A Search for Neutral Long-Lived Particles Decaying into Displaced Jets in the ATLAS Calorimeter in Proton-Proton Collisions at the LHC

Mason Proffitt

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2025

Reading Committee:
Gordon Watts, Chair
Henry Lubatti
Stephen Sharpe

Program Authorized to Offer Degree:
Physics

University of Washington

Abstract

A Search for Neutral Long-Lived Particles Decaying into Displaced Jets in the ATLAS Calorimeter in Proton-Proton Collisions at the LHC

Mason Proffitt

Chair of the Supervisory Committee:
Gordon Watts
Department of Physics

A search for neutral long-lived particles (LLPs) decaying in the ATLAS hadronic calorimeter using 140 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV delivered by the LHC is presented. The analysis targets pair-produced LLPs where at least one LLP is produced with sufficiently low boost that its decay products can be resolved as separate jets. No excesses of events relative to the background predictions are observed. Higgs boson branching fractions to pairs of hadronically decaying neutral LLPs larger than 1% are excluded at 95% confidence level for proper decay lengths in the range of 30 cm to 4.5 m depending on the LLP mass, a factor of three improvement on previous searches in the hadronic calorimeter.

Contents

List of Figures **vii**

List of Tables **xv**

1 Introduction **1**

2 Theory **3**

2.1 Quantum field theory 3

2.1.1 Lie groups and Poincaré symmetry 3

2.1.2 Projective representations 4

2.1.3 Gauge symmetry 5

2.1.4 Path integral formalism 7

2.2 The Standard Model 8

2.2.1 Electroweak sector 8

2.2.2 Quantum chromodynamics 11

2.2.3 Fermions 11

2.2.4 Limitations 15

2.3 Long-lived particles 17

2.3.1 Particle lifetimes 17

2.3.2 Models 18

3 **Experimental apparatus** **20**

3.1 The Large Hadron Collider 20

3.2 The ATLAS detector 27

3.2.1 Inner detector 28

3.2.2 Calorimetry 32

3.2.3 Muon spectrometer 39

3.2.4 Luminosity measurement 44

3.2.5 Trigger and data acquisition 45

3.2.6 Physics object reconstruction 47

3.2.7 Simulation 53

4 **Analysis** **55**

4.1 Simulated samples 55

4.1.1 Signal samples 55

4.1.2 Background simulation 56

4.2 Triggers 56

4.2.1 Level 1 signature-driven CalRatio triggers 57

4.2.2 CalRatio trigger at HLT 58

4.2.3 Efficiency 60

4.3 Data samples 61

4.4 Jet selection 62

4.4.1 Jet cleaning 62

4.4.2 LLP jet tagging neural network 64

4.5 Event selection 68

4.5.1 Preselection 68

4.5.2 Event cleaning 69

4.6 Background estimation 73

4.7 Lifetime extrapolation 76

4.8 Systematic uncertainties . . . 78

4.9 Statistical interpretation . . . 82

5 Results 84

6 Conclusions 87

A MATHUSLA 90

A.1 Proposed experiment . . . 90

A.2 Test stand . . . 92

A.2.1 Scintillation counters . . . 93

A.2.2 Resistive plate chambers (RPCs) . . . 94

A.2.3 Electronics, trigger, and data acquisition . . . 96

A.2.4 Coordinate system . . . 97

A.2.5 Timing calibration . . . 98

A.2.6 Track reconstruction . . . 99

A.2.7 Detector efficiency . . . 100

A.2.8 Simulation . . . 101

A.2.9 Results . . . 104

A.2.10 Conclusions . . . 109

B IRIS-HEP 110

B.1 Analysis description languages . . . 111

B.1.1 ADL benchmarks . . . 111

B.1.2 FuncADL . . . 112

B.2 ServiceX . . . 117

Bibliography 119

List of Figures

2.1 Two possibile shapes for a real potential of quadratic and quartic terms [1]. If the quadratic term is nonnegative (a), the minimum value is at $\phi = 0$. Otherwise (b), the minimum value is reached at a a non-zero absolute value of v. 9

2.2 Summary of the elementary particles of the Standard Model (Wikimedia Commons, 2024). . 14

2.3 Running of the strong coupling α_s as a function of energy scale Q [2]. The colored bars indicate various determinations of the coupling at particular scales, and the band of black lines is calculated by appling the β function to the average of these results. 15

2.4 Theoretical Standard Model cross section predictions vs. measurements by the ATLAS experiment for the production of many different states [3]. 16

2.5 Total energy content of the universe by source, based on observations by the Planck space telescope (ESA and the Planck Collaboration, 2013). "Ordinary matter" represents all energy attributable to the particles of the Standard Model. 16

2.6 Mean lifetime vs. mass for various Standard Model particles [4]. The lighter middle region roughly represents the lifetime range in which current particle physics experiments could directly detect displaced decays. 18

2.7 Feynman diagram for a generic hidden sector model with a scalar mediator Φ and a long-lived scalar s [5]. For the base model, Φ is the SM Higgs boson H. 19

