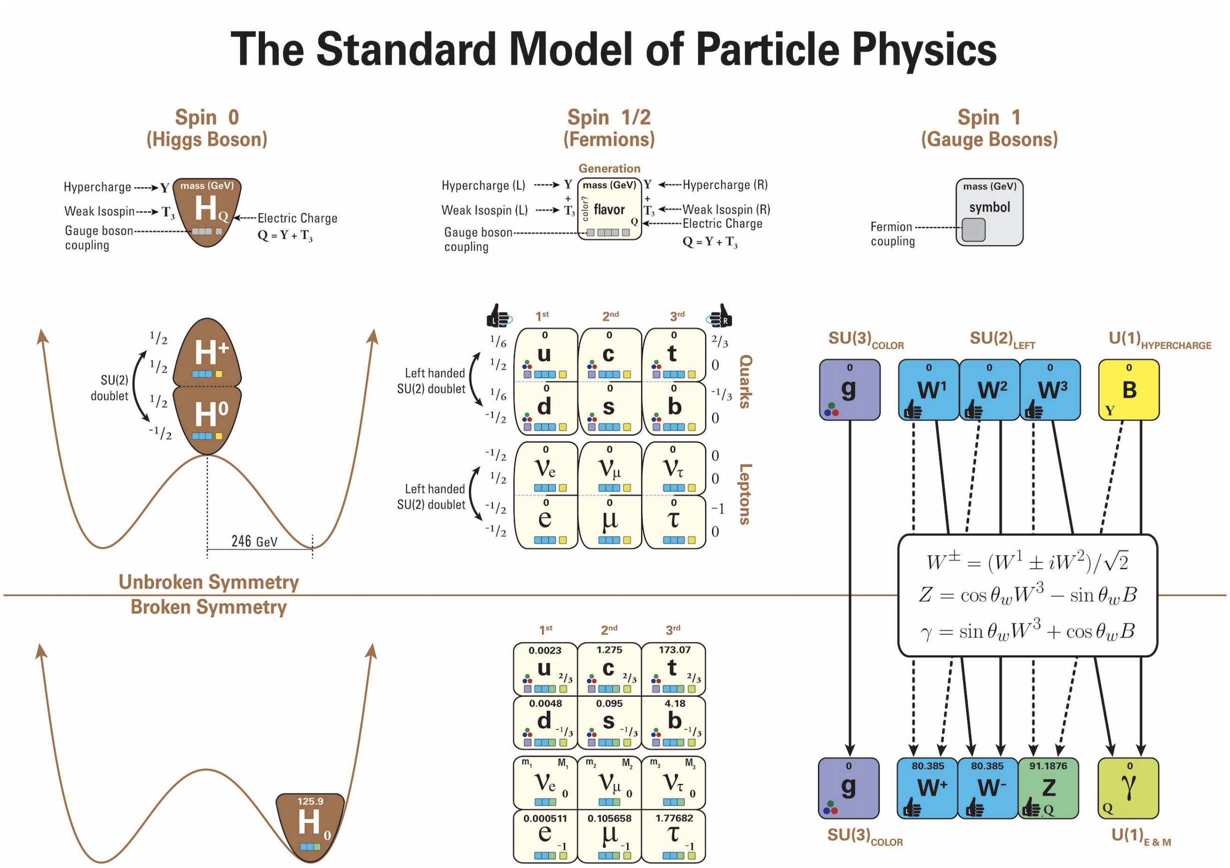


Figure 2.1: This diagram represents the particle content of the SM and the forces by which each particle interacts. The legend at the top specifies where to find each particle's name, mass, quantum numbers, gauge boson couplings, and fermion couplings. The upper pane depicts the unbroken symmetry configuration of the SM, while the lower pane depicts the SM after EWSB (see section 2.6). The left diagram shows the Higgs potential and the vacuum expectation value of the Higgs field. The mixing of electroweak gauge bosons is shown between the two panes on the right [6].

Fermions

Fermions are subdivided into two classes, quarks and leptons. Each class comes in three generations. Each generation has identical quantum numbers, except for their masses. A fundamental particle's species is also referred to as its *flavour*.

Quarks are particles with fractional electric charge, that come with $q = +\frac{2}{3}$ (up, charm, top) or $q = -\frac{1}{3}$ (down, strange, bottom). Quarks also carry a *colour charge* (see section 2.4) and appear to naturally exist only in *colourless* bound states; a quark and an anti-quark of the same colour (meson), or a triplet of differently-coloured quarks (baryons). Familiar baryons include protons (uud) and neutrons (ddu), and a common meson is the pion (e.g., $\pi^+ = u\bar{d}$). Together these bound states are known as hadrons. Quarks interact through the weak, electromagnetic, and strong forces.

Leptons are generally less massive than quarks. The charged leptons are electrons (e), muons (μ), and taus (τ), each of which carries unit electric charge. Charged leptons interact through the electromagnetic and weak forces. The neutral leptons, neutrinos, are the only electrically neutral fermions and are, therefore, relatively difficult to detect as they interact exclusively through the weak force.

A corresponding antiparticle exists for each particle type where every quantum number is reversed. Antiparticles have masses identical to their matching particles.

Gauge bosons

The spin-one gauge bosons are particles that mediate the forces experienced by fermions. Fermions interact through a fundamental force by the exchanging of relevant gauge bosons. The electromagnetic force is mediated by the *photon* (γ), which is massless and neutrally charged. The weak force is mediated by the charged W^+ and W^- and the neutral Z *electroweak bosons*, each of which is massive. The strong force is mediated by the eight *gluons* (g), which, like the photon, carry no mass and no electric charge, though they do carry the colour charge of the strong force.

Higgs boson

The only fundamental spin-0 particle in the SM is the scalar *Higgs boson*. This boson plays a distinctive role in the SM as it is responsible for imparting mass to the fundamental particles through the mechanism of electroweak symmetry breaking (section 2.6). The variation in masses between flavours is directly due to each flavour's different coupling strength to the Higgs field.

2.3 Standard Model Lagrangian

The Lagrangian density[1] $\mathcal{L}$ determines the dynamics and kinematics of the fields. The existing SM Lagrangian was developed over decades by a continuous exchange between theoretical formulation and experimental observation. At the heart of the SM Lagrangian is the idea of physical symmetries. The Lagrangian (or, more precisely, the action) must remain unchanged under some set of transformations. These transformations include the external transformations of relativistic Lorentz boosts to satisfy the constraints of special relativity, as well as internal transformations.

The most important of these internal symmetries are the *gauge symmetries* in which a different configuration of the fields results in identical observable quantities. The rich phenomena of particle physics appear to arise from the requirement that the Lagrangian remains invariant under local, continuous gauge transformations. This requirement prompts the appearance of *gauge fields* in the Lagrangian to preserve invariance. The structure of the gauge fields and how they couple to the fermionic fields define the interactions between particles of matter.

[1]Sometimes referred to only as the Lagrangian.

The SM is a non-abelian[2] gauge theory, invariant under the symmetry group

$$G = SU(3)_c \times SU(2)_L \times U(1)_Y. \tag{2.1}$$

The $SU(3)_c$ factor corresponds to quantum chromodynamics (QCD) which mediates the *strong* force. The combination of $SU(2)_L \times U(1)_Y$ corresponds to the electroweak interactions through the boson fields W_i ($i = 1, 2, 3$) and B. Through the process of electroweak symmetry breaking (EWSB), this group gives rise to the more familiar *weak* and *electromagnetic* forces (section 2.6). The c, L, and Y subscripts in eq. (2.1) do not have any group theory significance but instead indicate the physical circumstance in which they are applied. c refers to the conserved *colour charge* of the strong force, L to the fact that the $SU(2)$ coupling only acts on *left-chiral* fields, and Y indicates the conservation of the *weak hypercharge* quantum number.

2.4 Strong force

The strong force, or QCD, is the theory that describes how quarks and gluons interact with one another. Quarks and gluons carry a set of three quantum charges, distinct from the electric charge, which is called *colour charge*: r, g, b. Experimental observations find that particles only exist in nature with zero net colour charge, i.e., colour singlets. QCD is, therefore, *confining* so that the strong force's range is restricted to small distances, on the order of the size of a nucleus. Another consequence is that as two quarks separate, it becomes energetically favourable to break the colour field by producing new quarks.

The Lagrangian of the strong force is

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{q}(i\gamma^\mu D_\mu - m_q)q - \frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu}. \tag{2.2}$$

The summation is over all six quark flavours, $q = u, d, c, s, t, b$. γ^μ represents the Dirac gamma matrices, and m_q is the quark mass. G^a_μ is the gluon field strength tensor which dictates the gluon self-interaction strength. The covariant derivative D_μ is the ingredient necessary to keep this Lagrangian invariant under local $SU(3)$ transformations.

2.5 Electroweak force

In the SM, the weak force and the electromagnetic force are each aspects of the unified *electroweak* force. The weak force is notable in that it is the only force that can change the flavour of the fermion with which it interacts, and every fermion in the SM experiences weak force interactions. Furthermore, the mediators of the weak force are massive bosons, as opposed to the massless gluons and photons.

The electroweak $SU(2)_L \times U(1)_Y$ factor is *chiral* meaning the transformations act differently on left-chiral (ψ_L) and right-chiral (ψ_R) fermions. The weak force acts exclusively on left-chiral fermions, which transform as flavour doublets. In contrast, right-chiral fermions are singlets under the electroweak transformations and

[2] The composition of pairs of elements in an abelian group commute. In a non-abelian group, at least one pair does not.

thus unaffected by the weak force. In terms of the concrete SM fields, each of the three generations contains the following:

$$\begin{aligned} \text{Left-chiral doublets:} \quad & \begin{pmatrix} \nu^i \\ e^i \end{pmatrix}_L \equiv L_L, \quad \begin{pmatrix} u^i \\ d^i \end{pmatrix}_L \equiv Q_L \\ \text{Right-chiral singlets:} \quad & u^i_R, \;\; d^i_R, \;\; e^i_R, \;\; \nu^i_R? \end{aligned} \tag{2.3}$$

with u and d referring to the quark fields and e and ν referring to the lepton fields. The components of the left-chiral doublets are weak flavour eigenstates which transform into one another under $SU(2)$. Though it is not part of the current SM, a placeholder for ν_R is shown above as its existence is central to many of the proposed neutrino mass models (see section 3.3.3). If it were discovered, it would be listed among the right-chiral $SU(2)$ singlets.

The electroweak Lagrangian is typically written in four parts:

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \tag{2.4}$$

each of which is described below.

The electroweak interactions are described by

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{2.5}$$

which contains the field strength tensors W^i $(i = 1, 2, 3)$ for the $SU(2)_L$ symmetry, and B for the $U(1)_Y$ symmetry. Before EWSB, the bosons represented by these fields are not identified with the massive electroweak bosons.

$$\mathcal{L}_{\text{fermion}} = \sum_{f_L} \bar{\psi}_L i \not{D}^\mu \psi_L + \sum_{f_R} \bar{\psi}_R i \not{D}^\mu \psi_R \tag{2.6}$$

is the kinetic term for the SM fermions. The first sum is over all left-chiral doublets, and the second sum is over all right-chiral singlets. $\not{D}$ (written with Feynman slash notation), is the covariant derivative, which maintains invariance of the Lagrangian under $SU(2)$ and $U(1)$ transformations.

$$\mathcal{L}_{\text{Higgs}} = (D^\mu \phi)^\dagger D_\mu \phi - V(\phi) \tag{2.7}$$

contains the complex scalar Higgs field ϕ and its interactions with the gauge bosons and with itself. The potential term is

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{2.8}$$

μ and λ relate to EWSB as described in the following section.

$$\mathcal{L}_{\text{Yukawa}} = \sum_{\text{gen.}} -y_u \overline{Q}_L \tilde{\phi} u_R - y_d \overline{Q}_L \phi d_R - y_e \overline{L}_L \phi e_R + (h.c.) \tag{2.9}$$

describes the set of scalar-fermion *Yukawa* interactions. The first term represents up-quark-like interactions, the second, down-quark-like ones, and the third, charged lepton interactions. The factors y_u, y_d, and y_e are the coupling strengths for each interaction, and the summation is over the three generations of fermions. The conjugate Higgs field $\tilde{\phi}$ is defined as $\tilde{\phi} = i\sigma_2\phi^\dagger$, and $(h.c.)$ signifies the Hermitian conjugate terms. Here the left-chiral doublets and the right-chiral singlets couple to the Higgs doublet, which, after EWSB discussed below, is the source of fermion masses.

Notably, fermion masses do not yet appear in the SM Lagrangian. Due to the lack of left-right chiral symmetry in the $SU(2)_L$ factor of the SM gauge group, terms such as $m\bar{\psi}\psi$ would violate gauge invariance. This appears to be problematic as there is a long history of observations that the universe does indeed contain mass. Furthermore, the $W^\pm$ and Z bosons are observed to have relatively large masses, though they too are without a mass term in the above Lagrangian. Fortunately, the non-static nature of the Higgs potential in eq. (2.8) provides a neat solution by breaking a symmetry of the SM.

2.6 Electroweak symmetry breaking and the Higgs mass mechanism

As seen previously, it is not possible to add a mass term to the SM Lagrangian "by hand," in a way that leaves it invariant under the transformations of the gauge group. However, by allowing for the existence of a complex scalar doublet ϕ an effective mass term may be generated by spontaneous electroweak symmetry breaking (EWSB). In this process, the $SU(2)_L \times U(1)_Y$ gauge group breaks down into the weak and electromagnetic interactions.

The complex scalar doublet of $SU(2)_L$, which has become known as the Higgs field is

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \qquad \phi^+, \phi^0 \in \mathbb{C}, \tag{2.10}$$

which has four independent degrees of freedom. The superscripts indicate positive and neutral electric charges and $\mathbb{C}$ represents the set of complex numbers.

The lowest-energy state, called the vacuum expectation value (VEV), is given by $\langle 0|\phi|0\rangle \equiv \langle\phi\rangle$ where $|0\rangle$ is the vacuum state with zero particles. Its value is located at the minimum of the potential $V(\phi)$. Considering the Higgs potential of eq. (2.8), if $\mu^2 > 0$, the VEV is simply a stable global minimum at the origin, $\phi^\dagger\phi = 0$.

However, if the potential evolves so that $\mu^2 < 0$, as is hypothesized to have happened shortly after the big bang, the picture becomes quite different. In this case, the global minima form a hypersphere with the value:

$$\phi^\dagger\phi = \frac{-\mu^2}{\lambda} \equiv v^2 \tag{2.11}$$

such that

$$v = \sqrt{\frac{-\mu^2}{\lambda}}. \tag{2.12}$$

The consequence of this is that the corresponding field is no longer vanishing but has a finite VEV v with units of mass. The minima are degenerate along the hypersphere, but the axis in this four-dimensional space can be rotated by an $SU(2) \times U(1)$ transformation to specify a particular minimum. Without loss of generality, one can choose the axis so that only one of the four components of ϕ is left nonzero, namely the real part of the ϕ^0 component

$$\phi = \begin{pmatrix} 0 \\ H \end{pmatrix}. \tag{2.13}$$

The nature of the minima in the hypersphere can be visualized by depicting only the ϕ^0 component as seen in fig. 2.2. This choice of a definite ground state breaks the $SU(2)_L \times U(1)_Y$ invariance, though leaves the ground state invariant under a set of $U(1)_Q$ transformations corresponding to electromagnetism and conserving electric charge, Q.

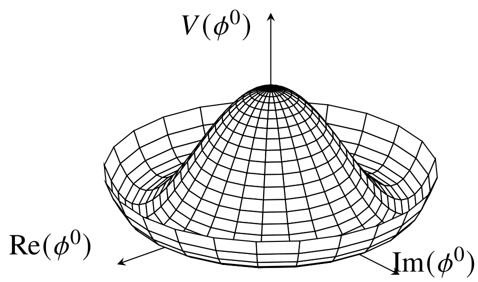


Figure 2.2: The “champagne” shape of the ϕ^0 component of the Higgs potential. The potential contains a circle of degenerate local minima. A choice of definite ground state breaks the $SU(2)_L \times U(1)_Y$ invariance and provides the Higgs field with a VEV [7].

Expanding the ground state of eq. (2.13) as small perturbations about the chosen minimum and working in the unitary gauge to remove the additional degrees of freedom, the scalar doublet becomes

$$\phi = \begin{pmatrix} 0 \\ v + h(x)/\sqrt{2} \end{pmatrix}, \tag{2.14}$$

where the $h(x)$ defines a Hermitian scalar field associated with the physical Higgs boson.

The scalar doublet of eq. (2.14) can then be substituted into the full SM Lagrangian. In the electroweak sector, the expansion of the scalar field in $\mathcal{L}_{\text{Higgs}}$ brings about the masses of the electroweak gauge bosons. W^i and B are no longer mass eigenstates, but they can be redefined by diagonalizing the mass matrix with the linear combinations

$$W^{\pm}_{\mu} = \frac{1}{\sqrt{2}}(W^1 \mp iW^2) \qquad \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix}, \tag{2.15}$$

where θ_W is the *weak mixing angle*. These represent the familiar, and now massive, electroweak gauge bosons $W^{\pm}$ and Z, and the massless photon field A.

Vitally, the expanded scalar Higgs doublet can be substituted into eq. (2.9) to generate mass terms. Considering for the moment only the electrons, the Yukawa terms of the Lagrangian take the form

$$\mathcal{L}_{\text{Yukawa},e} = -y_e\, \bar{e}_L (v + h/\sqrt{2}) e_R + (h.c.). \tag{2.16}$$

Therefore the mass term is $m_e \bar{e}_L e_R$ with $m_e = y_e v$ where v is the Higgs VEV. The coupling strength between the electron and the Higgs field h is proportional to the same Yukawa coupling constant y_e. This form of the mass term, which connects the two chiral components, is known as the Dirac mass. All fermions in the SM (perhaps except for neutrinos) gain mass by this Yukawa coupling and EWSB.

2.7 Successes and limitations of the Standard Model

2.7.1 Successes

The particle content shown in fig. 2.1 represents all visible matter in the universe. The relatively simple Lagrangian densities presented above describe the dynamics of this matter. This model is not only succinct, but it contains remarkable predictive power. The SM postulated the existence of several particles and their properties before experiments discovered evidence of them. Particles predicted and later discovered include the $W^{\pm}$ and Z bosons [8], the gluon [9], the top quark [10], the charm quark [11], and the Higgs boson [12].

Many decades of experimental research agree with the SM's theoretical predictions. For example, fig. 2.3 summarizes ATLAS measurements of cross sections for various production processes which agree with theoretical predictions over ten orders of magnitude [13].

However, the SM requires several experimental inputs to obtain the numerical values of its several free parameters. In total, 19 independent parameters are inputs to the SM: nine Yukawa couplings which dictate the particle masses; three mixing angles and one phase that defines the quark mixing matrix; three gauge couplings corresponding to the three gauge groups; the QCD vacuum angle; and the VEV and mass of the Higgs field.

2.7.2 Limitations

Despite its incredible successes, the SM still has some substantial experimental and theoretical deficiencies. For example, the lack of a satisfying explanation of gravity is seen as a considerable shortcoming of the theory. Though effective field theories exist that account for gravity, they still contain serious incompatibilities with the SM [14]. The observation of neutrino oscillations, discussed at length in chapter 3, is another phenomenon that the existing SM does not readily explain. One of the most prominent theoretical peculiarities is the so-called *hierarchy problem* in which quantum corrections to the Higgs mass require a high degree of fine-tuning in order to obtain the observed mass within the existing theoretical framework [15]. These problems point to the possibility that the existing SM is only the low-energy limit of a more fundamental theory. Below, selected observations inconsistent with the existing theory are discussed.

Dark matter

There is firm astrophysical evidence for the existence of a type of matter that may interact with known baryonic matter through gravitation alone. The most direct evidence comes from galactic rotation curves where the observed luminous matter differs significantly from the distribution of mass indicated by the stellar velocity profile. Other compelling evidence comes from gravitational lensing measurements, cosmic

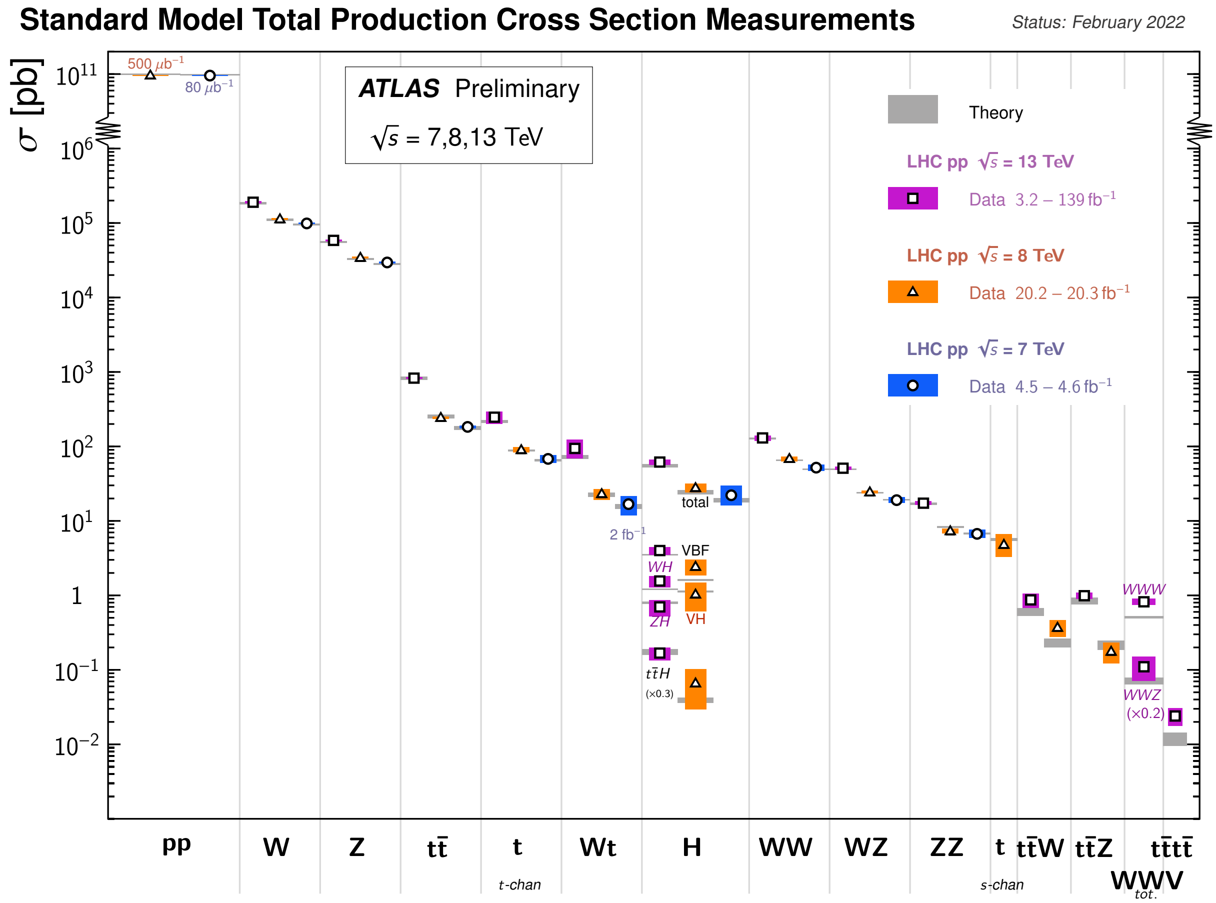


Figure 2.3: Summary of SM cross section measurements at the ATLAS detector as of February 2022. The measurements are corrected for branching fractions, compared to the corresponding theoretical expectations [13].

microwave background (CMB) radiation, and other sources. A comprehensive review of the evidence is available in reference [16]. Such observations indicate that dark matter outweighs baryonic in our universe by over a factor of 5:1 [17]. Despite concerted efforts to identify the source of dark matter, there has yet to be any direct observation of a dark matter candidate. This suggests that dark matter may interact with the known SM particles only very feebly, if at all [18].

Baryon asymmetry

The overabundance of (baryonic) matter in the universe, as opposed to antimatter, is another puzzling observation that remains unexplained by the SM. According to the prevailing cosmological theory, the universe produced in the *Big Bang* contained equal numbers of particles and antiparticles. However, two independent observations show that matter dominates over antimatter in the present universe.[3] Analyses of the power spectrum of the CMB and the products of big bang nucleosynthesis each reveal asymmetry values

[3]Other than the obvious local overabundance of matter over antimatter.

in agreement with one another [19].

In 1967, Sakharov proposed physical conditions which would be required to produce such an asymmetry. In his proposal, a suitable theory would have baryon number violation, C-symmetry and CP-symmetry violation, and out-of-thermal equilibrium dynamics leading to a degree of irreversibility. Certain theories, such as that of leptogenesis via a Majorana neutrino, lepton number violating processes in the early and hot universe could produce the observed baryon-antibaryon imbalance [20].

Chapter 3

Neutrino masses and heavy neutral leptons

The matter content of the SM contains right-handed and left-handed chiral states of each fermion except for the neutrino, which only interacts via its left-handed state. Since the right-handed neutrino ν_R does not couple to the SM gluons or electroweak gauge bosons, there is no evidence that it exists at all. Consequently, the current formulation of the SM is one in which neutrinos carry no mass, which was consistent with observation for many years.

However, since the late 1960s, observations began to show some anomalies that can now be interpreted as neutrinos transforming from one flavour state to another — a phenomenon only possible if neutrinos do, in fact, carry mass. Subsequent experiments have measured these oscillations with increasing sensitivity, and it is now clear that neutrinos have an exceptionally small but non-zero mass, on the order of $O(0.1\,\text{eV})$.

If the neutrino were to acquire mass the same way as other fermions, through electroweak symmetry breaking, the peculiar smallness of its coupling to the Higgs field might be puzzling. However, an alternative called the *seesaw mechanism* leads to appropriately small SM neutrino masses while predicting right-handed counterparts with relatively high mass. The following summary will present the relevant observations and theoretical framework to understand the basics of the neutrino seesaw.

3.1 Neutrino oscillation

Unique among the SM fermions, neutrinos have no colour or electric charge and are thus only detectable through their weak-force interactions with charged SM fermions. The SM neutrinos, ν_e, ν_μ, and ν_τ, are flavour states defined by the flavour of the charged lepton with which they interact. Observations have shown that these are not the fundamental free particle states, as evidenced by the phenomenon of *neutrino oscillation*. The following summarizes notable experiments which lay the observational foundation for such oscillations. A more complete review can be found in references [21–24].

3.1.1 A brief history of detection

Early in the 20th century, a critical violation of energy conservation appeared to be taking place inside radioactive β-decay. In the process where a nucleus emits an electron and increases its charge,

$$ {}^{A}_{Z}\mathrm{X} \rightarrow {}^{\;\;A}_{Z+1}\mathrm{X}' + e^-, \tag{3.1}$$

the electron energy was expected to equal the difference in nuclear masses Q. Instead, what was observed was a continuous spectrum like that in fig. 3.1 [25, 26]. Nearly 20 years later, Wolfgang Pauli, in a

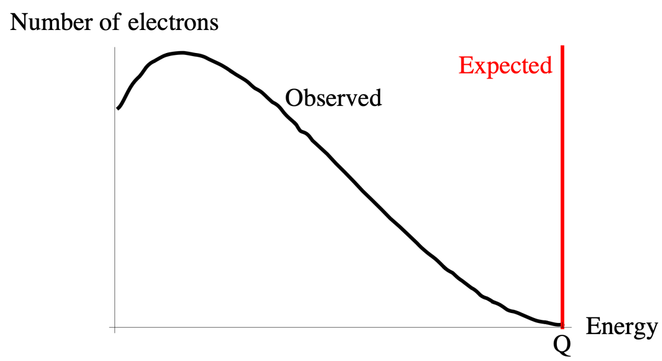


Figure 3.1: Electron spectrum of nuclear β-decay. This apparent violation of energy conservation prompted Pauli's suggestion of a "dark particle." [22]

"desperate remedy," proposed that a neutral and relatively light particle was also being produced and escaping undetected [27, 28]. If the sum of energies of the electron and the undetected particle was equal to the nuclear mass difference, energy conservation could be saved. Enrico Fermi, who was among the first to take Pauli's idea seriously, gave this neutral particle an Italian moniker meaning "little neutral one," and with it developed the theory of β-decay [29].

This interaction also allows for the absorption of neutrinos via the inverse beta process, $\bar{\nu} p \to n e^+$, though it was estimated to have an exceptionally low cross section: $\sigma_{\bar{\nu}} \leq 10^{-44}\,\text{cm}^2$ at $E_{\bar{\nu}} \sim 2\,\text{MeV}$. This would mean the neutrino has a mean free path of thousands of light-years in water.

In 1956 Reines and Cowen were the first to definitively detect neutrinos by placing large water-tank-based detectors near a nuclear reactor [30] (though, as scientists at Los Alamos, a fission bomb was originally the proposed source! [31]). The enormous neutrino flux from the reactor interacted with the protons in the water, producing a neutron and a positron which produced gamma rays in delayed coincidence.

Bruno Pontecorvo, in a 1958 paper, began to entertain the idea that oscillations may occur between neutrino and antineutrino states [32]. At that time, it was believed that only one flavour of neutrino existed. In the early 1960s, Maki, Nakagawa, and Sakata discussed that the neutrino associated with the electron ν_e might be a different particle than the neutrino associated with the muon [33]. As evidence appeared that the neutrino did come in at least two flavours, Pontecorvo, in a seminal 1968 paper, suggested they might oscillate into one another and that the evidence of such oscillations could be found by studying neutrinos produced in the sun [34]. It was not until 1998 that the Super-Kamiokande experiment announced the direct observation of neutrino oscillations by measuring atmospheric neutrinos [35].

3.1.2 Neutrino oscillation experiments

Methods to measure the oscillation of neutrino flavours comes in two signatures: first, the appearance of the "wrong" type of charged lepton from a source of known neutrino flavour and flux, and second, the disappearance of the "right" type of charged lepton. Both have been used in various experiments to measure neutrino flux from the sun, supernovae, the atmosphere, nuclear reactors, and accelerator beams.

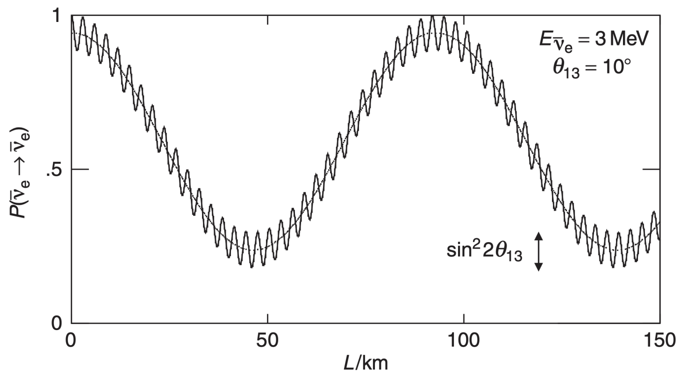


Figure 3.2: The $P(\overline{\nu}_e \to \overline{\nu}_e)$ survival probability as a function of distance for 3 MeV $\overline{\nu}_e$ assuming $\theta_{13} = 10°$. The two components of $\overline{\nu}_e$ flavour oscillation are clearly distinguishable [24].

Solar neutrinos are produced in the nucleosynthesis processes of the sun. The relatively rare solar process of boron-8 β-decay ($^8_5\text{B} \rightarrow ^8_4\text{Be}^* + e^+ + \nu_e$) produces electron neutrinos with energies up to 15 MeV, useful for earth-based detection experiments. Radiochemical experiments (Homestake [36], SAGE [37], and GALLEX [38]) measured a deficit in the ν_e flux, and later Čerenkov experiments (Super-Kamiokande [39]) confirmed this "solar neutrino problem." The Sudbury Neutrino Observatory (SNO) used heavy water to separately detect the ν_e and total neutrino flux from the sun. It found that the total flux matched predictions but detected a substantial $\nu_\mu + \nu_\tau$ component. Since only ν_e is produced substantially in the sun, the SNO experiment showed clear evidence of neutrino flavour transformation [40].

Atmospheric neutrinos are created when cosmic rays interact with the atmosphere, mainly via the production of pions. Consider the dominant decay process of a π^- or π^+ meson:

$$\begin{aligned} \pi^- &\rightarrow \mu^- + \overline{\nu}_\mu \rightarrow (e^- + \overline{\nu}_e + \nu_\mu) + \overline{\nu}_\mu \\ \pi^+ &\rightarrow \mu^+ + \nu_\mu \rightarrow (e^+ + \nu_e + \overline{\nu}_\mu) + \nu_\mu \end{aligned} \tag{3.2}$$

It is clear that (for sufficiently low energies)[4] that the ratio of muon-type neutrinos to electron-type neutrinos from the atmosphere is 2:1. Since the 1980s, this ratio has been measured by several experiments including several calorimeter-like detectors (NUSSEX [41], Frejus [42], Soudan 2 [43], and MACRO [44]), and water Čerenkov detectors (Kamiokande [45] and IMB [46]). Some of these experiments indicated that the flux of muon-type neutrinos was significantly less than expected. In 1998 the Super-Kamiokande experiment announced results that demonstrated beyond any reasonable doubt that muon neutrinos were indeed disappearing along their path from the atmosphere to the detector [35]. Furthermore, it found that only the ν_μ flux was much lower if it came from beneath the detector (i.e., from the atmosphere on the other side of the earth). This observation was strong evidence for the oscillation of muon neutrinos into other types.

Further experiments revealed that neutrino oscillation occurs on two different length scales thus many experiments come in "short-baseline" and "long-baseline" types. **Reactor neutrinos** are produced at high

[4] An increase in the energy of the muons decreases the fraction of those muons that decay in flight.

flux from the decay of radioisotopes produced in fission reactors. The short-wavelength component, $O(1\,\text{km})$, is often measured by reactor neutrino experiments (Daya Bay [47], RENO [48]). **Accelerator neutrinos**, produced in a beam by the decay of accelerated muons, are used to measure the long-wavelength component, $O(100\,\text{km})$, in experiments such as MINOS [49].

3.2 Mass and weak eigenstates

From these observations, it is clear that neutrinos oscillate to different flavours of charged lepton with which they interact. Just as is done for quarks, one can identify the *mass eigenstates* as the stationary states of the free-particle Hamiltonian, and the *weak eigenstates* (or *flavour eigenstates*) as the states interacting with the respective charged lepton. There is no reason why mass eigenstates need to be identical to flavour eigenstates, and neutrino oscillation indicates that they are not. The following “plane wave” derivation presents a simplified model involving two neutrinos which can be easily extended to the three-neutrino case.

3.2.1 Two-neutrino system

First, consider a system consisting of two neutrinos where the weak eigenstates, ν_e and ν_μ, are coherent linear superpositions of the mass eigenstates, ν_1 and ν_2. The mass eigenstates have rest mass m_1 and m_2 respectively and propagate as plane waves like

$$\begin{aligned} |\nu_1(t)\rangle &= |\nu_1\rangle e^{i(\mathbf{p_1}\cdot\mathbf{x}-E_1 t)} = e^{-ip_1\cdot x} \\ |\nu_2(t)\rangle &= |\nu_2\rangle e^{i(\mathbf{p_2}\cdot\mathbf{x}-E_2 t)} = e^{-ip_2\cdot x} \end{aligned} \tag{3.3}$$

where $p\cdot x = Et - \mathbf{p}\cdot\mathbf{x}$ is the Lorentz-invariant phase.

A 2×2 unitary matrix relates the mass and weak eigenstates and can be expressed with a single mixing angle θ by

$$\begin{pmatrix}\nu_e\\ \nu_\mu\end{pmatrix} = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\nu_1\\ \nu_2\end{pmatrix} . \tag{3.4}$$

So a neutrino produced in the process $\mathrm{u}\to\mathrm{d}\mathrm{e}^+\nu_e$ will propagate as the coherent superposition of the mass eigenstates ν_1 and ν_2,

$$|\nu_e(\mathbf{x})\rangle = \cos\theta|\nu_1\rangle e^{-ip_1\cdot x} + \sin\theta|\nu_2\rangle e^{-ip_2\cdot x} . \tag{3.5}$$

The proportion of mass eigenstates then depends on both the mixing angle θ and the difference between the phases $\Delta\phi_{12} = (p_1\cdot x) - (p_2\cdot x)$. To find the probability that the state that was initially produced as ν_e will interact at some distance L as ν_μ to produce a muon is found by contracting with the $\langle\nu_\mu|$ state:

$$\langle\nu_\mu|\nu_e(L)\rangle = \sin^2(2\theta)\sin^2\left(\frac{\Delta\phi_{12}}{2}\right) . \tag{3.6}$$

At this point, some assumptions are made that, although reasonable, are not rigorous. By assuming that $m_\nu \ll E_\nu$, that $p_1 \approx p_2 = E$, and that the mass eigenstates propagate with the same velocities (namely

$v \approx c$), we find that the $\nu_e \to \nu_\mu$ *transition probability* is

$$P(\nu_e \to \nu_\mu) = \langle \nu_\mu | \nu_e(L) \rangle = \sin^2(2\theta)\sin^2\left(\frac{\Delta m_{12}^2}{4E}L\right). \tag{3.7}$$

It follows that the *survival probability* of an electron neutrino can be found using the conservation of probability, $P(\nu_e \to \nu_e) = 1 - P(\nu_e \to \nu_\mu)$. Ultimately the flavour oscillation depends on the mixing angle θ and the mass difference of the mass eigenstates Δm_{12}^2.

A more rigorous derivation accounts for the velocity difference in mass eigenstates by treating them as wave packets. An even more formal treatment uses the framework of QFT and effectively treats the neutrino as a virtual particle exchanged between the interactions at the source and the detector. Fortunately, each of these approaches gives the same result as the plane wave approximation, eq. (3.7).

3.2.2 Three-neutrino system and the PMNS matrix

Extending the same mass-flavour eigenstate relationship to three neutrinos is done using the 3×3 unitary Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix U_{PMNS} [33].

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}. \tag{3.8}$$

The unitarity of the PMNS matrix ensures $U^{-1} = U^\dagger \equiv (U^*)^T$, and $U^\dagger U = I$.

If U_{PMNS} were real it could be described by three rotation angles θ_{12}, θ_{13}, and θ_{23}. However, it is complex and needs an additional six complex phases. Five of these can be absorbed into the neutrino and charged lepton states with no physical consequences, and so the full description of U_{PMNS} needs only a single complex phase δ in addition to the rotation angles.[5]

The full expression for the survival probability of an electron neutrino is then

$$P(\nu_e \to \nu_e) = 1 - \cos^4(\theta_{13})\sin^2(2\theta_{12})\sin^2\left(\frac{\Delta m_{21}^2}{4E_\nu}L\right) - \sin^2(2\theta_{13})\sin^2\left(\frac{\Delta m_{32}^2}{4E_\nu}L\right), \tag{3.9}$$

where $\Delta m_{ij}^2 \equiv (m_i^2 - m_j^2)$.[6] As experiments have found that $\Delta m_{21}^2 \ll \Delta m_{32}^2$, this quantifies the disappearance of ν_e with oscillations on two different length scales. The short-wavelength component depends on Δm_{32}^2 while the long-wavelength component depends on Δm_{21}^2. The mass splittings can be measured by observing neutrinos at different distances or varying the neutrino production energy. The rate of oscillation in the three-flavour scenario thus depends on six independent parameters: the three mixing angles, one complex phase, two mass differences.

KamLAND (a long-baseline reactor experiment) and MINOS (a long-baseline muon-beam experiment)

[5]In the case of three Majorana neutrinos, no independent right-handed field exists, and only three phases can be absorbed by the fields. In this case, the U_{PMNS} needs three complex phases instead of one. See reference [21] for a more complete description.

[6]The oscillations depend on only two differences in m^2 of the mass eigenstates as Δm_{21}^2 and Δm_{32}^2 ($\Delta m_{31}^2 = \Delta m_{32}^2 + \Delta m_{21}^2$).

have measured the mass differences to < 5% precision [21, 24]:

$$\Delta m^2_{21} = (7.6 \pm 0.2) \times 10^{-5}\text{eV}^2, \quad |\Delta m^2_{32}| = (2.3 \pm 0.1) \times 10^{-3}\text{eV}^2 \; . \tag{3.10}$$

Note that only the absolute value of Δm^2_{32} has been measured, which gives rise to the question of *neutrino mass hierarchy*. The differences in neutrino masses have been precisely measured, but not the absolute values. It is possible that either $m_3 > m_{1\approx 2}$ or $m_3 < m_{1\approx 2}$ and no experiment to date has determined the correct ordering [50]. The upper bound on the sum of neutrino masses is found from cosmological observations to be $\sum_i m_{\nu_i} < 0.2\,\text{eV}$ [21].

Results from several experiments including SNO, KamLAND, Super-Kamiokande, MINOS, Daya Bay, Reno, and Double Chooz have together measured the following values for the mixing angles [21, 24]:

$$\sin^2(2\theta_{12}) = 0.87 \pm 0.04, \quad \sin^2(2\theta_{23}) > 0.92, \quad \sin^2(2\theta_{13}) = 0.10 \pm 0.01 \tag{3.11}$$

Only in 2012 was θ_{13} found to be non-zero. Finally, the phase δ has yet to be measured and is the focus of next-generation experiments. Its existence has consequences for charge-parity (CP) violation in neutrino oscillations [51].

3.3 Mass mechanisms

There are two common ways to extend the SM to produce a neutrino mass. The first predicts that the neutrino is a Dirac fermion, in which the corresponding mass mechanism is straightforward, though arguably unsatisfying. The second predicts that the neutrino may be a Majorana fermion and that the small neutrino masses are due to the presence of much heavier neutrinos that have yet to be detected. This section will focus exclusively on the electron generation, but the argument extends naturally to the other generations of leptons. The following combines derivations found in references [4, 21, 22, 24].

3.3.1 Dirac mass

Consider the left-chiral lepton doublet L_L (see eq. (2.3))

$$L_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} . \tag{3.12}$$

A simple extension to the SM is to introduce a new right-handed neutrino ν_R that transforms as a singlet under all SM gauge groups. This field is distinct from the existing right-handed lepton singlet e_R. The Yukawa factor in the Lagrangian of eq. (2.9) then contains the term

$$\begin{aligned} -\mathcal{L}_{\text{Yukawa}} &\supset y_\nu \overline{L}_L \tilde{\phi} \nu_R + (h.c.) \\ &= y_\nu (v + h/\sqrt{2}) \bar{\nu}_L \nu_R + (h.c.) \end{aligned} \tag{3.13}$$

where the second line is the result of EWSB of the Higgs scalar doublet $\tilde{\phi}$ (see section 2.6). Therefore, the neutrino acquires a Dirac mass,

$$m_D = y_\nu v \tag{3.14}$$

where v is the VEV of the Higgs field. This is the same mechanism by which the up-type quark acquires mass when coupled to the quark doublet Q_L in eq. (2.9). The Dirac mass term for the neutrino is then

$$-\mathcal{L}_D = m_D(\overline{\nu_R}\nu_L + \overline{\nu_L}\nu_R) \tag{3.15}$$

Aside from the Yukawa interaction with the Higgs doublet, this right-handed neutrino would have no other interactions via the SM since it is a singlet of the complete SM gauge group.

This results in a valid mass term and implies that a right-handed chiral neutrino exists, yet this mechanism is not entirely satisfying. The neutrino masses (and therefore the Yukawa coupling constant y_ν) are smaller than the other fermion masses by an order of 10^6, which could be seen as suspicious. Perhaps there is another mechanism responsible for the smallness of the neutrino mass.

3.3.2 Majorana mass

An alternative mechanism by which neutrinos can acquire mass is through a *Majorana mass term*. First, let the CP conjugate field for the right-handed neutrino be defined as

$$\nu_R^c = \hat{C}\hat{P}\nu_R \tag{3.16}$$

which corresponds to a left-handed antineutrino $\overline{\nu_L}$. $\hat{C}$ is the charge conjugate operator and the $\hat{P}$ is the parity transformation operator. The Majorana mass term, then, can be written as

$$-\mathcal{L}_M = \frac{1}{2}M(\overline{\nu_R^c}\nu_R + \overline{\nu_R}\nu_R^c)\ . \tag{3.17}$$

where M is the Majorana mass. Since this is formed from right-handed neutrino and left-handed antineutrino fields alone, both of which transform as singlets under the full SM gauge transformations, this mass term does not break the gauge invariance of the SM. It is worth highlighting that this proposed "bare" mass term is unique among SM fermions, the rest of which are not singlets under all SM gauge groups and therefore do not allow this type of bilinear coupling [24].

3.3.3 Type-I seesaw mechanism

It is now possible to re-express the Dirac and Majorana mass terms to combine both masses in matrix form. After some algebra, it is found that the masses of the physical states notably relate to one another.