3.1 Illustration of the location of the Large Hadron Collider [6]. 21

3.2 Layout of the arcs and insertion regions of the LHC (not to scale) [7]. 22

3.3 Diagram of the full CERN accelerator complex in 2018 [8]. LHC protons began in LINAC2 and passed through the PS Booster, the PS, and the SPS before being injected into the LHC ring. Heavy ions are accelerated by LINAC3, LEIR, and the SPS for injection into the LHC. 23
3.4 Left: View of dipole magnets in the LHC tunnel during installation [9]. Right: Cross sectional schematic of an LHC dipole magnet [10]. 24
3.5 Schedule of LHC and HL-LHC operation as of January 2025. 27
3.6 Cut-away view of the ATLAS detector. 28
3.7 ATLAS cavern prior to installation of the detector. 29
3.8 Cut-away view of the ATLAS inner detector. 30
3.9 Half of a pixel barrel layer (left) and one completed pixel endcap (right). The layout of the individual modules can be seen wrapping around where the beam pipe would be. For the endcap layers, the other half of the pixel modules are on the back side of the disks. 31
3.10 SCT barrel module (left) and an SCT endcap disk (right). The barrel modules are arranged in cylindrical layers around the beam pipe in a pattern similar to the pixel barrel modules. The endcap disk has a middle ring of modules on the other side, between the outer and inner rings shown here. For geometrical acceptance reasons, not all SCT endcap disks have all three rings of modules. 32
3.11 Left: A four-plane TRT endcap unit during assembly with the radial straw tubes visible. Right: End view of the TRT barrel. 33
3.12 Central solenoid magnet. 34
3.13 Overview of the calorimeter system. 35
3.14 Schematic of the electromagnetic barrel calorimeter structure. 36
3.15 Electromagnetic barrel (left) and endcap (right) calorimeters during assembly. 36
3.16 Left: Schematic of a tile calorimeter submodule. Right: Segmentation of the tile calorimeter readout for the central barrel and the extended barrel. 37
3.17 Illustration of a HEC module (left) and one completed endcap of the HEC (right). 37

3.18 Left: Schematic detail of tube structure for the hadronic FCal modules, showing the tungsten slugs surrounding each tube. Right: Completed FCal assembly being insertedinto one of the endcap cryostats. 38

3.19 Muon spectrometer overview. 39

3.20 Left: Barrel toroid before installation of detectors. Right: An endcap toroid in its cryostat with the end open. 40

3.21 RPC chamber structure (left) and layout (right). 41

3.22 Left: TGC chamber structure. Right: Structure of TGC triplet and doublet units. 41

3.23 Left: Tube structure for the MDTs. Right: MDT Chamber layout. 42

3.24 Left: Internal CSC structure. Right: Layout of large and small CSCs. 43

3.25 Muon detector layout. Left: Azimuthal RPC/MDT sectors in the barrel. Right: Detector layers arranged by r and η. RPC are in white, TGCs in purple, MDTs in green, and CSCs in yellow. Note that there is an extra endcap wheel (labeled EEL here) of MDTs between the inner and middle wheels to cover the part of the angular acceptance of the inner wheel that does not project onto the outer wheel. 43

3.26 LUCID-2 configuration for one side of ATLAS in 2015. The LED calibration system was replaced with additional ^{207}Bi sources in 2016 and 2017. The support cylinder that four of the groups of photomultipliers are attached to has a diameter of 22 cm. 45

3.27 Schematic of trigger tower clusters used in the L1Calo trigger algorithms [11]. The transverse energy of a trigger tower is obtained by summing the energy deposited in each layer of the calorimeter in a $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ slice, and clusters are formed by adding the energy of adjacent trigger towers together. Electron/photon and tau triggers are based on 0.2×0.2 clusters, and jet triggers are based on 0.4×0.4 or 0.8×0.8 clusters. 46

3.28 ATLAS trigger and data acquisition system in Run 2. The arrows illustrate data flow between components. 47

3.29 Diagram of the superclustering algorithm for electrons and photons. Seed clusters are shown in red, satellite clusters in blue. 52

4.1 Simulated trigger efficiency as a function of the LLP decay decay position x-y plane for LLPs decaying in the barrel ($|\eta| < 1.4$). A comparison is shown for the three versions of the CalRatio triggers for one of the low-mass signal samples (left) and one of the high-mass signal samples (right). In each of the plots, the light filled area corresponds to the CalRatio trigger seeded by L1_TAU60, open markers correspond to the CalRatio trigger seeded by the 2016 version of L1_LLP-NOMATCH, and filled markers to the 2018 version of L1_LLP-NOMATCH. . . . 61

4.2 Simulated trigger efficiency as a function of the LLP decay decay position long the longitudinal direction for LLPs decaying in the endcap ($|\eta| > 1.4$). A comparison is shown for the three versions of the CalRatio triggers for one of the low-mass signal samples (left) and one of the high-mass signal samples (right). In each of the plots, the light filled area corresponds to the CalRatio trigger seeded by L1_TAU60, open markers correspond to the CalRatio trigger seeded by the 2016 version of L1_LLP-NOMATCH, and filled markers to the 2018 version of L1_LLP-NOMATCH. 62

4.3 Simulated trigger efficiency as a function of the LLP p_{T}. A comparison is shown for the three versions of the CalRatio triggers for one of the low-mass signal samples (left) and one of the high-mass signal samples (right). In each of the plots, the light filled area corresponds to the CalRatio trigger seeded by L1_TAU60, open markers correspond to the CalRatio trigger seeded by the 2016 version of L1_LLP-NOMATCH, and filled markers to the 2018 version of L1_LLP-NOMATCH. 63

4.4 Low-E_T per-jet NN BIB score, in the dijet control region, for an network with no adversary (left) and with an adversary (right). 66

4.5 High-E_T per-jet NN BIB score, in the dijet control region, for an network with no adversary (left) and with an adversary (right). 66

4.6 Diagram of how the adversary interacts with the main network. The main network, which discriminates between signal, QCD, and BIB, is shown with the inputs moving through the network as black arrows. The inputs to the adversary, which classifies jets as being from MC simulation or from data, are shown moving through the network as orange arrows. The whole network tries to minimize the loss function as shown here, which it does by minimizing the main network loss function, learning to discriminate signal from BIB and QCD, while minimizing the negative of the adversary loss function, learning not to discriminate between simulation and data. 67

4.7 The output signal score (confidence of a jet being signal) shown for low mass (left) and high mass (right) training. 68

4.8 The output BIB score (confidence of a jet being BIB) shown for low mass (left) and high mass (right) selections. 68

4.9 The output QCD score (confidence of a jet being QCD) shown for low mass (left) and high mass (right) selections. 69

4.10 Distribution of $\sum \Delta R_{\text{min}}$. 70

4.11 Stability of $\sum \Delta R_{\text{min}}$ (jet, tracks) against pileup. 71

4.12 Top: Distribution of the time of the first signal jet candidates. The left plot is for the jet with the highest low-E_T NN score, and the right is for the jet with the highest high-E_T NN score. Bottom: Distribution of $\log_{10}(E_\text{H}/E_\text{EM})$ for the same jets. 72

4.13 ABCDisCo NN output distribution 74

4.14 First BIB jet candidate time versus η from the BIB data sample in three different intervals of the NN output, increasing from left to right. The top row of plots shows the high-ET BIB jet candidate, and the bottom row shows the low-ET BIB jet candidate. 75

4.15 First BIB jet candidate ϕ from the BIB data sample in three different intervals of the NN output, increasing from left to right. The top row of plots shows the high-ET BIB jet candidate, and the bottom row shows the low-ET BIB jet candidate. 76

4.16 Distributions of events in the ABCD plane. The upper plots show data from the main dataset (left) and the BIB dataset (right). The lower plots show the distribution for two different signal samples. 77

4.17 The extrapolated signal efficiency in region A as a function of mean LLP lifetime. The bands represent statistical uncertainties only. The blue and orange curves are obtained from the same model, where events were generated using two different mean lifetime assumptions. The dashed vertical lines represent the originally generated mean lifetimes. The lower-lifetime sample is shown in blue, while orange represents the higher-lifetime sample. 79

4.18 CR+2J validation regions. The left plots demonstrate how $\mathrm{VR}_{\mathrm{BD}}$ and $\mathrm{VR}_{\mathrm{CD}}$ are defined and show the distribution of events within them. $\mathrm{VR}_{\mathrm{BD}}$ only includes events in the main dataset with NN values less than 7.61, and $\mathrm{VR}_{\mathrm{CD}}$ only includes main dataset events with $\sum \Delta R_{\mathrm{min}}$values less than 0.71. The plots on the right compare the observed and expected number of events in A′ in $\mathrm{VR}_{\mathrm{BD}}$ and $\mathrm{VR}_{\mathrm{CD}}$ by applying the standard ABCD procedure while varying cut values between the subregions. The vertical dashed lines indicate the nominal boundaries between the VR subregions, shown in the 2D plots and used when varying the cut in the other variable. These were set to be the boundaries between the nominal ABCD regions where possible and to be the median values in the respective validation region otherwise. . . 80

4.19 Ratio of observed and expected events in region A′ of $\mathrm{VR}_{\mathrm{CD}}$. The left plot includes only statistical uncertainties, while the right includes an additional 5% non-closure systematic. . . 80

5.1 Signal efficiencies for HSS signals. Extrapolation of efficiencies to new mean LLP lifetimes is done via the event reweighting procedure. 84

5.2 Expected and observed upper limits for HSS signal models. The “2 CalRatio jet analysis” lines refer to the search for two CalRatio jets performed in [5]. 86

6.1 95% confidence level upper limits on the SM Higgs boson decaying to a pair of neutral scalars as a function of $c\tau$ for a hidden sector model where $m_H = 125$ GeV and m_s is in the range of 55 to 60 GeV. In this plot, the long-lived scalars are then assumed to decay to b-quarks with 100% branching fraction. This results of this analysis are repesented by the line labeled "Calorimeter" in the legend. 88

A.1 Location of proposed MATHUSLA detector at the CMS site. 91

A.2 Diagram of MATHUSLA location and potential LLP decays. 91

A.3 Sensitivity of MATHUSLA and other experiments to hadronically decaying LLPs produced in exotic Higgs boson decays. 92

A.4 (a) 3D model of the MATHUSLA test stand. (b) Photo of the final assembled structure installed above the ATLAS IP. The green dots identify the two scintillator layers used for triggering, while the red dots mark the three RPC double-layers used for tracking. 93

A.5 Schematic [12] and photo of a DØ forward muon scintillation trigger counter. 94

A.6 Layout of the top (a) and bottom (b) scintillator planes, each of which has a 2.5 × 2.5 m^2 active area. 95

A.7 (a) Schematic of the cross-section of an ARGO-YBJ chamber [13]. (b) Strip panel used for readout. The total size of each chamber is 285 × 126 × 4.7 cm^3. 95

A.8 Layout of one of the RPC double layers illustrating the support structure that ensures overlap of individual chambers to avoid gaps in coverage. 96

A.9 Distribution of hits for examples of a downward-going track (top) and an upward-going track (bottom) from data. Left: Event display of the scintillation counters and RPC pads comprising the track hits. Right: Plot of the z-coordinate versus time for each hit. The red line represents the fitted track in each case. 100