Rewriting the Dirac Lagrangian of eq. (3.15) as

$$-\mathcal{L}_D = \frac{1}{2}m_D(\overline{\nu_L}\nu_R + \overline{\nu_R^c}\nu_L^c)\ , \tag{3.18}$$

we can combine it with the Majorana Lagrangian to form the most general renormalizable Lagrangian containing neutrino masses:

$$-\mathcal{L}_{DM} = \frac{1}{2}(m_D \overline{\nu_L} \nu_R + m_D \overline{\nu_R^c} \nu_L^c + M \overline{\nu_R^c} \nu_R) \ . \tag{3.19}$$

This can be made more compact using a mass matrix as

$$-\mathcal{L}_{DM} = \frac{1}{2} \begin{pmatrix} \overline{\nu_L} & \overline{\nu_R^c} \end{pmatrix} \begin{pmatrix} 0 & m_D \\ m_D & M \end{pmatrix} \begin{pmatrix} \nu_L^c \\ \nu_R \end{pmatrix} \ . \tag{3.20}$$

By diagonalizing the mass matrix **M** one can obtain the physically observable states of the system. The masses of these states are obtained from the eigenvalue expression

$$\det(\mathbf{M} - \lambda I) = \lambda^2 - M\lambda - m_D^2 = 0 \tag{3.21}$$

which then gives the masses

$$m_\pm = \lambda_\pm = \frac{M \pm M\sqrt{1 + 4m_D^2/M^2}}{2} \ . \tag{3.22}$$

In the circumstance that the Majorana mass is much larger than the Dirac mass, $M \gg m_D$, the masses reduce to

$$|m_-| \approx \frac{m_D^2}{M} \equiv m_\nu \quad \text{and} \quad m_+ \approx M \equiv m_\mathcal{N} \tag{3.23}$$

where $m_\mathcal{N}$ and m_ν correspond to heavy and light neutrino states respectively. The physical states of this system obtained from the mass matrix are

$$\nu \approx (\nu_L + \nu_L^c) - \frac{m_D}{M}(\nu_R + \nu_R^c) \quad \text{and} \quad \mathcal{N} \approx (\nu_R + \nu_R^c) + \frac{m_D}{M}(\nu_L + \nu_L^c). \tag{3.24}$$

What is fascinating about this result is that now the Dirac mass terms for the neutrinos can be of the same order as those for the other fermions, $\mathcal{O}(\text{GeV})$. The small mass of the light neutrino states is a direct consequence of the large Majorana mass M, hence the "seesaw." We no longer need to contend with a six-order-of-magnitude variation in the Yukawa coupling among fermions.

Finally, the new light neutrino states ν couple to the weak charged-current interaction effectively the same as the existing SM neutrinos since the ν states are almost entirely left-handed. Similarly, the heavy states $\mathcal{N}$ are almost entirely right-handed and would rarely interact via the charged-current or neutral-current interactions if they exist. $\mathcal{N}$ is referred to as a Heavy Neutral Lepton (HNL) for its relatively large mass or a *sterile neutrino* for its non-interaction with the rest of the SM particles.

3.3.4 Searches for Majorana fermions

If the Majorana mass term (eq. (3.17)) is nonzero, then neutrinos are Majorana fermions meaning that the neutrino is its own antiparticle: $\nu = \overline{\nu}$. For instance, if electrons were Majorana particles, the transition $e^+ \leftrightarrow e^-$ would be allowed, violating charge conservation. Interestingly, for neutrinos, the observable effects of this feature would be rather difficult to detect.

To illustrate the faint effect, consider the following process: A pion decays, producing $\pi^+ \to \mu^+\nu_\mu$, where the ν_μ then interacts via inverse-β decay, $\nu_\mu n \to \mu^- p$. The number of leptons minus the number of antileptons, or lepton number L, remains constant. If, on the other hand, the neutrino was its own antiparticle, then it could interact as a right-handed chiral state and produce $\nu_\mu n \to \mu^+ p$. What remains would be two antimuons, breaking lepton number conservation by two units. Due to the small mass of the neutrino, it will travel near light speed, making its helicity state and chirality state nearly identical. Therefore, this process is suppressed by $O(m_\nu^2/m_\mu^2) \sim 10^{-9}$ and is virtually impossible to measure.

An alternative method of investigation is to search for signs of *neutrinoless double β-decay* ($0\nu\beta\beta$) which is a process that occurs only if the neutrinos are Majorana particles. Even-even nuclei, kinematically forbidden from undergoing normal β-decay, can decay via simultaneous β-decay processes to lower-energy even-even nuclei. The energy spectrum of the electrons produced this way would be continuous due to the escape of the neutrinos, similar to that shown in single β-decay in fig. 3.1.

If the neutrino were a Majorana particle, then the process shown in fig. 3.3 could take place and would produce no external neutrinos. The process would produce mono-energetic electrons, each carrying half the mass difference of the parent and daughter nuclei. Unfortunately, this double β-decay process has a half-life of 10^{19}–10^{25} years, so observing these processes is already quite rare. To date, there is no confirmed discovery of neutrinoless double β-decay [22, 24].

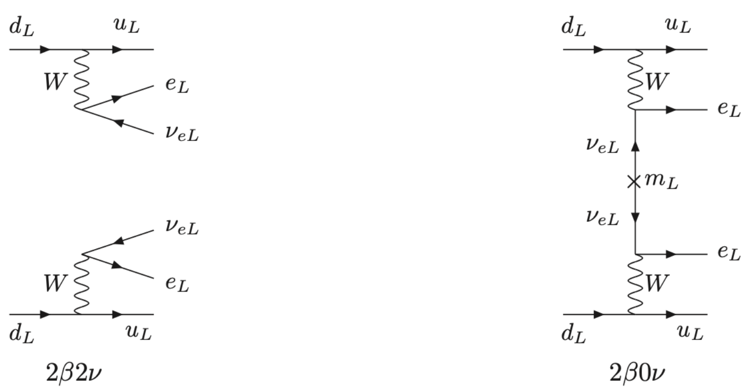


Figure 3.3: Normal double β-decay (left). Neutrinoless double β-decay (right) [22]. If the neutrino is a Majorana particle, it is its own antiparticle, and the process on the right would be allowed. This is one of the most promising methods of identifying whether a neutrino is a Majorana particle.

3.3.5 The Weinberg operator and other types of seesaw

Another way to express the Majorana term starts with the combination $\overline{L}_L\tilde{\phi}$ — a contraction of the conjugate lepton doublet and conjugate Higgs doublet — which is a singlet under all SM gauge groups. The contraction

$$-\mathcal{L} \supset \overline{L}_L\tilde{\phi}\alpha C(\overline{L}_L\tilde{\phi})^T + (h.c.) \tag{3.25}$$

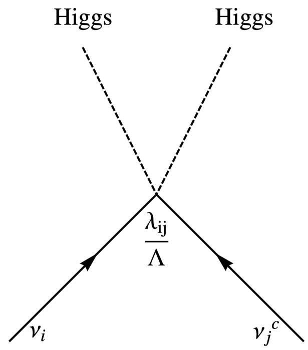


Figure 3.4: The Weinberg operator is a dimension-5 operator that couples a neutrino-antineutrino pair to two Higgs fields [22].

is gauge invariant and becomes

$$-\mathcal{L} \supset \overline{\nu}_L \alpha v^2 C \overline{\nu}_L^T + (h.c.) \tag{3.26}$$

under EWSB. Equation (3.25) is known as the *Weinberg operator* and eq. (3.26) is the same Majorana mass term that was shown in eq. (3.17). Figure 3.4 demonstrates the Weinberg operator as a coupling between a pair of Higgs bosons and a pair of Majorana neutrinos.

As fermion fields have mass dimension $[\psi] = \frac{3}{2}$ and the Higgs scalar has mass dimension $[\phi] = 1$, the Weinberg operator is a dimension-5 operator. The Lagrangian term can be kept at dimension-4 by setting the coupling α to have mass dimension $[\alpha] = -1$. By expressing α in terms of a dimensionless operator λ and some new physics scale Λ (unrelated to the electroweak scale[7]), the SM neutrino mass is then

$$m_\nu = \frac{\lambda v^2}{\Lambda}. \tag{3.27}$$

Neutrino masses are then suppressed by a factor of v/Λ compared to charged fermions. This should begin to look familiar. By taking the new physics scale to be that of the Majorana mass in eq. (3.23) and setting it to be much larger than the electroweak scale ($\Lambda \gg v$), it is exactly the relationship seen in the suppression of the mass term for light neutrinos. Furthermore, the Weinberg operator violates lepton number conservation like the Majorana neutrino.

The Weinberg operator may behave for neutrino-Higgs interactions like the low-energy Fermi operator does for β-decay, where interactions really take place by the exchange of some massive particle as seen in fig. 3.5. These possible particles are classified and result in three types of seesaw mechanism, the first of which was detailed in section 3.3.3.

Type I	:	Heavy singlet fermions	N
Type II	:	Heavy triplet scalar	Δ
Type III	:	Heavy triplet fermions	Σ

[7]The EW scale refers to energies on the order of 100 GeV where EW theory is important in describing processes at that energy. The W, Z, and Higgs boson masses are on the order of the EW scale.

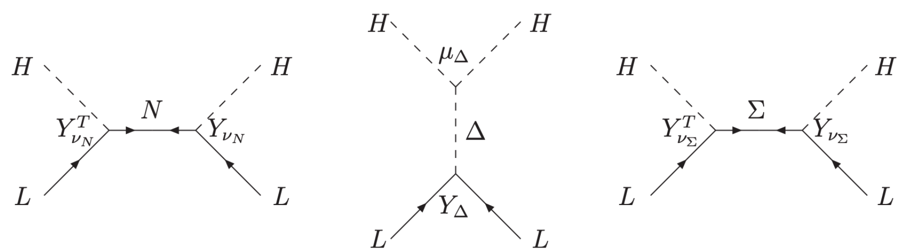


Figure 3.5: The three generic realizations of the seesaw mechanism, depending on the nature of the heavy field exchanged [52]. This is a "magnified" view of the Weinberg operator in fig. 3.4.

We can also see this from the perspective of an effective field theory. Physics beyond the SM characterized by a high energy scale Λ will induce lower-energy effects described by an effective field theory.

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{\alpha_i}{\Lambda} O_i^{d=5} + \sum_i \frac{\beta_i}{\Lambda^2} O_i^{d=6} + \dots \tag{3.28}$$

The lower the dimension, the more relevant the operator is at low energies, and as it so happens, the only allowed $d = 5$ operator that can be constructed from SM fields (and that obeys the necessary symmetries) is the Weinberg operator. It certainly does seem reasonable to expect that Majorana neutrino masses could be one of the first indications of BSM physics [22].

3.3.6 Are heavy neutral leptons beyond the Standard Model?

A familiar disagreement is whether neutrino oscillations constitute evidence of physics beyond the SM. The original $SU(2) \times U(1)$ Glashow-Weinberg-Salam model of electroweak interactions remains the basic framework of neutrino physics, and no evidence of theoretical extensions is yet verified [3].

As shown in section 3.3, the most straightforward extension would be to introduce a lepton-number conserving right-handed ν_R field, which couples to the active ν_L field with a Dirac mass. This extension requires no lepton-number-violating Majorana fields to enter the SM Lagrangian, and the consequent mass mechanism takes a familiar form. Nevertheless, a simple extension to the SM is still an extension.

Perhaps one of the sources of the confusion is that the Standard Model has no standard that all physicists agree is the authentic version. The closest example of a definitive text may be the Particle Data Group's "Review of Particle Physics," which has become the most-cited publication in the field. The most recent publication [21] contains a review of neutrino masses, mixing, and oscillations which explicitly states, "We conclude that within the SM neutrinos are precisely massless. Consequently one must go beyond the SM in order to add a mass to the neutrino."

This dissertation will assume that the SM presented in section 2.1 does not contain any mechanism which allows for neutrinos with non-zero mass, and consequently, that observations of neutrino oscillations are strictly BSM phenomena.

3.4 Motivation for HNL searches

Though nearly completely right-chiral, the theoretical HNL contains a small admixture of an active neutrino flavour: ν_e, ν_μ, ν_τ. This allows it to participate in weak interactions, albeit suppressed by a dimensionless mixing angle U_α, where $\alpha = e, \mu, \tau$. The discovery of such a particle, or class of particles, has the potential to resolve three major open physics questions (see section 2.7.2).

First, a right-handed neutrino would provide a mechanism by which the active neutrinos gain their non-zero masses. Moreover, if the HNL is a Majorana fermion (its own antiparticle), the mass may be imparted by a mechanism different than all other SM fermions (see section 3.3). Second, if at least one HNL were found to be in the $\mathcal{O}$(keV) mass range, it would constitute a potential dark matter candidate. The HNL would be a type of decaying dark matter, putting an upper bound on its mixing with active neutrinos [53]. Finally, the baryon asymmetry of the universe — the excess of matter over antimatter — could be explained by the possible Majorana nature of the HNL. The Majorana mass term allows neutrinos to become antineutrinos, allowing for a natural mechanism for generating a net lepton number in the hot and early universe, aka *leptogenesis*. Along with *sphalerons* furnishing the means to convert baryons to antileptons (or the conjugate inverse), the HNLs may be the fundamental source for the observed baryon asymmetry [20, 54, 55].

Clearly, finding direct evidence of an HNL would have substantial impacts on physics and astronomy, and many searches for such evidence have been performed over decades. References [56–58] present an extensive review of HNL experiments and constraints, both existing and proposed. Typically, experiments to date assume that only one HNL is kinematically accessible and that other HNLs are outside the experiment's sensitivity. Thus, only two independent variables must be considered: the mass of the HNL and the coupling strength $|U_\alpha^2|$.

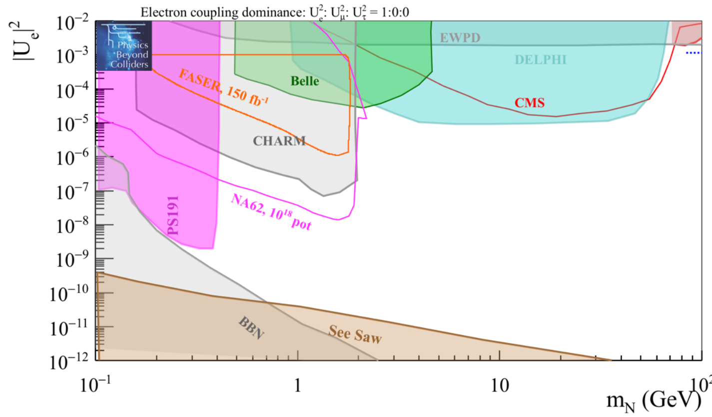
Stringent constraints have been placed on sterile-active neutrino mixing in a wide mass range from $\mathcal{O}$(eV) to $\mathcal{O}$(TeV) [59]. Data from oscillation experiments exclude masses of $M_N < 1\,\text{eV}$. Constraints on mixing with ν_e are set by $0\nu\beta\beta$ experiments and β-decay energy spectra for $10\,\text{eV} < M_N < 1\,\text{MeV}$. Constraints on ν_e and ν_μ mixing in the range $1\,\text{MeV} < M_N < 1\,\text{GeV}$ are set by searching for peaks in the resulting energy spectrum of leptonic pion and kaon decays. Decay products from beam-dump experiments have set complementary limits on mixing to all active neutrino flavours. Lower bounds on the mixing angles are set by considerations of Big Bang Nucleosynthesis [60].

3.4.1 GeV-scale HNLs and displaced vertices

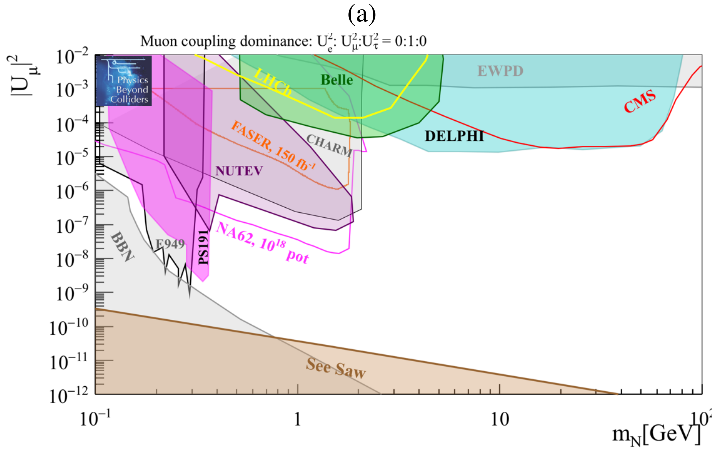
Of particular relevance to LHC searches GeV- to TeV-scale HNLs with masses up to the the TeV scale being probed in the Majorana limit using prompt leptons while dedicated searches looking for displaced vertices probe the GeV range. In pp colliders, searches for Majorana HNLs in this mass range often focus on a "smoking gun" lepton-number violating (LNV) signature of same-sign, same-flavour dilepton plus jets with no missing transverse energy: $pp \to W^* \to N\ell^\pm \to \ell^\pm\ell^\pm jj$ [59]. There is a distinct interest in searching for HNLs with masses below the electroweak scale (below the W, Z, and Higgs boson masses). In a GeV-scale seesaw model, HNLs can simultaneously explain the masses of active neutrinos and the observed baryon asymmetry. A simple extension to the SM with three HNLs below $\mathcal{O}(100\,\text{GeV})$ including a dark matter candidate at O(keV) is known as the neutrino Minimal Standard Model (νMSM) [61, 62]. Depending on the exact masses of the three HNLs, νMSM has the potential to solve all three of the above SM limitations without involving a new energy scale. GeV-scale HNLs may also play a role in flavour violation, lepton universality violation, and $0\nu\beta\beta$ (see reference [56] and references therein for details).

A broad range of experiments probes HNLs in the range $100\,\text{MeV} < M_N < 500\,\text{GeV}$. $0\nu\beta\beta$ searches, peak searches in meson decays, beam dump experiments, rare LNV decays of mesons, Z boson decays, electroweak precision tests, and direct collider searches are all used to place complementary limits on the sterile-active mixing angles. Readers are directed to reference [59] for a detailed overview of these searches.

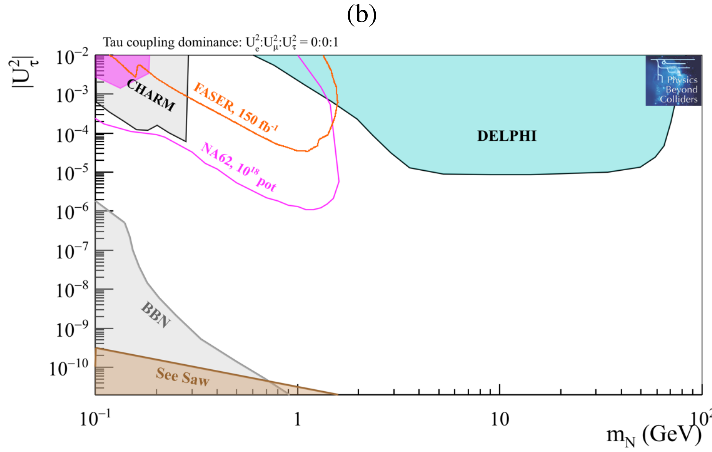
One method to constrain theoretical HNL properties is to search for their decay products since the HNL's mixing with active neutrinos allows it to participate in weak interactions. This enables (if kinematically allowed) HNL decays via charged-current (CC) and neutral-current (NC) processes, resulting in visible final states, including charged leptons, pions, or kaons. Collider experiments, in particular, make use of this avenue of discovery. The DELPHI experiment at LEP1 used neutrinos from Z boson decays, which until the LHC experiments, provided the tightest constraints in the HNL mass range 2–75 GeV [63]. HNLs are probed in hadron colliders using W rather than Z boson decays due to the higher production cross section and certain trigger requirements. Graphical overviews of the limits referenced above are shown in fig. 3.6.



(a)



(b)



(c)

Figure 3.6: Limits on the mixing between an active neutrino and a single HNL in the mass range 0.1–100 GeV. The active neutrino is ν_e in (a), ν_μ in (b), and ν_τ in (c). For details see reference [58].

Chapter 4

Experimental apparatus

To test the full range of particles and interactions in the SM, exotic forms of matter must be created and their behaviour measured. Many of the particles of interest are relatively massive, so large amounts of energy must be imparted to each interaction for high-mass particles to be produced. Furthermore, many processes of interest occur with a very low probability, so an enormous number of interactions must be produced and recorded to enable a statistically robust measurement. Colliding beam experiments are currently the best-suited method to provide large numbers of the highest-energy collisions in a controlled environment.

The LHC is a hadron collider that delivers primarily proton-proton (pp) collisions. At $\sqrt{s} = 13$ TeV, the produced collisions have the highest centre-of-mass (COM) energy ever achieved by humans at the time of this writing. The LHC also operates in alternative *heavy-ion* modes which studies lead-lead or proton-lead collisions. The major experiments located at the LHC collision points are ATLAS and CMS, general-purpose particle detectors, LHCb, which focuses on b-hadron physics, and ALICE, which studies heavy-ion collisions.

The following chapter describes the experimental apparatus that enables the analyses presented in chapters 6 and 7. Section 4.1 describes the LHC and its relevant parameters. Section 4.2 presents a summary of the components of the ATLAS detector. Although the LHC is capable of colliding heavy ions, discussion in this chapter is restricted to the proton-proton collisions.

4.1 The Large Hadron Collider

The LHC is the world's largest and highest-energy particle accelerator and collider [64]. By most metrics, it is the largest machine ever built [65]. The subterranean collider is, on average, 100 m underground in the 26.7 km tunnel initially constructed for the Large Electron-Positron Collider (LEP) collider. The LHC, a synchrotron accelerator, is designed to be capable of $\sqrt{s} = 14$ TeV centre-of-mass (COM) collisions by colliding two separate proton beams, each sectioned into bunches.

Each beam travels in an ultrahigh vacuum beam pipe, with pressures reaching as low as 10^{-11} mbar. The collider is separated into eight straight sections and eight arcs. The straight sections, each roughly 528 m long, contain experiments and collimation, acceleration, and beam dumping utilities. Figure 4.1 shows a schematic of the sections of the LHC and the four major experiments.

Acceleration of the beam takes place in one utility segment using a radio frequency (RF) cavity system. As the acceleration occurs, powerful dipole magnets steer the beams, and quadrupole magnets focus them, repeating the circuit several million times until the beams cross at the four interaction points where the experimental detectors are positioned. The following sections will describe the LHC beam parameters and injection chain in the Run 2 data-taking period.

4.1.1 Accelerator and injection chain

For the proton beams to reach the highest energies, they undergo a sequence of accelerations in an *injection chain* [64]. The protons within the LHC beam begin their journey supplied by an unassuming 10 L tank of hydrogen [66].[8] The hydrogen molecules are stripped of electrons, and the resulting bare protons are accelerated to 50 MeV by the Linac2 linear accelerator.[9] The beam feeds into the Proton Synchrotron Booster (PSB), where the protons accelerate to 1.4 GeV, then are injected into the Proton Synchrotron (PS). The PS, CERN's first synchrotron in operation since 1959, boosts the energy of the protons to 25 GeV. The next destination for the proton beam is the Super Proton Synchrotron (SPS) which brings each proton to 450 GeV. Finally, the beam is ready for injection into the LHC.

Until this point, all protons travel together, circulating in the same direction. The SPS injects protons into the two separate beam pipes of the LHC so that the resulting beams travel in opposing directions. A complete fill of the LHC takes a minimum of 16 minutes. Only after this filling process does the acceleration stage begin. On each pass through the accelerating segment of the LHC, each beam passes through an RF cavity operating at 400 MHz. Accelerating the beam from 450 GeV to the maximum energy of 7 TeV takes approximately 20 minutes.

Each machine in the injection chain was developed initially to enable several other lower-energy experiments before they were re-commissioned to supply high-energy protons to the LHC. It is a testament to CERN's re-use of existing infrastructure that many of the chain's components are upgraded versions of decades-old accelerators. Not surprisingly, the proposed Future Circular Collider (FCC) would likely use the LHC as the final injection synchrotron [69].

4.1.2 Beam characteristics

Physics studied in the ATLAS detector depends on the nature of particles produced by collisions in the LHC. It is vital to understand the properties of the beams and the frequency with which protons collide inside the detector.

Bunches

Instead of a continuous beam, the protons are separated into individual packets called *bunches*. The complex procedure for separating a continuous beam into bunches occurs at the PS stage for ultimate injection into the LHC ring. When filled, each LHC beam contains 2808 bunches spaced by 25 ns, with each bunch measuring 1.06 ns [64].[10] Each bunch contains on the order of one hundred billion protons.

[8] The number of protons used by the LHC is relatively small, and the 10 L hydrogen bottle would well outlast the lifetime of the LHC. In reality, the bottle is changed roughly three times per year as other experiments at CERN consume more protons than the LHC [67].

[9] In 2019, the Linac2 was replaced by the Linac4, which accelerates protons to 160 MeV and allows for a more luminous beam [68].

[10] The beams are travelling so close to light speed that length is measured in units of seconds. To convert to metres, multiply by the speed of light.

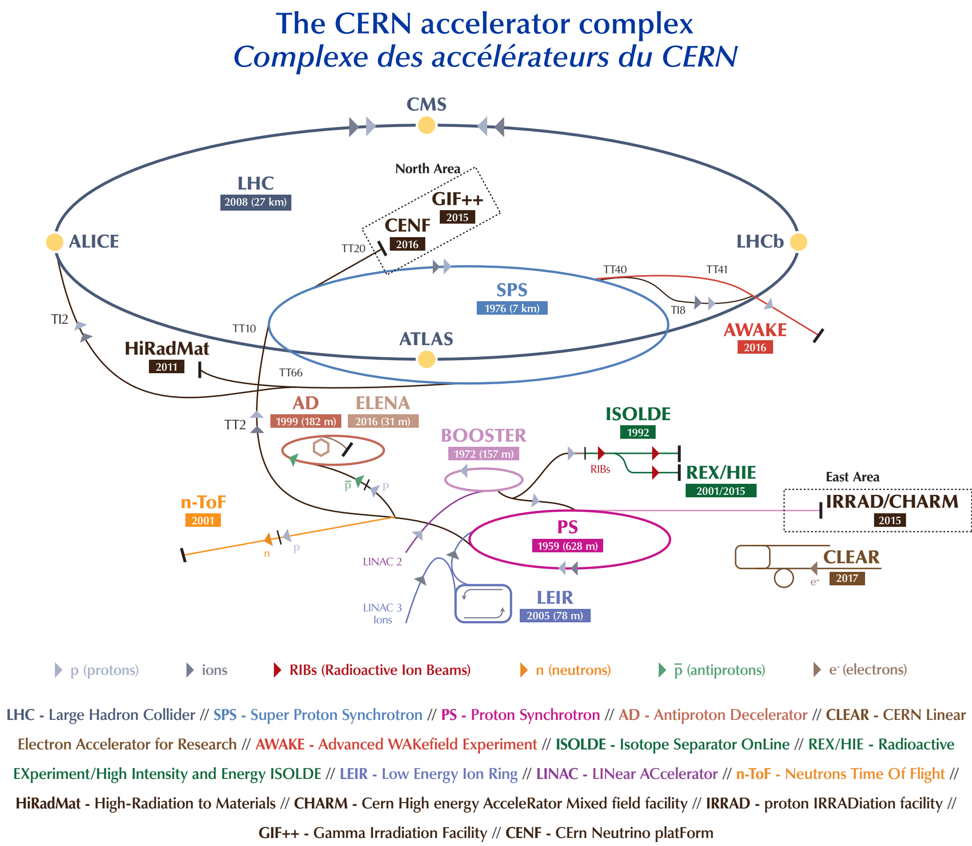


Figure 4.1: Schematic view of CERN's accelerator complex [70].

Energy

The possible physics processes ultimately depend on the COM energy of each pp collision $\sqrt{s}$. The LHC was designed so that each beam could accelerate protons to 7 TeV, reaching a COM energy of $\sqrt{s}$ = 14 TeV. In Run 1, from 2010–2013, the LHC provided 7 TeV, then 8 TeV collisions. Run 2 from 2015–2018 provided collisions of 13 TeV. In July of 2022, Run 3 began an achieved collisions up to 13.6 TeV [71]. The main limiting factor in achieving higher energies is the high magnetic field required of the superconducting dipole magnets used to steer the protons along their trajectory [72].

Luminosity

The collision rate of a collider is called the *luminosity* [73]. *Instantaneous luminosity* $\mathcal{L}$ is effectively the outgoing flux of particles per unit area and time after a collision of proton bunches. For a Gaussian beam distribution, the instantaneous luminosity is

$$\mathcal{L} = \frac{N^2 n_b f_{\text{rev}}}{4\pi\sigma_x\sigma_y} \cdot S \tag{4.1}$$

where N is the number of particles in each bunch, n_b is the number of bunches in one beam, f_{rev} is the revolution frequency of the beam, $\sigma_{x,y}$ represents the beam width in the transverse axes, and S is a geometrical factor accounting for the angle of the bunch crossing.

Integrated luminosity L measures the collisions over a period of time and is the traditional metric for the output of collider experiments. The number of occurrences of a particular process N_i is a product of the luminosity and the cross section of the process being studied σ_i:

$$N_i = \sigma_i \int \mathcal{L}\, dt = \sigma_i L. \tag{4.2}$$

Cross section is conventionally measured in units of femtobarns (fb) and conversely, integrated luminosity is measured in units of inverse femtobarns (fb^{-1}). One inverse femtobarn is approximately equivalent to 10^{12} pp collisions [74]. The luminosity of the beam in each fill is not constant but decays due to degradation of beam intensity and beam shape or *emittance*. Considering all these factors, the beam's luminosity reduces by a factor of $1/e$ in approximately 15 hours.

Since the processes of interest are often relatively rare, achieving a high luminosity is essential for recording a sufficient number of events. The LHC is designed for a peak instantaneous luminosity $\mathcal{L}$ =10^{34} $\text{cm}^{-2}\,\text{s}^{-1}$ at the high-luminosity collision points. In 2016 it achieved this rate and had doubled the peak luminosity by 2017 [75, 76]. In ATLAS, primary luminosity measurements are obtained using the LUCID-2 detector [77].

Pileup

Unfortunately, a high peak luminosity comes at a price. When a bunch crossing occurs, several distinct pp interactions take place. The resulting particles will overlap in the detector and complicate the identification of physics objects originating from the process of interest. The mean number of inelastic pp interactions is called the pileup parameter μ, and the time-averaged quantity is $\langle\mu\rangle$. This is also called in-time pileup. The higher the luminosity, the higher the pileup, and the more crowded the energy deposits in the detector [78].

Additionally, because the timing between bunch crossings is so short (25 ns), detector systems are sensitive to particles from collisions that happen after and before the triggered event. This is known as out-of-time pileup and primarily affects systems with long signal integration time, such as calorimeters (section 4.2.3).

Over the entirety of Run 2 the average pileup was $\langle\mu\rangle = 33.7$ and reached a maximum of $\langle\mu\rangle = 70$ [76]. Figure 4.2 shows the distributions of pileup in each data-taking year of Run 2.

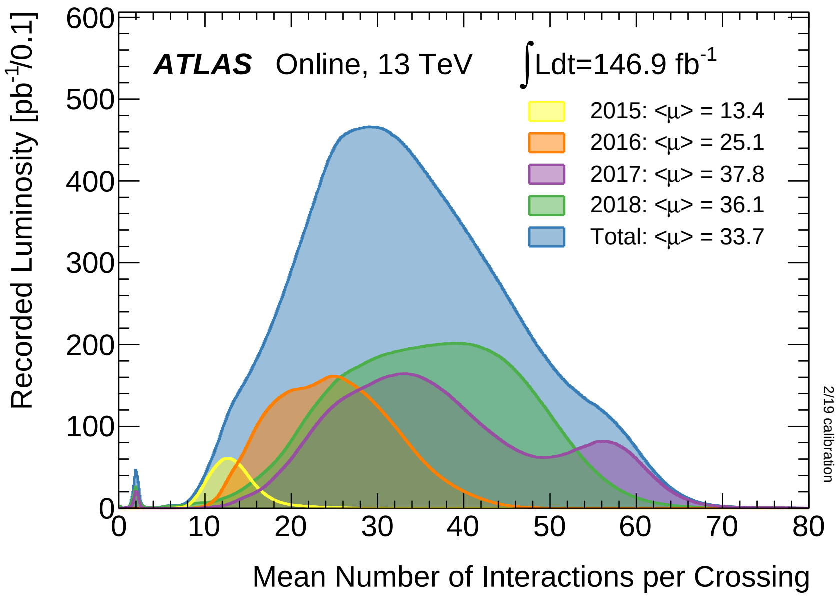


Figure 4.2: Shown is the luminosity-weighted distribution of the mean number of interactions per crossing for the 2018 pp collision data at $\sqrt{s} = 13$ TeV COM energy. All data recorded by ATLAS during stable beams is shown, and the integrated luminosity and the mean μ value are given in the figure [76].

4.1.3 Run 2

Because several of the processes of interest are relatively rare, one must record many collisions to make statistically significant measurements. The LHC and the associated detectors operate for several years to obtain sufficient data.

The LHC operates in multi-year *runs* in which data is taken more-or-less continuously. Scheduled short and long shutdowns allow for maintenance, repair, and upgrades to the accelerator and detectors. A regular winter shutdown and spring ramp-up period conveniently segments each data-taking period (roughly) into calendar years.

In the initial operational phase, called Run 1, collisions were produced in 2010–2011 at $\sqrt{s} = 7$ TeV and 2012 at $\sqrt{s} = 8$ TeV delivering an integrated luminosity of 28 fb^{-1}. This data was sufficient for discovering a particle consistent with the SM Higgs boson [79].

Run 2 took place from 2015–2018 with collisions of $\sqrt{s} = 13\,\text{TeV}$. It was possible in 2017 and 2018 to increase the luminosity to twice the nominal specification [80] with the LHC delivering, in total, $156\,\text{fb}^{-1}$ of 13 TeV data. The analysis presented in this dissertation uses data collected in Run 2.

Run 3 began in 2022 and plans to deliver an additional $290\,\text{fb}^{-1}$ of data from collisions at $\sqrt{s} = 13.6\,\text{TeV}$ through 2025. The high-luminosity LHC program will substantially increase the instantaneous luminosity of the collisions while maintaining the COM energy, obtaining an integrated luminosity $3000\,\text{fb}^{-1}$. Thus, the data recorded so far from the LHC is only 1/20 of its ultimate lifetime output.

4.2 The ATLAS detector

Recording exotic physics processes several million times per second requires a powerful camera. The ATLAS (A Toroidal LHC ApparatuS) detector, aside from a strained acronym, is one of the most complex and powerful "cameras" ever constructed. It is a general-purpose cylindrical detector situated about one of the four LHC collision points. Like its titanic namesake, ATLAS is an enormous machine measuring 46 metres in length and 25 metres in height, weighing 7000 tonnes and containing nearly 3000 km of cables for its 100 million readout channels [81, 82].

The detector body is broadly separated into a central barrel and two end-caps, which provide a near-complete solid-angle coverage. It is formed of several concentric subdetector systems which specialize in measuring different parameters of various particles produced in pp collisions. The inner detector measures charged particle momentum, the electromagnetic and hadronic calorimeters measure particle energy, and the muon spectrometer identifies muons that pass through the rest of the detector. Two powerful magnet systems encase the detector so that the trajectory of charged particles can reveal their momentum. This design makes ATLAS an extraordinarily versatile detector, enabling several hundred distinct physics analyses to operate in parallel.

Figure 4.3 shows a cutaway schematic of the ATLAS detector with a clear view of its systems. After defining the coordinate system and commonly-used terminology, this section will describe in further detail the major systems.

4.2.1 Coordinate system

The ATLAS detector uses a right-handed coordinate system with the origin defined at the nominal interaction point (IP) in the centre of the detector. The x-axis points towards the centre of the LHC ring, the y-axis points away from the centre of the earth, and the z-axis is aligned with the LHC beam pipe. Cylindrical coordinates (r, ϕ, z) are used in definitions of detector elements and material distributions. The x-y plane, orthogonal to the beam axis, is referred to as the *transverse plane*.

Spherical coordinates (r, ϕ, θ) are useful in defining particle trajectories where the azimuthal angle ϕ is measured about the beam axis and the polar angle θ is measured from the beam axis. A common parameterization of the polar angle is *pseudorapidity*, defined as $\eta = -\ln\tan\theta/2$. Pseudorapidity is often preferred over θ as differences in pseudorapidity are Lorentz invariant under relativistic boosts (in the massless limit). Coordinates generally used to describe a particle's four-momentum are $(p_{\text{T}}, \eta, \phi)$ where

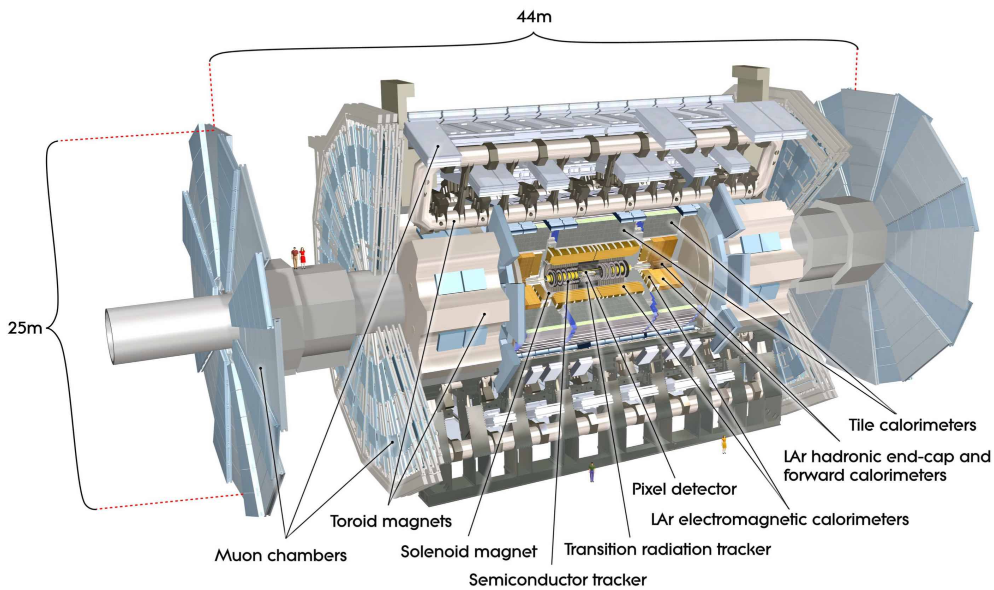


Figure 4.3: Cut-away view of the ATLAS detector [81].

p_{T} refers to the transverse momentum of a particle — the component of its momentum in the transverse plane. The angular separation $\Delta R = \sqrt{(\delta\eta)^2 + (\delta\phi)^2}$ quantifies the distance between particles. This quantity, defined in terms of pseudorapidity, is also Lorentz invariant.

4.2.2 Inner detector

The inner detector (ID) is the first system encountered by the debris of pp collisions. It is composed of three separate subsystems, which together are used to determine the direction and momentum of charged particles as well as identify the vertices from which they were produced. Together the barrel and end-cap sections of these detectors provide coverage up to $\eta < 2.5$. A thin solenoid magnet surrounds the full ID and produces a strong magnetic field along the z-axis, so charged particles are deflected along the ϕ-axis. The ID is a *non-destructive* detector — particles will pass through the material with minimal interference. How the energy deposits in the ID are interpreted as physical particles is detailed in section section 5.1.

Pixel detector

The pixel detector is the innermost part of the ID [83]. 1744 identical sensors contain over 80 million pixels in total, each with a separate readout channel. Each pixel is made from oxygenated n-type wafers, specifically designed to operate continuously in a high-radiation environment [81]. The nominal pixel size is $50\,\mu\mathrm{m}$ $(r\text{-}\phi) \times 400\,\mu\mathrm{m}$ (z) which provides a spatial resolution of $10\,\mu\mathrm{m}$ $(r\text{-}\phi) \times 115\,\mu\mathrm{m}$ (z). The three barrel layers are located at $r = 50.5$, 88.5, and 122.5 mm and the end-cap disks at $|z| = 495$, 580, and 650 mm. A detailed schematic is shown in fig. 4.5. This configuration provides individual three-dimensional measurements of a charged particle's trajectory with high precision.

Between Run 1 and Run 2, an additional layer of high-resolution pixels was inserted near the beam pipe

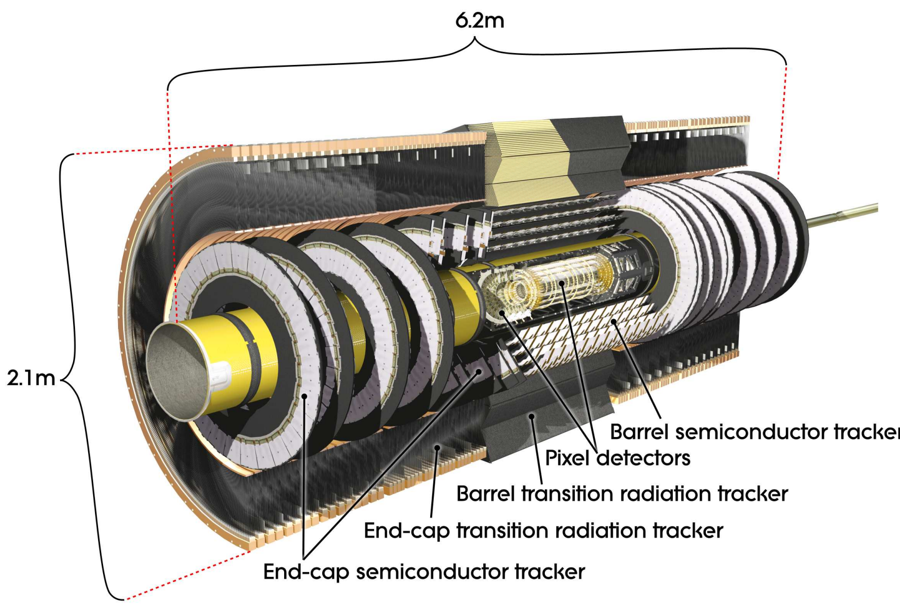


Figure 4.4: Cut-away of the ATLAS inner detector [81].

at 33.3 mm. This insertable B-layer (IBL) was installed to provide substantially improved vertexing and identification of displaced b-hadrons. The IBL consists of an additional 12 million pixels of dimensions 50 µm (r-ϕ) × 200 µm (z), corresponding to a spatial resolution of 8 µm (r-ϕ) × 40 µm (z). Of particular relevance to this dissertation, the insertion of the IBL improves the capacity to reconstruct secondary vertices, which are critical in the search for many exotic physics signatures with long-lived particles.

Semiconductor tracker

The semiconductor tracker (SCT) is the next system encountered by outwardly radiating particles. For reasons of cost and reliability, single-sided p-in-n sensors are used, different from that of the pixel sensors. Each layer of the SCT module holds two sensors on each side, each of dimensions 80 µm (r-ϕ) × 12 cm (z). The rectangular geometry of the SCT sensors leads to the often-used alias *strips*. The strips are offset by a stereo angle of 4 mrad which enables a more precise resolution in the z-axis: 17 µm (r-ϕ) × 580 µm (z). The barrel contains four layers at r = 299, 371, 443, and 514 mm, and the end-caps contain 9 disks each between $|z|$ = 854 and 2720 mm. The 4088 silicon strip modules are read out by approximately 6.3 million channels.

Transition radiation tracker

The final element of the ID is the transition radiation tracker (TRT). This section consists of gaseous drift tube detectors as opposed to the silicon-based detectors in the other two sections. It is composed of approximately 300,000 straw tubes, each 4 mm in diameter, with a gold-plated tungsten wire at the centre. A potential difference is kept between the wire and the tube wall. Each tube is filled with a Xe-based gas mixture[11] which, when ionized by charged particles, causes electrons to drift to the wire where they are amplified and read as signal. The tubes are embedded in a matrix of polypropylene fibres which causes incident particles

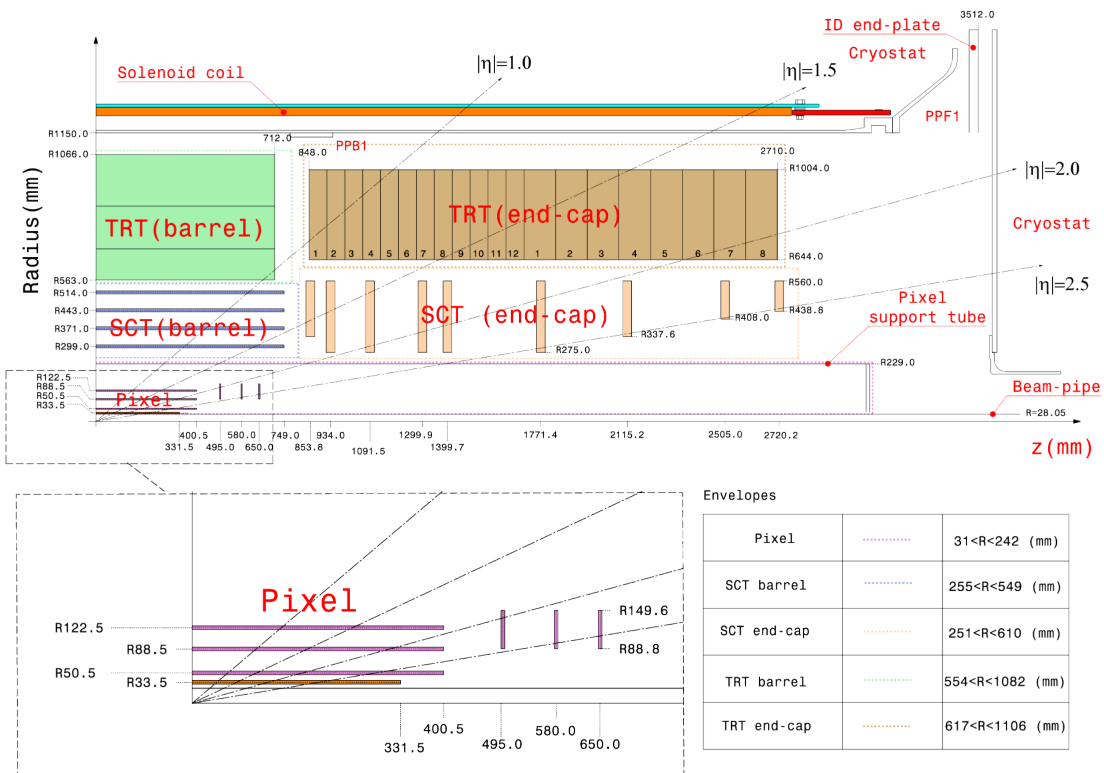


Figure 4.5: The r-z cross section view of the layout of a quadrant of the ATLAS inner detector for Run 2. The top panel shows the whole inner detector, whereas the bottom-left panel shows a magnified view of the pixel detector region. Compared to Run 1, the IBL (shown in red in the bottom-left panel) and its services, together with the new beam pipe, were added [83].

to produce transition radiation, further amplifying the signal.