A.10 Diagram with the geometry and material of the ATLAS cavern, the test stand, and its surroundings. 102

A.11 Distribution of reconstructed downward-going tracks as a function of the zenith angle (left) and the azimuthal angle (right). Data events are shown as black markers. The yellow area corresponds to the downward cosmic ray simulation, normalized to data. 105

A.12 Distribution of reconstructed upward-going tracks as a function of the zenith angle (left) and the azimuthal angle (right). Data events with no beam are shown as upward-pointing black triangles. The green area corresponds to the cosmic ray inelastic backscattering simulation, with the statistical uncertainty shown by the hatched area. This distribution is normalized by multiplying it by the ratio of downward tracks in simulation to downward tracks in data without beam. Downward-going tracks from data events with no beam, normalized using the $R_{\text{up}-\text{to}-\text{down}}$ factor, are included in the plot as downward-pointing blue triangles. 106

A.13 Distribution of reconstructed upward-going tracks as a function of the zenith angle (left) and the azimuthal angle (right). Data events in runs with beam are shown as black markers. Overlaid is a simulation of particles coming from the ATLAS IP in orange. The blue distribution corresponds to cosmic ray inelastic backscattering, showing the downward-going tracks from data runs with beam, normalized using the $R_{\text{up}-\text{to}-\text{down}}$ factor. The hatched area shows the combination of the uncertainties of the IP muon simulation and the cosmic ray inelastic backscattering prediction. 107

A.14 Distribution of the number of reconstructed tracks as a function of the ATLAS integrated luminosity during each one-hour test stand run. Left: Downward tracks. Right: Upward tracks (black circles), including tracks with a zenith angle $(\theta) > 6^\circ$ (blue squares) and tracks with a zenith angle $< 4^\circ$ and absolute value of azimuthal angle $(\phi) < 90^\circ$ (red triangles). . . 108

B.1 Projected computing resource needs for ATLAS for CPU (left) and disk (right). 110

B.2 Diagram of the components of ServiceX and how they interact. The user only needs to interact with the Jupyter Notebook interface. 118

List of Tables

5.1 Sequential impact of each requirement on the number of events passing the selection. For the example signal models considered, the columns represent the cumulative fraction of events passing each step. 85

5.2 Application of the likelihood-based ABCD method to the final selections. The columns A, B, C and D contain the number of observed events in each region. "Estim. A" refers to the estimated contribution of background in the search region A assuming no signal, as calculated using the full statistical model. As described in Section 4.9, the a priori estimate refers to the pre-unblinding case and the a posteriori estimate refers to the post-unblinding case, including the observed data in region A into the background-only global fit. Only statistical uncertainties are included in the quoted error. 85

A.1 Summary of cross-sections, acceptances, and expected rates at the test stand from various LHC pp collision processes at $\sqrt{s} = 13$ TeV. The rates are normalized to an instantaneous luminosity of 10^{34} cm^{-2} s^{-1} and do not include test stand detector inefficiencies. 104

B.1 Implemented benchmark tasks. In each case, `ds` is the `EventDataset` object that contains the events. Only the FuncADL query is provided rather than the plot, which is a histogram of the elements in the result. These plots can be seen in the GitHub repository. 114

Chapter 1

Introduction

The theory that is the crown jewel of particle physics, the Standard Model, is perhaps simultaneously one of the greatest triumphs and greatest embarrassments of modern science. While it can predict the properties of apparently elementary particles such as the electron to a precision of less than one part in a trillion [14], it fails to explain something as familiar as gravity. Even the majority of matter in the universe does not seem to be accounted for by the Standard Model. With such glaring holes in our fundamental understanding of the physical world, and with the discovery in 2012 of the final particle of the Standard Model, the Higgs boson [15, 16], the hunt is on for new physics.

One potential avenue for new physics is in the existence of new long-lived particles. Until fairly recently, this was an overlooked idea, and so most large experiments today were not designed with this possibility in mind. Particle physics analyses tend to focus on particles that are either completely stable (like protons or electrons) or so extremely unstable that they essentially decay instaneously after being produced (like the Higgs boson). Long-lived particles fall in between these extremes, so a signature of their existence could be sudden unexpected activity in a particle detector far away from where any known collisions are happening.

The primary topic of this work is a search for electrically neutral long-lived particles beyond the Standard Model that are produced in pairs and decay in or near the calorimeter of the ATLAS experiment during high-intensity collisions of protons at the Large Hadron Collider. This follows a paper that was published by the ATLAS collaboration on both this analysis and a parallel search for long-lived particles produced in combination with leptons [5]. The data for this analysis was collected during 2015–2018, with a total integrated luminosity of 140 fb^{-1} at a center-of-mass energy of 13 TeV.

The remainder of this dissertation is structured as follows. Chapter 2 develops the current theoretical basis of particle physics, the Standard Model, as well as possible extensions. In Chapter 3, the experimental setup of the Large Hadron Collider and ATLAS is described. The analysis methods for the long-lived particle search mentioned above are explained in detail in Chapter 4. The results of this analysis are presented in Chapter 5. Finally, Chapter 6 discusses the results and provides a summary and outlook for this research.

Chapter 2

Theory

In this chapter, the Standard Model of particle physics will be introduced, along with its known shortcomings, starting from the foundations of quantum field thoery. Possible extensions of this model will then be discussed, including in particular those involving additional long-lived particles.

2.1 Quantum field theory

Quantized fields form the basis of nearly all modern theories of particle physics. There may be a different underlying true theory of nature (e.g, string theory), but at least in a low energy limit, it must reduce to a theory of quantum fields [17]. Quantum field theories (QFTs) take these fields to be the fundamental objects, and particles are quantized excitations of the fields. Special relativity dictates that these fields must transform in a Lorentz-covariant manner. More specifically, each field must tranform under a projective representation of a Lie group known as the Poincaré group.

2.1.1 Lie groups and Poincaré symmetry

A Lie group is a group that can be smoothly parameterized by a set of continuous real parameters $\{\alpha_a\}$ [18]. The neighborhood of a group element can be treated as a vector space under infinitesimal transformations [19]. The basis vectors $\{X_a\}$ are called the generators of the group. If we associate the identity element with $\alpha_a = 0$, then we can build a representation of the group directly from the generators:

$$O(\alpha) = e^{i\alpha_a X_a}$$

where exponentiation is defined by a Taylor series expansion, and the Einstein convention is used to indicate summation for repeated indices. The algebra of the generators of a Lie group is called a Lie algebra, which is completely defined by the commutation relations:

$$[X_a, X_b] = i f_{abc} X_c$$

where f_{abc} is known as a structure constant.

Examples of Lie groups important in particle physics include $U(N)$, the group of $N \times N$ unitary matrices, and $SU(N)$, the group of $U(N)$ matrices with a determinant of 1. Also critical is the Lorentz group, which is defined by the transformations ${\Lambda^\mu}_\nu$ such that:

$$\eta_{\mu\nu} {\Lambda^\mu}_\rho {\Lambda^\nu}_\sigma = \eta_{\rho\sigma}$$

where $\eta_{\mu\nu}$ is the diagonal Minkowski spacetime metric, with $\eta_{00} = -1$ and $\eta_{11} = \eta_{22} = \eta_{33} = 1$. The Poincaré group consists of the part of the Lorentz group connected to the identity ($\det {\Lambda^\mu}_\nu = 1$ and ${\Lambda^0}_0 > 0$). The generators of the Poincaré group are P^μ, J^i, and K^i. P^0 or H is the Hamiltonian, which describes the evolution of states over time, as in classical and quantum mechanics. The generators that commute with H are associated with conserved quantities. P^μ represents four-momentum and generates spacetime translations. J^i represents angular momentum and generates spatial rotations. K^i generates boosts but does not commute with the Hamiltonian, and so is not associated with a conservation law. A general Poincaré transformation can be constructed from this set of generators. The Poincaré group represents the full group of symmetries respected by special relativity.

2.1.2 Projective representations

If $O(g)$ is the operator that acts on an object ϕ to apply the transformation associated with the group element g, ϕ is said to transform in a projective representation if, for any group elements g_1, g_2:

$$O(g_2)O(g_1)\phi = C(g_2, g_1)O(g_2 g_1)\phi$$

where $C(g_2, g_1)$ is a real-valued function. If $C = 1$, then this is an ordinary representation. For the Poincaré group, its topology dictates that $C = \pm 1$.

Now assume that $|\phi\rangle$ is a one-particle state that, without loss of generality, has three-momentum normal to the xy-plane and is in a eigenstate of J_3 (any other states of the particle can be produced by rotations and superposition). If we rotate around the z-axis by π twice, a tranformation that corresponds to the identity element of the Lorentz group, we have:

$$e^{i\pi J_3} e^{i\pi J_3} |\phi\rangle = \pm |\phi\rangle$$

$$e^{i2\pi\sigma} |\phi\rangle = \pm |\phi\rangle$$

where σ is the z-component of the particle's spin. Thus $\sigma = \frac{n}{2}$ for some integer n. Particles for which spin is a half-integer are called fermions, and particles with integer spin are called bosons.

The linear combinations $\frac{1}{2}(J^i \pm iK^i)$ of the generators of the Lorentz group form two independent subalgebras with the same commutation relations as $SU(2)$. Thus general projective representations of the Poincaré group can be labeled by two non-negative numbers, (j_1, j_2), where each is either an integer or a half-integer. These represent two components that may transform differently under Lorentz transformations. The components are exhanged by a parity transformation, so these are identified as having left-handed and right-handed chirality. A representation can contain components with total spins of $|j_1 - j_2|$, $|j_1 - j_2| + 1$, ..., $j_1 + j_2$. The representations that a state transforms under determine which types of interactions with other states are allowed or forbidden by symmetry principles.

2.1.3 Gauge symmetry

Poincaré symmetry is an example of a global symmetry, where the coefficients $\{\alpha_a\}$ of the group generators in the definition of $O(\alpha)$ above are constants across spacetime. The most important other type of symmetry in QFT is local symmetry, or gauge symmetry. A local symmetry extends the transformation properties to the case where α_a is a function of spacetime. In other words, the value of a field at every spacetime point individually transforms under the same representation of the group. For example, let the field ϕ transform under a local symmetry with Hermitian generators $\{X_a\}$:

$$\phi(x) \to O(\alpha)\phi(x) = e^{i\alpha_a(x)X_a}\phi(x)$$

for any set of real-valued functions $\{\alpha_a\}$. Any dynamic theory will depend on derivatives of the fields, but these do not transform under a projective representation of the gauge symmetry:

$$\partial^\mu \phi \to \partial^\mu \left(e^{i\alpha_a(x)X_a}\phi\right) = i\partial^\mu \alpha_a(x) X_a \phi + e^{i\alpha_a(x)X_a}\partial^\mu \neq O(\alpha)\partial^\mu \phi$$