The straw tubes are 144 cm long in the barrel region, aligned with the z-axis, and 39 cm in the end-cap region, radially arranged perpendicular to the z-axis. Because of their length, hits in the straws only provide 2-dimensional trajectory information. However, with 130 μm resolution and dense packing, the TRT complements the silicon detectors with continuous tracking for particles within $|\eta| < 2.0$. Furthermore, transition radiation has an additional benefit: the magnitude of the radiation is related to the incident particle's mass. Therefore, the TRT can be used for particle identification, mainly to distinguish between electron and charged hadron (e.g., pion) tracks.

4.2.3 Calorimeters

Beyond the ID and solenoid magnet lies the calorimetry system. The ATLAS calorimeters destructively measure the energy of the particles produced in the pp interaction. They effectively absorb nearly all the energy from SM particles, except for muons and neutrinos, stopping them in dense material. The systems

[11]In Run 1, the tubes were only filled with the Xe-based mixture, but due to large irreparable leaks, many channels were converted to the much cheaper Ar-based mixture. The TRT's tracking performance is largely unaffected by this change; however, particle identification efficiency is impacted [84].

are separated by the type of particles they are designed to absorb: the electromagnetic calorimeter (ECAL) effectively absorbs energy from electrons and photons, while the hadronic calorimeter (HCAL) is designed to absorb charged and neutral hadrons.

The calorimetry system has a coverage of $|\eta| < 4.9$. This near-hermetic coverage, as well as the thickness of the calorimeters, is important to fully capture the total energy of the jets. The ECAL and HCAL systems both use a sampling calorimeter design. This type of calorimeter alternates dense *absorbing* layers with *active* layers which measure the deposited energy. The layout of the calorimeter subsystems can be seen in fig. 4.6.

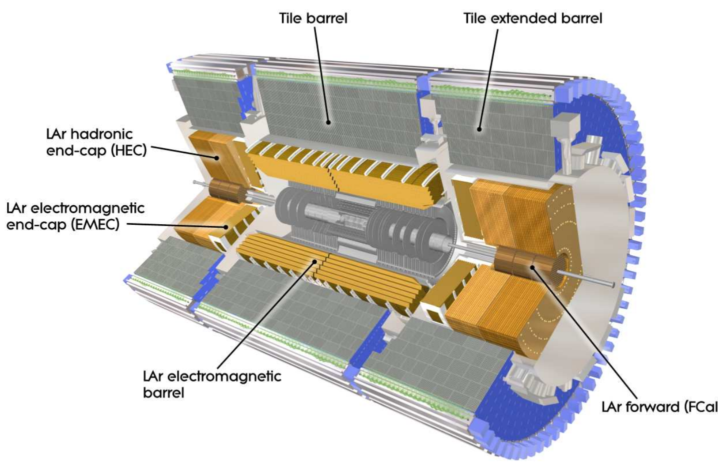


Figure 4.6: Cut-away view of the ATLAS calorimeter system [81].

Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) [85] measures the energy of electrons and photons while also preventing the resulting particle showers from reaching the subsequent detector layers. As a sampling calorimeter, the ECAL is composed of alternating absorbing and active layers in an accordion arrangement to provide total coverage in ϕ. The absorber is made of lead plates which produce electromagnetic showers when struck by incident electrons and photons. The active material is cryogenically cooled liquid argon (LAr), which produces electron-ion pairs when ionized by electromagnetic showers. These pairs drift to charged electrodes, where they are digitized and read out.

Composed of a barrel and two end-caps, The ECAL covers the region of $|\eta| < 3.2$. Resolution varies across the ECAL, but is at its finest granularity in the $|\eta| < 2.5$ region: 0.025×0.025 $(\eta \times \phi)$. In this region, the tracking detector provides complimentary charged particle information. The thickness is conventionally measured in radiation lengths X_0 where one radiation length corresponds to the distance in which the energy of an electron will be reduced to a factor of $1/e$ due to bremsstrahlung interactions [86]. The ECAL has a thickness of more than 22 X_0 in the barrel and more than 24 X_0 in the end-caps. Approximately 173,000 channels read data from this system.

Hadronic calorimeter

The hadronic calorimeter (HCAL) is located directly outside the ECAL envelope and measures the energy of charged and neutral hadrons [87]. Hadronic showers are generally more complex than electromagnetic showers since the hadrons are subject to strong force interactions. Hadronic fragmentation and escaping stable particles leave some of the incident hadronic energy invisible to the detector. Therefore, the total energy of the hadronic interaction must be estimated from the recorded energy.

One barrel and two extended-barrel region sections cover $|\eta| < 1.7$ and are composed of sampling steel-plastic tile calorimeters. Charged and neutral hadrons form cascading particle showers when interacting with the steel tiles. These particles then interact with the scintillating polystyrene active layers. The generated light is transmitted via fibre optic cables to photomultiplier tubes, and then the signal is digitized and read out. Compared to EM showers, hadronic showers have a greater lateral spread; therefore, a lower resolution is necessary for the HCAL systems. The finest spatial resolution in the tile calorimeter is 0.1×0.1 $(\eta \times \phi)$. The end-cap HCAL is a LAr calorimeter very similar to the ECAL system. It covers the range from $|\eta| > 1.5$ to $|\eta| < 3.2$ and also has a minimum spatial resolution of 0.1×0.1 $(\eta \times \phi)$. A forward calorimeter (FCal), set within the end-caps close to the beampipe, provides further coverage for $3.1 < |\eta| < 4.9$. Two separate sections of the FCal are optimized for electromagnetic and hadronic calorimetry.

4.2.4 Muon spectrometer

The outermost system of ATLAS, the Muon Spectrometer (MS), detects charged particles that escape the calorimeter system [88, 89]. The momentum calculation is based on the magnetic deflection of muon tracks by superconducting toroid magnets. These magnets are arranged to provide a field roughly orthogonal to the muon trajectory to maximize deflection. The system is subdivided into separate trigger and high-precision tracking chambers.

High-precision tracking is primarily achieved by Monitored Drift Tube chambers (MDT) [90], which are chosen for their simplicity and high accuracy. Each chamber, composed of three to eight layers of drift tubes, achieves an average resolution of 35 µm per chamber and covers a pseudorapidity range $|\eta| < 2.7$. To complement the MDTs, Cathode Strip Chambers (CSC) [91] are used in the innermost tracking layer for the region $2.0 < |\eta| < 2.7$. These multiwire proportional chambers are used in this region for their higher rate and time resolution capabilities. The resolution of each chamber is 40 µm x 5 mm $(r \times \phi)$.

The muon fast trigger system provides track information within tens of nanoseconds of particle crossing. Resistive Plate Chambers (RPC) [92] in the barrel region and Thin Gap Chambers (TGC) [93] in the end-cap region serve simultaneous purposes. They are capable of fast momentum determination necessary for L1 trigger selection (see section 4.2.6) and with a 15–25 ns time resolution, they are able to identify individual bunch crossings. The RPC and TPC coordinates in the non-bending plane complement the high-precision coordinates from the MDTs in the bending plane.

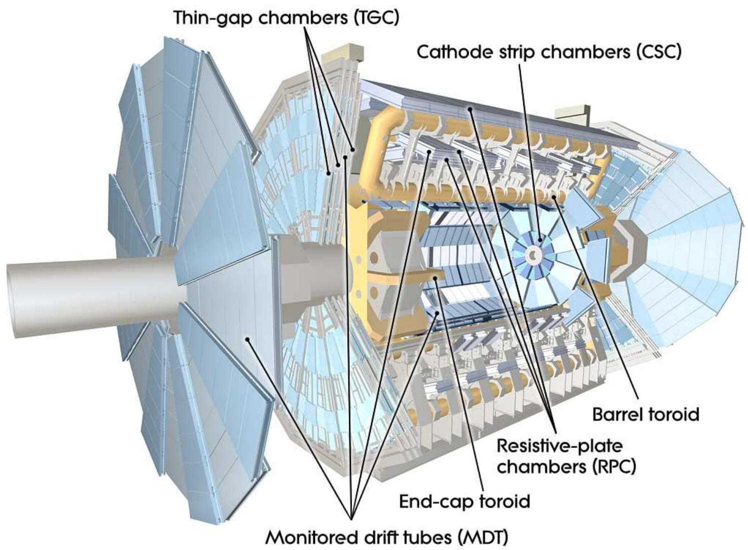


Figure 4.7: Cut-away view of the ATLAS muon system [81].

4.2.5 Magnet system

The magnet system is a defining feature of the ATLAS detector due to its strength and unique geometry. It is composed of a hybrid system of four large superconducting magnets — three toroids and one solenoid. The entire system is 22 m in diameter by 26 m in length, and stores an energy of 1.6 GJ.

The central solenoid magnet provides a 2 T axial magnetic field through the inner detector. Its layout is carefully optimized to reduce material thickness upstream of the calorimeter cells, achieving approximately 0.66 radiation lengths at normal incidence. The diameters of the solenoid are 2.46 m (inner) and 2.56 m (outer) and is 5.8 m long. In total, the coil mass is 5.4 tonnes and stores 40 MJ of energy at the nominal 7.7 kA operational current.

Beyond the calorimeter envelope, the toroid magnets produce fields for momentum measurements in the MS system. The barrel toroid consists of eight coils encased in vacuum vessels with diameters of 9.4 m (inner) and 20.1 m (outer) and length of 25.3 m. It produces a field of 1.5 T and stores 1.1 GJ of energy at its nominal operating current of 20.5 kA. Each end-cap toroid is a single cold mass unit containing eight square coil units. They produce an average field of 1 T, storing 0.25 GJ of energy each.

4.2.6 Trigger

pp interactions of interest that may reveal new physics are rare, and the vast majority of them may be safely discarded without impacting most ATLAS analyses. Furthermore, bandwidth and storage capacity limitations make it impractical to record the full readout of the detector systems at the 40 MHz rate of particle collisions. The Trigger and Data Acquisition system (TDAQ) is a combined hardware and software system that determines the events of interest to physics analyses in real-time and stores the relevant detector response data [94].

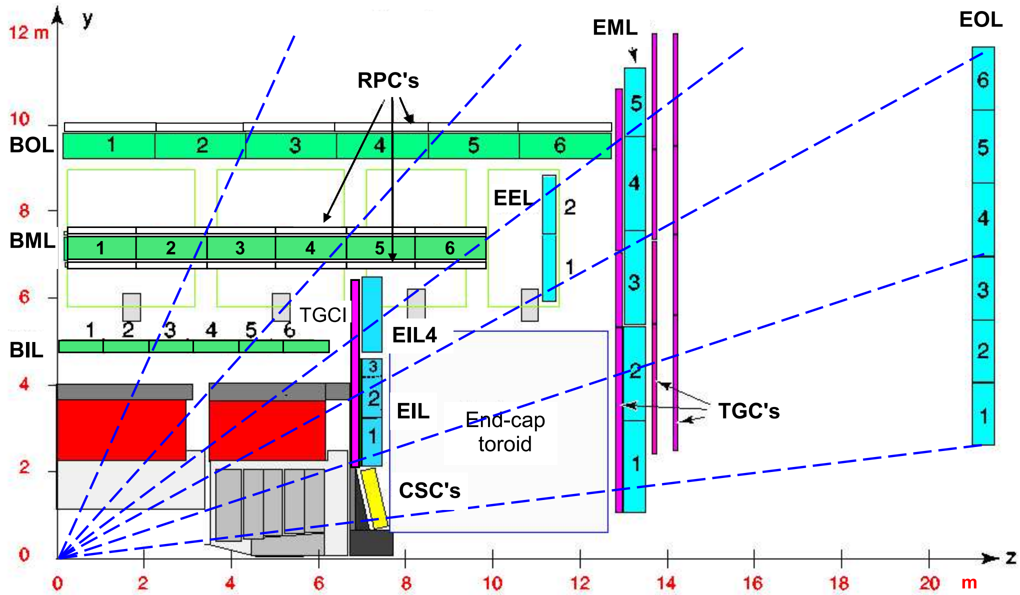


Figure 4.8: Cross section of the muon system in a plane containing the beam axis (bending plane). Infinite-momentum muons would propagate along straight trajectories which are illustrated by the dashed lines and typically traverse three muon stations [81].

The Level-1 (L1)[12] trigger uses a hardware-based system to rapidly process reduced-resolution signals from the calorimeters (L1Calo) and muon detectors (L1Muon). Combining information from several dedicated custom electronics, the Central Trigger Processor (CTP) makes a pass/no-pass decision for the event within 2.5 µs of the bunch crossing. The L1 trigger seeks out high-momentum objects (leptons, photons, jets), high $E_{\mathrm{T}}^{\mathrm{miss}}$ (see chapter 5), and events with overall high energy deposits. The rate of event readout is reduced by a factor of 400 to an upper limit of 100 kHz. In addition, it specifies Regions-of-Interest (ROIs) in η and ϕ, which are to be examined in more detail by the next trigger stage.

The High-Level Trigger (HLT), the second stage of the trigger system, is a software-based system that performs a more precise object reconstruction similar to the offline reconstruction discussed in chapter 5. The HLT requests detector information at full resolution from the data readout buffers within the ROIs defined in L1. This more computationally-expensive selection reduces the event rate by another factor of 100 to a rate of 1 kHz.

The precise decision criteria are determined by a *trigger menu* which, for Run 2, contained roughly 1500 individual event selections. These menus specify the physics signatures, selection algorithms, event rate, and bandwidth. Each trigger menu is determined by the physics goals for a data collection period, accounting for the instantaneous luminosity, data transfer, and storage constraints. For events which pass both the L1 Trigger and HLT systems, the complete detector information is read out (approximately 1.5 MB/event) and sent to the local Tier 0 computing centre for permanent storage and full event reconstruction.

[12]In Run 1, the HLT consisted of a L2 trigger and Event Filter stage, but these were replaced by a single HLT system in Run 2 to reduce complexity and optimize performance.

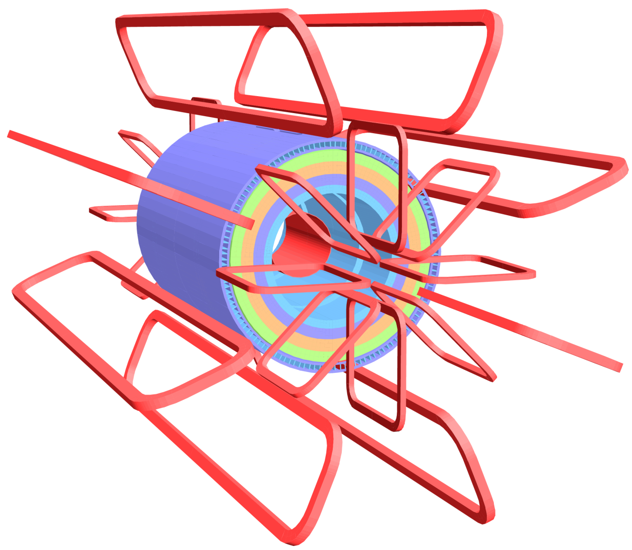


Figure 4.9: Geometry of magnet windings [81].

Figure 4.10: Bare central solenoid in the factory after completion of the coil winding [81].

Chapter 5

Object reconstruction

The entire detector response information is read to permanent storage for collision events that successfully pass a trigger. However, at this stage, an event is simply a raw collection of voltages. A tremendous computation is undertaken in which the detector response is combined with geometrical information and physical domain knowledge to produce a representation of reality. The final output of the several reconstruction algorithms is a collection of *physics objects*, which are digital descriptions of the physical entities produced in the collision. These objects are stored in so-called `xAOD` format, which is a custom event data model (EDM) design used in ATLAS [95]. Physics objects in the EDM are represented as C++ classes in the `xAOD` namespace such as `xAOD::Muon` or `xAOD::Electron`. The framework used in the reconstruction of these objects is the Athena software package [96, 97]. Additional metadata is stored which records the detector conditions for each data-taking segment. Data is recorded in continuous chunks called *luminosity blocks*.

Figure 5.1 shows a schematic representation of how fundamental particles interact with the ATLAS subdetectors. In the ID, charged particles produce tracks which are then combined to produce vertices, each of which is described in section 5.1. The ECAL records the electromagnetic jets produced by electrons and photons, the reconstruction of which is detailed in section 5.2. Charged and neutral hadrons produce collimated showers of particles in the HCAL. These showers are called jets and are described in section 5.4. Muons, as explained in section 5.3, can pass through the entire detector, leaving signals in the ID and the MS. Finally, the ATLAS detector does not see neutrinos directly, but their existence is inferred by measuring the transverse energy that is missing from an event. This is called $E_{\mathrm{T}}^{\mathrm{miss}}$, and its reconstruction is detailed in reference [98].

Simulated physics data is heavily used in ATLAS, and the same reconstruction algorithms are applied to real and simulated data. After applying calibrations, the remaining efficiency differences in simulation with respect to data are corrected by applying *scale factors*.

5.1 Tracks and Vertices

The trajectories of charged particles, or tracks, are calculated using energy deposit information from the ID. Figure 5.2 depicts a visualization of the detector elements used in the tracking of charged particles. The primary vertex (PV) corresponding to the hard-scatter interaction is determined from the tracks. The standard reconstruction of tracks focuses on *prompt* tracks, defined as those originating from nearby the PV. As detailed in chapter 6, an additional pass to reconstruct tracks and vertices from secondary particle decays may also be performed. Tracking is computationally intensive due to a high density of charged particles in each pp bunch crossing. A sequence of complex and carefully optimized algorithms is required

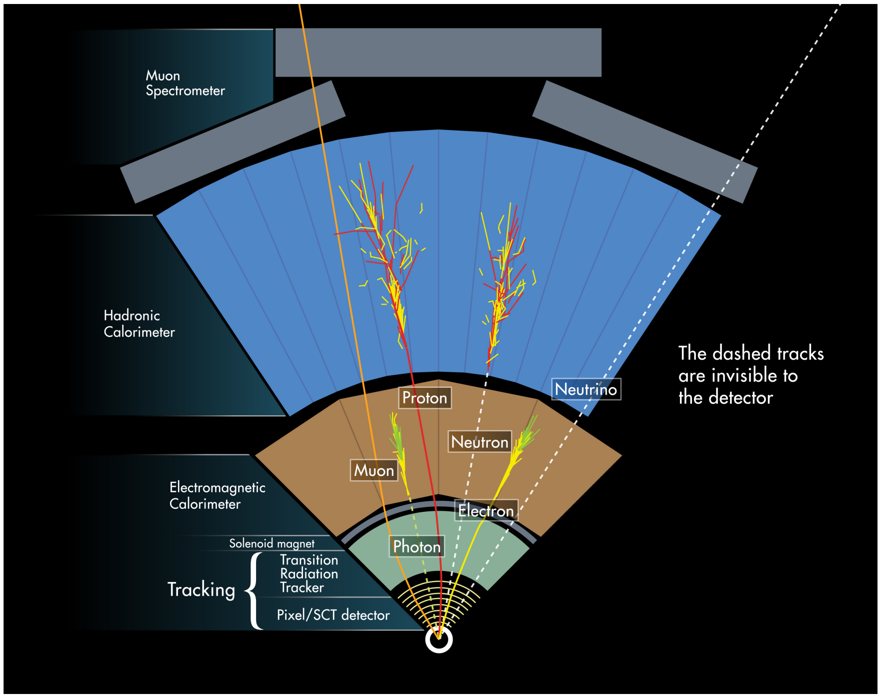


Figure 5.1: A schematic slice of the ATLAS detector in the barrel region. Different detector systems are sensitive to different particle types [99].

to reconstruct the tracks within a reasonable computational budget.

5.1.1 Reconstruction

Track reconstruction is performed by "connecting the dots" of detector deposits in an iterative process. Each step reduces the number of potential candidates, so more intensive fits are only performed on likely tracks. After the primary *inside-out* tracking pass is completed, an additional *outside-in* pass is performed using the remaining deposits.

Five parameters and a reference point describe each track in a perigee representation. These parameters are depicted in fig. 5.3. The transverse impact parameter d_0 is defined as the track's point of closest approach to the reference point in the transverse plane, while the longitudinal impact parameter z_0 is the point of closest approach in the longitudinal plane. The reference point used is the average position of the pp interactions — the *beam-spot* position. The remaining three parameters are the azimuthal angle ϕ and the polar angle θ of the track momentum, and the ratio q/p is the charge of the reconstructed track divided by the magnitude of its momentum.

Clusterization

The first step in charged particle reconstruction is to combine ID energy deposit information into spatially distinct clusters. In the pixel and SCT detectors, deposits from adjacent sensors are combined to form three-dimensional space-points, or *hits*, which approximate the point where the charged particle traversed the ID

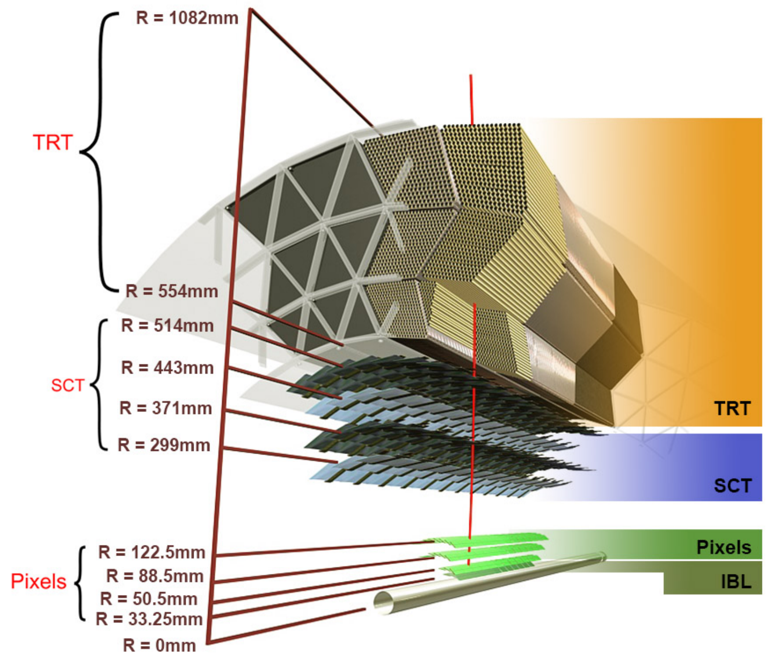


Figure 5.2: A 3D visualization of the structure of the barrel of the ID. The beam pipe, the IBL, the pixel layers, the four cylindrical layers of the SCT and the three layers of TRT barrel modules consisting of 72 straw layers are shown [100].

material. Three-dimensional hits are formed from individual layers in the pixel detector or by combining both layers, axial and stereo, of an SCT module. In dense environments, clusters can be formed by contributions from more than one charged particle and each cluster can be used to form multiple tracks.

Inside-out tracking

The primary pass used in track reconstruction is the inside-out sequence. Initially, track *seeds* are formed from three hits in separate layers. Seeds can be formed from three SCT hits, three pixel hits, or three hits from a combination of the two detectors, and the degree of confidence in a seed's purity follows that order. This approach allows for a large set of space-point combinations while providing a rough track trajectory and momentum estimate. Seeds must pass a loose selection which requires that the charged particle originates near the IP. In particular, the impact parameters of the track seed, d_0 and z_0, are required to be 10 mm and 250 mm, respectively, and the seed momentum is required to be above 500 MeV. If these criteria are not met, the potential track is discarded. A seed can be *confirmed* by replacing its inner- or outermost hit with one from the next available silicon layer, then checking if the new trajectory is compatible with the old. A confirmed seed's reconstruction is prioritized over those which cannot be confirmed.

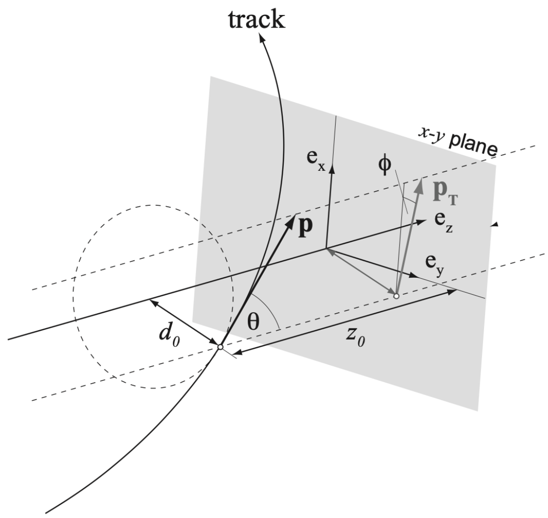


Figure 5.3: Five parameters — d_0, z_0, ϕ, θ, q/p — define a track in a perigee representation [101].

Candidate formation

Once a track seed is established, there is sufficient trajectory information to extend it to a *track candidate* by selectively adding hits in additional detector layers. This is accomplished using a *combinatorial Kalman filter* [102, 103]. This algorithm iteratively predicts the location of the subsequent hit along the track trajectory and then, if found, updates the trajectory calculation and track parameters. The filter creates multiple track candidates if more than one space-point is compatible. As the track parameters are more accurate at this stage, additional selections are applied to the candidates. Candidates are rejected if they fail to meet the criteria in table 5.1.

Ambiguity resolution

The collection of fully-formed candidates contains all potentially valid combinations of hits, meaning many tracks overlap, share hits, or contain incorrect hits. The *ambiguity solving* stage attempts to resolve these duplications and inconsistencies in the candidate collection. Figure 5.4 illustrates a set of tracks in the SCT with ambiguous hits.

The ambiguity resolution begins by assigning each candidate a score — a higher score is more likely to correctly represent the trajectory of a charged primary particle. In order of decreasing score, candidates are fitted and accepted into the output container. Critically, the hits that compose the accepted track are recorded, and restrictions are imposed on whether they can be reused in subsequent candidates. Each track can have a maximum of two shared hits. Each hit can be used by a maximum of two tracks. Preference is given to tracks already accepted by the ambiguity solver, which is the purpose behind the initial score-ordering. If a cluster causes either the candidate or an accepted track to fail the criteria, it is removed from the candidate in question, which is then re-scored and returned to the ordered list of remaining candidates. If this removal causes the candidate to fail any of the criteria in table 5.1 it is rejected.

parameter	criteria
min. p_T	400 MeV
max. $\lvert\eta\rvert$	2.5
min. si. hits	7 (12 are expected)
max. shared hits	1 pix or 2 SCT
max. si. holes	2
max. pix holes	1
max. $\lvert d_0^{BL}\rvert$	2 mm
max. $\lvert z_0^{BL}\sin(\theta)\rvert$	3 mm

Table 5.1: This set of minimum requirements is imposed on each track candidate used in the inside-out tracking pass. A *hole* is defined as an intersection of the track trajectory that contains a sensitive detector element but does not contain a cluster. Quantities with *si.* refer to silicon — the combined pixel and SCT detectors. d_0^{BL} is the transverse impact parameter calculated with respect to the measured beam-line position. z_0^{BL} is the longitudinal impact parameter measured from the primary vertex, and θ is the polar angle of the track [104].

In dense environments, it is possible that multiple charged particles will traverse the same sensors. In this case, it is appropriate for two or more tracks to share hits without being penalized by the ambiguity solver. Identifying these *merged* clusters is vital to achieving good reconstruction efficiency in environments with high track multiplicity. A neural network (NN) is used to identify merged clusters. It uses as inputs the measured charge (proportional to the deposited energy), the relative position of pixels in the cluster, and the particle's incident angle, which is provided by the track candidate. The NN identification efficiency is 90% for clusters formed from two charged particles and 85% for three or more charged particles. The NN does not distinguish between three or more than three contributing particles. The rate at which the NN misidentifies single-particle clusters as having been formed from more than one particle is a few percent [104].

The score used in candidate ordering is calculated with a reward/penalty strategy based on simple measures of track quality. Each cluster in the candidate increases the score while each hole decreases it. Measurements from the more precise pixel detector are given a stronger scoring weight than measurements from the SCT. The fit quality is considered by applying the χ^2 per number of degrees of freedom to the score. Energetic tracks gain an increase in score by applying the logarithm of the p_T of the candidate. This assignment also suppresses the tracks with incorrectly assigned clusters which generally have a low p_T.

Track fitting

After applying the selections listed in table 5.1, a high-resolution fit is performed on each track by minimizing a global χ^2 value. The position and uncertainty of each cluster used in the fit are estimated using a NN. A set of linear equations is solved via matrix inversion using these positions, uncertainties, and additional parameters to account for material effects. This calculation is among the most computationally intensive in the track reconstruction process and is therefore delayed until necessary for ambiguity resolution. Fitted tracks which pass all selections after ambiguity resolution are added to the output track collection.

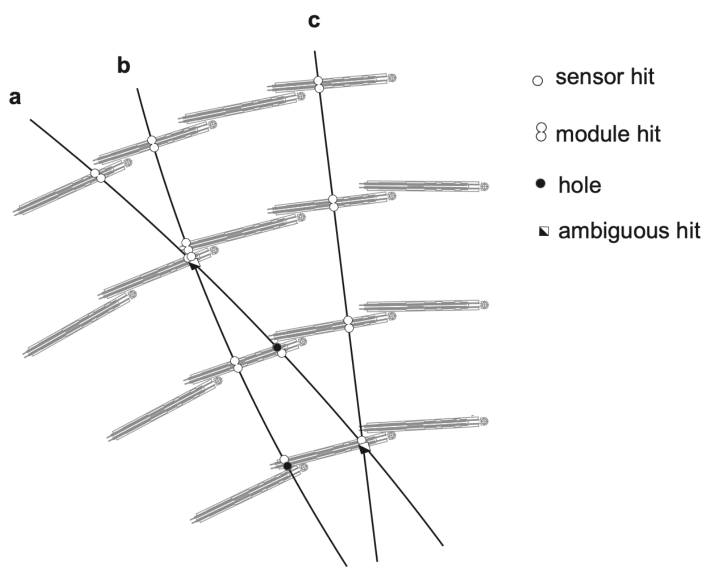


Figure 5.4: A simplified model of the track ambiguity problem in the SCT barrel. Track candidates **a**, **b**, and **c**, share hits. A score-based ambiguity resolution procedure is performed in order to distinguish between true and fake tracks. Hits improve the candidate's score while holes penalize it. The candidate with the higher score takes the ambiguous hits [105].

TRT extension

After ambiguity resolution and fitting, the algorithm attempts to extend the track candidates to the TRT by incorporating compatible measurements in the TRT straws without modification to the silicon-only tracks. The collection of extension candidates is refit and scored using the same framework as the ambiguity solver. Tracks in the output collection are only extended if their score improves with the addition of TRT information.

Outside-in tracking

An additional reverse tracking sequence is performed to recover tracks missed by the inside-out sequence. Outside-in tracking begins by forming standalone TRT segments in regions of interest determined by the ECAL. Any hits not already used by the inside-out pass are incorporated into the TRT segment if compatible with the candidate trajectory, then fitted. Looser requirements are imposed on these track candidates than on the inside-out tracks, particularly reducing the minimum required silicon hits.

This sequence is designed to account for tracks originating from secondary decay vertices inside the ID, including K^0_S mesons and other undiscovered signatures. Photon conversions with insufficient silicon detector activity may also fail the inside-out pass but can be recovered with the outside-in sequence. The resulting tracks from the inside-out and outside-in sequences form the final collection of reconstructed standard tracks.

5.1.2 Performance

To measure track reconstruction performance, the following track classifications are used. *Truth-matching probability* (p_{match}) is a contrived quantity used in simulation to define the likelihood that a reconstructed track is formed from an individual charged particle. The clusters used to reconstruct a single track object may originate from one or more charged particles. p_{match} is the ratio of the weighted number of measurements originating, at least in part, from the same simulated particle to the weighted number of all measurements used in a track [104]. This quantity is not available in data as there is no record of the origin of the detector clusters. The weights reflect the average number of measurements in each subdetector. Pixel hits have a weight of ten, SCT hits a weight of five, and TRT hits a weight of one. Typically for a track to be considered properly reconstructed, it must have a p_{match} of at least 0.5. This definition is used in the present dissertation.

Conversely *fake* tracks are those with a p_{match} below 0.5. These are formed from combinations of hits which do not originate primarily from a single charged particle. The high detector occupancy that leads to large numbers of fakes is often due to high pileup.

An additional source of fake tracks is due to the mechanics of the ATLAS detector simulation. After the event generation and showering procedures, the particle trajectory and detector response are simulated using Geant4 [106–108]. As charged particles interact with detector material, they radiate energy in the form of bremsstrahlung and subsequent pair production, forming *secondary* tracks. To reduce computational costs, the Geant4 simulation ceases to keep track of the parentage of secondary particles once they fall below a certain momentum threshold. It is a matter of definition and preference whether to count these secondaries as fake tracks. Low-momentum objects are generally poorly reconstructed, and subsequently, these secondary tracks rarely contribute to physics of interest and should be suppressed by the reconstruction algorithm. For this reason, combinatorial fake tracks and low-momentum secondary particles are classified without distinction in the remainder of this dissertation.

Additional selection criteria are imposed to define standard *Tight* and *Loose* working points, for which the specific criteria can be found in reference [109]. *Tracking efficiency* ε_{trk} is defined as

$$\varepsilon_{\text{trk}} = \frac{N_{\text{rec}}^{\text{matched}}}{N_{\text{gen}}}. \tag{5.1}$$

In this ratio, $N_{\text{rec}}^{\text{matched}}$ is the number of reconstructed tracks matched to true charged particles and N_{gen} is the number of true charged particles generated in simulation. Tracking efficiency is a quantity estimated using only simulated events since the true number of particles is required for the measurement. The efficiency as a function of p_{T} and η are seen in fig. 5.5.

5.1.3 Primary vertex reconstruction

Closely connected to the reconstruction of tracks is the identification and reconstruction of the *primary vertex* (PV), which is defined as the point at which a pp interaction occurred. The tracks used in PV reconstruction are selected with tight requirements to reduce the rate of fake tracks or tracks from secondary interactions. The full set of constraints can be found in reference [110]. Vertices are then reconstructed in the following

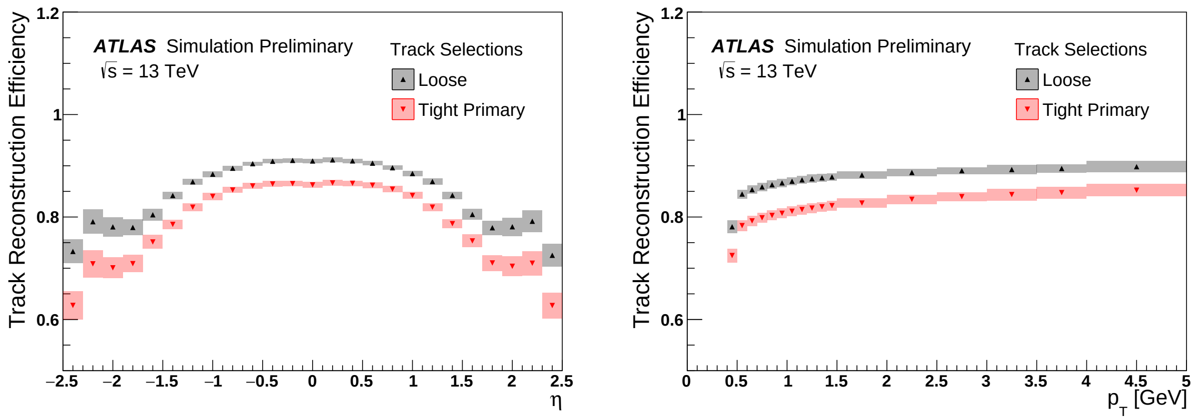


Figure 5.5: Run 2 tracking efficiency as a function of η and p_T is shown for *Tight Primary* and *Loose* working points [109].

steps:

- A seed position is selected using the centre of the beam spot and the mode of the z-coordinates of the tracks.
- An iterative annealing fit is performed, which repeatedly down-weights less compatible tracks and refits the vertex position.
- Tracks incompatible with the determined vertex are removed and allowed to be used to find additional vertices.
- The procedure is repeated using the tracks not already belonging to a vertex.

The output of this procedure is a set of three-dimensional vertex positions and their covariance matrices.

Vertex reconstruction efficiency is determined from data by taking the ratio between events with a reconstructed vertex and events with at least two reconstructed tracks [111]. The efficiency reaches above 99% for events with five or more tracks. In high-pileup conditions, vertex reconstruction efficiency is reduced due to the merging of vertices from separate interactions.

5.2 Electrons and photons

5.2.1 Reconstruction

Electron reconstruction combines information from energy deposits in the ECAL and charged particle tracks in the ID by matching the $\eta \times \phi$ coordinates of each [112, 113]. Reconstruction of electrons is restricted to $|\eta| < 2.47$ to benefit from the high-granularity region of the ECAL. A schematic of the path of an electron through the detector elements is shown in fig. 5.6.

First, energy deposits in the ECAL and HCAL are combined into EM *topo-clusters* using the '4-2-0' threshold (see reference [114] for a detailed description). Electrons lose significant energy to bremsstrahlung

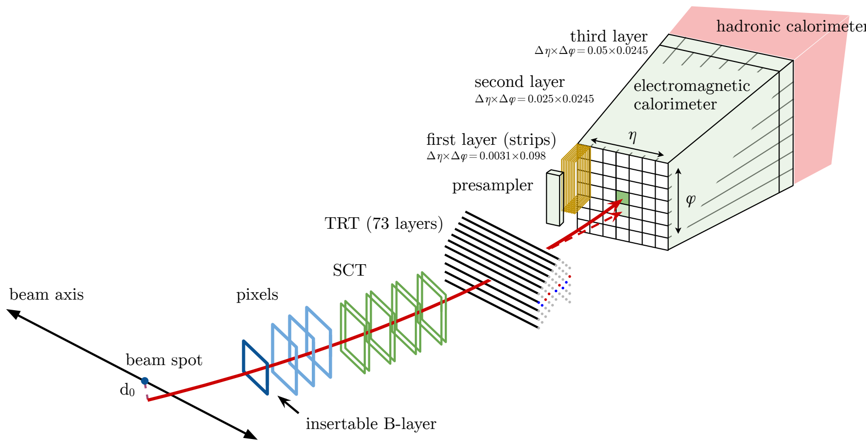


Figure 5.6: An illustration of the path of an electron through the elements of the ATLAS detector. The dashed red line indicates the trajectory of a photon produced by the electron's interaction with material in the ID [112].

interactions with the detector material, often forming electron-positron pairs via the emitted photons. These collimated emissions are typically reconstructed as a single cluster. Clusters with measured energy over 400 MeV are considered for matching with tracks from the ID. Tracks that are loosely matched in η and ϕ to these topo-clusters undergo a fitting procedure using a *Gaussian-sum filter* (GSF). The GSF is a nonlinear version of the Kalman filter, which better accounts for bremsstrahlung losses [115].

Together, these GSF tracks and EM topo-clusters are used to form dynamic, variable-size *superclusters*. High-energy *seed* topo-clusters combine with nearby *satellite* topo-clusters, so the resulting supercluster includes bremsstrahlung emissions. An electron object is defined by the supercluster and its matched track or tracks. Calibrations are applied to the electron energy and resolution to account for differences in real and simulated data [116].

5.2.2 Identification

Reconstructed electron candidates are subject to a likelihood-based (LH) identification to separate the prompt electron signal from backgrounds. These background sources include jets, photon conversions, and non-prompt electrons from heavy-flavour hadron decays. The electron LH is based on the product of several probability distribution functions of discriminating variables determined to differ between signal and background sources. These discriminants, listed in full in reference [112], include quantities related to hadronic leakage, EM shower width, track parameters, and energy deposit ratios. Simulation of $Z \to ee$ and $J/\psi \to ee$ processes are used to obtain distributions for signal variables, and simulation of two-to-two (dijet) events are used for background distributions. The distributions are corrected to accurately model data.

The final discriminant is defined so that a larger value represents a more electron-like candidate and a lower value a more background-like one. Electrons are grouped by quality according to a set of *working points* (WPs), each with a specific minimum discriminant value that the candidate must pass. Different WPs allow for a trade-off between background rejection and reconstruction efficiency. The WPs relevant to this dissertation are *LHTight*, *LHMedium*, *LHLoose*, *LHVeryLoose*, and a specialized *LHVeryVeryLoose* WP designed for the analysis in chapter 7. The working points are designed so that an electron meeting the requirements of a tighter WP will also pass all looser WPs. (e.g., an *LHMedium* electron will also count as an *LHLoose* electron.)

5.2.3 Isolation

Prompt electrons are expected to be spatially well-separated from other produced particles. By measuring the track and calorimeter activity in the vicinity of the electron candidate, the *isolation* metric is used to reject background processes.

Calorimeter-based isolation sums the transverse energy of topo-clusters within a cone around the electron's centre. Track-based isolation, used in this dissertation, is calculated as the transverse momentum scalar sum of tracks within a cone centred around the electron trajectory, excluding those tracks matched to the electron. Only tracks with $p_T > 1$ GeV are considered. The isolation variable is labelled according to the radius of the cone defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. For example, the $\Sigma p_T^{(0.3)}$ quantity measures the activity near the electron within a radius of $\Delta R = 0.3$.

5.2.4 Photons

Photons produced in LHC collisions are detected similarly to electrons and reconstructed using the same framework. There are two major classes of centrally-produced photons: *converted* and *unconverted*. Converted photons are those which interact with some material in the ID and produce an electron-positron pair. If both charged particles are distinguishable, they form a *conversion vertex*. A supercluster matched to a conversion vertex is used to identify the photon and its trajectory. Unconverted photons do not interact with the ID but only appear as energy deposits in the calorimeters. These are distinguished by a supercluster with no matching tracks or conversion vertices.

Photons undergo calibration, identification, and isolation like electrons but with distinct algorithms and working points. Because of the similarity of their signatures, it is necessary to resolve ambiguities between candidates that resemble both electrons and photons. Details of these procedures are given in detail in reference [113].

5.3 Muons

5.3.1 Reconstruction

In ATLAS, muons appear as minimum-ionizing particles that produce some combination of tracks in the MS, tracks in the ID, and distinctive energy deposits in the calorimeters.

In the ID, muon reconstruction begins with the tracking algorithm described in section 5.1. In the MS, nearby hit patterns are combined into straight-line "segments" within each muon chamber. A segment-seeded combinatorial search is performed to form track candidates by selecting for fit quality and relative orientation of nearby segments. These segments are loosely fit with a first-order parabolic trajectory approximation.

The precision measurements in the bending plane are combined with measurements from the muon trigger detectors (RPCs and TPCs) to provide a three-dimensional track candidate. To calculate precise trajectory, a global χ^2 fit of the candidate is executed using the hits from each segment, then iteratively refit by removing outliers and including unassociated hits.

Muon signatures are sorted into four classes depending on which detectors contribute to their identification: *combined*, *segment-tagged*, *calorimeter-tagged*, and *MS extrapolated*. Combined muons link independently-formed tracks from the ID and MS which are then globally refit. Segment-tagged muons are formed from ID tracks matched to one or more track segments in the MS. This accounts for muons with low momentum or which pass through MS regions of low acceptance. Calorimeter-tagged muons are ID tracks that can be matched to a calorimeter deposit consistent with a minimum-ionizing particle signature. MS extrapolated muons extend the $|\eta|$ region uncovered by the ID and are formed solely from candidates in the MS. Though they have no ID or calorimeter matching requirement, extrapolated muon trajectories are loosely constrained to originate from the IP. The search presented in this dissertation uses only combined and MS extrapolated muons. The reconstruction efficiency over most of the covered phase space is close to 99%. Calibrations are applied to the muon momentum scale and resolution to account for differences in real and simulated data [117]. After reconstruction, isolation is calculated for muons using the same definitions as those used for electrons (section 5.2.3).

5.3.2 Identification

Further quality requirements are applied to distinguish centrally-produced muons from backgrounds, mainly from pion and kaon decays. They are categorized into four WPs: *Loose*, *Medium*, *Tight*, and *High-*p_T. The Medium WP is the default muon definition in ATLAS and the WP used in the analysis presented in this dissertation.