We require an object called a covariant derivative that does transform in such a way:

$$D^\mu \phi \to O(\alpha) D^\mu \phi$$

If we introduce a vector gauge field A_a^μ and set:

$$D^\mu \phi = \partial^\mu \phi - igA_a^\mu X_a \phi$$

where g is a coupling constant, then this will be covariant if:

$$A_a^\mu X_a \to O(\alpha) A_a^\mu X_a O(\alpha)^{-1} + \frac{1}{g}\partial^\mu \alpha_a(x) X_a$$

Without the second term, this would simply say that A_a^μ transforms exactly as the generators X_a do. This is called the adjoint representation of the group. Thus A_a^μ transforms in the adjoint representation under global transformations (where $\partial^\mu \alpha_a = 0$) but not local transformations. There is an unphysical degree of freedom that represents a choice of gauge, as in classical electrodynamics. Any observables must be gauge invariant. A covariant object for the dynamics of the gauge field itself can conveniently be constructed by

$$\begin{aligned}[D^\mu, D^\nu] &= -ig\left(\partial^\mu A_a^\nu - \partial^\nu A_a^\mu\right) X_a - g^2 A_a^\mu A_b^\nu [X_a, X_b] \\ &= -ig\left(\partial^\mu A_a^\nu - \partial^\nu A_a^\mu\right) X_a - g^2 A_a^\mu A_b^\nu i f_{abc} X_c \\ &= -igF_a^{\mu\nu} X_a\end{aligned}$$

where

$$F_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial^\nu A_a^\mu + g f_{bca} A_b^\mu A_c^\nu$$

The eigenvalues of X_a in a representation are called charges and are quantum numbers that characterize each particle. The global symmetry means that these charges are conserved. A field that transforms in the trivial representation, in which all group elements are represented by the identity, has zero charge. The gauge field only interacts with charge-carrying fields. If the structure constants are all zero, and so the commutators above are all zero, then the group is called Abelian. Gauge fields associated with Abelian symmetries have zero charge and therefore have no direct self interactions. Non-Abelian theories, however, do have charged gauge fields that can interact with themselves.

2.1.4 Path integral formalism

Most QFTs that have been studied, including the Standard Model, can be characterized by a Lagrangian density $\mathcal{L}$. The action S is the spacetime integral of the Lagrangian density:

$$S = \int \mathcal{L}\, d^4x$$

Generally the most fundamental theoretical prediction that can be tested in a particle physics experiment is the transition probability between a prepared initial state and a measurable final state. The intial and final states are composed of stable or metastable particles separated by macroscopic distances and can be considered not to be interacting with each other. The interaction is localized to a microscopic volume of spacetime, and thus the known initial and final states can be considered to be the states at times $t = -\infty$ and $t = \infty$, respectively. This describes a scattering process, and the transition probability amplitudes between each initial state and each final state make up what is known as the S-matrix. S-matrix elements can be calculated from the path integral:

$$\int e^{iS[\phi]/\hbar}\, \mathcal{D}\phi$$

where ϕ represents all fields, and the integral is over all possible paths of field configurations between the initial and final states. This is a genaralization of the Hamilton's principle of stationary action. In the classical limit that $S \gg \hbar$, the only significant contribution to the path integral comes from the stationary point (where $\delta S = 0$ under infinitesimal perturbations to the path) due to interference from the rapidly oscillating phase away from this path.

The path integral can be evaluated order by order by expanding the exponential series of the action. Each term in this expansion is associated with a Feynman diagram, which consists of lines and vertices determined by the fields that are present in each term. A factor is ascribed to each line and vertex, and the diagram is equivalent to the product of these factors. Vertices represent interactions between particles via the interaction terms of the Lagrangian, and internal lines are associated with four-momenta carried by the particles. External lines represent initial and final state particles. The types of diagrams that are allowed and the factors that are associated with them can then be summarized in a set of Feynman rules. Once these rules have been derived for a quantum field theory, calculating S-matrix elements can be reduced to determining all possible diagrams that contribute to a scattering process up to a given order.

2.2 The Standard Model

The Standard Model (SM) of particle physics constitutes the state-of-the-art understanding of our universe at the most fundamental level as of 2024 [20, 2]. It consists of quantum fields with interactions characterized by the symmetry group $SU(3) \times SU(2) \times U(1)$. These fields are associated with 17 types of particles (and their antiparticles), which have all been experimentally observed. The particles can be grouped into four categories: quarks, leptons, gauge bosons, and scalar bosons. The quarks and leptons are arranged into three generations. In each generation, there are two flavors of quarks and a charged lepton and an uncharged lepton. The gauge bosons mediate fundamental forces: strong, weak, and electromagnetism. The strong force is mediated by gluons and is what allows bound states of quarks, known as hadrons, to form. The weak force is mediated by the $W^{\pm}$ and Z bosons and is responsible for many slow decay processes, such as beta decay of nuclei. Electromagnetism is mediated by photons and involves the interactions of electrically charged particles.

2.2.1 Electroweak sector

Electroweak interactions are mediated by the gauge fields of an $SU(2) \times U(1)$ symmetry. Let T_i and Y be the generators of the $SU(2)$ and $U(1)$ groups, respectively, where $i \in \{1, 2, 3\}$. The corresponding gauge fields are labeled A_i^{μ} and B^{μ}. The charges associated with the $SU(2)$ symmetry are called weak isospin, and the $U(1)$ charge is called weak hypercharge. There is one complex scalar field ϕ, which transforms in the defining representation of the $SU(2)$ as a doublet with a normalized weak hypercharge of 1/2.