Only combined and MS extrapolated muons are used in the Medium selection, and the latter only in the $2.5 < |\eta| < 2.7$ region. For a combined muon to be considered to be medium, it must have ≥ 3 hits in ≥ 2 MDT layers, except for those in the $|\eta| < 0.1$ region. An additional condition on the ID and MS momentum compatibility is required in the form of *q/p significance* used to suppress hadrons misidentified as muons. q/p significance is the absolute difference between the ratio of the charge and momentum of the muons measured in the ID and MS divided by the sum in quadrature of the corresponding uncertainties. Medium muons must have a q/p significance < 7.

For muons with p_{T} between 4 and 100 GeV, the Medium identification efficiency is 96% with less than 0.4% misidentification of hadrons as muons as measured in simulated samples of $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ events.

5.4 Hadronic jets

QCD confinement is the feature of the strong force that restricts free quarks from propagating, requiring they exist only in colour singlet combinations. Quarks are often produced in LHC pp collisions, but due to confinement, they form showers of secondary quarks and gluons. These showers hadronize, forming colour-neutral hadrons, mainly in the form of pions with some contribution from kaons and light baryons [24]. These high-momentum showers are relativistically boosted into collimated *jets*. Jets deposit their energy in the calorimeters and form tracks in the ID. Algorithms combine information from these detectors to reconstruct the energy and trajectory of the centrally-produced particle.

5.4.1 Reconstruction

Jets are reconstructed using a clustering algorithm with four-vectors from topologically associated energy deposits and charged-particle tracks. ATLAS uses topologically connected calorimeter cell signals known as *topo-clusters*. The algorithm which constructs the topo-clusters explores the three-dimensional spatial distribution of energy deposits in nearby calorimeter cells. The procedure is detailed in reference [114].

The primary observable that controls the cluster formation is the cell signal significance $\varsigma_{\text{cell}}^{\text{EM}}$ defined as the ratio of the cell signal amplitude to the cell's expected noise:

$$\varsigma_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}. \tag{5.2}$$

First, a seed cell with high significance is selected, then neighbouring cells with sufficient significance are iteratively added.

Once the topo-clusters are formed, they are clustered using an anti-k_t algorithm (described in detail in reference [118]) as implemented in the FastJet software package [119]. The anti-k_t algorithm is a sequential recombination procedure that uses the distance d_{ij} between every pair of topo-clusters. The distance is defined as

$$d_{ij} = \min\left(p_{\text{T},i}^{-2}, p_{\text{T},j}^{-2}\right) \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R^2}, \qquad d_{iB} = p_{\text{T},i}^{-2} \tag{5.3}$$

in terms of the transverse momenta p_{T}, rapidities y, and azimuthal angles ϕ. R is a tunable parameter of the algorithm which controls the size of the output jet. d_{iB} is the distance of the object from the beam. In order of increasing distance, objects i and j will be combined into a new object by adding their four-momenta. If, for some object j, the distance $d_{ij} > d_{iB}$, then object i is classified as a jet, removed from the list of objects, and the remaining distances recalculated.

High-momentum objects cluster with nearby low-momentum objects before low-momentum objects cluster amongst themselves. This results in conical jets of radius R with edges resilient to soft radiation but flexible to hard radiation.

Chapter 6

Large-radius tracking

Many theories of physics beyond the Standard Model (BSM) contain particles with lifetimes that are long compared to SM particles at the weak scale (see references [120, 121] and references therein). Long-lived particle (LLP) signatures are a promising yet relatively new domain for discovering new physics at the LHC. However, as searches at colliders have typically focused on low-lifetime phenomena, the reconstruction algorithms have been optimized to search for charged particles originating near the interaction point (IP).

The default ATLAS track reconstruction algorithm, in particular, imposes strict pointing requirements that lower the sensitivity to charged particle tracks originating some macroscopic distance from the IP. An additional pass of the tracking algorithm with relaxed pointing requirements can be applied to improve sensitivity to LLPs in the ID (referred to simply as LLPs for the remainder of this chapter). The large-radius tracking (LRT) pass used in Run 2 analyses provided a high track reconstruction efficiency, but it also produced many incorrectly reconstructed (fake) track candidates that need to be removed downstream by individual analyses. Additionally, this large number of fake tracks made the LRT algorithm computationally expensive, and therefore only a small subset of the recorded data could be reconstructed using LRT tracking, significantly affecting the sensitivity of many LLP searches.

In preparation for Run 3, a small team within ATLAS undertook a substantial retuning of the standard and LRT tracking algorithms focused on reducing the computational cost while maintaining a high track-reconstruction efficiency. These improvements enable the LRT pass to be run by default on all ATLAS events, streamlining the workflows of LLP searches and expanding their sensitivity.

This chapter presents the LRT algorithm and data preselection workflow used in Run 2 analyses. Improvements to the algorithm for Run 3 and Run 2 data reprocessing are then discussed as well as potential techniques for further refinement. Metrics of performance on simulation, data, and resources are presented, and the impact on future analysis workflows is estimated.

6.1 Motivation

Many BSM models contain predictions of neutral particles with long lifetimes.[13] These particles may have macroscopic decay lengths when travelling near light speed. The proper lifetime τ of a particle is calculated as the inverse of its decay width, which in the framework of QFT is

$$d\Gamma \sim \frac{1}{m}|\mathcal{M}|^2 d\Pi \tag{6.1}$$

[13] Some particles in the SM also have long lifetimes, such as charged and neutral kaons, charged pions, and B mesons.

where m is the particle's mass, $\mathcal{M}$ is the matrix element of the decay, and $d\Pi$ is the available phase space. Suppression of the matrix element, kinematic suppression of the available phase space, or a low mass will increase the lifetime of a decaying particle. One such example of an LLP search is that presented in chapter 7 of this dissertation, which describes how the presence of a massive right-handed neutrino with low coupling strength produces a displaced leptonic vertex.

Because of the exponential decay times of LLPs, the highest density of decays occur close to the IP. For decays with lifetimes between $O(.01)$–$O(1.0)$ ns, the ATLAS ID provides substantial coverage for LLPs in such models. In recent years, several ATLAS analyses have published results which use LRT tracks to seek direct evidence of BSM phenomena [122–128]. The reconstruction of displaced tracks is crucial for achieving sensitivity to these signatures.

6.2 Run 2 LRT

The LRT configuration and procedure described here are used in the analysis presented in chapter 7. A detailed presentation of the algorithm and its performance can be found in reference [129], from which text and figures in this section are drawn.

In the default reconstruction procedure, stringent requirements are placed on the transverse and longitudinal impact parameters of the reconstructed candidate tracks relative to the primary vertex (d_0 and z_0, respectively). This is intended to keep the computation time reasonable while still being very efficient in reconstructing primary particles with high purity. The specific constraints, referred to as standard tracking, are discussed in section 5.1. The standard tracking procedure comprises two tracking passes, inside-out and outside-in, referring to the portion of the ID where track seeds are formed.

LRT is not a separate algorithm but an additional pass of the inside-out tracking algorithm with substantially loosened selections on track impact parameters and the number of hits. The most significant differences in these selections are listed in table 6.1. These differences allow for the reconstruction of tracks originating further from the IP. Crucially, the only hits used in the LRT pass are those that have not been associated with any tracks already reconstructed by standard tracking. The number of *leftover hits* depends on the reconstruction efficiency and fake rate of the previous passes.

Table 6.1: Important selections that differ between the standard and the LRT setups in the Run 2 configuration [129].

Track parameter	Standard	Large-radius
Max. $\lvert d_0\rvert$	10 mm	300 mm
Max. $\lvert z_0\rvert$	250 mm	1500 mm
Max. track $\lvert\eta\rvert$	2.7	5.0
Min. track p_T	500 MeV	900 MeV
Min. Si hits	7	7
Min. un-shared Si hits	6	5
Max. shared silicon modules	1	2
Seed extension	Combinatorial	Sequential

The new requirements are applied to all stages of the tracking algorithm: seeding, track finding, ambiguity resolution, and TRT extension. The extension of seeds into track candidates uses a sequential Kalman filter as opposed to the combinatorial Kalman filter used in standard tracking, preventing the creation of multiple candidates from a single seed. In Run 2, the resulting LRT tracks are merged with the standard tracks to form the final track collection. This direct merging is possible without overlap removal since LRT tracks, by construction, do not share hits with standard tracks.

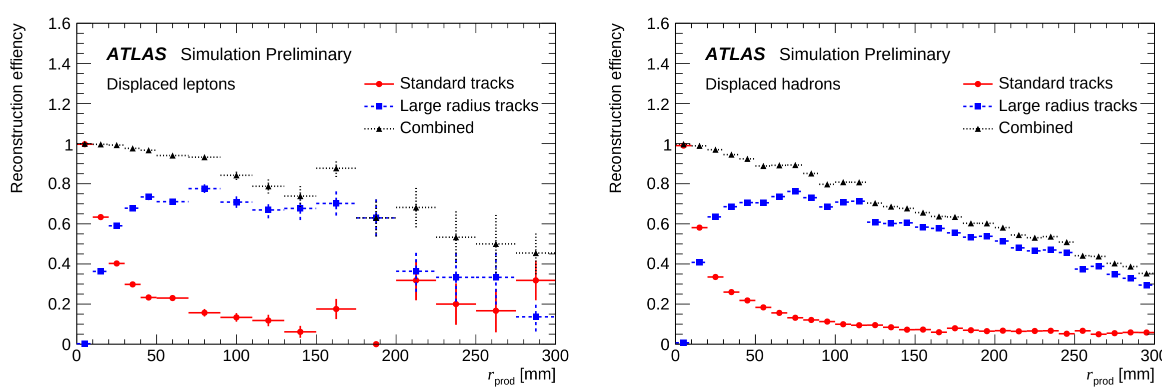
6.2.1 Performance

To gain insight into the LRT performance, the reconstruction efficiency is measured for two BSM physics processes, one supersymmetry (SUSY) model producing displaced leptons (via squark pair production) and another SUSY model producing displaced hadrons (via gluino pair production). These two processes were chosen to quantify the leptonic and hadronic performance of LRT, covering the common final-state signatures in ATLAS searches. The following studies use the definition of truth-match probability (p_{match}) as defined in section 5.1.2. The efficiencies are measured as a function of the production radius of displaced particles (r_{prod}).

In fig. 6.1, the inclusive reconstruction efficiency is defined as the ratio of reconstructed tracks matched to a truth particle ($p_{\text{match}} \geq 0.5$) over the number of truth particles:

$$\text{reconstruction efficiency} = \frac{\text{\# signal truth particles matched to reconstructed tracks}}{\text{\# signal truth particles}}. \qquad (6.2)$$

For both the numerator and denominator, the truth particles are subject to simple fiducial selections: $p_{\text{T}} > 1\,\text{GeV}$, $|\eta| < 2.5$, $r_{\text{prod}} < 440\,\text{mm}$, charge = ± 1, from LLP decay (signal).



(a) Displaced leptons

(b) Displaced hadrons

Figure 6.1: Inclusive track reconstruction efficiency for displaced charged particles produced by the decay of long-lived signal particles. The efficiency is shown as a function of r_{prod}, with truth particles and tracks subject to fiducial selections. The efficiency for standard and LRT is additive, and the total efficiency is the sum of both [129].

At low r_{prod}, standard tracking is very efficient, and few hits are left over to form tracks in the LRT step. Tracks which are produced far from the interaction point are unlikely to be reconstructed by the standard tracking; without LRT, fewer than 20% of the tracks with $r_{\text{prod}} > 50\,\text{mm}$ originating from the LLP decay would be reconstructed. Adding the LRT step substantially increases the reconstruction efficiency beyond r_{prod} of 50 mm for both benchmark samples, maintaining above 50% efficiency to nearly 300 mm.

Notably, the average number of fake tracks per event (fake rate) is not reported in the Run 2 optimization of LRT [129]. It is found that an average of 70% of LRT tracks in Run 2 are fake.

6.3 Improvements for Run 3

ATLAS's Run 2 LRT approach is highly efficient in reconstructing the trajectories of displaced tracks, but the relaxed constraints allow for more combinatorial fake track candidates to be formed (see section 5.1.2 for a definition of fakes). Consequently, the time needed to run the complete ATLAS event reconstruction, including LRT, increases by about a factor of 2.5 compared to the standard configuration [129]. A corresponding increase in required storage space is also seen. The ATLAS computing budget does not allow the entire recorded dataset to undergo this secondary reconstruction, and so only a subset of the data is processed in this way. Which data is processed through the LRT pass is selected using a filtering scheme described in section 6.3.1.

Given the unavoidable inefficiencies of such filters and the human development and validation effort required, our LRT optimization team undertook the task of including LRT by default in preparation for the Run 3 data-taking period. The primary aim of this effort was to reduce the LRT pass's computational load so that it could be run on the full ATLAS dataset while staying within the CPU and storage budget, removing the need for filters for most analyses.[14]

While designing the HNL analysis presented in chapter 7 it became clear that the deficiencies in the existing LRT strategy caused the analysis sensitivity to suffer. Not only was the filter removing a substantial portion of relevant data, but the large number of fake tracks reconstructed by the existing LRT configuration led to poor-quality signal vertices and a large background originating from random-track-crossings (see section 7.5.5). A motivating factor in undertaking the LRT optimization was to improve the signal sensitivity for subsequent HNL analyses.[15]

In parallel, a similar campaign was conducted to optimize the standard tracking pass, expanding the flexibility of the computing budget allocated for tracking. A full description of the standard tracking optimizations is discussed in reference [130]. Incidentally, the optimization of standard tracking reconstructs fewer tracks (both real and fake), leaving more hits for the LRT pass. If LRT were not also optimized, the additional leftover hits would translate to an increase in fake LRT tracks.

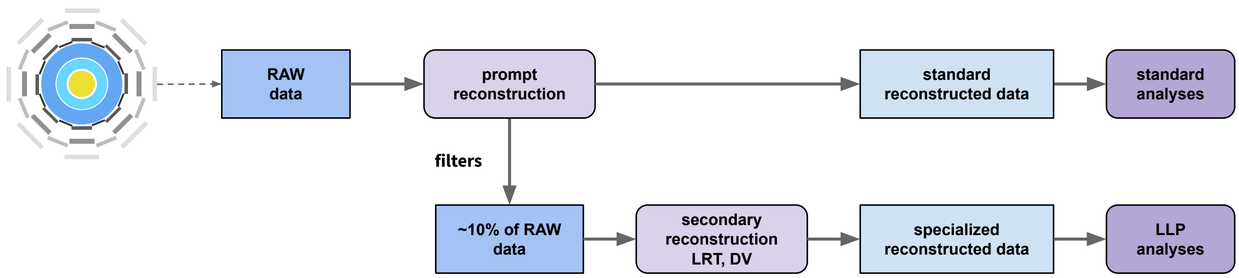
6.3.1 Filters

Current Run 2 LLP analyses rely on a set of filters to select a subset of events to be further processed. A schematic of the procedure is shown in fig. 6.2(a). The complete Run 2 dataset undergoes a standard reconstruction, from which the filters select events of interest. Approximately 10% of the raw data is selected. The reconstruction algorithms are then run again on this smaller subset, but with LRT and displaced vertex

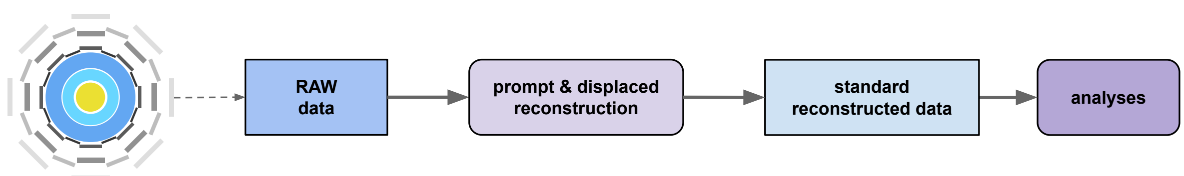
[14] The filter scheme is still retained in Run 3 for a few analyses which require secondary reconstruction algorithms beyond LRT.
[15] Indeed, an effort is currently underway to re-analyze the Run 2 dataset with the new LRT configuration.

finding (DV) enabled, allowing for use by LLP analyses. The filters are defined for each analysis in a way that optimizes the selection of signal events while keeping the total fraction of filtered events low. For example, the filter for the HNL analysis defined in section 7.2.4 selects approximately 50% of the simulated signal events but only approximately 2.6% of all recorded data.

Though passable, this method of filtering and reprocessing the data is problematic. It is cumbersome, computationally expensive, requires substantial development and validation labour, and is prone to errors.[16] As well as being burdensome, the filters discard potentially relevant data. This difference is significant for searches looking for only a few signal events.



(a) Run 2



(b) Run 3

Figure 6.2: In Run 2 (a), LRT tracking and displaced vertex reconstruction were only run on a subset of data selected by applying specialized filters to the standard reconstruction. This scheme requires substantial computing and human resources and is prone to errors. In Run 3 (b), the optimization of LRT tracking allows it to be run by default in the event reconstruction. LLP analyses can be performed using the standard data stream, allowing for more flexible analysis designs. The full ATLAS dataset will be available for future analyses without the need for filters.

6.3.2 Selection criteria

Track reconstruction takes place in sequential steps, with roughly increasing computational complexity. In order to minimize the time spent reconstructing fake tracks, they are removed as early as possible in the sequence and suppressing them at any point before final storage reduces the impact on disk space. Below is a summary of the carefully chosen selections applied at each step which optimize the LRT algorithm for Run 3. The numerical stages in the following descriptions refer to those in fig. 6.3.

[16]In preparing the data used in this dissertation, a configuration bug caused the primary reconstruction used by the filter to specify the incorrect ID alignment geometry. The months-long reconstruction effort had to be discarded, causing a significant delay for nearly a dozen analyses. Not to mention a substantial consumption of energy and resources from the LHC computing grid.

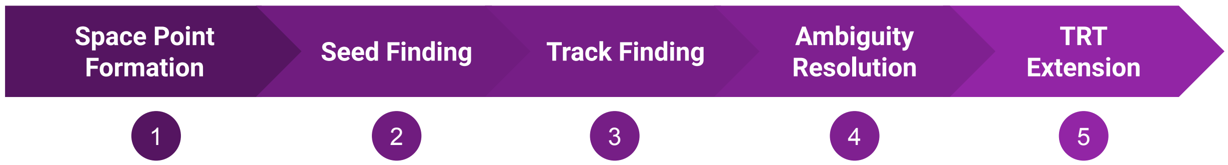


Figure 6.3: LRT tracks follow the same reconstruction sequence as standard tracks. By removing fake tracks as early as possible in the sequence, less time is spent reconstructing them in the later (and generally more computationally expensive) stages.

- **SCT seeding (2)** — Only SCT hits are used to produce three-hit seeds, as opposed to SCT or pixel hits.
- **Seed ranking (2)** — The seed ranking is modified. This ranking decides the order in which tracks are processed and prioritizes signal-like tracks.
- **Seed confirmation (2)** — Seeds, formed from three space-points, must be *confirmed*. This confirmation is achieved by replacing the outermost hit with one from the next silicon layer. If the old and new seed parameters are compatible with one another, the seed in confirmed. The confirming hit can be a pixel or SCT space-point.
- **Tightened cuts (2-4)** — Several cuts are tightened compared to the Run 2 configuration. Impact parameters, minimum required hits, maximum allowed holes, and fit quality are all tuned to minimize the rate of fake tracks. A complete description of changes is listed in table 6.2.
- **Neural network cut (4)** — Though not enabled by default, a neural network model is trained and implemented to substantially reduce fake tracks in the ambiguity resolution step.

6.3.3 DNN discriminant

In the standard and LRT tracking passes, several requirements are applied in the ambiguity-solving stage to remove fake tracks. These direct cuts are effective but one-dimensional. Multivariate discriminants are powerful tools useful for separating collections with correlations in multiple dimensions.

As part of the LRT optimization effort, a deep neural network (DNN) was implemented to differentiate between *good* and *fake* tracks (see section 5.1.2 for definitions). The inputs to this network are eight carefully chosen track parameters: z_0, d_0, η, number of SCT hits, number of SCT holes, number of silicon hits, number of degrees of freedom, and an engineered variable named $\Delta\theta$. $\Delta\theta$, as shown in fig. 6.4 is defined as the angular difference between the primary and secondary particle's flight directions, assuming the track is produced in a secondary decay. Each of these quantities is available after the global χ^2 fitting of the tracks. Many other quantities are available, but an effort was made to exclude quantities known to have poor agreement between data and MC (e.g., χ^2, pixel holes), as the DNN may exploit features of the simulation it was trained on, which may hinder its performance on real particle data.

Table 6.2: Differences introduced for Run 3 optimization of LRT. The following changes drastically reduce the fake rate while maintaining high reconstruction efficiency of displaced tracks.

Cut name	Run 2 LRT	Run 3 LRT
Min. p_T	0.9 GeV	1 GeV
Max. $\|\eta\|$	5	3
Max. $\|z_0\|$	1500 mm	500 mm
Min. secondary p_T	0.5 GeV	1 GeV
Min. clusters	7	8
Min. Si. not shared	5	6
Max. SCT hole	2	1
Max. double holes	1	0
Max. holes gap	2	1
Max. χ^2	15	9
Max. χ^2 no add.	35	25
Road width	20	5
Min. weighted clusters	6	8
Max. d_0 on seeds	1000 mm	300 mm
Do Z boundary	false	true
Strip seeding only	false	true
Require confirmed seed	false	true
Seed score	d_0	$\Delta\theta$

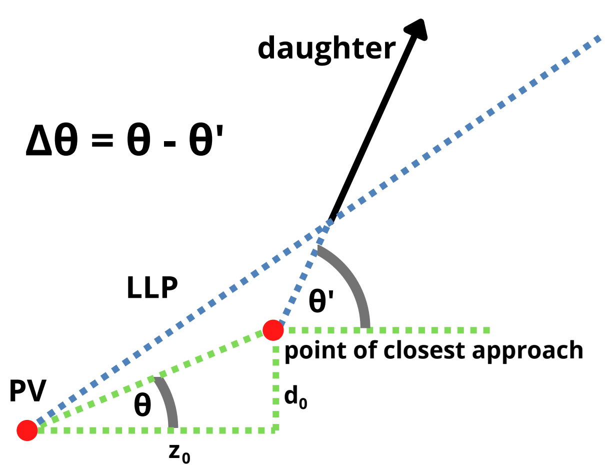
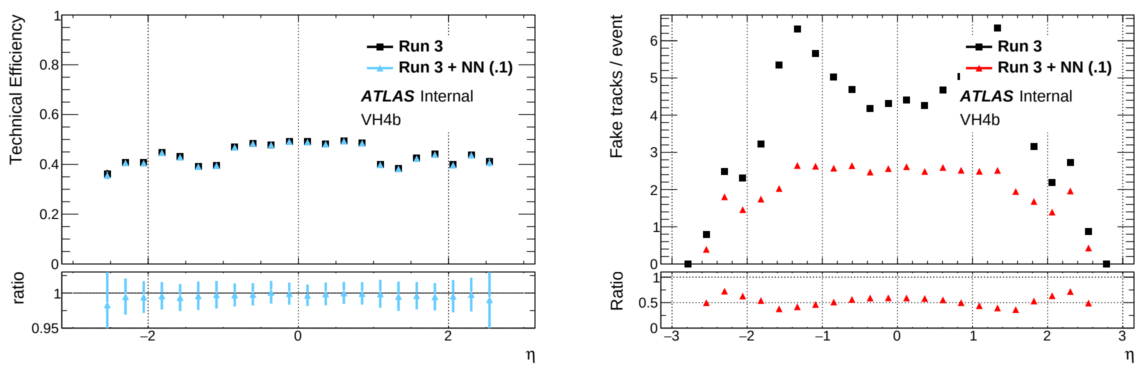


Figure 6.4: $\Delta\theta$, an engineered variable used in the DNN discriminant, is the angular difference between the primary and secondary particle's flight directions, assuming the track is produced in a secondary decay. This is used to estimate how "LLP-like" a track is.

A model was produced using Keras [131] with TensorFlow [132] backend. The input layer has eight input parameters, followed by four hidden layers with 24, 18, 8, and 8 inputs, each using a *relu* activation function. The output layer has one output and uses a *sigmoid* activation function. The model is trained using a *binary cross-entropy* loss function.

The DNN was trained on a collection of tracks from MC simulation samples, extracted and evaluated at the end of the ambiguity solver stage after fitting, as shown in fig. 6.6. These are the tracks that remain after all other optimizations from section 6.3 are applied. The samples were split into training, testing, and validation sets, and each track was categorized as good or fake. The categories were balanced, the features normalized, and the data weighted to avoid preferential treatment of the over-represented low-p_{T} tracks. The network's output is a value between zero and one, with a higher value representing the similarity of the input track to a good track. By implementing a cut at an arbitrary threshold, tracks that the DNN considers similar to fake tracks are removed.

By selecting 0.1 as the DNN score threshold, it is possible to remove 50% of fakes while only removing 2% of good tracks. Figure 6.5 shows the effect of this cut for a VH4b sample (see section 6.4.1). As the DNN is applied relatively late in the track reconstruction pipeline, this additional selection aims to reduce the required storage space rather than the reconstruction time. However, the evaluation of the DNN discriminant is very fast, and the removal of fake tracks from the TRT extension actually reduces the total time spent on track reconstruction.



(a) Technical efficiency

(b) Fake rate

Figure 6.5: Implementing a cut on the trained DNN discriminant can reduce fake LRT tracks a further 50% beyond the already-implemented optimizations. This reduction comes at the cost of only a 1–2% reduction in signal track reconstruction.

This trained model was implemented in the ATLAS reconstruction pipeline using the *lwtnn* (Lightweight Trained Neural Network) interface [133]. However, it is not used in default reconstruction and must be manually enabled. One reason for this is that the previous optimizations to LRT were sufficient to reduce the computing time to within acceptable limits. Another reason was that there remains skepticism from the track reconstruction experts that this network would not rely upon features poorly modelled in simulation. For example, the detector simulation takes an optimistic view of the damaged pixel sensors in the ID. The DNN may assign a high penalty to the number of silicon holes in each track, and as the pixel sensors are damaged by radiation throughout Run 3, the performance of the DNN may fall or introduce bias.

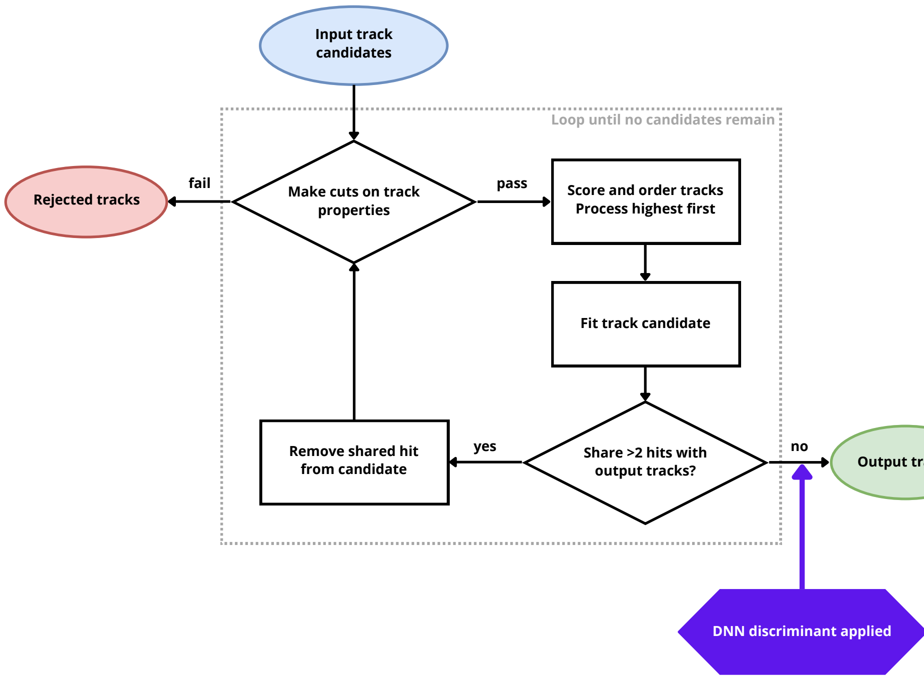


Figure 6.6: A diagram of the track ambiguity-solving procedure. The DNN discriminant is applied immediately before the output tracks are stored. This is also the point at which tracks are extracted to train the DNN.

6.4 Performance study on simulation

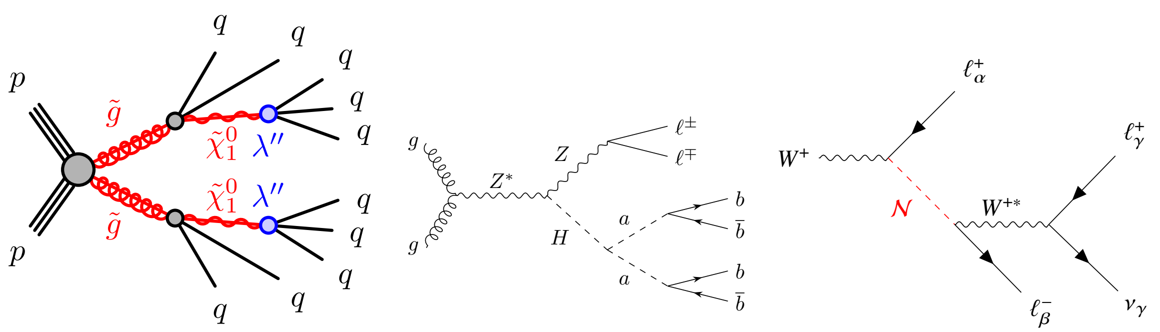
Performance studies were conducted on real and simulated data to quantify the improvements to the LRT parameters. The studies were used to iterate on the tuning and demonstrate the maintained reconstruction efficiency to the physics analysis groups most likely impacted by the changes.

6.4.1 Samples

Three sets of simulated BSM physics signals are studied to analyze the performance of the LRT pass on signatures of interest. Each of these signals is characterized by the production of a long-lived neutral particle which subsequently decays into charged SM particles. These signals have a high probability of producing charged particles from vertices originating in the ID, producing a high number of displaced tracks. Figure 6.7 shows the models with displaced vertices (DVs) used in the performance studies. Data and $t\bar{t}$ samples are also used in certain studies.

- **SUSY** (massive DV) — A supersymmetric (SUSY) long-lived neutralino model is used to study hadronic DVs with high mass and a high multiplicity of low-momentum tracks. This signal is simulated similarly to that in reference [126] but with coupling parameters tuned such that the vertex is displaced.

- **VH4b** (light DV) — A Higgs-portal model is used to study hadronic DVs with lower mass, producing



(a) SUSY (b) VH4b (c) HNL

Figure 6.7: Three different DV models used to optimize and characterize the LRT algorithm for Run 3. Figures from [1, 123, 126].

a few tracks in their decay vertices. The simulation of these samples is detailed in reference [123]. In the following performance plots, this model is labelled as VH4b (vector boson to Higgs to four b quarks).

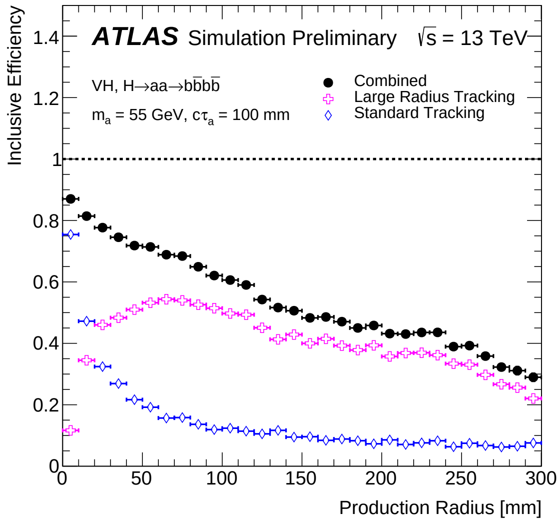
- **HNL** (lepton DV) — A heavy neutral lepton (HNL) model which produces a di-muon DV is used to study leptonic decay vertices. This is the same model as studied in chapter 7 and reported in reference [1].
- **t$\bar{\text{t}}$** (backgrounds) — To simulate common backgrounds which appear in dense environments, top-antitop ($t\bar{t}$) MC simulation is used with both top quarks decaying hadronically. The parameters of the simulation are reported in reference [134].
- **Data** — $\sqrt{s}$ = 13 TeV data from the 2018 data-taking period is used to measure the accuracy of the simulation. An unbiased sample of events (zero bias) is used, in which data is collected by triggering randomly filled bunches.

6.4.2 Efficiency

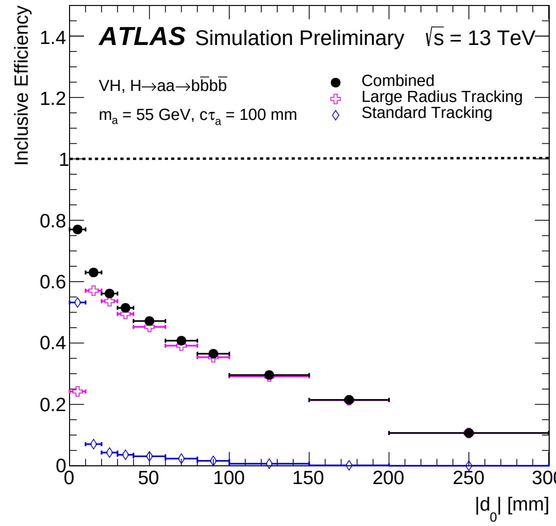
The efficient reconstruction of charged-particle tracks is important, but with the revised LLP analysis strategy, a sacrifice in efficiency is unavoidable. In tuning the LRT pass, the goal is to keep the reconstruction efficiency as high as possible in key regions of phase space.

Truth particles are subject to the same fiducial selections described in section 6.2.1. Distributions of the standard and LRT tracking efficiency, as well as their combination, are shown in fig. 6.8. The performance for both hadronic and leptonic signals is quite high. As in the performance studies for the Run 2 configuration, LRT recovers the substantially diminished tracking efficiency at high displacement.

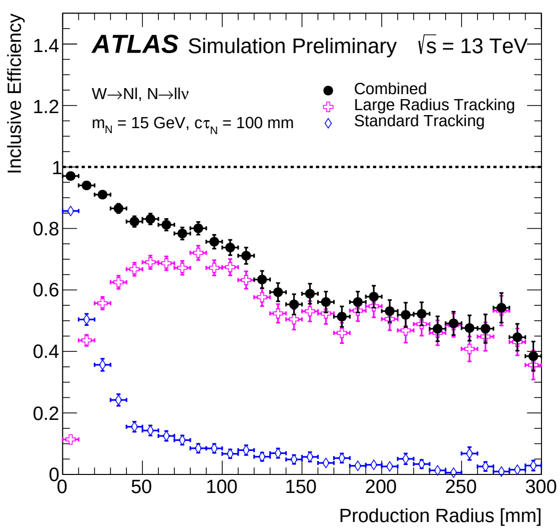
A *technical efficiency* is also defined to separate the algorithmic efficiency from restrictions arising from the detector limitations. The technical efficiency is used to determine the reconstruction performance for truth particles expected to leave enough hits in the detector to be reconstructed by the tracking algorithms employed. Selections similar to the track reconstruction requirements are applied to the truth particles in the



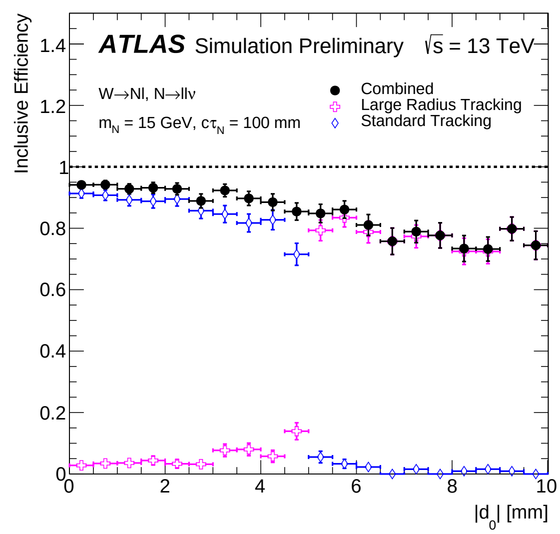
(a) Displaced hadrons



(b) Displaced hadrons



(c) Displaced leptons

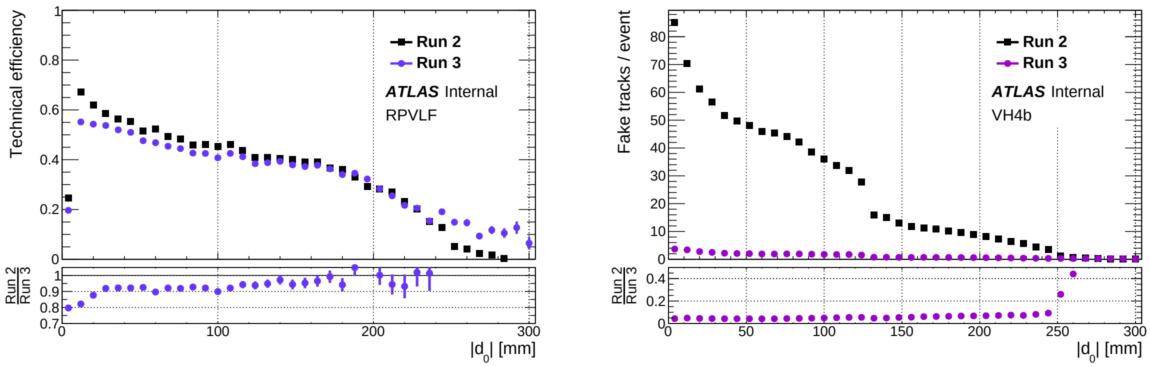


(d) Displaced leptons (zoomed)

Figure 6.8: Inclusive track reconstruction efficiency for displaced charged particles produced by the decay of long-lived scalar particles, a (figs. (a) and (b)) or by the decay of long-lived heavy neutral leptons, N (figs. (c) and (d)). The efficiency is shown as a function of the radius of production of displaced particles (figs. (a) and (c)) and the absolute value of the transverse impact parameter, $|d_0|$ (figs. (b) and (d)). As the HNL decay products are highly boosted, the signature has lower $|d_0|$ tracks, so fig. (d) focuses on the transition from standard tracking to LRT around $|d_0| = 5\,\text{mm}$ [135].

denominator. The selections are: $p_\text{T} > 1\,\text{GeV}$, $|\eta| < 5$, $r_\text{prod} < 300\,\text{mm}$, number of silicon hits ≥ 7. This measures the efficiency of the tracks that *could* have been reconstructed.

A direct comparison of technical reconstruction efficiency between Run 2 and Run 3 configurations is shown in fig. 6.9(a). The hadronic and leptonic signatures show a 10–15‰ reduction in track reconstruction efficiency. This was seen as acceptable by the physics analysis groups in light of the tremendous reduction in fake tracks and the ability to gain access to the full recorded dataset.



(a) (b)

Figure 6.9: The technical efficiency of LRT tracks for a SUSY signal sample compared between Run 2 and Run 3 configurations is shown in (a). For most of the d_0 range, over 85% of the Run 2 efficiency is retained. The fake rate comparison for the same signal sample is shown in (b). The rate of fake tracks is reduced by 95% on average.

6.4.3 Fake rate

Fake tracks include both combinatorial fakes and secondary fakes. *Combinatorial fakes* are defined as reconstructed tracks that have a $p_{\text{match}} < 0.5$, where p_{match} is a weighted measure of how many hits are matched to a single truth particle. *Secondary fakes* are tracks with no matching truth link in simulation and generally from material interactions. Both are discussed in more detail in section 5.1.2.

A useful measure of optimization is the *signal sensitivity* defined as the signal count divided by the square root of the background count, $s = S/\sqrt{B}$ [5].[17] For any given analysis design, one can simulate the number S of expected signal events, and the expected number B of background events. The Run 3 optimizations described above reduce the fake rate by 95% in both signal (HNL, VH4b, SUSY) and background (hadronic $t\bar{t}$) compared to the Run 2 configuration. Figure 6.9(b) shows the drastic reduction in the fake rate between Run 2 and Run 3 configurations. This equates to an increase in signal sensitivity of 400%! By enabling the DNN discriminant, the fake rate can be decreased by an additional 50% at the cost of only 1–2% efficiency reduction or a sensitivity increase of 550%. This reduction in fake tracks also has downstream effects on vertex creation.

[17] A general statement of the significance s of the hypothetical signal observation is then given by the number of Gaussian standard deviations of the background distribution to which the signal corresponds, $s = S/\sigma(B)$. If the systematic uncertainty on the background can be neglected, this is given by $s = S/\sqrt{B}$, where we have assumed a Poisson distribution for the background event counts, thus making the standard deviation equal to the square root of the mean [5].

6.4.4 Vertexing

Many LLP analyses use additional secondary vertex reconstruction algorithms (VSI), with vertices often formed from LRT tracks [136]. (See section 7.3.3 for a summary of the VSI algorithm). A high number of fake tracks contaminating the event translates to a large fake vertex background. This background particularly impacts light LLP signatures, for which fake vertices can be difficult to reject. As such, the purity of the track collection (ratio of good to fake tracks) is crucial in reconstructing high-quality vertices.

The impact of the considerable reduction in fake LRT tracks is found to reduce the number of fake secondary vertices by up to a factor of ten. Not only are the fake vertex backgrounds reduced, but the reconstruction efficiency improves due to fewer fake tracks interfering with the vertex reconstruction. Figure 6.10 shows the vertexing efficiency and fake rate for a VH4b signal comparing LRT reconstruction with Run 2 and Run 3 configurations. The definition of reconstruction efficiency and fake vertices is found in reference [136].

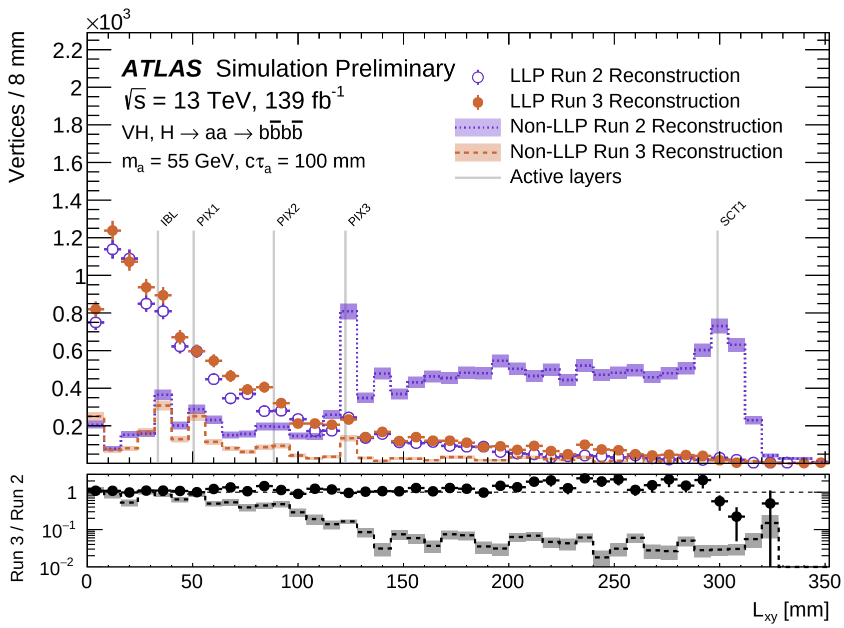


Figure 6.10: A comparison of the radial distributions of reconstructed secondary vertices in a simulated LLP sample using the Run 2 and Run 3 track reconstruction configurations is shown. The circular markers represent reconstructed vertices matched to truth-level LLP decay vertices. The dashed lines represent reconstructed vertices not matched to truth-level LLP decay vertices (non-LLP). The four pixel layers and the first SCT layer are shown in the figure with vertical lines, and charged-particle interactions with this material are responsible for spikes in the distributions. The Run 3 reconstruction shows a modest increase in the number of LLP vertices and considerably fewer non-LLP vertices due to the significant increase in purity of the LRT algorithm [135].

6.4.5 Pileup robustness

The average pileup is expected to increase to $\langle \mu \rangle = 50$ in Run 3 (with extremes up to 90) compared to between 20–40 in Run 2, further motivating the need to optimize LRT and other reconstruction algorithms [137]. In anticipation of this increase, it is helpful to characterize the LRT algorithm's response to pileup and ensure that the number of fake tracks does not increase at an unsuitable rate. The difference in slopes in fig. 6.11 indicates that high-pileup environments increase the fake rate more for LRT tracks than for standard tracks. Regardless, pileup sensitivity for LRT tracks is improved by a factor of six from the Run 2 configuration.

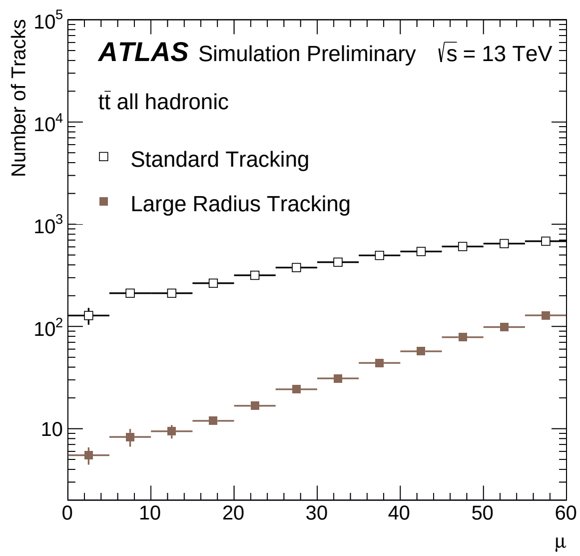


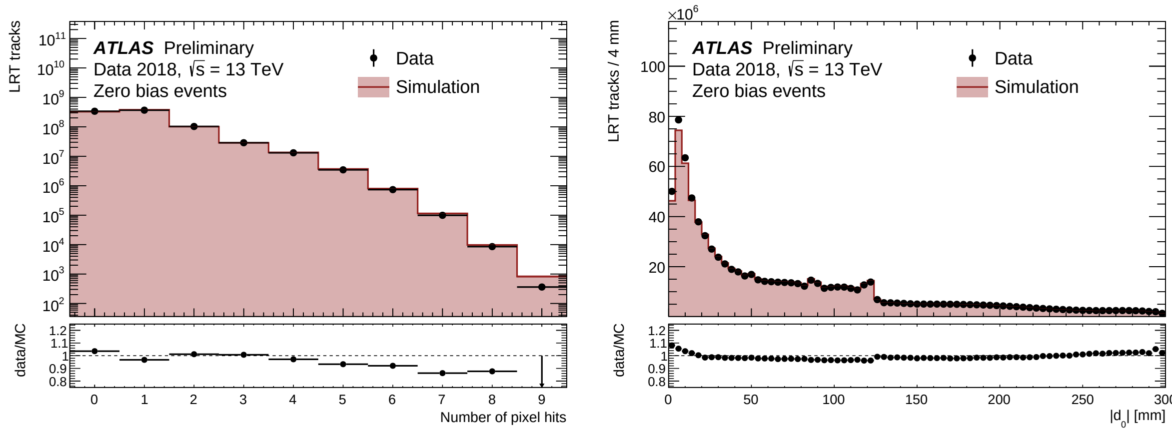
Figure 6.11: The mean number of reconstructed tracks per event as a function of μ in simulated $t\bar{t}$ events for the standard and LRT track reconstruction (Run 3 configuration). This quantity reflects the additional tracks reconstructed by the LRT pass for typical SM processes. The number of real tracks is expected to scale linearly with this quantity since it is related to the number of charged particles produced in the collisions. The number of random combinations is expected to scale with a higher power since the enhanced combinatorics allow for more such candidates to be formed. The number of tracks for the LRT reconstruction increase faster as a function of μ which indicates that it is more impacted by random combinations at high pileup [135].

6.5 Comparison to data

The reconstruction algorithms must produce similar results in data and MC simulation as many analyses rely heavily on simulation for background estimates and expected signal efficiency. A comparison of several distributions is made to ensure that the modifications to the LRT algorithm do not inadvertently produce unacceptable data/MC differences. Zero bias data (i.e., data collected from pseudo-randomly selected events[18]) from the 2018 data-taking period is compared to simulated samples of inelastic pp scattering events, both reprocessed using the Run 3 reconstruction software [130]. Simulated events are reweighted such that the number of interactions per bunch crossing is consistent with that of the data, with a mean value $\langle\mu\rangle = 34$. The distributions of track transverse momenta are found to be consistent between data and simulation [138]. Figure 6.12 shows examples of quantities used to characterize tracks, with MC and data in

[18]The zero bias data stream selects events gathered one complete beam revolution after a high-p_T trigger is fired. This is intended to be a "random" bunch collision but proportional to the bunch's instantaneous luminosity.

good agreement.



(a) (b)

Figure 6.12: A comparison of track quantities in zero bias data (black markers) and simulated inelastic pp collisions (filled red histogram) is shown. The simulation is normalized to the number of tracks in data. Figure (a) shows the distribution of the number of hits in the pixel detector, and figure (b) shows the distribution of the absolute value of the transverse impact parameter, $|d_0|$ [138].

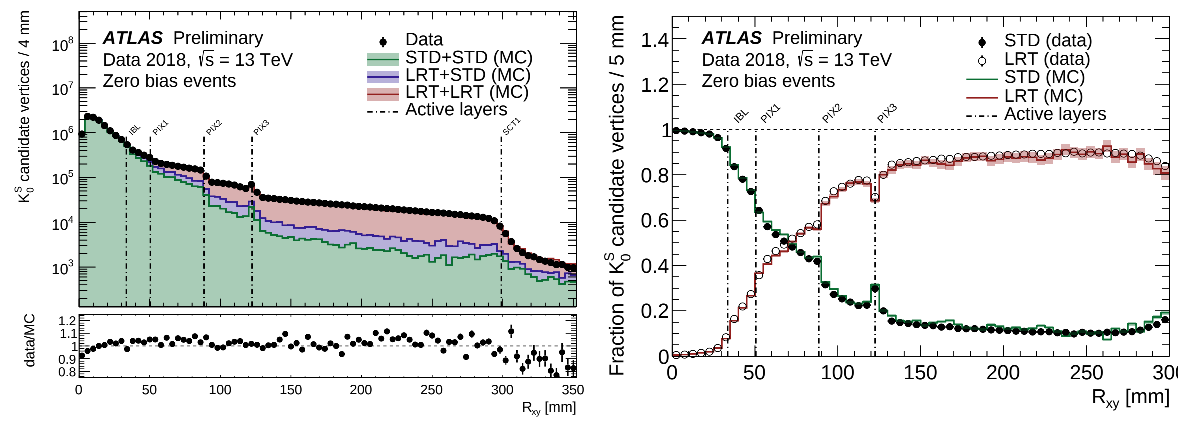
As displaced vertices are often the final product of LRT tracks, a "standard candle" in data is needed. K^0_S mesons are good candidates as they have a mean decay length of $c\tau = 27$ mm and are produced in abundance in LHC pp collisions. K^0_S decays are identified by reconstructing two-track vertices using both standard and LRT tracks and selected by requiring $472\,\text{MeV} < m_{\text{DV}} < 522\,\text{MeV}$. Simulation is normalized to data by the total number of DVs. Figure 6.13 shows that LRT efficiency in data and MC are in good agreement and that LRT tracks are vital for efficiently reconstructing vertices produced beyond the first layer of the pixel detector.[19] K^0_S decays are important for estimating displaced vertexing and tracking uncertainties as discussed in section 7.6.4.

6.6 Performance study on computing resources

The optimization of LRT for Run 3 is best seen in the context of the parallel optimizations of standard tracking. Improvements to both algorithms drastically reduce the CPU time and required disk space. Both result from the reduction of fake tracks in the reconstruction pipeline.

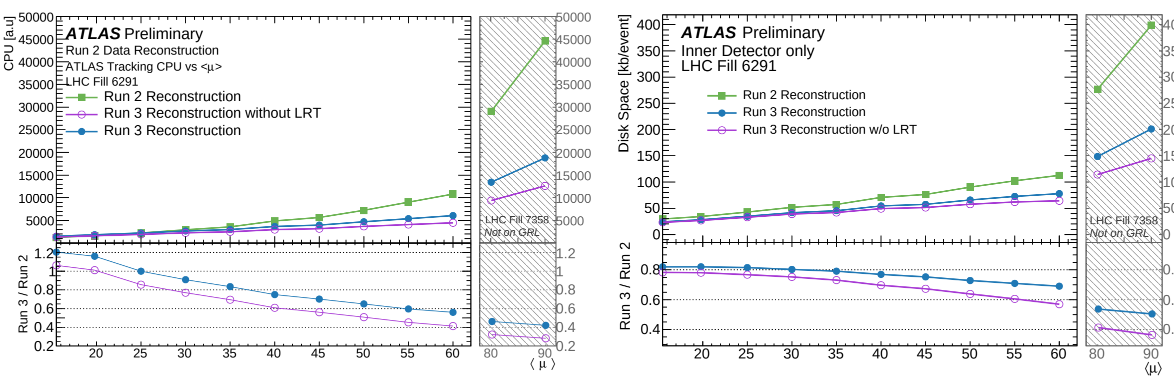
The time taken for track reconstruction in the software release used during Run 2 compared to the new release prepared for Run 3 is shown in fig. 6.14(a) as a function of $\langle\mu\rangle$ [139]. The reconstruction time is shown for both standard-only and standard+LRT tracking. Even with LRT, the total time needed for tracking reconstruction is reduced substantially compared to Run 2, up to 40% at $\langle\mu\rangle = 60$. The new tracking optimization also scales less sharply with $\langle\mu\rangle$. In terms of disk space, similar reductions are seen in fig. 6.14(b).

[19]In these figures, R_{xy} represents the same quantity as r_{DV} in chapter 7: the DV's reconstructed distance from the IP in the x-y plane.



(a) (b)

Figure 6.13: Figure (a) shows the radial distribution of reconstructed K_S^0 candidate vertices in 2018 zero bias data (black markers) and simulated inelastic pp collisions (stacked histogram). Vertices in simulation are separated into three categories based on the number of LRT tracks included in the vertex fit. The total number of vertices in simulation is normalized to the number of vertices in data. The four pixel layers and the first SCT layer are shown in the figure with dashed vertical lines. Figure (b) shows the fractional distribution of the same K_S^0 candidates in data and MC comparing vertices reconstructed with two standard (STD) tracks and at least one LRT track. Both figures indicate that LRT efficiency in data closely matches MC [138].



(a) CPU (b) Disk space

Figure 6.14: Processing time (a) and disk space (b) taken per event versus average pileup $\langle\mu\rangle$ to reconstruct the same data events, comparing the Run 2 (green squares) and Run 3 (purple unfilled circles) reconstruction software. The Run 3 numbers are also presented, including the impact of the additional LRT tracking pass (blue filled circles). The bottom panels depict the time taken as a fraction of the Run 2 result. The disk space values reflect the output of the ID reconstruction in the ATLAS EDM format [130].

A back-of-the-envelope calculation can estimate the additional savings by removing filtered and reprocessed data. In the default reconstruction scheme (without LRT), for an estimated 25×10^9 events in Run 3 at 400 kB/event, the expected AOD size is ~ 10 PB. Storing LRT tracks for each event increases the size by an additional 11.7% for events with $\langle\mu\rangle$ between 40 and 90 (a conservative estimate). Approximately an additional 1.17 PB of data and 2.17 PB of MC will be required by enabling LRT. However, if the filter scheme is carried on into Run 3, a duplicated 2.40 PB of raw and 2.22 PB of reconstructed data must be stored. Additionally, the derived AODs (copies of AODs with only analysis-relevant data retained) are expected to reduce in size by 80% due to the removal of 95% of fake tracks, translating to approximately 1.44 PB in savings. Summing these debits and credits, it is possible to recover on the order of 2.72 PB of space by storing LRT information in the standard stream instead of in a duplicate data stream.

6.7 Impact on analyses

For a sample of LLP analyses, the LRT optimization team collected the estimated impact of switching from a Run 2+filters configuration to a Run 3 configuration without filters. Each analysis generally sees the signal reconstruction efficiency increase or remain constant. The feedback reported here is taken from an intermediate optimization with roughly 5% reduced LRT efficiency, except for the HNL signal, which reports the final configuration with 10-15% reduced LRT efficiency. An attempt was made to separate the impact of only the LRT optimization and the removal of the LLP filter.

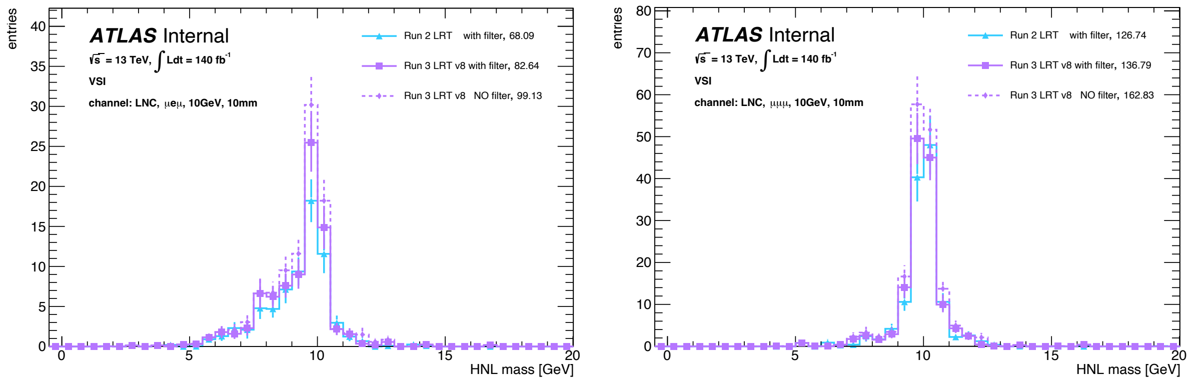
In a sample HNL signal with a di-muon displaced vertex (see chapter 7), the reconstruction efficiency increases by 8% over the Run 2 configuration simply by optimizing LRT to reduce fake tracks. By removing the filter, that gain increases to 28%. For an HNL sample with an $e\mu$ displaced vertex, LRT optimization shows gains of 21% and removing the filter increases that to 45%. This shows that for the HNL analysis, the optimized LRT pipeline would correspond to an improved sensitivity equivalent to gaining an additional 63 fb^{-1} of data. Distributions of the reconstructed HNL mass and the total yields are shown in fig. 6.15.

A VH4b sample of displaced hadrons shows nearly a 100% increase in signal efficiency by using the Run 3 LRT with no filter. The signal remains effectively constant for a displaced τ lepton sample while fakes are reduced by 80%. A SUSY DV analysis generally sees more DVs in regions where m_{DV} is less than 800 GeV.

In addition to the signal gains, backgrounds are heavily suppressed for each of these analyses. For example, the dominant background in the search for HNLs is due to two-track vertices formed by the random crossing of unrelated tracks (see section 7.5.5). It is suspected that the majority of these are due to the crossing of fake tracks. For signals with multi-track vertices, fake tracks attached to the vertex can distort the calculation of the invariant mass and other quantities. Run 3 is expected to show tremendous gains in sensitivity and accuracy due to the increased track purity.

6.8 Inclusion in Run 3 and Run 2 reprocessing

After performance studies on the efficiency and computing load were presented, the collaboration agreed to include the LRT pass in the default ATLAS reconstruction pipeline. The standard and LRT track collections are stored as separate entities in the ATLAS EDM, but a *virtual view* is implemented, allowing analyzers to access the merged collections easily. This implementation is preferred as it minimizes disruption to ongoing analyses that would otherwise be suddenly inundated with displaced tracks if stored in a single collection.



(a) $\mu - e\mu$

(b) $\mu - \mu\mu$

Figure 6.15: Optimizing the LRT algorithm improves the yield of two types of displaced HNL decay vertices by reducing the number of fake reconstructed tracks and thus improving the displaced vertex reconstruction performance. Vertices produced with the Run 2 reconstruction software are plotted with blue triangles. Vertices produced with the Run 3 LRT configuration are purple squares and show increased signal reconstruction efficiency. Removing the filter further improves the efficiency, and these vertices are shown with a dashed purple line. In this figure, v8 refers to the final version of the LRT configuration ultimately included in the default reconstruction software. The quantity on the x-axis is the reconstructed HNL mass as defined in chapter 7.

Subsequent reconstruction algorithms which rely on tracks, particularly leptons, use these views to create a "standard" and a "LRT" version of the reconstructed objects.

With LRT now able to run by default on all recorded data, LLP analyses can follow the same pipeline as standard analyses. LRT, as well as many other reconstruction optimizations, will be part of reprocessing the raw data from the full Run 2 dataset and will be used for processing all data from the upcoming Run 3. Not only will considerable time and effort be freed, but the remaining 90% of the ATLAS dataset will be available for the LLP analyses. Access to the full dataset, higher-purity displaced tracks, and more accurately modelled background processes are all expected to increase these analyses' signal sensitivity.

Chapter 7

Heavy neutral lepton analysis

As described in detail in chapter 3, the mass of the neutrino provides an alluring window into BSM physics. The theoretical right-handed component of the neutrino carries no SM gauge charges and is therefore also known as a *sterile neutrino*, referring to its non-interaction with the other SM particles. However, the seesaw mechanism (section 3.3.3) postulates the existence of a heavy neutrino state $\mathcal{N}$ that contains a small admixture of an active neutrino flavour: $\nu_e, \ \nu_\mu, \ \nu_\tau$. This allows $\mathcal{N}$ to participate in weak interactions, albeit suppressed by a dimensionless mixing coefficient U_α, where $\alpha = e, \ \mu, \ \tau$.

With all the necessary ingredients presented in previous chapters, this chapter will detail a direct search for heavy neutral leptons (HNLs) using the ATLAS detector [1]. In the decay channels studied, the HNL, produced via the decay of a W boson, decays into a neutrino and two charged leptons,[20] which form a displaced vertex. Feynman diagrams of the relevant decays are shown in fig. 7.2. The search is performed using 139 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data.

For HNLs with low mass and low mixing angles, the decay vertex can occur at some macroscopic distance from the interaction point. These *displaced vertex* (DV) signatures provide sensitivity to low coupling strengths with very little background contamination. Historically, searches for new particles (such as for the Higgs boson) assume the particle decays very quickly, leaving a vertex within millimetres of the interaction point. However, without any discovery of new physics, an effort is underway in high-energy physics to probe unexplored regions of phase space by searching for *long-lived particles* (LLPs).

The defining feature of this analysis is that it searches for HNLs that are *long lived*, decaying some macroscopic distance from the primary vertex (PV). HNLs may produce displaced vertices due to a small coupling between the HNL and the states to which it decays. The quantity $c\tau$ is often used to characterize the decay length of LLPs. This is a measure of distance defined as the product of the LLP's proper lifetime τ and the speed of light c, which is a reasonable approximation of the speed of the decay products. $c\tau$ is the *proper decay length* of the particle which can be mapped to how far the particle will travel before it decays [5]. In this dissertation, and in much of the literature, $c\tau$ is sometimes referred to as the "lifetime" of the LLP. This analysis searches for HNLs that decay between 4 and 300 mm, which is inside ATLAS's inner detector volume.

7.1 Models

Experiments to date have presented results that assume only one HNL is kinematically accessible and that other HNLs are outside the experiment's sensitivity. These experiments place exclusion limits on the HNL

[20] In this chapter the term *leptons* refers to electrons or muons unless otherwise specified.

mass m_N and the mixing angle $|U_\alpha|^2$.[21] This is referred to as a "one-HNL model with single-flavour mixing" (1SFH). Though useful for deriving generic bounds and benchmark measurements, a 1SFH model is not phenomenologically viable as it does not account for the two observed neutrino mass splittings and flavour oscillations. A minimal model that can account for these splittings is that of "two quasi-degenerate HNLs" with close masses and couplings (2QDH).[22]

The interactions of each HNL via the weak force are suppressed by the mixing angles $U_{\alpha I}$ between an active neutrino flavour $\alpha = e, \mu, \tau$ and a specific HNL I. Experimentally, only the cumulative mixing can be measured, so the following definitions are applied:

$$|U_\alpha|^2 \equiv \sum_I |U_{\alpha I}|^2 \quad \text{and} \quad |U_{\text{tot}}|^2 \equiv \sum_{\alpha,I} |U_{\alpha I}|^2 \tag{7.1}$$

where I is the number of HNLs in the model. In 1SFH models, as reported in other experiments, U_α and $U_{\alpha I}$ are equivalent, so no distinction is made, but the clarification is important in defining the 2QDH models. The fractional mixing x_α is defined as

$$x_\alpha = |U_\alpha|^2/|U_{\text{tot}}|^2. \tag{7.2}$$

In the following analysis, two 1SFH models and two 2QDH models are used to interpret the observed data. The 1SFH models are the common benchmarks in which the HNL mixes either completely with the electron neutrino ($|U_\alpha|^2 = |U_e|^2$) or completely with the muon neutrino ($|U_\alpha|^2 = |U_\mu|^2$). The possible realistic values of the HNL mixings in the 2QDH models are constrained by neutrino oscillation data as discussed in reference [140]. Two values are selected for the 2QDH parameters — one consistent with the normal hierarchy (NH) of neutrino masses and one for the inverted hierarchy (IH). A ternary diagram showing the oscillation data fits from NuFIT [141] and the four selected mixings is shown in fig. 7.1.

Furthermore, for each benchmark, this analysis considers the case in which the HNL is a Majorana particle or a Dirac particle. In the Majorana limit, the HNL decay has an equal chance of being lepton-number violating (LNV) or lepton-number conserving (LNC), whereas in the Dirac case only LNC is allowed. A complete summary of the models used in the interpretation of results can be found in table 7.1.

7.2 Samples

7.2.1 Monte Carlo samples

To study the signal sensitivity, Monte Carlo (MC) signal samples are generated using Pythia 8.212 [142] with the A14 set of tuned parameters [143] and the NNPDF2.3lo PDF set [144]. The impact of multiple pp interactions per bunch crossing (pileup) are modelled by adding simulated minimum-bias events generated with Pythia 8.210 using the A3 tune [145] and the NNPDF2.3lo PDF set. Particles are propagated through a detector simulation [107] based on Geant4 [108].

[21] In a direct search such as this, the mixing angle dictates both the production and decay rates, and therefore only the square of the coupling constant can be measured.

[22] While two HNLs can account for active neutrino masses and matter-antimatter asymmetry of the universe, a third HNL at $\mathcal{O}$(keV) scale could be a viable dark matter candidate.

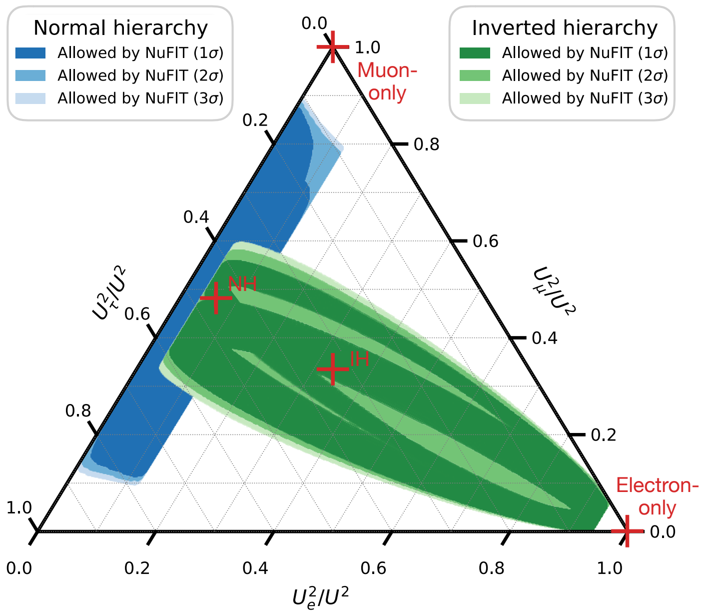


Figure 7.1: The allowed combinations of $|U_\alpha|^2/|U_{\mathrm{tot}}|^2$ that are consistent with NuFIT fits to neutrino oscillation data [141]. The allowed regions under normal (blue) and inverted (green) hierarchy are shown. The red cross markers represent two benchmark models that are chosen and agreed upon by the community [140].

Table 7.1: A summary of the eight models that are used in the interpretation of the analysis results. Models are defined according to the number of HNLs (one or two), the fraction of LNC and LNV decays, and the relative strength of the mixing between the lepton generations. The inverted (IH) and normal (NH) hierarchy benchmarks are used for 2QDH models, while only the 1SFH model is used for the single-flavour mixing interpretation.

Model	Fraction of LNV decays	Fraction of LNC decays	(x_e, x_μ, x_τ) $x_\alpha = |U_\alpha|^2/|U_{\mathrm{tot}}|^2$	Channels
1SFH muon-only in 'Majorana-limit' 1SFH muon-only in 'Dirac-limit'	1/2 0	1/2 1	(0 , 1 , 0)	$\mu-\mu\mu$, $\mu-\mu e$, $\mu-ee$
1SFH electron-only in 'Majorana-limit' 1SFH electron-only in 'Dirac-limit'	1/2 0	1/2 1	(1 , 0 , 0)	$e-ee$, $e-e\mu$, $e-\mu\mu$
2QDH IH mixing in 'Majorana-limit' 2QDH IH mixing in 'Dirac-limit'	1/2 0	1/2 1	(1/3 , 1/3 , 1/3)	$\mu-\mu\mu$, $\mu-\mu e$, $\mu-ee$, $e-ee$, $e-e\mu$, $e-\mu\mu$
2QDH NH mixing in 'Majorana-limit' 2QDH NH mixing in 'Dirac-limit'	1/2 0	1/2 1	(0.06, 0.48, 0.46)	$\mu-\mu\mu$, $\mu-\mu e$, $\mu-ee$, $e-ee$, $e-e\mu$, $e-\mu\mu$

In this search, six decay modes (channels) are generated whose combinations define each model. These channels are defined as $\alpha-\beta\gamma$ according to the flavour of the prompt lepton (α) and the flavours of the two charged leptons produced in the HNL decay ($\beta\gamma$). The relevant channels using this nomenclature are shown in fig. 7.2. For each $\alpha-\beta\gamma$ final state, signal samples are generated with HNL masses in the range $3\,\mathrm{GeV} < m_N < 20\,\mathrm{GeV}$ and proper decay lengths $c\tau_N = 1, 10,$ and 100mm.[23]

[23]In this chapter, a distinction is made between the HNL's mass m_N used in signal generation and the reconstructed mass m_{HNL} used in the event selection (see section 7.3.4).

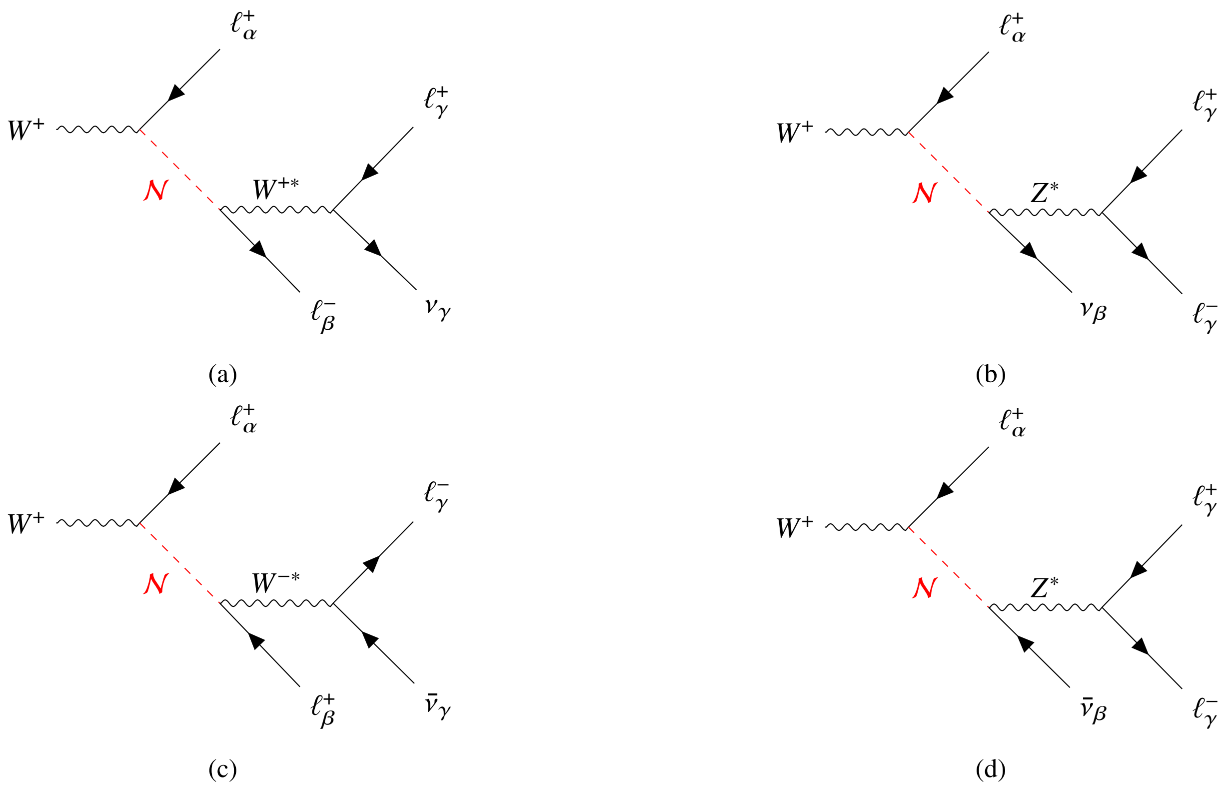


Figure 7.2: Feynman diagrams for the HNL production and decay modes targeted in this analysis. The flavors of the leptons in the diagrams, labeled by α, β, and γ, are either muons or electrons. If the charged leptons in the HNL decay have the same flavour, then both the diagrams with the virtual W (a,c) and virtual Z (b,d) contribute to the process. Lepton number conserving (a,b) and lepton number violating (c,d) processes are shown. Equivalent processes are also valid for an initial state W^- boson [1].

Cross section

To compute the cross section, two of the HNL's three interdependent properties — mass, mixing angle, and lifetime — must be specified. The branching ratio (BR) for the W boson decaying to an HNL ($\mathcal{N}$) and a charged lepton (ℓ) is directly proportional to the coupling strength $|U_\alpha|^2$ and can be expressed as [146]

$$BR(W \to \ell_\alpha \mathcal{N}) = BR(W \to \ell_\alpha \nu) \cdot |U_\alpha|^2 \left(1 - \frac{m_{\mathcal{N}}^2}{m_W^2}\right)^2 \left(1 + \frac{m_{\mathcal{N}}^2}{2m_W^2}\right). \tag{7.3}$$

The total HNL cross section $\sigma_{\mathcal{N}}$ for a specific $\alpha-\beta\gamma$ lepton channel is then

$$\sigma_{\mathcal{N}} = \sigma(pp \to W) \cdot BR(W \to \ell_\alpha \mathcal{N}) \cdot BR(\mathcal{N} \to \ell_\beta \ell_\gamma \nu) \tag{7.4}$$

$$= \sigma(pp \to W) \cdot BR(W \to \ell_\alpha \nu) \cdot x_\alpha \cdot |U_{\text{tot}}|^2 \cdot \left(1 - \frac{m_{\mathcal{N}}^2}{m_W^2}\right)^2 \left(1 + \frac{m_{\mathcal{N}}^2}{2m_W^2}\right) \cdot BR(\mathcal{N} \to \ell_\beta \ell_\gamma \nu). \tag{7.5}$$

The cross section times BR for W boson production and subsequent leptonic decay ($\sigma(pp \to W) \cdot \mathrm{BR}(W \to \ell\nu)$) in $\sqrt{s} = 13$ TeV pp collisions is taken from the ATLAS measurement in reference [147]. It is measured as (20.6 ± 0.6) nb.

$|U_\alpha|^2$ is factored as $x_\alpha \cdot |U_{\mathrm{tot}}|^2$, allowing for the variation of the fraction x_α to define each model. The total mixing angle $|U_{\mathrm{tot}}|^2$ for one Majorana HNL particle is computed for each of the generated $m_{\mathcal{N}}$–$c\tau_{\mathcal{N}}$ samples using a direct Fermi theory computation described in reference [148]. The proper time $\tau_{\mathcal{N}}$ itself has a strong dependence on the $m_{\mathcal{N}}$ as well as the mixing angle $|U|^2$. The HNL decay is a three-body charged-current-mediated processes and so the lifetime scales as

$$\tau_{\mathcal{N}} \propto \frac{1}{m_{\mathcal{N}}^5 |U_{\mathrm{tot}}|^2}. \tag{7.6}$$

These computations include additional corrections due to the finite masses of the final-state fermions.

Individual HNL decay branching ratios $BR(\mathcal{N} \to \ell_\beta \ell_\gamma \nu)$ depend on $m_{\mathcal{N}}$ and the mixing fractions x_e, x_μ, and x_τ. In the case of displaced leptons of the same flavour ($\beta = \gamma$), the BR is modified by interference from a virtual Z boson decay shown in fig. 7.2 (b,d). The BR is calculated for each $m_{\mathcal{N}}$ in each HNL decay channel. The calculated mixing angles and BRs are used in eq. (7.5) to compute the cross section for a Majorana HNL with a given $m_{\mathcal{N}}$ and $c\tau_{\mathcal{N}}$ for each $\mathcal{N} \to \ell_\beta \ell_\gamma \nu$ decay. The cross sections for the remaining model benchmarks are calculated by scaling the appropriate contributions from each channel according to the x_α values in table 7.1. For each model, the Majorana limit and the Dirac limit are studied. In 1SFH models, this is simply a matter of only considering LNC decays (Dirac) or allowing for equal probabilities of LNC and LNV decays (Majorana). As presented in reference [140], 2QDH models are enhanced by a factor of two. In the 2QDH Dirac limit, all decays are LNC, and in the 2QDH Majorana limit, they are equal parts LNC and LNV. The modifications to the cross sections are summarized in table 7.2.

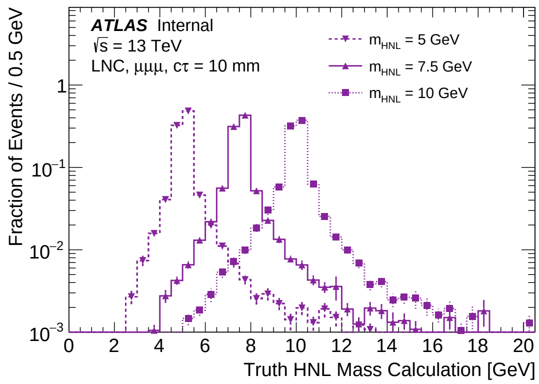
Table 7.2: The cross section for LNC and LNV decays for one HNL model in the Dirac or Majorana cases, a quasi-degenerate HNL pair in the Majorana case and a quasi-degenerate HNL pair in the Dirac case. The total cross section is the sum of the LNC and LNV contributions.

	One HNL Majorana-limit	One HNL Dirac-limit	Quasi-degenerate pair Majorana-limit	Quasi-degenerate pair Dirac-limit
LNC decays	σ_M	$2\sigma_M$	$2\sigma_M$	$4\sigma_M$
LNV decays	σ_M	0	$2\sigma_M$	0
Total	$2\sigma_M$	$2\sigma_M$	$4\sigma_M$	$4\sigma_M$

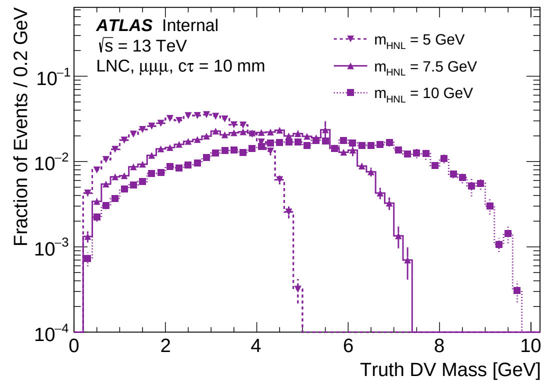
Signal kinematics

Kinematic distributions of truth-level[24] quantities are shown in fig. 7.3 for simulated HNL signals decaying in the μ–$\mu\mu$ channel. At truth level, kinematic differences between the signal channels are negligible since electrons and muons with p_{T} of order $\mathcal{O}$(GeV) can be treated as equivalent (massless) objects.

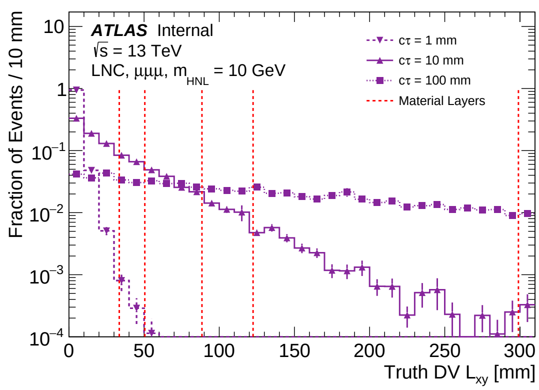
[24] Truth-level indicates quantities of the particles produced by the event generator, as opposed to those quantities reconstructed in the detector simulation.



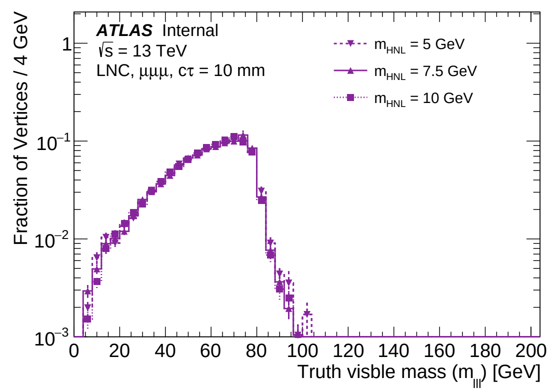
(a) HNL mass m_{HNL}



(b) DV mass m_{DV}



(c) DV transverse displacement r_{DV}



(d) Tri-lepton mass m_{lll}

Figure 7.3: Kinematic distributions of truth-level quantities from simulated HNL signal in the μ–$\mu\mu$ channel. Plots are shown for a 10 GeV HNL with proper lifetimes $c\tau_N$ = 1, 10, 100 mm. Each of these quantities is used in the analysis selection.

Spin correlations

To properly simulate spin correlations between W boson decay products, which are not accounted for in Pythia 8 [149–151], events are weighted to reproduce the angular distributions obtained with MadGraph5_aMC@NLO 2.9.3 [152] using the HeavyN model [153, 154]. The weighting procedure is validated by comparing the momentum spectra of each of the charged-lepton flavours and the neutrino between the weighted Pythia 8 and MadGraph5_aMC@NLO samples. Further details of the reweighting procedure can be found in reference [155].

7.2.2 Data samples

This analysis uses 139 fb^{-1} of $\sqrt{s}$ = 13 TeV pp collision data recorded by the ATLAS detector in the 2015–2018 data-taking period. The recorded data is subjected to a few quality requirements, then considered valid for use in the analysis [156].

- **Good Run Lists (GRLs)** — GRLs specify the luminosity blocks in which the beam conditions were stable, and the relevant detector subsystems were operational. Luminosity blocks which are not present in the GRLs are discarded.

- **Event cleaning** — Further cleaning takes place on the individual event level. Events may be flagged if there are problems in the liquid argon system, tile calorimeter system, SCT inner detector system, or if the event is incomplete. Such problematic events are discarded.
- **Primary vertex** — Each event is required to contain a PV formed from two or more associated tracks with $p_{\mathrm{T}} > 500\,\mathrm{MeV}$.
- **Trigger** — Events must have passed at least one of the triggers described in section section 7.2.3.
- **Filter** — Events must pass at least one RPVLL filter as described in section 7.2.4.

7.2.3 Trigger

The W boson decay produces a high-p_{T} prompt lepton (one produced close to the IP) which provides an ideal trigger for the analysis. The trigger is responsible for signalling that an entire event is of interest and should be stored (see section 4.2.6 for details). The triggers used to select events in this analysis are the lowest p_{T} unprescaled single muon and single electron triggers from 2015–2018 [157, 158]. The list of HLT triggers is shown in table 7.3, each of which is seeded by a specific L1 trigger not mentioned here. The overall trigger efficiency for the HNL signal is approximately 30–40%. The trigger requirement is ignored in the selection of events for the inverted prompt lepton validation region (section 7.4.4). For that region, any event that passes any trigger and any RPVLL filter is considered.

Table 7.3: Single μ and single e triggers from 2015–2018 are used to collect signal region data. mu20 and mu26 signify that the trigger requires a ≥20 GeV muon and a ≥26 GeV muon, respectively. Similarly, e24 and e26 correspond to a ≥24 GeV and a ≥26 GeV electron.

	Trigger name	2015–2016	2017	2018
μ	HLT_mu20_iloose_L1MU15	×		
	HLT_mu26_ivarmedium	×	×	×
e	HLT_e24_lhmedium_L1EM20VH	×		
	HLT_e26_lhtight_nod0_ivarloose	×	×	×

7.2.4 Filter

As this analysis searches for displaced vertices, it relies on the reconstruction of large-radius tracks (LRTs) (see chapter 6) that can not be run on every event as it is computationally intensive. A dedicated data stream is used so only a subset of data is reconstructed using LRT. A set of filters to select this data subset is defined by several analyses focused on LLP signatures. A initial prompt reconstruction is performed on the entire dataset, and events which pass at least one filter are saved in their detector level (`RAW`) format. This data subset is internally referred to as `DRAW_RPVLL` since it is "derived" from the complete "`RAW`" dataset and was originally used for supersymmetric R-parity violating long-lived (RPVLL) particle analyses. The `DRAW_RPVLL` data is then fully reconstructed including the LRT tracking, and the output is stored in `DAOD_RPVLL` format, analogous to the standard ATLAS `xAOD` format.[25]

Approximately 10% of the full Run 2 dataset successfully passes the complete set of `DRAW` filters.[26] The signal region of the present analysis uses only events that pass one of four custom HNL filters. Each filter requires at least one prompt lepton with $p_T > 28$ GeV and one displaced lepton with $p_T > 5$ GeV. Additional requirements such as quality and isolation are also applied to these objects. A complete set of the filter requirements is shown in table 7.4. The real and simulated data are both subject to these filter requirements, however, a subtle discrepancy arises as discussed in section 7.6.5.

Table 7.4: The four HNL filters included in the `DRAW` filter in the `DRAW_RPVLL` data stream.

<table>
<tr><th></th><th></th><th>Prompt μ displaced μ filter</th><th>Prompt e displaced μ filter</th><th>Prompt e displaced e filter</th><th>Prompt μ displaced e filter</th></tr>
<tr><td>HLT trigger</td><td></td><td>Single μ trig.</td><td colspan="2">Single e trig.</td><td>Single μ trig.</td></tr>
<tr><td rowspan="5">Prompt lepton</td><td>p_T</td><td colspan="4">> 28 GeV</td></tr>
<tr><td>$|\eta|$</td><td colspan="4">< 2.5</td></tr>
<tr><td>Muon type</td><td>Combined</td><td colspan="2">-</td><td>Combined</td></tr>
<tr><td>Electron ID</td><td>-</td><td colspan="2">Loose</td><td>-</td></tr>
<tr><td>Isolation</td><td colspan="4">$\Sigma p_T^{(0.3)}/p_T < 0.05$</td></tr>
<tr><td rowspan="6">Displaced lepton</td><td>p_T</td><td colspan="4">> 5 GeV</td></tr>
<tr><td>$|\eta|$</td><td colspan="4">< 2.5</td></tr>
<tr><td>Muon type</td><td colspan="2">Combined, MS extrapolated</td><td colspan="2">-</td></tr>
<tr><td>Electron ID</td><td colspan="2">-</td><td colspan="2">None</td></tr>
<tr><td>Isolation</td><td colspan="4">$\Sigma p_T^{(0.3)}/p_T < 1.0$</td></tr>
<tr><td>d_0</td><td colspan="2">> 0.1 mm for combined muon with $\chi^2/n_{dof} < 5$</td><td colspan="2">> 1 mm</td></tr>
</table>

7.3 Reconstruction and object definition

This search requires the reconstruction of the prompt lepton expected from the W boson decay and a displaced dilepton vertex in the ID. The reconstruction of displaced tracks is discussed in detail in chapter 6. In this analysis, leptons are reconstructed using both standard and LRT tracks so modifications to their definitions are necessary. The following sections summarize the relevant definitions and details of muon, electron, and DV reconstruction. The invariant mass of the dilepton DV is calculated as well as the mass of the HNL using an estimate of the neutrino trajectory.

[25] A default secondary vertexing (VSI) is also run in this stream, but this analysis instead uses a custom VSI configuration executed in a later derivation.

[26] This is by construction as 10% is what the computing budget allows.

7.3.1 Muons

In this analysis, the only muons used in event selection are of combined type, formed by matching ID and MS muon tracks and refitting as a single particle. A detailed description of muon reconstruction is presented in section 5.3. Additionally, only *Medium* quality muons are considered, and the standard calibrations are applied to account for known discrepancies between data and simulation.

The **prompt muon** is defined as the highest-p_T muon in the event with $|d_0| < 3$ mm and $|(z_0 - z_{PV}) \sin\theta| < 5$ mm that does not overlap with any tracks in the DV. z_0 is the track's longitudinal impact parameter, and z_{PV} is the z coordinate of the PV. The candidate is also required to have a $p_T > 3$ GeV. θ is the polar angle described in section 4.2.1. The `DRAW` filter used to select events in this analysis also requires that each prompt muon has a track-based isolation of $\Sigma p_T^{(0.3)}/p_T < 0.05$ (see section 5.2.3).

A **displaced muon** is defined as a muon object whose associated ID track is part of a DV. The displaced muon is also required to have a $p_T > 3$ GeV. The `DRAW` filter requires that each displaced muon has a track-based isolation of $\Sigma p_T^{(0.3)}/p_T < 1.0$.

7.3.2 Electrons

The reconstruction of electrons combines tracks from the ID and energy deposits from the ECAL. Section 5.2 presents a detailed description of the electron reconstruction process.