The terms in the SM Lagrangian that depend only on ϕ and not its derivatives or other fields are:

$$\mathcal{L}_\phi = -\frac{\mu^2}{2}\phi^\dagger\phi - \frac{\lambda^2}{4}\left(\phi^\dagger\phi\right)^2$$

where μ and λ are constants. We can interpret this as the negative of the potential for the scalar field. That is, $\mathcal{L}_\phi = -V[\phi]$. In order for the potential to be bounded below, we must have $\lambda^2 > 0$. With $\mu^2 \geq 0$, the minimum of this potential would be at $\phi = 0$ as in Figure 2.1(a). The Standard Model takes μ^2 to be negative, so that the minimum of this potential occurs at $\phi^\dagger\phi = v = -\mu^2/\lambda^2$, the vacuum state, shown in Figure 2.1(b). By a gauge transformation, we can set:

$$\phi(x) = \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$

where H is real. This vacuum state is said to spontaneously break the $SU(2) \times U(1)$ symmetry of the Lagrangian.

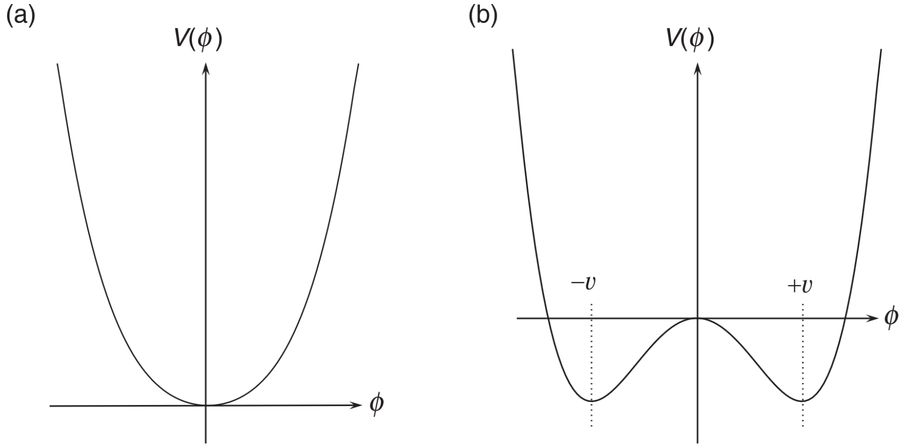


Figure 2.1: Two possibile shapes for a real potential of quadratic and quartic terms [1]. If the quadratic term is nonnegative (a), the minimum value is at $\phi = 0$. Otherwise (b), the minimum value is reached at a a non-zero absolute value of v.

Now we have:

$$\mathcal{L}_\phi = \mu^2 H^2 - \lambda^2 v H^3 - \frac{\lambda^2}{4}H^4 + \text{const.}$$

The real scalar field H is called the Higgs field, and its associated particle has a mass of $m_H = \sqrt{-\mu^2} = |\mu|$.

The kinetic terms of the Lagrangian for the electroweak sector are:

$$\mathcal{L}_{\text{kinetic}}^{\text{EW bosons}} = -\frac{1}{4}A_{i\mu\nu}A_i{}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \left(D^\mu\phi\right)^\dagger\left(D_\mu\phi\right)$$

where

$$A_i^{\mu\nu} = \partial^\mu A_i^\nu - \partial^\nu A_i^\mu + g\varepsilon_{ijk} A_j^\mu A_k^\nu$$

$$B^{\mu\nu} = \partial^\mu B_i^\nu - \partial^\nu B_i^\mu$$

where ε is the Levi-Civita symbol, the totally antisymmetric tensor such that $\varepsilon_{123\ldots n} = 1$ for rank n. The covariant derivative is given by

$$D^\mu \phi = \left(\partial^\mu - iA_i^\mu T_i - iB^\mu Y\right)\phi$$

We can break this up into

$$D^\mu \phi = D^\mu \left(\langle\phi\rangle + \begin{pmatrix} 0 \\ H \end{pmatrix}\right) = D^\mu \langle\phi\rangle + D^\mu \begin{pmatrix} 0 \\ H \end{pmatrix}$$

Expanding the first term:

$$\begin{aligned} D^\mu \langle\phi\rangle = \frac{-i}{2\sqrt{2}}\left(gA_i^\mu \sigma_i + g' B^\mu\right)\begin{pmatrix} 0 \\ v \end{pmatrix} &= \frac{-i}{2}\begin{pmatrix} gA_3^\mu + g'B^\mu & g(A_1^\mu - iA_2^\mu) \\ g(A_1^\mu + iA_2^\mu) & -gA_3^\mu + g'B^\mu \end{pmatrix}\begin{pmatrix} 0 \\ v \end{pmatrix} \\ &= \frac{-iv}{2\sqrt{2}}\begin{pmatrix} g(A_1^\mu - iA_2^\mu) \\ -gA_3^\mu + g'B^\mu \end{pmatrix} \end{aligned}$$