Electrons are classified by quality using a likelihood-based identification. This analysis only uses *Medium*, and a customized working point (WP) which is looser than all of the standard WPs. The *LHVeryVeryLoose* WP was designed to accommodate displaced electrons by removing requirements that constrain the track to originate near the IP. The d_0-related requirements are removed, along with requirements on the number of Pixel and SCT hits (however, tracks that are included in the DV must pass their own minimum hit requirements). Since the custom *LHVeryVeryLoose* WP is non-standard and has no collaboration-wide calibrations, the scale factors corresponding to the loosest low-p_T WP available (*LHLooseBLayer* with *Gradient* isolation) are applied to the electron collection.[27] Scale factors similar to the *LHVeryVeryLoose* WP in the $p_T \geq 20$ GeV regime were produced for another analysis and found to be very similar to those for the *LHLooseBLayer* WPs. It was thus assumed that applying the *LHLooseBLayer* scale factors to the *LHVeryVeryLoose* electron collection was appropriate.

The kinematic requirements for an electron to be considered the **prompt electron** are nearly identical to the prompt muon, except that the minimum p_T requirement is 4.5 GeV. The prompt electron is required to pass the *LHMedium* quality WP. The kinematic requirements for a **displaced electron** are similar to those for a displaced muon except that the candidate must have $p_T > 4.5$ GeV. Displaced electrons are required to pass the custom *LHVeryVeryLoose* quality WP.

[27]Producing and validating data/MC calibrations is a substantial undertaking that dedicated groups in ATLAS generate only for widely-used and critical WPs.

7.3.3 Displaced vertices

Vertices from displaced decays are reconstructed using the displaced vertexing algorithm known as `VrtSecInclusive` (VSI) described in the ATLAS public note [136]. This algorithm uses the combined collection of standard and large-radius tracks to determine the position of vertices formed of two or more tracks that originate at a macroscopic distance from the primary vertex. The present analysis uses a customized version of VSI optimized for the displaced two-lepton decays found in the displaced HNL signal models. Several of the optimizations used are inspired by the tuning of the LRT algorithm discussed in chapter 6. For instance, tightening the d_0 and z_0 criteria for the tracks used for vertexing substantially decreased the reconstruction of fake vertices. More details about the definition and optimization of the custom VSI configuration can be found in reference [155].

The main steps of the VSI algorithm are summarized in fig. 7.4. First, tracks are preselected for vertex seeding with a series of quality cuts. For the custom configuration, this track preselection also requires that the tracks are matched to muon or electron objects. After the track selection, there is an initial two-track vertex finding step, which forms vertices with all compatible two-track pairs. Only the pairs of tracks that are deemed loosely compatible are kept. These two-track seed vertices are then combined into n-track vertices using an incompatibility graph method. Then vertices that share multiple tracks are merged together. At this point, all tracks (lepton matched and non-lepton matched tracks) that are *not* initially preselected for vertex seeding are not associated with any DV. The final step in this reconstruction chain attempts to attach all tracks to compatible vertices. The attached tracks are required to meet basic quality checks (p_T, χ^2 cuts) and are required to be nearby the candidate vertex, both in d_0 and z_0 significance.

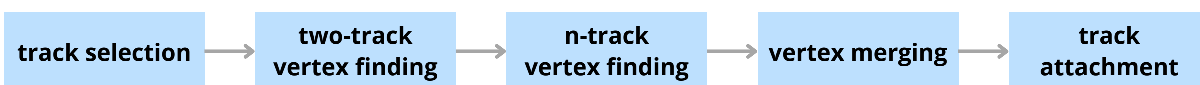


Figure 7.4: The main steps in the VSI vertexing algorithm. A full description of the algorithm can be found in reference [136].

To evaluate the vertex reconstruction efficiency on the signal MC samples, a truth matching procedure is used to determine if LLP decays are reconstructed as DVs. This truth matching procedure follows the same definitions adopted in reference [136]. The total efficiency $\varepsilon_{\text{total}}$ is given by

$$\varepsilon_{\text{total}} = \mathcal{A} \cdot \varepsilon_{\text{seed}} \cdot \varepsilon_{\text{core}} \tag{7.7}$$

where the acceptance $\mathcal{A}$ is the fraction of signal decays with reconstructed tracks in the fiducial region, $\varepsilon_{\text{seed}}$ is the fraction of two-track signal decays that pass selection requirements, and $\varepsilon_{\text{core}}$ is the fraction of signal decays that match a reconstructed vertex.

The vertexing efficiency for three signal types is shown in fig. 7.5 as a function of radial decay position, $r_{\text{DV}} = \sqrt{x^2 + y^2}$. The total efficiency peaks between 0.15–0.25 depending on the signal sample. Total efficiency at very low values of r_{DV} is suppressed in part due to a minimum $|d_0|$ cut of 1 mm for at least one track and at high values of r_{DV} due to requirements on the number of hits per track.

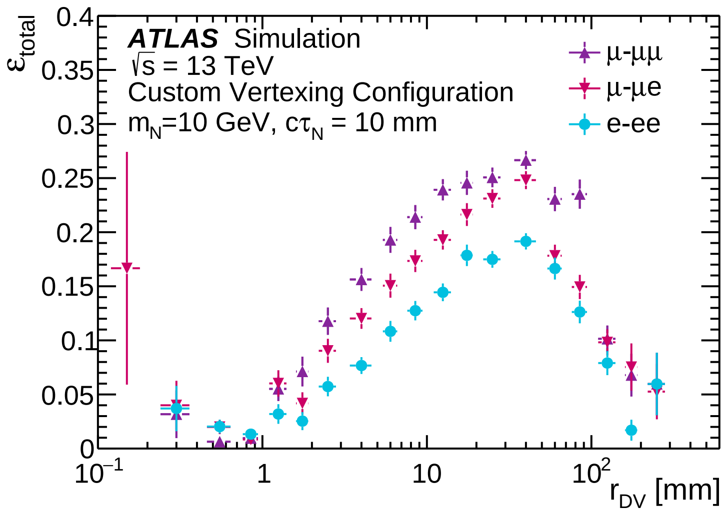


Figure 7.5: The total displaced vertex reconstruction efficiency ($\varepsilon_{\mathrm{total}}$) as a function of displaced vertex radius (r_{DV}) for the custom configuration used in this analysis. The definition of the displaced vertex reconstruction efficiency can be found in reference [136]. The efficiency is shown for $\mu{-}\mu\mu$, $\mu{-}\mu e$, and $\mu{-}ee$ signals with $m_N = 10\,\mathrm{GeV}$ and $c\tau_N = 10\,\mathrm{mm}$.

7.3.4 HNL mass and DV mass

The displaced vertex mass m_{DV} is calculated as the invariant mass of the two tracks in the vertex. The invariant mass of the DV is reconstructed using the track kinematics (p_{T}, η, ϕ) calculated with respect to the DV position and using a pion mass assumption. Since the neutrino is invisible to the detector, the HNL signal is reconstructed using only the two charged leptons in the vertex, thus m_{DV} will always underestimate the true HNL mass. The resulting distribution of the DV mass in signal is a broad distribution that cuts off at the simulated HNL mass, as can be seen in fig. 7.3. A variable cut on m_{DV} is used to suppress backgrounds from metastable meson decays.

By using known quantities of the proposed signal signature to constrain the neutrino four-momentum, the invariant mass of the HNL can be calculated. Reference [155] contains an in-depth overview of this variable's calculation. The calculation applies four-momentum conservation in the W and HNL decays using the kinematics of the charged leptons, the known W mass, and the approximation that the leptons and neutrino are massless. The flight direction of the HNL is given by the vector connecting the PV and DV and constrains two of the three degrees of freedom from the unknown neutrino momentum, while the final degree of freedom is constrained by the W boson mass assumption. This calculation yields a quadratic equation with two solutions, of which the positive-radical solution is used to define the invariant mass of the HNL candidate (m_{HNL}). The resulting distribution of m_{HNL}, as seen in fig. 7.3, nicely peaks around the true HNL mass. m_{HNL} is used as the final discriminating variable to define signal and control regions.

7.4 Analysis design

Through the reconstruction step, nearly all data processing is designed and executed by multiple large teams within the collaboration. The trigger design, reconstruction, calibration, simulation etc., is less of an exploratory exercise, and the collaboration uses stringent coding guidelines and computational resource management practices up to the level of the `xAOD` physics object (in this case, the analogous `DRAW_RPVLL`) format. This centralized effort serves the dozens of ATLAS analyses that can readily access relevant datasets, trusting their provenance. Each analysis team designs and implements a set of algorithms and data structures to perform their specified search or measurement. Recommendations and shared frameworks are common, but the team ultimately determines the approach.

As this analysis is focused on identifying long-lived particles, its structure incorporates several new or unconventional techniques. Though every ATLAS analysis is a complex enterprise, according to reference [5], "Long-lived particle searches are without doubt the hardest analyses to perform, since they often require a complete rethink of how object reconstruction works." A new discriminating variable, clever background estimation techniques, and unique systematic uncertainties arising from displaced tracking are all aspects of this analysis that add to its complexity.

7.4.1 Analysis framework

A multi-step pipeline is used to ultimately produce a calibrated, concise, and correct set of events and observables with which a statistical analysis is performed. Figure 7.6 shows a schematic diagram of the workflow and data structures used. Once passing through the trigger and filter, `RAW` detector signals are reconstructed into representations of physical objects as discussed in chapter 5. Additionally, the relevant signal must be simulated and reconstructed as discussed in section 7.2.1.

The displaced HNL analysis framework, depicted in fig. 7.6, can be divided into eight algorithmic steps:

1. **Trigger** — The trigger system operates in real-time to decide which data is to be recorded by the detector. The system is described in section 4.2.6 and this analysis uses standard single-lepton triggers to select data events section 7.2.3.

2. **DRAW Filter** — The data is required to pass one of the `DRAW` filters (see section 7.2.4).

3. **Reconstruction** — Events that are selected by the `DRAW` filter are then reconstructed using a set of reconstruction algorithms which use LRT and displaced vertexing (VSI). In this step, only the default VSI configurations are run. The output of this reconstruction step is a collection of `DAOD_RPVLL` files in the `xAOD` format. For more details about the reconstruction, see chapter 5.

4. **Derivation** — A derivation step is used to skim away information that is not required for this displaced HNL analysis. This derivation step also reruns VSI to reconstruct the custom lepton-only DVs (section 7.3.3). The output is a derived `xAOD` format called a `DAOD` and the derivation name used for this analysis is `EXOT29`.[28]

[28] Nearly all ATLAS analyses use some form of *derived* data with superfluous information and irrelevant events skimmed away.

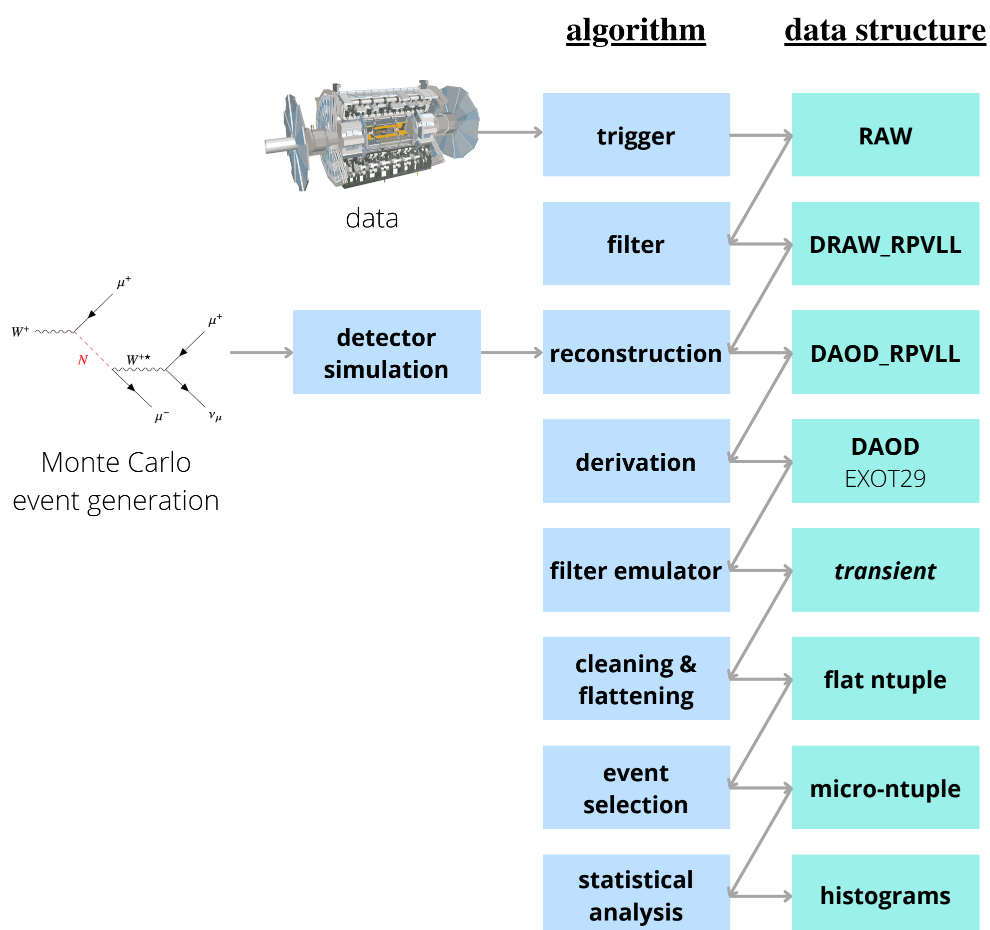


Figure 7.6: A schematic outline of elements of the analysis framework pipeline and successive data structures. Data and Monte Carlo have different entry points but are otherwise processed identically.

5. **Filter emulator** — As no information is stored in data about which filter(s) selected a given event, a filter emulator is run on the secondary-reconstructed data in the ntuple making code. This is used to approximate the event selection applied by the four HNL filters. This is run on both data and MC to ensure that both collections are subject to the same selection criteria.

6. **Cleaning & flattening** — The `DAOD` files are processed using "DHNL Algorithm" producing flat *ntuples*. Ntuples are array-like structures that contain relevant and necessary information about each event, implemented as ROOT `TTrees` [159]. In this analysis, existing lepton object and DV information and engineered quantities are recorded. This is also the step where calibrations are applied to the physics objects, and basic data event cleaning is performed.

7. **Event selection** — Event selection code called "DHNL Ntuple Analysis" consumes the ntuples produced in the previous step and applies the main event selection described in the sections below. This step saves *micro-ntuples* at specified selection levels, which are fed into the statistical analysis framework.

8. **Statistical analysis** — The final step in the analysis pipeline is the statistical analysis of selected signal and background events. The details of the statistical analysis framework, implemented with TRExFitter, are described in section 7.8.

One unique feature about the `DRAW_RPVLL` data stream is that there is not an equivalent `DRAW` filter step for the signal Monte Carlo. The MC is only reconstructed once with LRT enabled vs. data where the filter is applied to events reconstructed using the default reconstruction. This can be seen visually in fig. 7.6 where the signal Monte Carlo only enters the analysis pipeline at the reconstruction step. Furthermore, the data filter uses a preliminary set of detector conditions, resulting in slightly different reconstructed quantities. The differences introduce a systematic uncertainty which is discussed in section 7.6.5.

Code-base for the analysis steps

Several considerations go into selecting a software stack to analyze the data, including processing speed, development speed, language familiarity, and existing tools. As the initial `DAOD_RPVLL` dataset is extremely large, fast execution is vital. However, as the data footprint is reduced, differences in execution time become less consequential. What is important is the analysis team's ability to explore the data using familiar languages with well-supported libraries.

The ntuple production code, called DHNLAlgorithm, is based on the open-source framework xAODAnaHelpers (xAH) [160]. xAH, written in C++, acts as a minimal wrapper for the existing ATLAS EDM processing tools. Individual analyses can extend the classes of xAH to augment the output ntuples. DHNLAlgorithm (and its intermediary DVAnalysisBase) is an extension of xAH that calculates and records non-standard quantities such as the DV invariant mass, the track-to-lepton matching, and the `DRAW` filter decision. The primary execution code is written in C++ but is steered using a Python configuration file. In

To keep track of the dozens of derivation formats, a naming schema is used and `EXOT29` refers to this analysis falling under the *ATLAS Exotics* working group.

this step, fast execution time and existing tools made a C++-based algorithm the natural choice. However, aside from rudimentary data cleaning, very little event selection is performed by DHNLAlgorithm, whose primary function is to produce calibrated, flat ntuples.

The event selection functionality is implemented in the Python-based DHNLNtupleAnalysis. With its emphasis on code readability, Python allows for rapid exploration of datasets, allowing physicists to identify relationships and fine-tune analyses. As Python rapidly grows in popularity for applications in data science [161], a community of developers is providing a Python ecosystem for particle physics data analysis [162]. DHNLNtupleAnalysis uses the *uproot* package to read and process the ntuples produced in the previous step.

The statistical analysis presented in section 7.8 is implemented in TRExFitter — a framework, written in C++, for binned template profile likelihood fits in a HistFactory [163] format. The statistical modelling is expressed using RooFit [164] with additional classes provided by RooStats [165] for interval estimation and hypothesis testing. Inference performed with TRExFitter uses the methods discussed in reference [166].

7.4.2 Preselection

The displaced HNL event selection relies on reconstructing two types of physics objects: a prompt lepton and a DV. The event selection can be broken down into two distinct sections split roughly the same way. First, the event preselection imposes conditions on overall event properties and prompt lepton quantities. The top panel of table 7.5 summarizes the event selection criteria. Criteria are chosen based on their ability to suppress background while maintaining an acceptable signal reconstruction efficiency. In addition to the event cleaning (section 7.2.2), trigger (section 7.2.3) and filter (section 7.2.4), the following criteria are used:

- **Primary vertex** — Consistent with standard ATLAS data selection criteria, each event is required to contain a PV formed from two or more associated tracks with $p_{\mathrm{T}} > 500\,\mathrm{MeV}$. If multiple PVs are reconstructed, the one with the highest total p_{T} is selected.

- **Prompt lepton quality** — Selected events are required to have one prompt lepton (either a muon or an electron depending on the channel) as defined in section 7.3. Prompt muons (electrons) must be of *Medium* (*LHMedium*) quality.

- **Trigger-matched lepton** — A record is maintained of which lepton (or leptons) was responsible for firing the single-lepton trigger, enabling objects in the analysis to be *trigger-matched*. This selection requires that at least one *Medium* muon or *LHMedium* electron is trigger-matched and has $p_{\mathrm{T}} > 27\,\mathrm{GeV}$. The trigger-matched lepton can either be the prompt lepton or one of the two displaced leptons. In approximately 10% of events, leptons from the HNL decay have enough momentum to pass single lepton trigger requirements resulting in a small acceptance to signal events with prompt leptons (from the on-shell W boson decay) whose p_{T} is below the single lepton trigger thresholds.

- **Prompt lepton impact parameters** — The prompt lepton must have impact parameters $|d_0| < 3\,\mathrm{mm}$ and $|(z_0 - z_{\mathrm{PV}}) \sin\theta| < 5\,\mathrm{mm}$.

Table 7.5: Description of the selection criteria used in the signal region. The lepton quality selection *LHMedium* for prompt electrons is defined in reference [116] and *Medium* for muons is defined in reference [167]. The *Displaced* lepton quality requirement for displaced electrons is described in reference [128]. In the Z mass veto, α and β refer to flavours of the prompt and one of the two displaced leptons, respectively.

Selection	Value
Preselection	
Event cleaning	Standard ATLAS event cleaning
Trigger	Pass at least one single muon or electron trigger
Primary vertex	At least one (standard ATLAS selection)
Filter emulator	Pass any HNL filter
Prompt ℓ quality	*Medium* (muons) or *LHMedium* (electrons)
Trigger-matched ℓ	At least one lepton with *Medium* (or *LHMedium*) quality
Prompt ℓ impact parameters	$\lvert d_0\rvert < 3$ mm and $\lvert(z_0 - z_{\text{PV}})\sin\theta\rvert < 0.5$ mm
SR selection	
Cosmic veto	$\sqrt{(\Sigma\eta)^2 + (\pi - \Delta\phi)^2} > 0.05$
Displaced lepton-only vertex	At least one
Number of tracks in DV	2
Fiducial volume	$4 < r_{\text{DV}} < 300$ mm
	$\lvert z_{\text{DV}}\rvert < 300$ mm (only for ee DVs)
DV charge	Opposite-sign (OS) tracks
Prompt + displaced ℓ charge	OS leptons (1SFH in Dirac-limit)
DV type	ee, $e\mu$ or $\mu\mu$ vertex
Displaced ℓ quality	*Medium* (muons), *Displaced* (electrons)
Material veto	Applied for ee DVs only
Metastable particle veto	$\mu\mu$ DVs: $m_{\text{DV}} > 5.5$ GeV
	ee or $e\mu$ DVs: m_{DV}-r_{DV} correlation cut
Z mass veto	$m(\ell^{\pm}_{\alpha}\ell^{\mp}_{\beta}) < 80$ GeV or $m(\ell^{\pm}_{\alpha}\ell^{\mp}_{\beta}) > 100$ GeV, if $\alpha = \beta$
Tri-lepton mass	$40\text{ GeV} < m_{\text{DV}+\ell} < 90$ GeV
HNL mass	$m_{\text{HNL}} < 20$ GeV

7.4.3 Signal region selection

This section describes the variables used to discriminate between signal and background beyond the preselection requirements. The selections are summarized in the lower panel of table 7.5.

- **Cosmic veto** — DVs must also pass a veto that requires the two leptons within the DV to not be back-to-back, defined as $\sqrt{(\Sigma\eta)^2 + (\pi - \Delta\phi)^2} < 0.05$ [127]. This cosmic veto reduces backgrounds from single cosmic-ray muons passing through the ATLAS detector and reconstructed as two back-to-back muons. This cut results in no signal loss.

- **Displaced lepton-only vertex** — Events are required to include a DV that is reconstructed using the custom lepton-only vertex configuration described in section 7.3.3. The DV is required to have **exactly two tracks** matched to displaced leptons. The DV is also required to be within the **fiducial volume** defined by $4\text{ mm} < r_{\text{DV}} < 300$ mm. ee DVs must have a longitudinal impact parameter $z_{\text{DV}} < 300$ mm

- **DV charge** — As HNLs, being neutral, will decay into particles whose total charge is zero, the DV will contain two opposite-sign (OS) leptons. Only OS DVs are selected for the signal region.

- **DV type** — Depending on the flavour of the prompt lepton and the particular model considered, the leptons in the DV can have three different combinations of flavours: $\mu\mu$, $e\mu$ or ee. Six channels can thus be defined: $\mu{-}\mu\mu$, $\mu{-}\mu e$, $e{-}e\mu$, $e{-}ee$, $\mu{-}ee$, $e{-}\mu\mu$, where the first lepton is the prompt lepton and the second two leptons in the channel name correspond to the flavour of the leptons in the DV.

- **Prompt + displaced ℓ charge** — For HNL decays that conserve lepton number, the charge of the prompt and displaced lepton that couples to the HNL must have opposite charges. In the notation of fig. 7.2 (a), ℓ_α and ℓ_β must have opposite charge. Thus, for the models where only LNC decays are possible (Dirac limit), requiring these leptons to have the opposite charge keeps 100% of the signal and should remove approximately half of the expected background events. Experimentally, it is not possible to determine which lepton in the DV couples directly to the HNL unless the lepton flavours in the DV are different and only single-flavour mixing is considered. In such a case, an additional charge selection is applied between the charge of the same-flavour lepton pair in the event. This charge selection can only be applied for $e{-}e\mu$ and $\mu{-}\mu e$ channels in the 1SFH Dirac limit models.

- **Displaced ℓ quality** — Muons in the DV must be of *Medium* quality, and the electrons must pass the customized *LHVeryVeryLoose* WP.

- **Material veto** — DVs can be formed from particle interactions with detector material. This background, primarily leading to ee DVs, is suppressed by disallowing DVs produced in detector regions with substantial material. See section 7.5.1 for details.

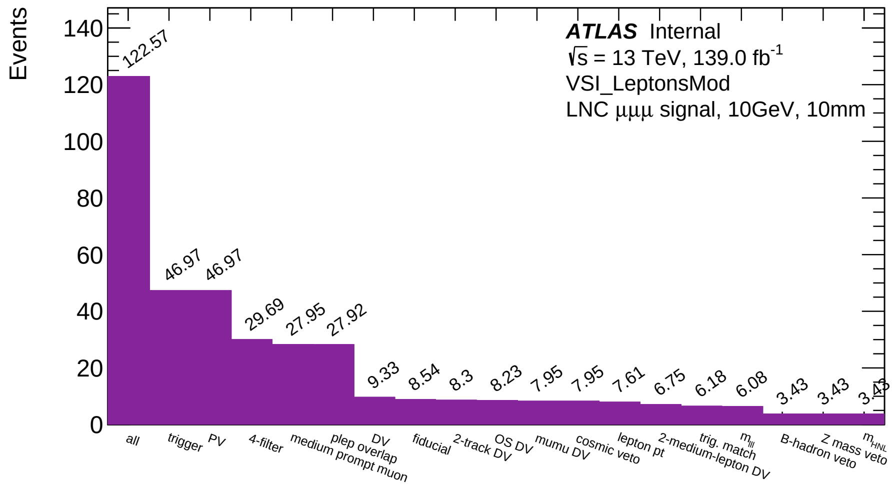
- **Metastable particle veto** — As will be discussed in section 7.5.2, metastable particles produced in LHC pp collisions can have lifetimes beyond 4 mm that contaminate the low-r_{DV} parameter space. Dedicated cuts in the m_{DV}-r_{DV} plane are imposed to remove this background.

- **Z mass veto** — $Z \to \ell\ell$ decays may mimic the displaced HNL signal if one of the leptons from the Z is identified as prompt and the other forms a random crossing vertex with a third, unrelated lepton. If the prompt and displaced leptons have the same flavour and opposite sign, the invariant mass of these two particles is calculated. Events are vetoed if this invariant mass is between 80 and 100 GeV. Further details can be found in section 7.5.3.

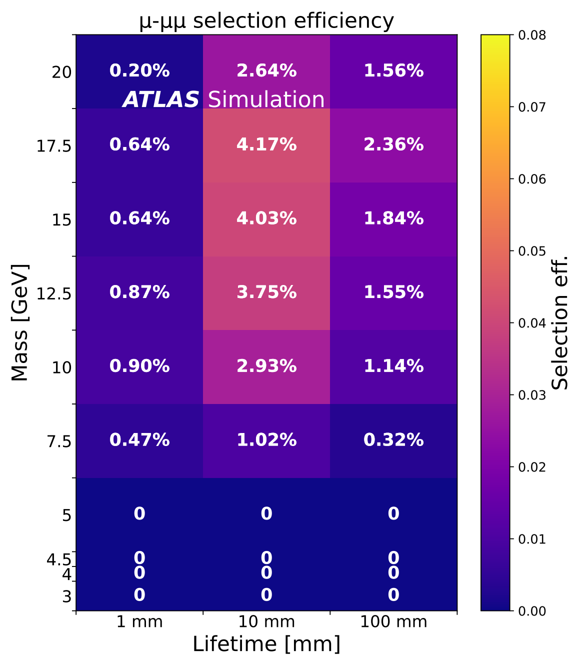
- **Tri-lepton mass** — The tri-lepton mass m_{lll} is the invariant mass of the prompt lepton and the two displaced leptons. Similar to the DV mass, the tri-lepton reconstructs three out of four of the particles from the W boson decay (only the neutrino is missing). As a result, the m_{lll} variable underestimates the true W boson mass.

- **HNL mass (m_{HNL})** — By reconstructing the neutrino four-momentum, the invariant mass of the HNL can be calculated, as discussed in section 7.3.4. This analysis requires that $m_{\mathrm{HNL}} < 20$ GeV as it is not sensitive to higher-mass signals. Distributions of m_{HNL} for signal and background are shown in fig. 7.19.

The selection efficiency (the fraction of generated signal events remaining after applying all selection criteria) ranges from 0% to 4.2% for the m_N–$c\tau_N$ samples considered. The efficiency depends strongly on



(a) cutflow



(b) efficiency

Figure 7.7: (a) HNL signal cutflow for analysis preselection requirements and signal region selection. The cutflow shown corresponds to a $\mu-\mu\mu$ HNL signal with $m_{\mathrm{HNL}} = 10\,\mathrm{GeV}$, $c\tau = 10\,\mathrm{mm}$. (b) The selection efficiency for all m_N–$c\tau_N$ points for $\mu-\mu\mu$ HNL samples [1]. Each bin in (b) reflects the final bin in the corresponding cutflow. The cutflow shown uses an out-of-date weighting scheme, hence the difference in value between the final bin and the efficiency map.

the m_N and $c\tau_N$ of the signal. The highest selection efficiency is for samples with a 10 mm lifetime and m_N between 10 and 20 GeV.

7.4.4 Data blinding strategy

In order to mitigate bias, the data in regions expected to be populated by signal are not observed until the analysis design is complete and approved by other collaboration members. Designing a *blind* analysis in this way relies on comparing simulation to data in final states similar to the signal region (SR) but in which any actual signal is unlikely to be found. *Control regions* (CRs) are defined as regions that are used in the final HNL mass fit but are not the SR. *Validation regions* (VRs) are any other regions that are not directly used in the final fit. This analysis defines three orthogonal regions: one SR, one CR, and one VR.

Control region

The CR is defined using the same SR selections as in table 7.5, except that the HNL mass cut is inverted to require that $20\,\text{GeV} < m_{\text{HNL}} < 50\,\text{GeV}$. Since HNLs with $m_N > 20\,\text{GeV}$ have very short proper decay lengths, they fail the r_{DV} requirements, resulting in negligible signal contamination in the CR. This region is used to study the relative abundance of OS and same-sign (SS) DVs in data similar to signal. The number of observed OS events in the CR constrains the background in the final m_{HNL} fit, as discussed in section 7.8.

Validation region

The VR is used for data-driven background modelling and evaluation of systematic uncertainties. The VR comprises events that passed one of the `DRAW` filters, underwent LRT reconstruction, and do *not* contain a prompt lepton. The DVs in the VR must satisfy the r_{DV} requirements and pass the cosmic-ray muon veto. The detector material veto is also applied to ee DVs. The VR selection is defined in table 7.6.

This VR has higher statistics than the SR and is used to study background sources that produce OS DVs. This region is also used to select a collection of random crossing two-lepton DVs that are used to estimate random crossing backgrounds (see section 7.5 for more details). It is also used in the estimation of the systematic uncertainty in displaced vertexing reconstruction (section 7.6.4). The signal contamination in the VR is found to be negligible.

Table 7.6: Definition of the VR event selection.

Level	Selection	Value
Preselection	Event cleaning	Standard ATLAS event cleaning
	Trigger	Pass any trigger
	Primary vertex	At least one (standard ATLAS selection)
	`DRAW` filter	Pass *any* `DRAW` filter
	Prompt lepton veto	0 prompt leptons
DV selection	Displaced lepton-only vertex	At least one
	Cosmic veto	$\sqrt{(\Sigma\eta)^2 + (\pi - \Delta\phi)^2} > 0.05$
	Number of tracks	2
	Fiducial volume	$4 < r_{\text{DV}} < 300\,\text{mm}$
	DV type	ee, $e\mu$, or $\mu\mu$ vertex

7.5 Backgrounds

All non-signal processes which mimic the signature of the relevant HNL decay are referred to as backgrounds. Backgrounds from five sources are known to produce two-lepton DVs which satisfy the preselection criteria.

1. DVs from particle interactions with detector material.
2. Decays of metastable SM particles.
3. $Z \to \ell\ell$ decays with a random crossing.
4. Cosmic-ray muons.
5. DVs from random crossings of lepton tracks.

The SR selection is designed to retain high signal efficiency and suppress the first four types of background to negligible levels, with random crossing remaining the only background that must be estimated.

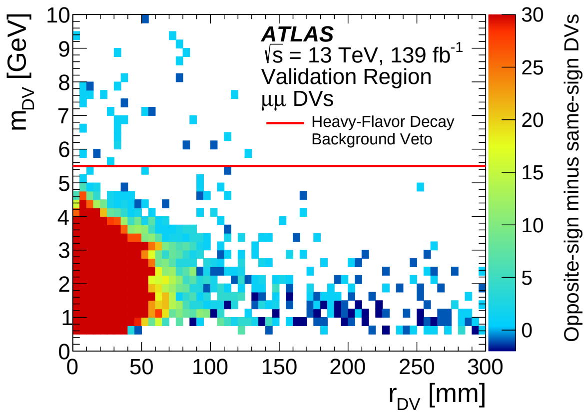
7.5.1 Material interactions

ee DVs have the most background from particle interactions with detector material, both from active layers and support structures. Many of the material interactions that produce ee DVs have a low reconstructed DV mass and are removed by requiring that m_{DV} is at least 2 GeV. However, some material interactions have $m_{\text{DV}} > 2$GeV, and because it is difficult to distinguish between vertices from material interactions and HNL signal decays, a *material veto* is used. The material veto is based on a pair of three-dimensional maps of the ID material [125]. The first map, which extends in radius to $r_{\text{DV}} < 150$ mm, is constructed from low-mass vertices in data (with selections to remove low-mass SM resonances). The second map, which extends in radius from 150 mm $< r_{\text{DV}} <$ 300 mm, is reconstructed using vertices in simulation, to create a more complete map than statistics will allow in data. Both maps extend out to $|z|$ = 300 mm and so the material veto also requires that the DV has $|z| <$ 300 mm.

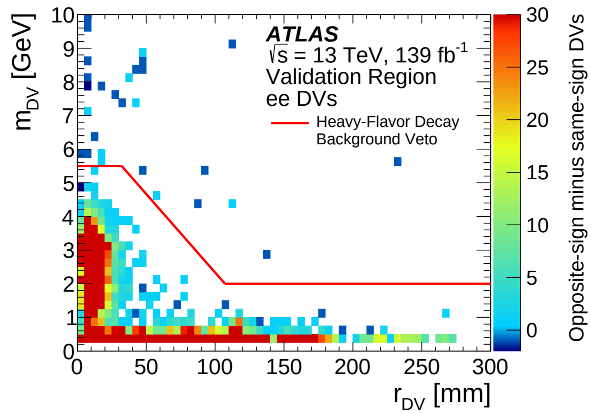
7.5.2 Metastable SM particles

Certain heavy-flavour SM particles produced in LHC pp collisions have long enough lifetimes that some decay within the SR beyond 4 mm. B-hadrons can have decay lengths of a few mm when they are highly boosted or in the tail of the exponential decay distribution. These particles have masses of the order 5–6 GeV, thus encroaching on the signal region. J/ψ mesons have masses centred at 3.1 GeV and can be displaced when other heavy-flavour particles decay to J/ψ.

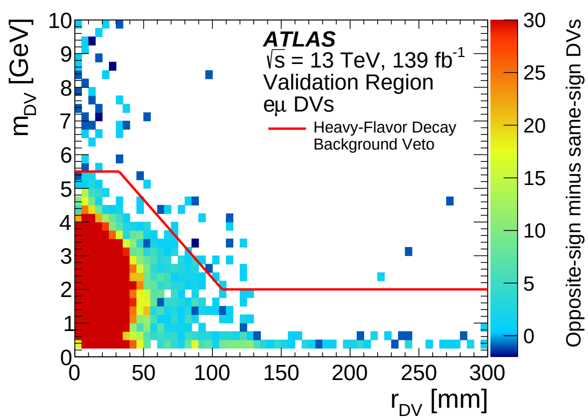
These contributions are removed from the signal region using dedicated cuts on m_{DV} and r_{DV}. Since these backgrounds are not expected to have an associated prompt lepton, cuts are designed by studying the inverted prompt lepton VR and removing regions with more OS DVs than SS DVs. For $\mu\mu$ DVs, $m_{\text{DV}} > 5.5$ GeV is required. The heavy-flavour contamination is smaller for $e\mu$ and ee DVs, allowing looser requirements that exploit correlations between r_{DV} and m_{DV}. The selections that are applied are described in table 7.7 and seen graphically in fig. 7.8.



(a) $\mu\mu$ DVs



(b) ee DVs



(c) $e\mu$ DVs

Figure 7.8: The two-dimensional distribution of m_{DV} and r_{DV}. Shown is the difference between the OS- and SS-vertex distributions in the VR. The z-axis is saturated at 30 to provide a scale that distinguishes between small positive (more OS- than SS-DVs) and small negative (more SS- than OS-DVs) entries. The solid red line represents the DV selection applied in each channel to remove backgrounds from heavy-flavour decays [1].

Table 7.7: Definition of the metastable particle veto.

DV type	Selection
$\mu\mu$	$m_{\mathrm{DV}} > 5.5\,\mathrm{GeV}$
ee or $e\mu$	$m_{\mathrm{DV}} > -\frac{7}{150} r_{\mathrm{DV}} + 7$, if $2 < m_{\mathrm{DV}} < 5.5\,\mathrm{GeV}$ else, $m_{\mathrm{DV}} > 5.5\,\mathrm{GeV}$

7.5.3 $Z \to \ell\ell$ decays with a random crossing

Studies of the data-driven SS background estimate (section 7.5.6) revealed that $Z \to \ell\ell$ decays can contaminate the SR by misreconstruction. Though the Z decay is prompt, the HNL signature can be mimicked if one of the leptons from the Z decay is identified as the prompt lepton and the other forms a DV with a third, unrelated lepton. This can occur for channels where the prompt lepton and at least one of the displaced leptons have the same flavour (i.e., $\mu-\mu\mu$, $\mu-\mu e$, $e-e\mu$, $e-ee$). It is assumed that the contamination would produce similar numbers of OS and SS vertices since the unrelated lepton may be of either charge and cross the path of either prompt lepton. A Z-mass veto is thus applied to remove these types of events from the SR. Events are discarded when the invariant mass of the prompt lepton and the displaced lepton with the same flavor (i.e., $\alpha = \beta$) and opposite charge satisfies $80\,\text{GeV} < m(\ell^{\pm}_{\alpha}\ell^{\mp}_{\beta}) < 100\,\text{GeV}$.[29]

7.5.4 Cosmic ray muons

As discussed in section 7.4.3, cosmic-ray muons, which can be reconstructed as two back-to-back muons in a DV, are rejected by requiring the two displaced tracks to satisfy $\sqrt{(\Sigma\eta)^2 + (\pi - \Delta\phi)^2} > 0.05$ [127].

7.5.5 Random track crossings

Random track crossings (RTCs) occur when two random uncorrelated tracks cross in the inner detector so close that the vertexing algorithm associates them to a common DV. While the probability of a two-lepton RTC is relatively low, the m_{DV} of these vertices is unconstrained, making them difficult to distinguish from the HNL signal. The previous *reducible* backgrounds can be suppressed by judicious selection criteria without substantially impacting the signal sensitivity. Conversely, the RTC background is considered to be *irreducible*, and its contribution to the SR must be estimated.

In this search, a data-driven approach estimates the number of background events from RTCs in the signal region. A full description of the data-driven estimate for RTCs using a prompt lepton shuffling method is described below.

7.5.6 Data-driven estimate

The RTC background is not well modelled by simulation; therefore, a simulated sample cannot be used to estimate the number of events in the SR. Instead, two methods use data to estimate the impact of the backgrounds which were developed and tested on data outside the SR to avoid accidental unbinding.

Same-sign

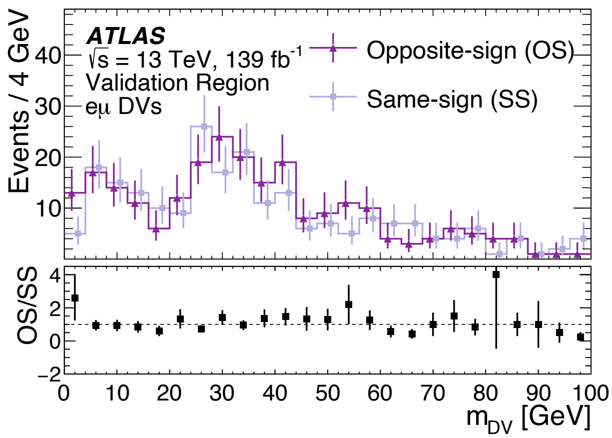
This background estimation method uses the number of random crossings with two same-sign (SS) tracks to estimate the number of random crossings with two opposite-sign (OS) tracks in the signal region. It is assumed that the probability for two tracks to form a random crossing vertex is independent of the charge of the two tracks, and therefore, the number and distributions of OS and SS DVs are similar. Inverting the OS DV requirement of the SR ensures an unbiased estimate with no HNL signal contamination.

This method is statistically limited, and as a result, it is not used as the final background estimate in this analysis. However, the SS region has been used to study the backgrounds and inform selections. For instance, residual contamination from heavy-flavour decays was discovered by comparing the OS and SS

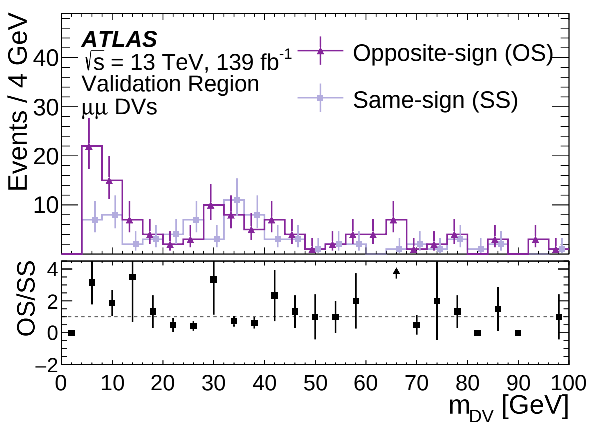
[29]The mass of the Z boson is 91 GeV [21].

vertices in the inverted prompt lepton VR prompting the development of the metastable particle veto (see fig. 7.8).

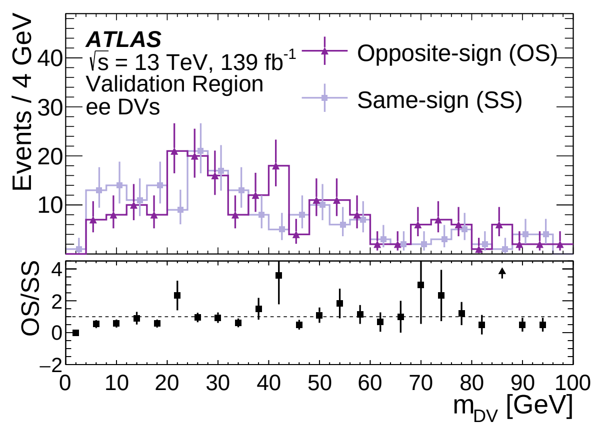
Once suppressing the reducible backgrounds, vertices are compared in the VR to validate the OS≈SS assumption. The resulting distributions are shown fig. 7.9. Agreement between the two distributions is generally good, indicating that the assumption is sound and that other background sources have been suppressed. Any disagreement between SS and OS vertices in the VR is taken into account later by the DV background systematic (section 7.7).



(a) $e\mu$ DVs



(b) $\mu\mu$ DVs



(c) ee DVs

Figure 7.9: The m_{DV} distributions of OS and SS DVs in the validation region, for (a) $e\mu$, (b) $\mu\mu$, and (c) ee DVs. These demonstrate that the SS and OS DV distributions generally agree well, in line with the assumption that the dominant residual background for the search is that of random-crossing vertices, which is then estimated by a data-driven method. The disagreements visible here result partly from the remaining nonrandom-crossing backgrounds, which are accounted for with a systematic uncertainty (see section 7.7) [1].

Prompt lepton shuffling

A method beyond the statistically-limited SS background estimate is required to obtain a smooth distribution of background events across m_{HNL}. The *shuffled event* method creates an artificial collection of events by combining DVs and prompt leptons from separate events. This sample is created by combining each OS DV in the VR with each prompt lepton found in non-VR events that contain a SS DV satisfying loose requirements: $m_{\mathrm{DV}} > 1\,\mathrm{GeV}$, with no lepton identification criteria imposed on its displaced leptons. For each channel, the shuffled sample has at least 2×10^3 times the number of events in the "unshuffled" data sample, in which the DV and the prompt lepton are from the same event. As with an unshuffled event, a shuffled event is classified as $m_{\mathrm{HNL}} < 20\,\mathrm{GeV}$ (SR) or $20\,\mathrm{GeV} < m_{\mathrm{HNL}} < 50\,\mathrm{GeV}$ (CR).

The increase in statistics is limited by the number of prompt leptons and random crossing DVs in the data sample. In order to have a sufficient number of initial events to draw from, the requirements on m_{HNL}, m_{lll}, and the Z-mass veto are removed, and that on m_{DV} is loosened to $m_{\mathrm{DV}} > 2\,\mathrm{GeV}$. The standard selection samples have between 36 and 187 events in each channel, and the samples used in the shuffling with these looser selections are more than 50 times larger. In all analysis regions with three leptons, the shuffled and unshuffled distributions in m_{HNL} and m_{lll} are in good agreement, and the shuffled distribution is then used to estimate the expected number of background events observed in the SR.

The shuffled-event background model assumes the absence of correlations between the randomly crossing tracks and the prompt lepton. The validity of this assumption is checked by comparing the kinematic distributions of shuffled events and unshuffled events in various analysis regions as described below.

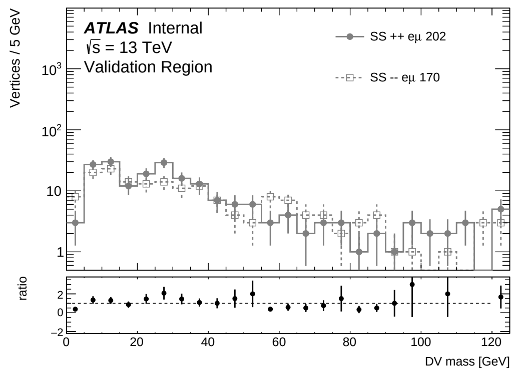
7.5.7 Background estimate validation

Both the SS and prompt lepton shuffling background estimation assume that the background events in the SR are dominated by random crossings of tracks combined with the association of a prompt lepton. A set of tests on the available background samples is performed to validate this claim.

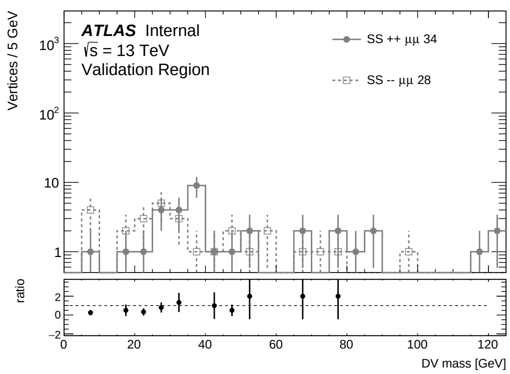
In the VR, it is assumed that DVs with two SS leptons have kinematic distributions similar to OS DVs, except in the low m_{DV} region (below 5.5 GeV) where B-hadron decays dominate. In the VR, OS and SS indeed show good agreement (fig. 7.9).

It is also assumed that there is no dependence on track charge in creating random vertices; therefore, SS DVs in the VR should show no asymmetry between $\ell^+\ell^+$ and $\ell^-\ell^-$ DVs. As shown in fig. 7.10, no significant asymmetry is observed in the SS DVs in the $e\mu$ and ee channels. The statistics in the $\mu\mu$ channel are limited though there is no physics motivation why the conclusions should differ.

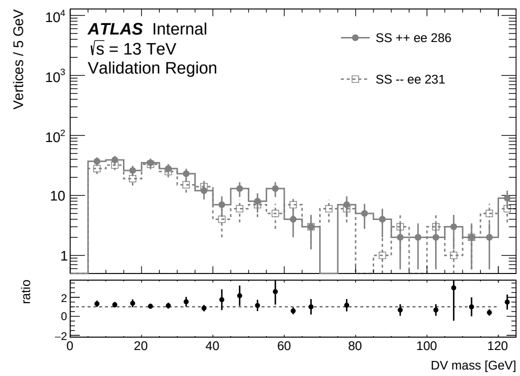
Finally, it is assumed that the pairing of a prompt lepton with a DV is random. Unshuffled SS events from data are compared to the statistics-enhanced shuffled event collection. Comparisons of the m_{HNL} distribution are shown for each channel in fig. 7.11. A Kolmogorov-Smirnov (KS) goodness-of-fit test is performed to compare the shapes of the two distributions. The results of the KS test for both the m_{HNL} and m_{lll} distributions indicate that the original and the shuffled SS events are compatible. This suggests that the pairing of a prompt lepton with an accidental crossing of a two-lepton displaced vertex is indeed random. Altogether, these validations indicate that the SS and the prompt lepton shuffling are valid background



(a) $e\mu$ DVs



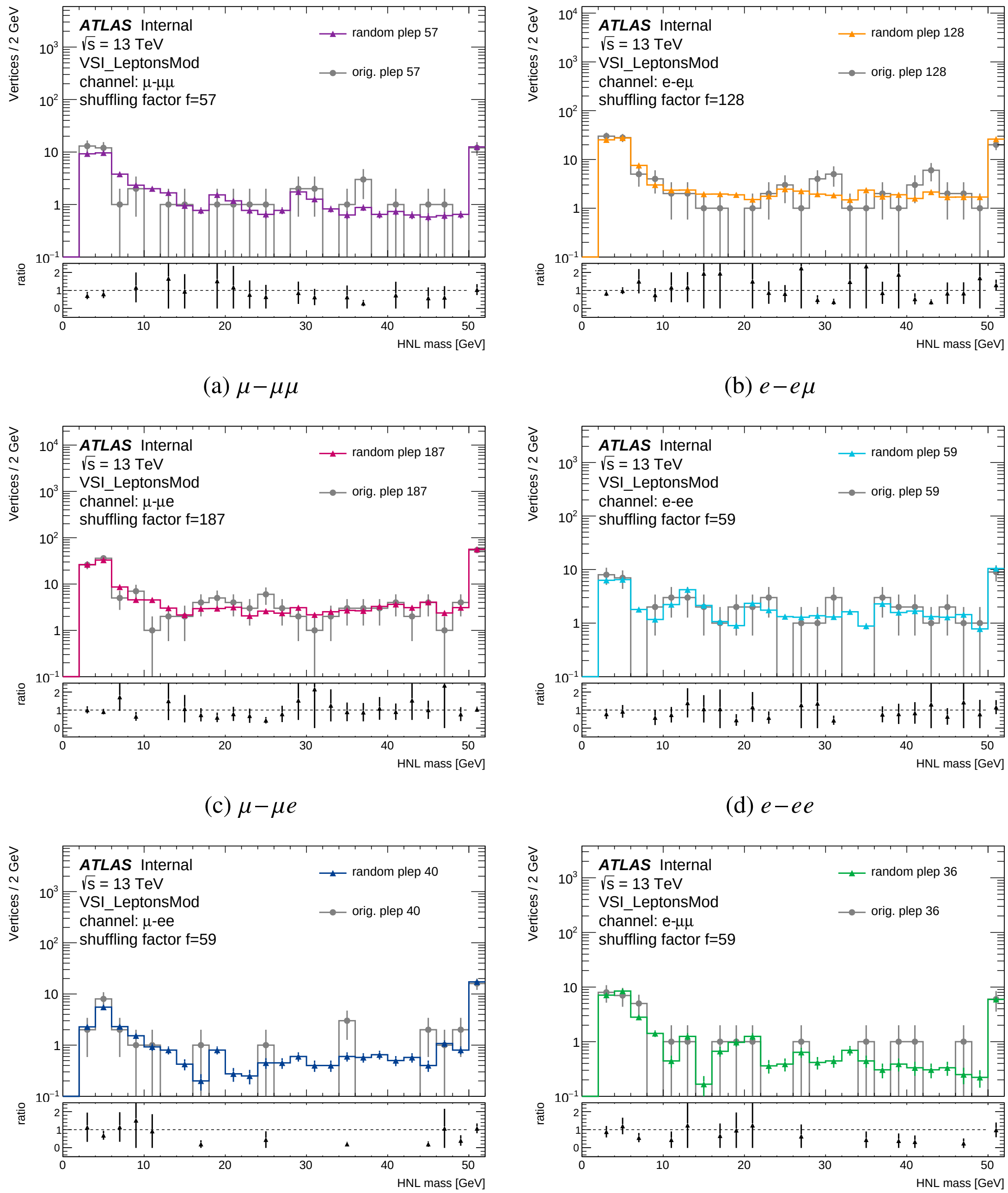
(b) $\mu\mu$ DVs



(c) ee DVs

Figure 7.10: SS DVs in the VR with the metastable particle veto and material veto for ee DVs applied. SS DVs are separated by DV charge in $\ell^+\ell^+$ and $\ell^-\ell^-$ DVs. The event yields are shown in the legend of each plot. No significant differences in the two types of DVs are seen.

approximation methods.



(a) $\mu-\mu\mu$ (b) $e-e\mu$

(c) $\mu-\mu e$ (d) $e-ee$

(e) $\mu-ee$ (f) $e-\mu\mu$

Figure 7.11: m_{HNL} distribution of SS background events, where the DV and prompt lepton are from the same event (gray circles) and the distribution obtained by taking SS DV and randomly shuffling each vertex with a prompt lepton from a different event in the SR (coloured triangles). A KS goodness-of-fit test indicates that the distributions are compatible. The shuffling factor f (ratio of the number of shuffled events to the number of original SS background events) is shown on each plot.

7.6 Signal systematics

The predicted abundance and distribution of signal events are susceptible to systematic uncertainties. Quantifying these uncertainties is essential for performing the necessary fits in the limit-setting procedure. This section describes the uncertainties, or *systematics*, identified in the analysis. Several sources of systematics impact the predicted signal yield, typically arising from differences between data and simulation.

In addition to the standard systematics present in many ATLAS analyses, the non-standard reconstruction techniques introduce several non-standard uncertainties. As this special reconstruction is not used widely, no combined performance groups have produced centrally available uncertainty recommendations. Each systematic estimates the impact of the uncertainty on the final signal selection efficiency. Some are applied as a uniform uncertainty while others vary across one or more variables such as track p_T or η.

Luminosity and pileup systematics account for uncertainties in ATLAS's measurement of the number of pp collisions. The centrally-provided tracking systematic, derived for low-$|d_0|$ tracks, is applied to all tracks, while an additional uncertainty accounts for high-$|d_0|$ tracks and displaced vertices simultaneously. To estimate the latter, a comparison of displaced vertices in MC and data is performed. As displaced dilepton vertices are very rare, displaced $K_S^0 \rightarrow \pi^+\pi^-$ vertices are used as an analog to measure DV reconstruction efficiencies. This is done under the assumption that the VSI algorithm performs similarly regardless of whether it uses tracks matched to hadrons or leptons. An additional lepton identification systematic derived for this analysis accounts for the uncertainty introduced by leptons produced some distance beyond where the standard uncertainties remain valid. The theory and signal modelling systematics account for uncertainties in the cross sections and branching ratios used to scale the signal strength. A summary of the relevant uncertainties is shown in table 7.8.

Table 7.8: Table summarizing systematic uncertainties impacting reconstruction efficiencies. The values shown in the table refer to the impacts on the signal yields for each channel. The total uncertainty is the sum in quadrature of all systematic components (the root of the sum of the squares).

Uncertainty Source	Maximum Selection Efficiency Uncertainty [%]					
Channel	$\mu-\mu\mu$	$\mu-\mu e$	$\mu-ee$	$e-ee$	$e-e\mu$	$e-\mu\mu$
Integrated luminosity	2					
Pileup	3					
Filter emulation	3					
Standard tracking	3					
Displaced tracking and DV	11	21	19	20	28	9
Lepton d_0 extrapolation	5	7	7	7	6	4
Trigger efficiency	< 1	1	< 1	< 1	< 1	< 1
Lepton reconstruction and identification	4	9	12	17	15	2
W cross section and modeling	3					
HNL branching fractions and decay	5					
Total	8 – 33					

7.6.1 Standard systematics

Standard systematics refer to those which are calculated using centralized analysis tools provided by the ATLAS combined performance groups. As these systematics are relevant to dozens of analyses, the combined performance groups regularly provide calibrations and associated systematic uncertainties. They are typically made available as imported C++ libraries.

The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [75]. The uncertainty related to the modelling of the pileup conditions is computed using the Pileup Reweighting (PRW) Tool for Run 2 ATLAS analyses. The systematic uncertainties associated with the lepton calibration, identification, reconstruction, and trigger efficiencies are implemented using the central analysis tools provided by ATLAS. The uncertainties used are those recommended for full Run 2 analyses. Similar to the scale factors discussed above, a study of the systematics of the *LHVeryVeryLoose* and *LHLooseBLayer* WPs indicated the *LHLooseBLayer* uncertainties can be used as an estimate of the *LHVeryVeryLoose* ones used in the electron selection. As there are no scale factors provided without isolation requirements, the *Gradient* WP is chosen as changing the WP changes the systematic negligibly so we assume the uncertainties for isolation-free leptons are the same as those from the gradient WP. Because it is possible for a prompt or displaced lepton to trigger the event, the trigger-matched lepton is required to be *Medium* quality, and the systematic variation used is that associated with that WP.

7.6.2 Theory and signal modelling

The signal strength scales with the W boson production cross section and W decay to a signal lepton flavour. This cross section times branching ratio is obtained from a direct ATLAS measurement in $\sqrt{s} = 13\,\text{TeV}$ pp collisions [147] which quotes a 3% uncertainty. This is applied as a systematic to the signal yields.

There is also an uncertainty on the theoretical branching fraction of the HNL. This analysis uses an updated calculation from reference [148] that includes perturbative QCD corrections not included in MC simulation. A conservative 5% systematic uncertainty is assigned to the HNL branching fraction by comparing the updated BR calculation to existing MC simulations.

7.6.3 Tracking

To determine the impact of track reconstruction uncertainties, a study is performed to measure the impact on the final yield. For each available signal sample, the standard event selection is applied but with a fraction of tracks randomly removed from each event. This "track-dropping probability" is determined from the tracking performance group's central recommendations and varies across p_{T} and η. It is derived using low-$|d_0|$ tracks. The tracking uncertainty map, used as the track-dropping probability, is shown in fig. 7.12. The ntuples produced with and without dropped tracks are processed with the signal selection framework and the selection efficiencies are compared.

The resulting reduction in signal selection efficiency is evaluated at the *DV type* selection[30] (see table 7.5). The difference is symmetrized about the nominal efficiency as one can easily exclude existing tracks but not easily create additional valid tracks. No variation in impact is observed when performing this test on

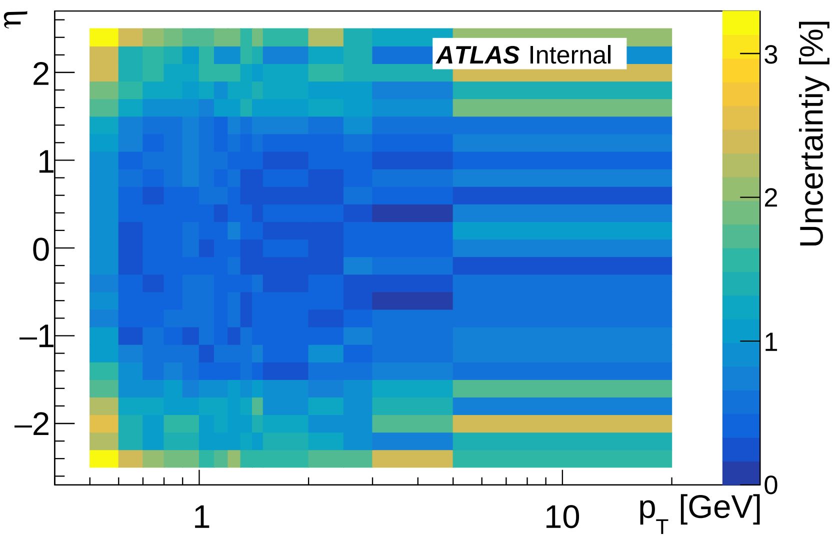


Figure 7.12: The overall tracking uncertainty from the tracking performance group's central recommendations. The value in each bin is used as the probability that a track will be randomly removed when calculating the impact on the signal selection efficiency.

samples with different masses, lifetimes, and decay channels. Figure 7.13 shows the combined reduction in efficiency is 2.6% at the DV type cut. A conservative estimate of 3% is chosen as the total impact of the tracking uncertainty on the final signal yield.

Distributions for the kinematic variables most relevant to the analysis are plotted and justify a flat systematic variation. No shape-dependant variation is seen in the distributions of the m_{DV}, r_{DV}, η, p_{T}, and m_{HNL}. Example distributions can be seen in fig. 7.14, and similar behaviour is seen in the other signal masses, lifetimes, and decay channels.

7.6.4 Displaced track and DV reconstruction

The efficiency of DV reconstruction can vary between data and simulation. To estimate the effect of these differences, $K_S^0 \to \pi^+\pi^-$ vertices are studied. K_S^0 mesons are produced in relative abundance at the LHC and have a relatively long mean lifetime ($c\tau = 26.84$ mm). This allows a data/simulation comparison of a well-known DV with a mean lifetime within the range of the HNL signal samples (1–100 mm).

K_{S}^0 vertices are selected from two-to-two parton scattering (dijet) simulation and data for which large-radius tracking and displaced vertexing have been used in the event reconstruction. The vertexing algorithm used is the same configuration as the custom vertexing described in section 7.3.3 except that it does not require lepton tracks but instead considers all charged tracks. To prevent contamination from the signal region, only vertices from the inverted prompt lepton validation region are used as described in section 7.4.4. The relative abundance of vertices in data and simulation (data/MC) is calculated in bins of p_{T} and production radius,

[30]The impact on the final signal yield is evaluated before the complete selection criteria are applied in order to observe a sufficient number of events and thus obtain a reasonably precise measurement of this uncertainty.

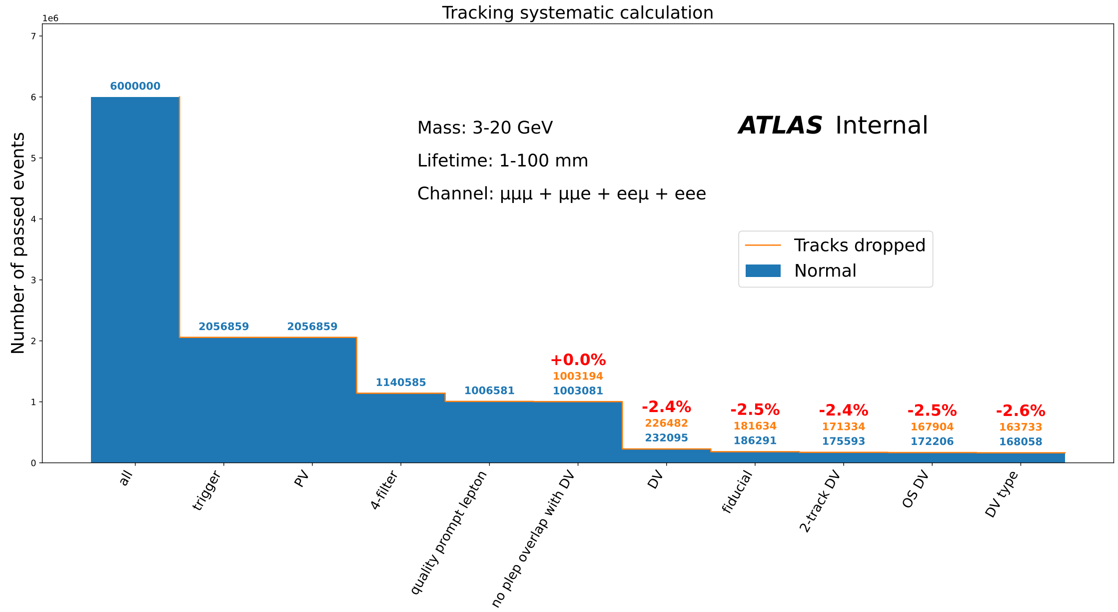
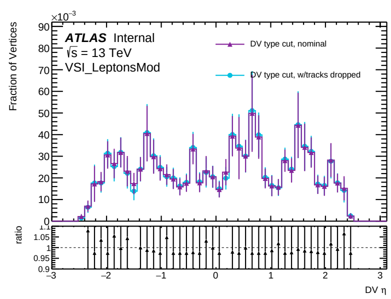
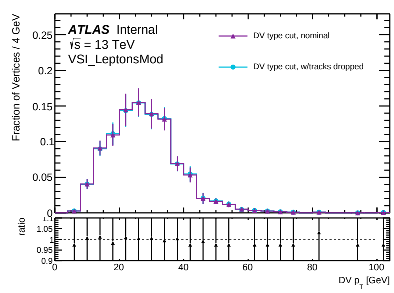


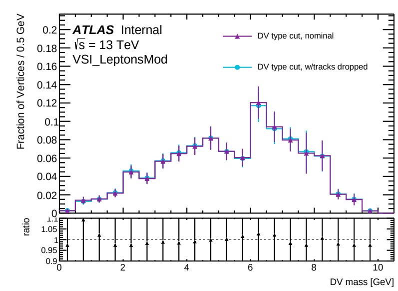
Figure 7.13: This cutflow shows the reduction in the number of vertices when randomly removing a proportion of tracks using the tracking performance group's recommendations. The average reduction in efficiency is 2.6% at the DV type cut.



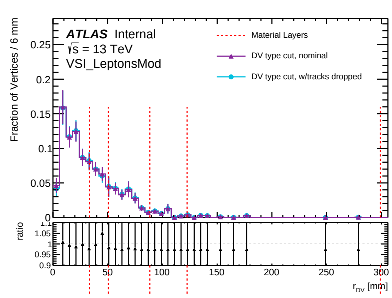
(a) η



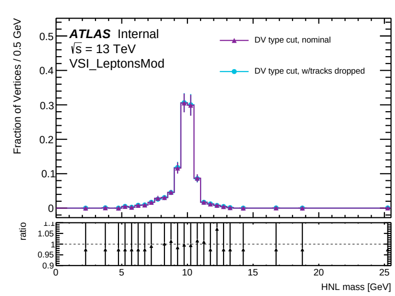
(b) p_{T}



(c) m_{DV}



(d) r_{DV}



(e) m_{HNL}

Figure 7.14: Kinematic distributions of several relevant variables illustrating the effect of randomly removing tracks from the sample. A small reduction in total events is seen, but no clear shape dependence. Sample shown: $\mu-\mu\mu$, $c\tau_N = 10\,\mathrm{mm}$, $m_N = 10\,\mathrm{GeV}$.

r_{DV}, chosen to provide coverage of the HNL signal distributions. These bins are scaled, so the data/MC ratio of vertices produced near the interaction point is unity. This is done under the assumption that the reconstruction of less-displaced vertices is more robust and that prompt tracking uncertainties are quantified by the tracking systematic. The uncertainty is applied to each vertex as a per-event systematic variation. This is shown graphically in fig. 7.15.

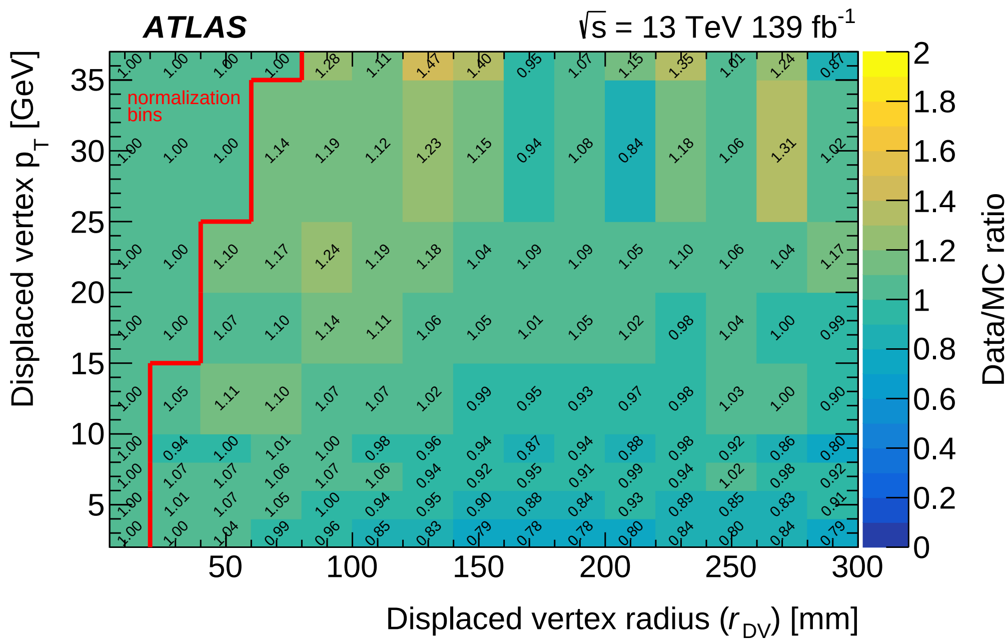


Figure 7.15: The dominant signal uncertainty is due to differences in the reconstruction of displaced vertices and tracks between data and MC produced with Pythia8.186. Only events in the VR are observed. The uncertainty is evaluated by comparing $K^0_S \to \pi^+\pi^-$ event yields in bins of p_{T} and r_{DV}. The data/MC ratio is normalized to the bin nearest the IP, where the tracking and vertex reconstruction algorithms are expected to be most robust. Higher-p_{T} vertices are grouped in larger low-r_{DV} "normalization bins" since fewer displaced vertices populate bins near the IP. The highest-p_{T} bin contains all vertices with $p_{\mathrm{T}} > 35$ GeV. The symmetrized difference from 1.0 is applied to each signal vertex as a per-event systematic variation [1].

7.6.5 Filter emulation efficiency

A `DRAW` filter is applied to data events before they are selected for secondary reconstruction (see section 7.2.4). An emulator is used on both MC and data to apply a filter emulation as close as possible to the real one. Due to differences mentioned below in the objects used by the `DRAW` filter and the emulated filter, there is a mismatch in the signal efficiency. The subtle difference arises from the filtering pipeline, as seen in fig. 7.16. The detector simulation outputs samples which have only undergone the specialized reconstruction, including LRT and with updated conditions. The full collection of tracks, including LRTs, is then used to produce physics objects such as electrons and muons. Each event passes through the filter emulator, which applies the same filter criteria as in the `RAW` → `DRAW_RPVLL` step but on these slightly different objects. On the other hand, real data is first reconstructed using standard physics objects in the `RAW` → `DRAW_RPVLL` step.

Data events passing this filter step continue to the secondary reconstruction step, which crucially *overwrites the standard reconstruction objects*. Data events then also pass through the filter emulator using the final electrons and muons constructed from both LRT and standard tracks. Effectively, data events are filtered twice, and MC events are filtered once, suggesting that the acceptance for data is always strictly less than for MC.

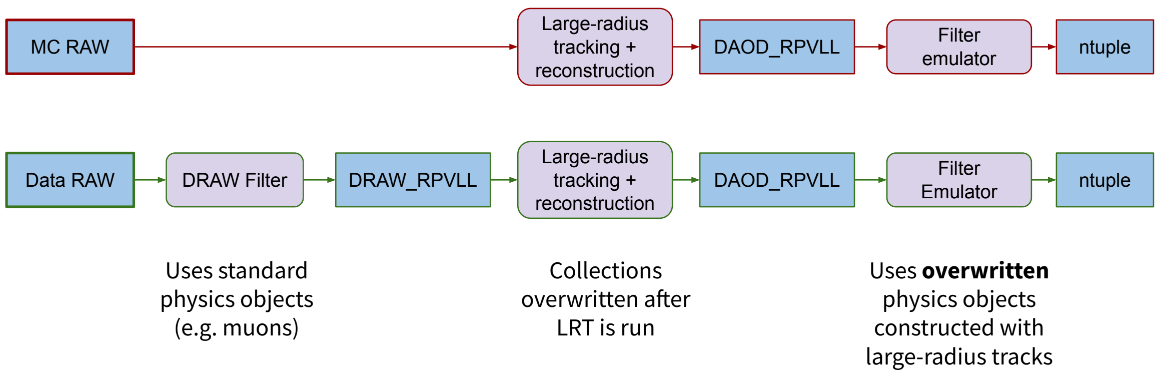


Figure 7.16: Real and simulated data are processed differently in the `DRAW` pipeline. MC has only filter emulator applied to LRT objects. Data has a filter applied to standard objects as well as a filter emulator applied to LRT objects. This introduces a discrepancy in the signal acceptance between data and MC.

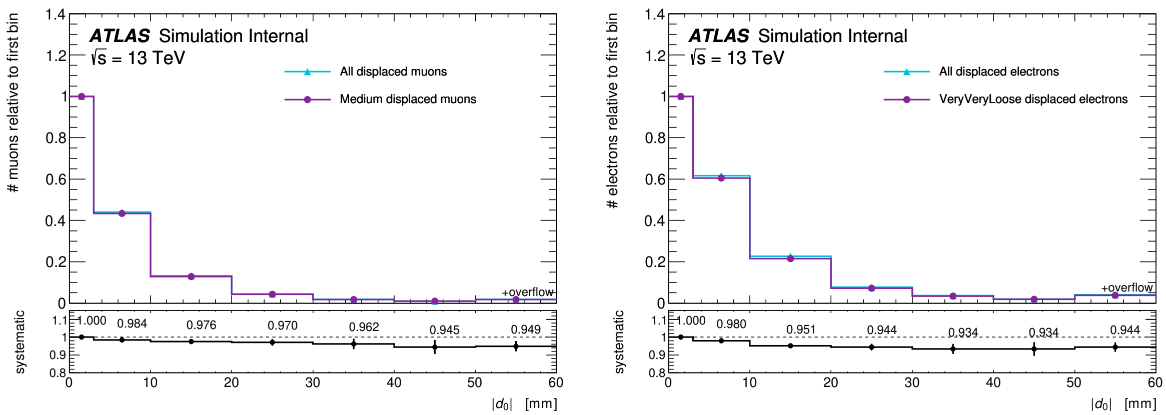
It is possible to estimate the size of this effect on the signal acceptance by using only simulated data. First, a subset of the MC events are reconstructed in two different pipelines, one where LRT is enabled (`DAOD_RPVLL`) and one where it is not (`AOD`). Events are matched between the the two pipelines and final yields are calculated for events which pass both filters (true positive) and those which pass the second but not the first (false positive), estimating the degree of undercounting of data events. A conservative estimate of 3% is applied as a systematic to account for this discrepancy.

7.6.6 Lepton d_0 extrapolation

The use of displaced leptons in secondary vertices introduces an additional uncertainty which is not captured using prompt lepton reconstruction uncertainties. In particular, the lepton identification efficiency is reduced when the lepton has a high impact parameter, $|d_0|$. This is expected to differ in simulation and data, but unfortunately, no pure sample of high-$|d_0|$ leptons are available in data to derive such a systematic. The following study uses only MC samples to estimate the uncertainty from lepton displacement.

To estimate this uncertainty, the effect of applying the lepton ID criteria to HNL signal DVs is explicitly measured. Events from each m_N–$c\tau_N$ point are taken after applying the full preselection, so they contain DVs of the specified lepton type (up to $\mu\mu$, ee, or $e\mu$ DV as shown in the selection cutflows in fig. 7.7). The lepton quality requirements are then applied to each lepton in the DV; *Medium* for muons, and *LHVeryVeryLoose* for electrons.

For each lepton type, the set of all reconstructed leptons is divided by the set of quality-selected leptons in bins of $|d_0|$. These bins are normalized to the 0–3 mm $|d_0|$ bin, where the ID efficiency in data and MC is expected to be most similar. The difference from one is symmetrized and used as the d_0 extrapolation uncertainty. Using the ratio of the histograms, effects from other selection efficiencies such as vertexing and tracking are factored out of this systematic. Figure 7.17 shows the calculated systematics in bins of $|d_0|$.



(a) $\mu-\mu\mu$ (b) $e-ee$

Figure 7.17: The displaced lepton ID systematic is derived by dividing all reconstructed leptons by all leptons that pass quality selection in bins of $|d_0|$. The systematic is calculated in relation to the prompt, 0–3 mm bin. Displaced muon ID (a) is less affected than displaced electron ID (b).

For each signal vertex, the uncertainty for each lepton depends on its type and its $|d_0|$. Since each vertex consists of two leptons, the systematic for each is multiplied, and the product is used as a systematic for that event. In many signal vertices, most leptons have small $|d_0|$; therefore, the lepton d_0 extrapolation uncertainty is often negligible.

The displacement and associated uncertainty can be substantial for signals with lower mass and higher lifetime. For muon channels, the average impact on the selection efficiency is 2.0% and the maximum, 5.5%. For electron channels, the average impact on the selection efficiency is 2.6% and the maximum 6.9%. The muon *Medium* WP definition is the same as is used for prompt muons, so the displacement effect is relatively small. However, electrons are identified using a likelihood trained with d_0 as a discriminating variable. Since the custom electron quality definition used in this analysis, *LHVeryVeryLoose*, has no maximum-d_0 requirement, the d_0 extrapolation uncertainty for electrons can depend quite strongly on displacement.

7.7 Background systematics

Systematic uncertainties in the background model, taken to be fully correlated between the CR and the SR, are evaluated for two sources. The first estimates the uncertainty from the assumption that non-random backgrounds are negligible and is estimated from differences between the m_{HNL} distributions of shuffled events created from SS and OS DVs. This uncertainty varies between 5% for the $e-e\mu$ channel and 79% for the $\mu-\mu\mu$ channel.

The second uncertainty accounts for statistical fluctuations in the m_{HNL} distribution of the shuffled sample due to the finite number of prompt leptons used. The limited number of prompt leptons used in the shuffling will introduce unphysical bumps in the kinematic properties of the shuffled DVs. The uncertainty is estimated from the differences between the m_{HNL} distributions for shuffled events of two types: in type 1 (2), the combined DV and prompt lepton originate from events in identical (different) DV channels (ee, $\mu\mu$, $e\mu$). This uncertainty is largest for the μ–μe channel, reaching 5%. The background systematics for each channel are summarized in table table 7.9. The total uncertainty is the quadrature sum of the two contributions. The uncertainties are treated as uncorrelated among the different channels.

Table 7.9: The relative systematic uncertainties on the background estimate, estimated by creating shuffled events from SS DVs or from prompt leptons taken from events with different DV channels [1].

Channel	Relative Uncertainty (%)		
	DV	Prompt ℓ	Total
$e-ee$	35	<1	35
$\mu-ee$	27	<1	27
$e-e\mu$	2	1	2
$\mu-\mu e$	5	5	7
$e-\mu\mu$	70	2	70
$\mu-\mu\mu$	79	3	79

7.8 Statistical methods and results

After finishing the analysis procedure, estimating the uncertainties and computing the expected sensitivity, then checking our assumptions by unblinding the CR, the SR is unblinded. Each MC signal sample corresponds to specific values of $|U_\alpha|^2$ vs. m_N, for which the efficiency is evaluated. A hypothesis test for a given model is performed at each point with 10^4 pseudoexperiments. The unblinded data shows no excess in any channels above the expected backgrounds. Exclusion limits are set for each theoretical HNL scenario. The statistical method used to interpret the observed data follows the standard ATLAS recommendations and is summarized below. A detailed presentation of the statistical foundations including the profile likelihood ratio and the "modified frequentist" CL_s method can be found in references [166, 168, 169].

7.8.1 Statistical methods

The most powerful test statistic for discrimination between the two hypotheses is the *likelihood ratio* [5, 170]. In the following definitions, the signal strength parameter μ is the parameter of interest (POI) and the systematic uncertainties are encoded as nuisance parameters (NPs) represented with the vector θ. Testing the different values of the signal strength parameter μ is done using a likelihood function $L(\mu, \theta)$.

A binned likelihood is represented as

$$L(\mu, \theta) = \underbrace{\prod_{i \in \text{bins}} \text{Pois}\left(n_i | \mu S(\theta) + B(\theta)\right)}_{\text{observed events}} \cdot \underbrace{\prod_{j \in \text{syst}} \text{Gaus}(\theta_j | \theta_j^0; \sigma_{\theta_j})}_{\text{systematics}} . \tag{7.8}$$

The probability distribution function of the observed events is modelled as a Poisson distribution, and that of the nuisance parameters is modelled as a Gaussian distribution. In this notation, n_i is the observed number of events in bin i, and j NPs represent the systematics. $\mu S(\theta) + B(\theta)$ is the expected number of signal and background events, with the signal yield scaled by the parameter μ, where $\mu = 1$ would represent the expected cross section. θ_j^0 is the central value of systematic uncertainty j, and σ_{θ_j} is its width. The nuisance parameters broaden L reflecting the loss of information about μ due to systematic uncertainties.

The profile likelihood ratio (PLR) is

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}_\mu)}{L(\hat{\mu}, \hat{\theta}_\mu)} \tag{7.9}$$

where hâts represent fitted values. In the numerator, μ is the fixed signal strength (e.g., $\mu = 0$ in the background hypothesis) and $\hat{\hat{\theta}}_\mu$ is the best-fit value of the NPs for fixed μ. In the denominator, the signal strength $\hat{\mu}$ and $\hat{\theta}_\mu$ are fit simultaneously. These values, which maximize each likelihood function, are known as maximum likelihood estimates (MLEs).

In this formulation, λ takes a value between 0 and 1, with higher values representing better agreement between the observed data and the hypothesized value of μ. The next technique is to take the negative logarithm of the likelihood ratio as the test statistic as minimizing the log-likelihood is equivalent to maximizing the likelihood while being easier to compute. The one-sided profile likelihood test statistic used for this analysis is defined as

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} . \tag{7.10}$$

In this definition, higher values of q_μ correspond to increased incompatibility between observed data and hypothesized μ.

The p-value quantifies the level of disagreement:

$$p_\mu = \int_{q_{\text{obs}}}^{\infty} f(q_\mu | \mu) dq_\mu . \tag{7.11}$$

In frequentist statistics, a model is rejected if this value is less than some threshold α (e.g., 95% confidence limits are set using a threshold of $\alpha = 0.05$). The CL_s method aims to reduce the ability to reject a model in the case where an experiment has very little sensitivity to the model. The CL_s quantity is defined as

$$\text{CL}_\text{s} \equiv \frac{p_{S+B}}{1 - p_B} \tag{7.12}$$

and takes the place of the p-value such that the signal+background ($S+B$) model will be rejected if $\mathrm{CL_s} < \alpha$. The reader is directed to reference [169] for a clear description of the $\mathrm{CL_s}$ method's motivations.

For low-background analyses such as this, it is useful to perform the integrals of eq. (7.11) numerically using "toy" techniques. In this analysis, distributions of the test statistic (eq. (7.10)) are created using randomly-generated pseudoexperiments (toys). The background (B) hypothesis is tested by generating toys from the likelihood distribution for $\mu = 0$. The $S + B$ hypothesis is tested by generating toys with a specific value of μ.

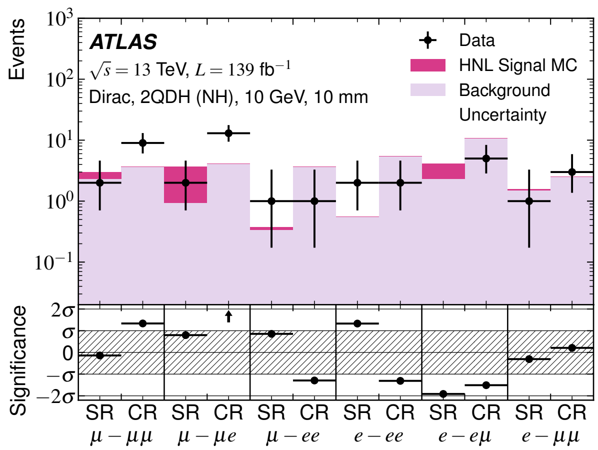
By calculating the p-values for a range of μ, the observed upper limit μ_{up} is determined where $\mathrm{CL_s}(\mu_{\mathrm{up}})$ drops below 0.05. The $\mathrm{CL_s}$ prescription is applied to each model in the m_N–$c\tau_N$ plane for each channel. Those models for which $\mu_{\mathrm{up}} < 1$ are excluded. Limit contours are then obtained by performing a cubic spline interpolation between the generated m_N–$c\tau_N$ points. *Expected limits* are obtained by performing the same procedure but instead setting the observed statistic to be the median of the B distribution. $\pm\sigma$ and $\pm 2\sigma$ bands are also shown by varying the expected value assuming B is Gaussian distributed.

7.8.2 Results

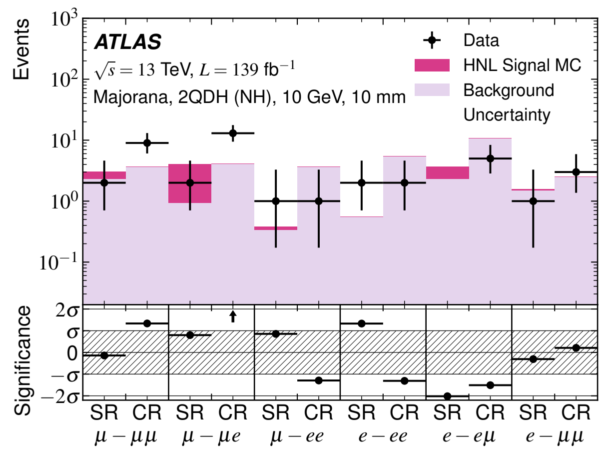
Limits are set at 95% confidence level (CL) on $|U|^2$ vs. m_N for each HNL model, using the $\mathrm{CL_s}$ prescription described above and implemented in the TRExFitter package [163–165]. Histograms of the m_{HNL} distribution in two bins (0–20 GeV, 20–50 GeV) are used as inputs to TRExFitter. The background histograms are constructed using the data-driven shuffled background estimate described in section 7.5.6 and scaled by a separate unconstrained normalization factor for each channel. The background normalization used to generate the toys is taken as that obtained by normalizing the OS shuffled backgrounds to the SS data in the CR. The signal histograms are generated from simulation and are assigned a single floating signal strength μ for all relevant channels, weighted according to each model (section 7.2.1). The inputs to the fit are the OS-event yields observed in the SR and CR. The inclusion of the CR in the fit constrains the predicted background yield in the SR.

Table 7.10 shows the post-fit estimated and observed yields in the SR and CR for all channels. These yields are consistent with the estimated backgrounds, indicating no excess signal. The observed yields in the CR are consistent with the CR background estimates. The pre-fit and post-fit yields are shown in fig. 7.18. All systematic uncertainties are included in the fit by using nuisance parameters, whose post-fit values do not show any significant pull or constraint. To visualize the distribution of predicted and observed yields, a more finely-binned display of the unblinded data is shown in fig. 7.19.

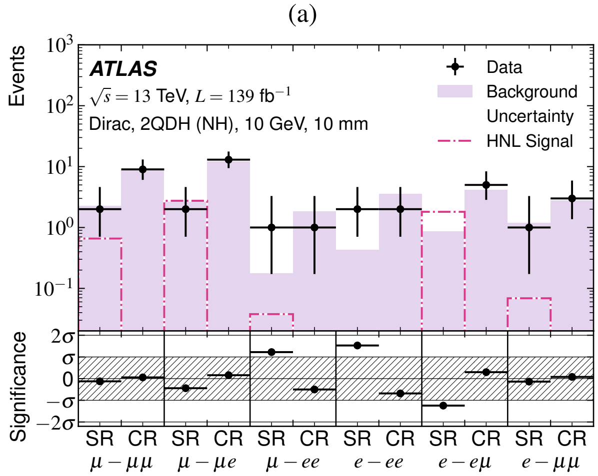
The various channels are combined in different ways to set limits for each of the four models presented in section 7.1. The Majorana limit and the Dirac limit of each model are treated separately. The observed 95% CL contour plots are shown in fig. 7.20 for all eight interpretations as well as the observed and expected limits for the 2QDH NH model with one- and two-sigma uncertainty bands. For a signal at a given m_N–$c\tau_N$ point in the plane, the expected event yield depends on the cross section and branching ratio predicted by each model, causing the limit contours of each to be different.



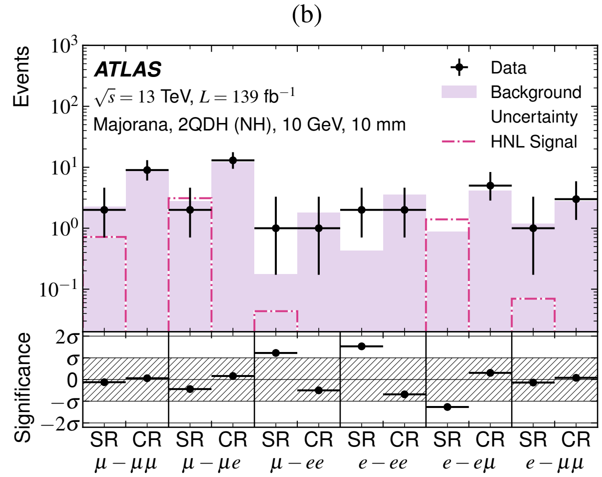
(a)



(b)



(c)

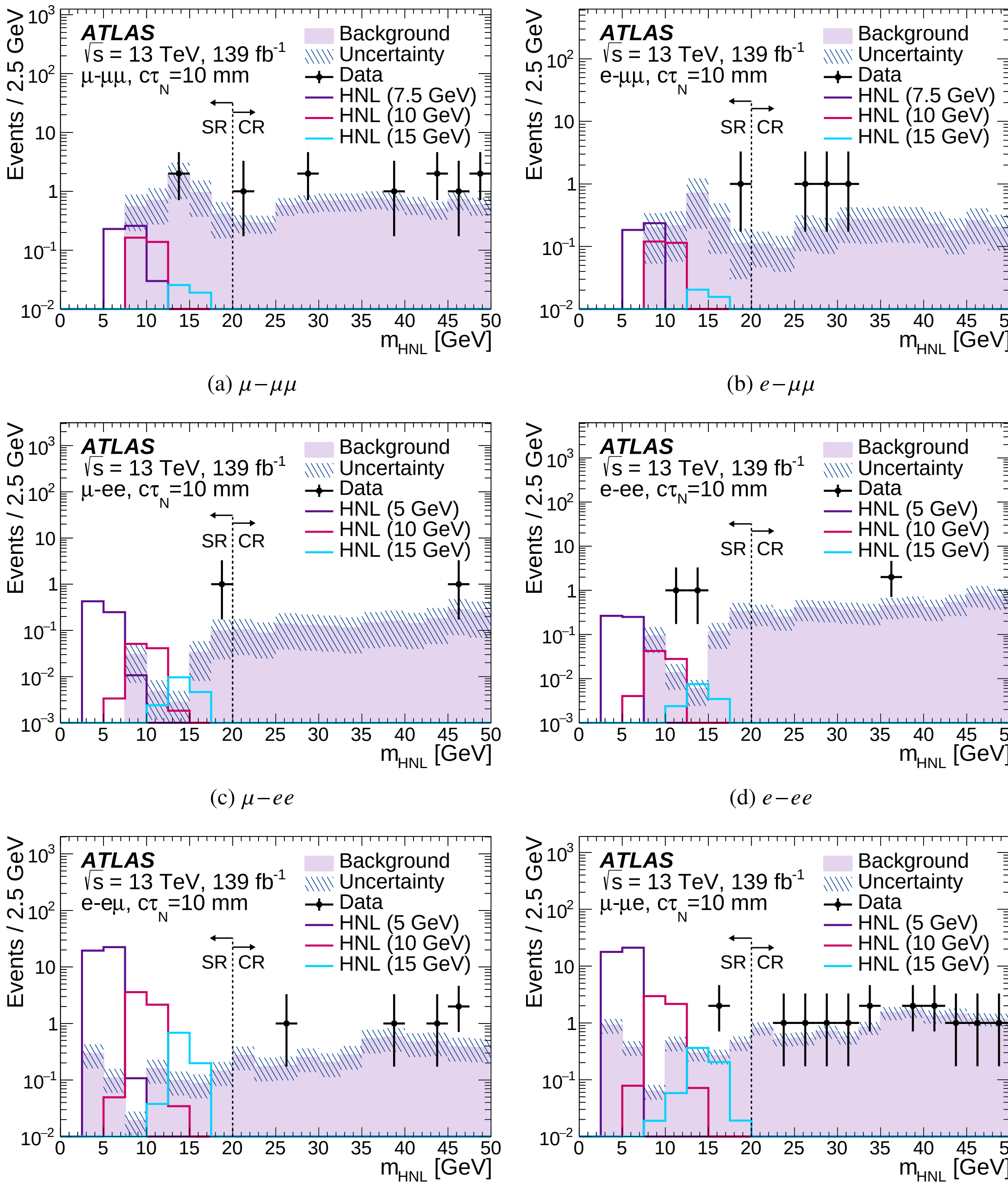


(d)

Figure 7.18: Summary plots for the 2QDH (NH) (two quasi degenerate HNLs, normal neutrino mass hierarchy mixing, $x_e = 0.06, x_\mu = 0.48, x_\tau = 0.46$) model in the Dirac-limit (a,c) and Majorana-limit (b,d). Shown are pre-fit plots (a,b) with background and HNL signal MC stacked, and post-fit plots (c,d) in which the pre-fit signal contribution is overlaid as a dashed line (post-fit signal contribution is zero). Figure from [1] supplemental material.

Table 7.10: Yields of estimated post-fit background events and of observed events in the signal and control regions for all final states. The background yield uncertainties include both systematic and statistical uncertainties. The last two rows show the 1SFH Dirac-limit, LNC configuration $\ell_\alpha^\pm$–$\ell_\alpha^\mp\ell_\gamma^\pm$ [1].

Channel	Signal region		Control region	
	Background	Observed	Background	Observed
$e-ee$	0.4 ± 0.3	2	3.6 ± 1.8	2
$\mu-ee$	0.2 ± 0.1	1	1.8 ± 1.3	1
$e-e\mu$	0.9 ± 0.4	0	4.1 ± 1.9	5
$\mu-\mu e$	2.8 ± 0.8	2	12.2 ± 3.2	13
$e-\mu\mu$	1.2 ± 0.9	1	2.8 ± 1.6	3
$\mu-\mu\mu$	2.2 ± 1.4	2	8.7 ± 2.9	9
$e^\pm$–$e^\mp\mu^\pm$	0.6 ± 0.3	0	2.4 ± 1.4	3
$\mu^\pm$–$\mu^\mp e^\pm$	1.9 ± 0.6	0	8.1 ± 2.6	10

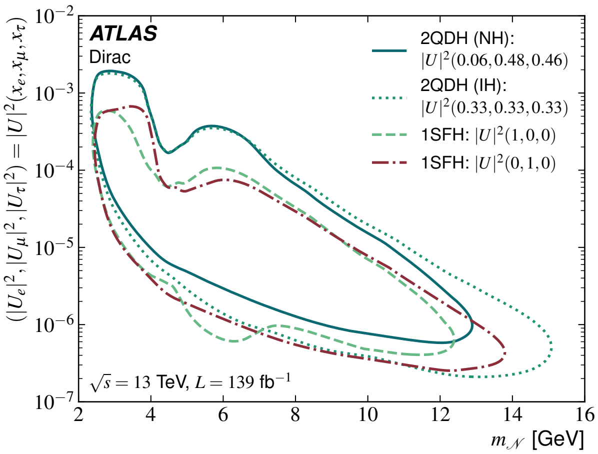


(a) $\mu-\mu\mu$ (b) $e-\mu\mu$

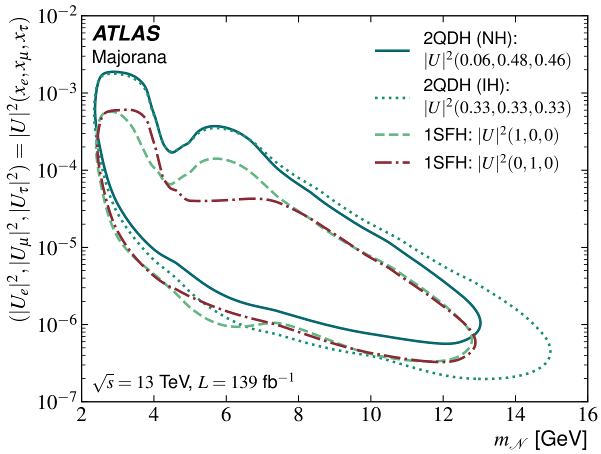
(c) $\mu-ee$ (d) $e-ee$

(e) $e-e\mu$ (f) $\mu-\mu e$

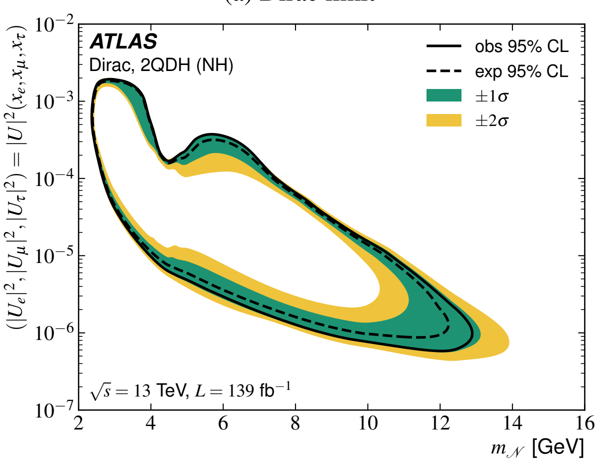
Figure 7.19: The m_{HNL} distribution in the signal (SR) and control (CR) regions for the observed data, the shuffled-event-model background normalized by the two-bin fit described in the text with its uncertainty, and simulated signals with a proper decay length $c\tau_N$ = 10 mm [1].



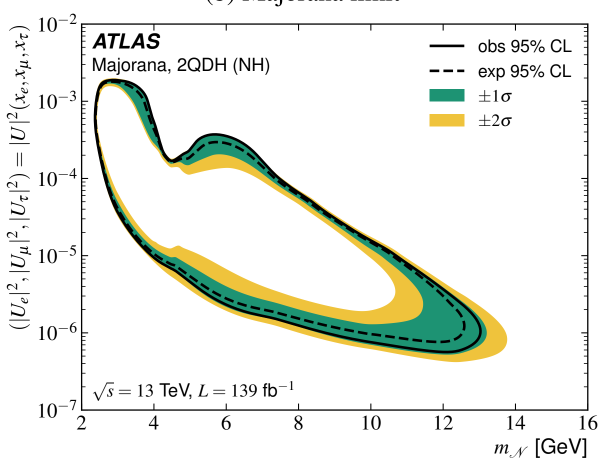
(a) Dirac limit



(b) Majorana limit



(c) Dirac limit



(d) Majorana limit

Figure 7.20: Summary of the observed 95% CL contours for the different HNL models in Dirac limit (a) and Majorana limit (b). The observed and expected 95% CL contours are shown for the 2QDH (NH) model in the Dirac (c) and Majorana (d) limits. The green and yellow bands show the one and two standard deviation (σ) spreads for the expected limits [1].

7.9 Discussion

The limits obtained by this search supersede those obtained in the previous ATLAS search for HNLs [122]. The previous search, which reported results for both prompt and displaced HNL decays, only analyzed the displaced muon decays using the $|U_\mu|^2$ mixing model. That search excluded models with mixing angles between $10^{-3} < |U_\mu|^2 < 10^{-6}$ for masses $4\,\text{GeV} < m_N < 10\,\text{GeV}$. The present search extends the $|U_\mu|^2$ sensitivity by nearly an order of magnitude to 10^{-7} in some mass ranges and substantially extends the HNL mass reach in both directions, setting limits on $2.5\,\text{GeV} < m_N < 14\,\text{GeV}$. This search set limits on the $|U_e|^2$ mixing model using displaced decays for the first time in an ATLAS publication.

Three particularly novel techniques are used in this analysis:

- The interpretation of the measured data is performed for the first time in any experiment using realistic HNL mixing models as well as the traditional single-flavour benchmark models.
- The reconstruction of the HNL mass by way of four-momentum conservation presented a new variable with which to bin the yields.
- The background from random track crossings is estimated using an event shuffling method.

The unusual shape of the exclusion limits is caused by the following. According to eq. (7.6), high-mass HNLs have a very short proper decay length, and the decay vertices are likely to fail the minimum r_{DV} requirement, limiting sensitivity beyond $m_N > 14\,\text{GeV}$. At $m_N < 2\,\text{GeV}$, the proper decay length is so long that too few HNLs decay within the ID to provide sufficient signal sensitivity. The "pinched" shape between 4–6 GeV is due to the strict cuts imposed to exclude backgrounds from SM metastable particle decays.

The 95% confidence levels are primarily limited in their sensitivity by low event yields and associated Poisson statistical uncertainties. There is very little change between the limits gained from the complete analysis and results obtained using only the statistical uncertainties. Fortunately, this suggests that Run 3 and high-luminosity LHC datasets will continue to improve the analysis sensitivity as more data is collected.

In addition to more data, further refinements to the analysis could be made. For example, even after statistical enhancement, the background estimation is limited by large systematic uncertainties. A better background estimate and more recorded data could allow the use of a more finely-binned m_{HNL} discriminant. The vertex reconstruction uncertainty, the largest of the signal systematic uncertainties, could be substantially reduced with better modelling of the detector response to displaced tracks and vertices (e.g., improved material maps and detector conditions). Future searches could improve signal sensitivity by including semi-leptonic final states — an HNL process in which the virtual W boson decays into a pair of hadrons instead of leptons.

Perhaps the largest improvement to the analysis procedure will be made by running the optimized LRT tracking in the default reconstruction stream. As discussed in section 6.7, optimizing the LRT algorithm improves the vertex reconstruction efficiency and allows the entire recorded dataset to be processed, as opposed to the 10% accessed using the `DRAW` filters. Not only would this increase the general signal sensitivity (nearly doubling the acceptance), but it would also allow for the improved modelling of background processes. For example, instead of the metastable particle veto described in section 7.5.2, mesons decaying at large

r_{DV} could be modelled and explicitly included in the background estimate. This would, in turn, increase the signal sensitivity, particularly at low m_{DV}, removing the pinched feature in the exclusion limit contours.

A recent publication from the CMS collaboration places limits on only the 1SFH mixing scenario using $138\,\mathrm{fb}^{-1}$ of $\sqrt{s} = 13\,\mathrm{TeV}$ pp collision data with results in those channels comparable with ATLAS [171]. CMS, however, uses a different track reconstruction workflow which effectively does not distinguish between standard and displaced tracks. This enables the modelling of SM backgrounds, which improves the signal sensitivity by allowing looser final selection. A notable consequence is in the low mass range limits set by CMS. The contours are not closed as they are in ATLAS, and the pinched feature is absent. The sensitivity is also slightly increased at high mass ranges extending the exclusion limits by approximately 1 GeV, presumably due again to looser selection criteria.

Chapter 8

Conclusion

Now, we've basically got it all worked out, except for small stuff, big stuff, hot stuff, cold stuff, fast stuff, heavy stuff, dark stuff, turbulence, and the concept of time.

–Zach Weinersmith, *Science: Abridged Beyond the Point of Usefulness*

The Standard Model of particle physics is incredibly successful at describing nearly all observed physical phenomena. Still, some conspicuous holes are present in our understanding, suggesting that physics beyond the Standard Model is yet to be discovered. The mystery of neutrino masses and oscillations is one such gap that requires imaginative theory and careful experimentation to address. Many theories that address these gaps predict metastable particles, which exist for finite periods before decaying into known matter. These phenomena are sought at colliders like the LHC and observed with detectors like ATLAS, where specialized algorithms are used to identify long-lived particles.

The search for Heavy Neutral Leptons (HNLs) presented in this dissertation used 139 fb^{-1} of pp collision data gathered by the ATLAS detector at the LHC. No excess of signal events was observed in the region of phase space explored. Exclusion limits are set for models with HNLs with mixing angles between 10^{-3}–10^{-7} and masses between 2.5–14 GeV, extending the limits set by the previous ATLAS analysis. These limits are comparable to, and consistent with, the recent CMS search for displaced HNLs. For the first time, limits are presented for realistic HNL mixing models motivated by neutrino flavour oscillation data.

Finding no new phenomena is typical of such a search in high-energy physics, yet with every null result we continue to accrue knowledge about our universe, which in turn helps theorists refine their imaginations with real-world measurements. This displaced vertex search complements other experiments searching for HNLs by probing a region of masses and coupling strengths yet to be investigated. Long-lived particle analyses are challenging as many standard assumptions and algorithms used within the collaboration must be altered.

The work in this dissertation is centred around the use and performance of ATLAS's large-radius tracking (LRT) algorithm, the use of which was crucial in the search for displaced HNL vertices among other searches. However, in conducting the analysis, it became clear that the configuration of the LRT algorithm was producing excessive numbers of fake tracks and limiting the signal sensitivity. The recognition that such a limitation impacted many current and future analyses prompted us to undertake a campaign to improve the use of LRT in future studies.

The refinements made to the LRT algorithm will deliver considerable benefits. By drastically reducing the rate of fake tracks at the expense of a small number of signal tracks, the LRT signal sensitivity will increase by 400% for future analyses. A DNN discriminant was developed, which, if enabled, can further enhance this improvement. By increasing the LRT track purity, the displaced vertexing algorithm produces higher-quality signal vertices and far fewer fake vertices.

Perhaps most importantly, the reduction in the computational cost of the LRT algorithm has allowed for its inclusion in the standard reconstruction pipeline. It can hardly be overstated how much this change will simplify future LLP analyses. Until now, ATLAS LLP searches that use tracking information were restricted to analyzing 10% of the complete dataset by selecting events with filters which required substantial development time and introduced discontinuities in the kinematic distributions. Including LRT in the default reconstruction removes the need for filters, making continuous prompt-to-displaced analyses possible, allowing analyses to develop well-informed control and validation regions, and amplifying signal sensitivity.

The next iteration of the displaced HNL analysis will use the improved data pipeline enabling the modelling of backgrounds from metastable particle decays, improving the sensitivity to low-mass HNLs. Because the low number of signal events limits the sensitivity of the present analysis, the reach of the HNL search will benefit from the inclusion of data from Run 3 of the LHC. Moreover, the analysis software pipeline will be substantially simplified, reducing development time, complexity, and chances for human error.

This dissertation comes to a close at a notable moment in the ATLAS collaboration. Two significant events are taking place the same week that these words are written. On July 4th 2022, CERN was alive with celebrations of the tenth anniversary of the Higgs boson discovery — a massive achievement in physics and a guiding force in the design of the LHC. The following day, the first $\sqrt{s} = 13.6$ TeV collisions of protons were recorded in Run 3, breaking the previous record and further advancing the energy frontier. These events mark the laudable achievements of the LHC, the extraordinary vision and capacity of CERN, and most of all, the curiosity and ambition of scientists who continue to push the bounds of knowledge beyond what was ever considered possible.

The end.

Bibliography

[1] ATLAS Collaboration, *Search for heavy neutral leptons in decays of W bosons using a dilepton displaced vertex in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, (2022), arXiv: 2204.11988 [hep-ex] (cit. on pp. vii, 62, 71, 74, 88, 91, 93, 101, 104, 107–110).

[2] ATLAS Collaboration, *Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC*, Eur. Phys. J. C **79** (2019) 375, arXiv: 1808.07858 [hep-ex] (cit. on p. vii).

[3] P. Langacker, The Standard Model and Beyond, Taylor & Francis, 2017, ISBN: 978-1-315-17062-6 (cit. on pp. 3, 23).

[4] C. P. Burgess and G. D. Moore, The Standard Model: A Primer, Cambridge University Press, 2006, ISBN: 978-0-511-81969-8 (cit. on pp. 3, 18).

[5] A. Buckley, C. White, and M. White, Practical Collider Physics, IOP Publishing, 2021, ISBN: 978-0-7503-2444-1 (cit. on pp. 3, 64, 71, 82, 104).

[6] L. Boyle, Standard model of particle physics, most complete diagram, 2019, URL: https://commons.wikimedia.org/wiki/File:Standard_Model_Of_Particle_Physics-Most_Complete_Diagram.png (cit. on p. 4).

[7] J. Riebesell, TikZ diagrams on Physics and Machine Learning, 2021, URL: https://tikz.netlify.app/higgs-potential (cit. on p. 9).

[8] S. Coleman, *The 1979 Nobel Prize in Physics*, Science **206** (1979) 1290 (cit. on p. 10).

[9] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191, ed. by J.-P. Hsu and D. Fine (cit. on p. 10).

[10] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652 (cit. on p. 10).

[11] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285 (cit. on p. 10).

[12] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, ed. by J. C. Taylor (cit. on p. 10).

[13] ATLAS Collaboration, *Standard Model Summary Plots February 2022*, tech. rep., 2022, URL: https://cds.cern.ch/record/2804061 (cit. on pp. 10, 11).

[14] J. F. Donoghue, *The effective field theory treatment of quantum gravity*, AIP Conf. Proc. **1483** (2012) 73, ed. by W. A. Rodrigues, R. Kerner, G. Pires, and C. Pinheiro, arXiv: 1209.3511 [gr-qc] (cit. on p. 10).

[15] F. Vissani, *Do experiments suggest a hierarchy problem?*, Phys. Rev. D **57** (1998) 7027, arXiv: hep-ph/9709409 (cit. on p. 10).

[16] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rept. **405** (2005) 279, arXiv: hep-ph/0404175 (cit. on p. 11).

[17] Planck Collaboration, *Planck 2013 results. I. Overview of products and scientific results*, Astron. Astrophys. **571** (2014) A1, arXiv: 1303.5062 [astro-ph.CO] (cit. on p. 11).

[18] G. Bertone and T. Tait M. P., *A new era in the search for dark matter*, Nature **562** (2018) 51, arXiv: 1810.01668 [astro-ph.CO] (cit. on p. 11).

[19] A. Pilaftsis, *The Little Review on Leptogenesis*, J. Phys. Conf. Ser. **171** (2009) 012017, ed. by J. Bernabeu, F. J. Botella, N. E. Mavromatos, and V. A. Mitsou, arXiv: 0904.1182 [hep-ph] (cit. on p. 12).

[20] M. Fukugita and T. Yanagida, *Baryogenesis Without Grand Unification*, Phys. Lett. B **174** (1986) 45 (cit. on pp. 12, 24).

[21] Particle Data Group Collaboration, *Review of Particle Physics*, PTEP **2020** (2020) 083C01 (cit. on pp. 13, 17, 18, 23, 92).

[22] P. Hernandez, "Neutrino Physics", 8th CERN–Latin-American School of High-Energy Physics, 2016 85, arXiv: 1708.01046 [hep-ph] (cit. on pp. 13, 14, 18, 21–23).

[23] A. de Gouvea, "TASI lectures on neutrino physics", 2004 197, arXiv: hep-ph/0411274 (cit. on p. 13).

[24] M. Thomson, Modern Particle Physics, Cambridge University Press, 2013, ISBN: 978-1-107-03426-6 (cit. on pp. 13, 15, 18, 19, 21, 52).

[25] C. Jensen, Controversy and Consensus: Nuclear Beta Decay 1911-1934, Birkhäuser Basel, 2000, ISBN: 978-3-0348-8444-0 (cit. on p. 13).

[26] J. Chadwick, *Intensitätsverteilung im magnetischen Spectrum der β-Strahlen von radium B + C*, Verhandl. Dtsc. Phys. Ges. **16** (1914) 383, URL: https://cds.cern.ch/record/262756 (cit. on p. 13).

[27] P. Gombás, *Collected Scientific Papers by Wolfgang Pauli*, Acta Physica Academiae Scientiarum Hungaricae **23** (1967) 251, ISSN: 0001-6705 (cit. on p. 14).

[28] L. M. Brown, *The idea of the neutrino*, Phys. Today **31N9** (1978) 23 (cit. on p. 14).

[29] E. Fermi, *Trends to a Theory of beta Radiation. (In Italian)*, Nuovo Cim. **11** (1934) 1 (cit. on p. 14).

[30] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, *Detection of the Free Neutrino: a Confirmation*, Science **124** (1956) 103 (cit. on p. 14).

[31] F. Reines and C. Cowan, *The Reines-Cowan experiments: Detecting the Poltergeist*, Los Alamos Sci. **25** (1997) 4 (cit. on p. 14).

[32] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Zh. Eksp. Teor. Fiz. **34** (1957) 247 (cit. on p. 14).

[33] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870 (cit. on pp. 14, 17).

[34] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, Zh. Eksp. Teor. Fiz. **53** (1967) 1717, URL: `http://www.jetp.ras.ru/cgi-bin/e/index/e/26/5/p984?a=list` (cit. on p. 14).

[35] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv: `hep-ex/9807003` (cit. on pp. 14, 15).

[36] B. T. Cleveland et al., *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, Astrophys. J. **496** (1998) 505 (cit. on p. 15).

[37] SAGE Collaboration, *Results from SAGE*, Phys. Lett. B **328** (1994) 234 (cit. on p. 15).

[38] GALLEX Collaboration, *GALLEX solar neutrino observations: Results for GALLEX IV*, Phys. Lett. B **447** (1999) 127 (cit. on p. 15).

[39] Super-Kamiokande Collaboration, *Measurement of the solar neutrino energy spectrum using neutrino electron scattering*, Phys. Rev. Lett. **82** (1999) 2430, arXiv: `hep-ex/9812011` (cit. on p. 15).

[40] SNO Collaboration, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, arXiv: `nucl-ex/0204008` (cit. on p. 15).

[41] NUSEX Collaboration, *Experimental study of atmospheric neutrino flux in the NUSEX experiment*, EPL **8** (1989) 611 (cit. on p. 15).

[42] Frejus Collaboration, *Study of Atmospheric Neutrino Interactions with the Frejus Detector*, Phys. Lett. B **227** (1989) 489 (cit. on p. 15).

[43] W. W. M. Allison et al., *Measurement of the atmospheric neutrino flavor composition in Soudan-2*, Phys. Lett. B **391** (1997) 491, arXiv: `hep-ex/9611007` (cit. on p. 15).

[44] MACRO Collaboration, *Measurement of the atmospheric neutrino induced upgoing muon flux using MACRO*, Phys. Lett. B **434** (1998) 451, arXiv: `hep-ex/9807005` (cit. on p. 15).

[45] Kamiokande-II Collaboration, *Observation of a small atmospheric muon-neutrino / electron-neutrino ratio in Kamiokande*, Phys. Lett. B **280** (1992) 146 (cit. on p. 15).

[46] D. Casper et al., *Measurement of atmospheric neutrino composition with IMB-3*, Phys. Rev. Lett. **66** (1991) 2561 (cit. on p. 15).

[47] Daya Bay, *Observation of electron-antineutrino disappearance at Daya Bay*, Phys. Rev. Lett. **108** (2012) 171803, arXiv: `1203.1669 [hep-ex]` (cit. on p. 16).

[48] RENO Collaboration, *Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment*, Phys. Rev. Lett. **108** (2012) 191802, arXiv: `1204.0626 [hep-ex]` (cit. on p. 16).

[49] MINOS Collaboration, *Measurement of Neutrino Oscillations with the MINOS Detectors in the NuMI Beam*, Phys. Rev. Lett. **101** (2008) 131802, arXiv: 0806.2237 [hep-ex] (cit. on p. 16).

[50] X. Qian and P. Vogel, *Neutrino Mass Hierarchy*, Prog. Part. Nucl. Phys. **83** (2015) 1, arXiv: 1505.01891 [hep-ex] (cit. on p. 18).

[51] S.-F. Ge, P. Pasquini, M. Tortola, and J. W. F. Valle, *Measuring the leptonic CP phase in neutrino oscillations with nonunitary mixing*, Phys. Rev. D **95** (2017) 033005, arXiv: 1605.01670 [hep-ph] (cit. on p. 18).

[52] F. Bonnet, M. Hirsch, T. Ota, and W. Winter, *Systematic study of the d=5 Weinberg operator at one-loop order*, JHEP **07** (2012) 153, arXiv: 1204.5862 [hep-ph] (cit. on p. 23).

[53] A. Boyarsky, O. Ruchayskiy, and M. Shaposhnikov, *The Role of sterile neutrinos in cosmology and astrophysics*, Ann. Rev. Nucl. Part. Sci. **59** (2009) 191, arXiv: 0901.0011 [hep-ph] (cit. on p. 24).

[54] A. Blondell, P. Charitos, and R. Jacobsson, Hunting for right-handed neutrinos: the new game in town, en, 2017, URL: https://ep-news.web.cern.ch/content/hunting-right-handed-neutrinos-new-game-town (cit. on p. 24).

[55] M. Shaposhnikov, *Baryogenesis*, J. Phys. Conf. Ser. **171** (2009) 012005 (cit. on p. 24).

[56] S. Alekhin et al., *A facility to Search for Hidden Particles at the CERN SPS: the SHiP physics case*, Rept. Prog. Phys. **79** (2016) 124201, arXiv: 1504.04855 [hep-ph] (cit. on pp. 24, 25).

[57] D. A. Bryman and R. Shrock, *Constraints on Sterile Neutrinos in the MeV to GeV Mass Range*, Phys. Rev. D **100** (2019) 073011, arXiv: 1909.11198 [hep-ph] (cit. on p. 24).

[58] J. Beacham et al., *Physics Beyond Colliders at CERN: Beyond the Standard Model Working Group Report*, J. Phys. G **47** (2020) 010501, arXiv: 1901.09966 [hep-ex] (cit. on pp. 24, 26).

[59] F. F. Deppisch, P. S. Bhupal Dev, and A. Pilaftsis, *Neutrinos and Collider Physics*, New J. Phys. **17** (2015) 075019, arXiv: 1502.06541 [hep-ph] (cit. on pp. 24, 25).

[60] D. Gorbunov and M. Shaposhnikov, *How to find neutral leptons of the νMSM?*, JHEP **10** (2007) 015, [Erratum: JHEP 11, 101 (2013)], arXiv: 0705.1729 [hep-ph] (cit. on p. 24).

[61] T. Asaka, S. Blanchet, and M. Shaposhnikov, *The νMSM, dark matter and neutrino masses*, Phys. Lett. B **631** (2005) 151, arXiv: hep-ph/0503065 [hep-ph] (cit. on p. 25).

[62] T. Asaka and M. Shaposhnikov, *The νMSM, dark matter and baryon asymmetry of the universe*, Phys. Lett. B **620** (2005) 17, arXiv: hep-ph/0505013 [hep-ph] (cit. on p. 25).

[63] DELPHI Collaboration, *Search for neutral heavy leptons produced in Z decays*, Z. Phys. C **74** (1997) 57, [Erratum: Z. Phys. C **75**, 580 (1997)] (cit. on p. 25).

[64] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001 (cit. on pp. 27, 28).

[65] L. Evans, The Large Hadron Collider: A Marvel of Technology, Epfl Press, 2009, ISBN: 978-1-4398-0401-8 (cit. on p. 27).

[66] R. Scrivens et al., *Overview of the Status and Developments on Primary Ion Sources at CERN*, Conf. Proc. C **110904** (2011) 3474, ed. by C. Petit-Jean-Genaz, URL: https://cds.cern.ch/record/1382102 (cit. on p. 28).

[67] F. Van der Veken, *How often is the LHC H2 source replaced, and who replaces it?*, 2016, URL: https://www.quora.com/How-often-is-the-LHC-H2-source-replaced-and-who-replaces-it (cit. on p. 28).

[68] R. Scrivens and M. Vretenar, *Linac2: The tale of a billion-trillion protons*, CERN Courier **58** (2018) 27, URL: https://cds.cern.ch/record/2654723 (cit. on p. 28).

[69] A. Abada et al., *FCC-hh: The Hadron Collider: Future Circular Collider Conceptual Design Report Volume 3*, Eur. Phys. J. ST **228** (2019) 755 (cit. on p. 28).

[70] A. Hoecker, "Physics at the LHC Run-2 and Beyond", 2016 European School of High-Energy Physics, 2016, arXiv: 1611.07864 [hep-ex] (cit. on p. 29).

[71] M. Rini, *Good Omens for Run 3*, APS Physics **15** (2022) 104 (cit. on p. 30).

[72] Restarting the LHC: Why 13 Tev?, 2014, URL: https://cds.cern.ch/record/1998739 (cit. on p. 30).

[73] W. Herr and B. Muratori, "Concept of luminosity", CERN Accelerator School and DESY Zeuthen, 2003 361, URL: http://doc.cern.ch/yellowrep/2006/2006-002/p361.pdf (cit. on p. 30).

[74] Inverse femtobarn | CERN writing guidelines, URL: https://writing-guidelines.web.cern.ch/entries/inverse-femtobarn.html (cit. on p. 30).

[75] 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 (cit. on pp. 30, 98).

[76] ATLAS Collaboration, *LuminosityPublicResultsRun2 Atlas Public TWiki*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (cit. on pp. 30, 31).

[77] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 30).

[78] G. Soyez, *Pileup mitigation at the LHC: A theorists view*, Phys. Rept. **803** (2019) 1, arXiv: 1801.09721 [hep-ph] (cit. on p. 31).

[79] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on p. 31).

[80] J. Wenninger, *Operation and Configuration of the LHC in Run 2*, tech. rep., CERN-ACC-NOTE-2019-0007, 2019, URL: https://cds.cern.ch/record/2668326 (cit. on p. 32).

[81] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 32–34, 36, 38–40).

[82] ATLAS Outreach, *ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC*, (2010), URL: https://cds.cern.ch/record/1457044 (cit. on p. 32).

[83] ATLAS Collaboration, *Study of the material of the ATLAS inner detector for Run 2 of the LHC*, JINST **12** (2017) P12009, arXiv: 1707.02826 [hep-ex] (cit. on pp. 33, 35).

[84] B. Mindur, *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*, Nucl. Instrum. Meth. A **845** (2017) 257, ed. by G. Badurek et al. (cit. on p. 35).

[85] ATLAS Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, (1996), URL: https://cds.cern.ch/record/331061 (cit. on p. 36).

[86] C. Grupen and I. Buvat, eds., Handbook of Particle Detection and Imaging, Vol. 1 and Vol. 2, Springer, 2012, ISBN: 978-3-642-13271-1 (cit. on p. 36).

[87] ATLAS Collaboration, *ATLAS tile calorimeter: Technical design report*, tech. rep., 1996, URL: https://cds.cern.ch/record/331062 (cit. on p. 37).

[88] ATLAS Collaboration, *ATLAS muon spectrometer: Technical design report*, tech. rep., 1997, URL: https://cds.cern.ch/record/331068 (cit. on p. 37).

[89] S. Palestini, *The muon spectrometer of the ATLAS experiment*, Nucl. Phys. B Proc. Suppl. **125** (2003) 337, ed. by F. L. Navarria, M. Paganoni, and P. G. Pelfer (cit. on p. 37).

[90] F. Bauer et al., *Construction and Test of MDT Chambers for the ATLAS Muon Spectrometer*, Nucl. Instrum. Meth. A **461** (2001) 17, ed. by G. Batignani, F. Cervelli, G. Chiarelli, and A. Scribano, arXiv: 1604.02000 [physics.ins-det] (cit. on p. 37).

[91] T. Argyropoulos et al., *Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance*, IEEE Trans. Nucl. Sci. **56** (2009) 1568 (cit. on p. 37).

[92] G. Aielli et al., *The RPC first level muon trigger in the barrel of the ATLAS experiment*, Nucl. Phys. B Proc. Suppl. **158** (2006) 11, ed. by S. K. Park, S. P. Ratti, and R. Santonico (cit. on p. 37).

[93] K. Nagai, *Thin gap chambers in ATLAS*, Nucl. Instrum. Meth. A **384** (1996) 219, ed. by F. Ferroni and P. Schlein (cit. on p. 37).

[94] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv: 2007.12539 [hep-ex] (cit. on p. 38).

[95] A. Buckley et al., *Implementation of the ATLAS Run 2 event data model*, J. Phys. Conf. Ser. **664** (2015) 072045 (cit. on p. 41).

[96] ATLAS Collaboration, Athena, 2021, URL: https://zenodo.org/record/4772550 (cit. on p. 41).

[97] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATL-SOFT-PUB-2021-001, 2021, URL: https://cds.cern.ch/record/2767187 (cit. on p. 41).

[98] 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] (cit. on p. 41).

[99] J. Pequenao and P. Schaffner, "How ATLAS detects particles: diagram of particle paths in the detector", 2013, URL: https://cds.cern.ch/record/1505342 (cit. on p. 42).

[100] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector in Run-2*, Eur. Phys. J. C **80** (2020) 1194, arXiv: 2007.07624 [hep-ex] (cit. on p. 43).

[101] T. G. Cornelissen et al., *Updates of the ATLAS Tracking Event Data Model (Release 13)*, tech. rep., CERN, 2007, URL: https://cds.cern.ch/record/1038095 (cit. on p. 44).

[102] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444 (cit. on p. 44).

[103] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Journal of Basic Engineering **82** (1960) 35, ISSN: 0021-9223 (cit. on p. 44).

[104] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex] (cit. on pp. 45, 47).

[105] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, and E. Moyse, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, tech. rep., 2007, URL: https://cds.cern.ch/record/1020106 (cit. on p. 46).

[106] ATLAS Collaboration, *ATLAS Computing: Technical Design Report*, ATLAS-TDR-17; CERN-LHCC-2005-022, 2005, URL: https://cds.cern.ch/record/837738 (cit. on p. 47).

[107] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on pp. 47, 72).

[108] GEANT4 Collaboration, S. Agostinelli, et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on pp. 47, 72).

[109] ATLAS Collaboration, *Early Inner Detector Tracking Performance in the 2015 Data at* $\sqrt{s} = 13\,TeV$, ATL-PHYS-PUB-2015-051, 2015, URL: https://cds.cern.ch/record/2110140 (cit. on pp. 47, 48).

[110] S. Boutle et al., *Primary vertex reconstruction at the ATLAS experiment*, J. Phys. Conf. Ser. **898** (2017) 042056, ed. by R. Mount and C. Tull (cit. on p. 47).

[111] F. Meloni, *Primary vertex reconstruction with the ATLAS detector*, JINST **11** (2016) C12060, ed. by C. Gemme and L. Rossi (cit. on p. 48).

[112] 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] (cit. on pp. 48, 49).

[113] 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] (cit. on pp. 48, 50).

[114] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex] (cit. on pp. 48, 52).

[115] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, URL: https://cds.cern.ch/record/1449796 (cit. on p. 49).

[116] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, JINST **14** (2019) P03017, arXiv: 1812.03848 [hep-ex] (cit. on pp. 49, 86).

[117] 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] (cit. on p. 51).

[118] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 52).

[119] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph] (cit. on p. 52).

[120] J. Alimena et al., *Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider*, J. Phys. G **47** (2020) 090501, arXiv: 1903.04497 [hep-ex] (cit. on p. 53).

[121] D. Curtin et al., *Long-Lived Particles at the Energy Frontier: The MATHUSLA Physics Case*, Rept. Prog. Phys. **82** (2019) 116201, arXiv: 1806.07396 [hep-ph] (cit. on p. 53).

[122] ATLAS Collaboration, *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*, JHEP **10** (2019) 265, arXiv: 1905.09787 [hep-ex] (cit. on pp. 54, 111).

[123] ATLAS Collaboration, *Search for exotic decays of the Higgs boson into long-lived particles in pp collisions at $\sqrt{s}$ = 13 TeV using displaced vertices in the ATLAS inner detector*, JHEP **11** (2021) 229, arXiv: 2107.06092 [hep-ex] (cit. on pp. 54, 62).

[124] ATLAS Collaboration, *Search for long-lived, massive particles in events with a displaced vertex and a muon with large impact parameter in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D **102** (2020) 032006, arXiv: 2003.11956 [hep-ex] (cit. on p. 54).

[125] ATLAS Collaboration, *Search for long-lived, massive particles in events with displaced vertices and missing transverse momentum in $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS detector*, Phys. Rev. D **97** (2018) 052012, arXiv: 1710.04901 [hep-ex] (cit. on pp. 54, 90).

[126] ATLAS Collaboration, *Search for R-parity-violating supersymmetric particles in multi-jet final states produced in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC*, Phys. Lett. B **785** (2018) 136, arXiv: 1804.03568 [hep-ex] (cit. on pp. 54, 61, 62).

[127] ATLAS Collaboration, *Search for displaced vertices of oppositely charged leptons from decays of long-lived particles in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **801** (2020) 135114, arXiv: 1907.10037 [hep-ex] (cit. on pp. 54, 86, 92).

[128] ATLAS Collaboration, *Search for Displaced Leptons in $\sqrt{s} = 13$ TeV pp Collisions with the ATLAS Detector*, Phys. Rev. Lett. **127** (2021) 051802, arXiv: 2011.07812 [hep-ex] (cit. on pp. 54, 86).

[129] ATLAS Collaboration, *Performance of the reconstruction of large impact parameter tracks in the inner detector of ATLAS*, ATL-PHYS-PUB-2017-014, 2017, URL: https://cds.cern.ch/record/2275635 (cit. on pp. 54–56).

[130] ATLAS Collaboration, *Software Performance of the ATLAS Track Reconstruction for LHC Run 3*, ATL-PHYS-PUB-2021-012, 2021, URL: https://cds.cern.ch/record/2766886 (cit. on pp. 56, 66, 68).

[131] F. Chollet et al., Keras, 2015, URL: https://keras.io (cit. on p. 59).

[132] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*, 2016, arXiv: 1603.04467 [cs.DC], URL: https://www.tensorflow.org/ (cit. on p. 59).

[133] D. Guest, Lightweight Trained Neural Network, 2022, URL: https://github.com/lwtnn/lwtnn (cit. on p. 60).

[134] ATLAS Collaboration, *Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Phys. Rev. D **98** (2018) 012003, arXiv: 1801.02052 [hep-ex] (cit. on p. 62).

[135] ATLAS Collaboration, *Expected Large-Radius Tracking Performance in Run 3*, 2021, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2021-003/ (cit. on pp. 63, 65, 66).

[136] ATLAS Collaboration, *Performance of vertex reconstruction algorithms for detection of new long-lived particle decays within the ATLAS inner detector*, ATL-PHYS-PUB-2019-013, 2019, URL: https://cds.cern.ch/record/2669425 (cit. on pp. 65, 80, 81).

[137] J. Wollrath, *Boosting the discovery potential for the LHC Run 3 Improved Track Reconstruction for prompt and long-lived particles in ATLAS*, tech. rep., CERN, 2022, URL: https://cds.cern.ch/record/2806787 (cit. on p. 65).

[138] ATLAS Collaboration, *Run 3 Large-radius tracking performance in data and simulation*, 2022, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2022-02/ (cit. on pp. 66–68).

[139] Z. M. Schillaci et al., *Improvements to ATLAS Inner Detector Track reconstruction for LHC Run-3*, tech. rep., CERN, 2021, URL: https://cds.cern.ch/record/2773196 (cit. on p. 67).

[140] J.-L. Tastet, O. Ruchayskiy, and I. Timiryasov, *Reinterpreting the ATLAS bounds on heavy neutral leptons in a realistic neutrino oscillation model*, JHEP **12** (2021) 182, arXiv: 2107.12980 [hep-ph] (cit. on pp. 72, 73, 75).

[141] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, *The fate of hints: updated global analysis of three-flavor neutrino oscillations*, JHEP **09** (2020) 178, arXiv: 2007.14792 [hep-ph] (cit. on pp. 72, 73).

[142] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 72).

[143] ATLAS Collaboration, *ATLAS Pythia 8 tunes to* 7 *TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419 (cit. on p. 72).

[144] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph] (cit. on p. 72).

[145] 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 (cit. on p. 72).

[146] M. C. Gonzalez-Garcia, A. Santamaria, and J. W. F. Valle, *Isosinglet Neutral Heavy Lepton Production in Z Decays and Neutrino Mass*, Nucl. Phys. B **342** (1990) 108 (cit. on p. 74).

[147] ATLAS Collaboration, *Measurement of $W^{\pm}$ and Z-boson production cross sections in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Lett. B **759** (2016) 601, arXiv: 1603.09222 [hep-ex] (cit. on pp. 75, 98).

[148] K. Bondarenko, A. Boyarsky, D. Gorbunov, and O. Ruchayskiy, *Phenomenology of GeV-scale Heavy Neutral Leptons*, JHEP **11** (2018) 032, arXiv: 1805.08567 [hep-ph] (cit. on pp. 75, 98).

[149] J.-L. Tastet and I. Timiryasov, *Dirac vs. Majorana HNLs (and their oscillations) at SHiP*, JHEP **04** (2020) 005, arXiv: 1912.05520 [hep-ph] (cit. on p. 76).

[150] R. Ruiz, *Quantitative study on helicity inversion in Majorana neutrino decays at the LHC*, Phys. Rev. D **103** (2021) 015022, arXiv: 2008.01092 [hep-ph] (cit. on p. 76).

[151] J.-L. Tastet, O. Ruchayskiy, and I. Timiryasov, *Reinterpreting the ATLAS bounds on heavy neutral leptons in a realistic neutrino oscillation model*, JHEP **12** (2021) 182, arXiv: 2107.12980 [hep-ph] (cit. on p. 76).

[152] 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] (cit. on p. 76).

[153] C. Degrande, O. Mattelaer, R. Ruiz, and J. Turner, *Fully automated precision predictions for heavy neutrino production mechanisms at hadron colliders*, Phys. Rev. D **94** (2016) 053002, arXiv: 1602.06957 [hep-ph] (cit. on p. 76).

[154] S. Pascoli, R. Ruiz, and C. Weiland, *Heavy neutrinos with dynamic jet vetoes: multilepton searches at $\sqrt{s}$ = 14 , 27, and 100 TeV*, JHEP **06** (2019) 049, arXiv: 1812.08750 [hep-ph] (cit. on p. 76).

[155] D. Trischuk, "Displaced vertex search for heavy neutral leptons using the ATLAS detector", PhD Thesis: University of British Columbia, 2022, URL: https://open.library.ubc.ca/soa/cIRcle/collections/ubctheses/24/items/1.0406559 (cit. on pp. 76, 80, 81).

[156] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det] (cit. on p. 76).

[157] 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] (cit. on p. 77).

[158] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv: 2004.13447 [physics.ins-det] (cit. on p. 77).

[159] I. Antcheva et al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, Comput. Phys. Commun. **180** (2009) 2499, arXiv: 1508.07749 [physics.data-an] (cit. on p. 84).

[160] xAODAnaHelpers (xAH), 2022, URL: https://github.com/UCATLAS/xAODAnaHelpers (cit. on p. 84).

[161] S. Raschka, J. Patterson, and C. Nolet, *Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence*, Information **11** (2020) 193, ISSN: 2078-2489, URL: https://www.mdpi.com/2078-2489/11/4/193 (cit. on p. 85).

[162] E. Rodrigues et al., *The Scikit HEP Project – overview and prospects*, EPJ Web Conf. **245** (2020) 06028, ed. by C. Doglioni et al., arXiv: 2007.03577 [physics.comp-ph] (cit. on p. 85).

[163] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep., New York U., 2012, URL: https://cds.cern.ch/record/1456844 (cit. on pp. 85, 106).

[164] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv: physics/0306116 (cit. on pp. 85, 106).

[165] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv: 1009.1003 [physics.data-an] (cit. on pp. 85, 106).

[166] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an] (cit. on pp. 85, 104), Erratum: Eur. Phys. J. C **73** (2013) 2501.

[167] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13\,TeV$*, Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex] (cit. on p. 86).

[168] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693 (cit. on p. 104).

[169] G. Cowan, *Statistics for Searches at the LHC*, tech. rep., 2013 321, arXiv: 1307.2487 [hep-ex] (cit. on pp. 104, 106).

[170] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, Phil. Trans. Roy. Soc. Lond. A **231** (1933) 289 (cit. on p. 104).

[171] CMS Collaboration, *Search for long-lived heavy neutral leptons with displaced vertices in proton-proton collisions at* $\sqrt{s}$ *=13 TeV*, (2022), arXiv: 2201.05578 [hep-ex] (cit. on p. 112).