So there is a term:

$$\mathcal{L}_m^{\text{EW bosons}} = \frac{g^2v^2}{4}\left[\frac{1}{\sqrt{2}}(A_{1\mu} + iA_{2\mu})\right]\left[\frac{1}{\sqrt{2}}(A_1^\mu - iA_2^\mu)\right] + \frac{v^2}{4}\left[\frac{1}{\sqrt{2}}(-gA_{3\mu} + g'B_\mu)\frac{1}{\sqrt{2}}(-gA_3^\mu + g'B^\mu)\right]$$

in the Lagrangian. Normally gauge fields cannot have mass terms of the form $-m^2 A_a^\mu A_{a\mu}$ since this is not gauge invariant if m is a constant, but here the non-zero expectation value of the Higgs field generates masses for these fields in the vacuum state following spontaneous symmetry breaking. We can now pick out the linear combinations of the fields that have acquired mass terms. There is a complex field and its conjugate:

$$W^\mu = \frac{1}{\sqrt{2}}\left(A_1^\mu \pm iA_2^\mu\right)$$

with mass $m_W = \frac{|g|v}{2}$. Setting $\theta = \tan^{-1}\frac{g'}{g}$, there is a real field:

$$Z^\mu = \cos\theta A_3^\mu - \sin\theta B^\mu$$

with mass $m_Z = \frac{v\sqrt{g^2+(g')^2}}{2}$. The orthogonal linear combination

$$A^\mu = \sin\theta A_3^\mu + \cos\theta B^\mu$$

remains massless, which means there is an unbroken $U(1)_{\text{EM}}$ symmetry generated by $Q \equiv T_3 + Y$. A^μ is the quantized electromagnetic vector potential, and Q is the generator associated with electric charge.

2.2.2 Quantum chromodynamics

The strong interactions of the SM are mediated by the gauge fields of an $SU(3)$ symmetry. The generators t_a of this symmetry commute with those of the electroweak sector, so the complete gauge symmetry group of the SM is $SU(3) \times SU(2) \times U(1)$. After spontaneous symmetry breaking, this reduces to $SU(3) \times U(1)_{EM}$. The conserved charges of the $SU(3)$ symmetry are called color charges, and so their interactions are also known as quantum chromodynamics (QCD). The gauge fields are labeled G_a^μ, and the particles associated with these fields are called gluons. The kinetic term is simply:

$$\mathcal{L}_{\text{kinetic}}^{\text{QCD bosons}} = -\frac{1}{4} G_{a\mu\nu} G_a{}^{\mu\nu}$$

where

$$G_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu + g_s f_{abc} G_b^\mu G_c^\nu$$

and g_s is the strong coupling constant and f_{abc} is a structure constant of $SU(3)$. The Higgs field carries no color charge, so gluons remain massless.

2.2.3 Fermions

The fermionic fields of the SM consist of quarks and leptons. The quark fields are Q_{Li}, U_{Ri}, and D_{Ri}, where $i \in \{1, 2, 3\}$. Q_{Li} is an $SU(2)$ doublet:

$$Q_{Li} = \begin{pmatrix} U_{Li} \\ D_{Li} \end{pmatrix}$$

U and D refer to up-type and down-type quarks, L and R refer to the left-handed and right-handed chirality of the fields, and the numeric index represents the generation number. Each quark field has three color components, and these transform under the defining represntation of the $SU(3)$ symmetry. Up-type quarks have an electric charge of $2/3$, and down-type quarks have a charge of $-1/3$. The lepton fields are L_{Li} and E_{Ri}, where again $i \in \{1, 2, 3\}$ and L_{LI} is an $SU(2)$ doublet:

$$L_{Li} = \begin{pmatrix} N_{Li} \\ E_{Li} \end{pmatrix}$$

N and E refer to neutrino-type and electron-type fields. Leptons do not carry any color charge. Neutrino-type leptons are electrically neutral, and electron-type leptons have a charge of -1. Using the format

$(R_{SU(3)}, R_{SU(2)})_{q_Y}$, where R_G is the representation that a field transforms in for group G and q_Y is the weak hypercharge, the tranformation properties of the fermion fields can be summarized as:

$$Q_{Li} : (3,2)_{1/6} \quad U_{Ri} : (3,1)_{2/3} \quad D_{Ri} : (3,1)_{-1/3} \quad L_{Li} : (1,2)_{-1/2} \quad E_{Ri} : (1,1)_{-1}$$

The defining reprentations of $SU(3)$ and $SU(2)$ are labeled 3 and 2, respectively, and 1 is the trivial representation.

The full covariant derivative is

$$D^\mu = \partial^\mu - ig_s G_a^\mu t_a - igA_i^\mu T_i - ig' B^\mu Y$$

The kinetic terms for the fermions are:

$$\mathcal{L}_{\text{kinetic}}^{\text{fermions}} = i\overline{\psi_a} \not{D} \psi_a$$

where $\not{D} = \gamma^\mu \partial_\mu$ and ψ_a runs over all of the fermionic fields above.

Each fermion can couple to the scalar field, which is called a Yukawa interaction:

$$\mathcal{L}_{\text{Yukawa}} = -Y_{ij}^U \varepsilon_{ab} \overline{Q_{Li}}_a U_{Rj} \phi_b^* + Y_{ij}^D \overline{Q_{Li}}_a D_{Rj} \phi_a + Y_{ij}^E \overline{L_{Li}}_a E_{Rj} \phi_a + \text{h.c.}$$

where Y_{ij}^U, Y_{ij}^D, and Y_{ij}^E are the elements of 3×3 complex matrices, and a and b are indices for the two components of the $SU(2)$ doublets. The second term must include the complex conjugate field ϕ^* to preserve gauge invariance, since the $SU(2)$ components are exchanged by the Levi-Civita symbol. Like the gauge bosons, the fermions cannot have mass terms with constant coefficients directly in the Lagrangian. Because all the left-handed fields transform differently from right-handed fields (known as a chiral symmetry) and they all carry weak hypercharge, such a mass term would not be gauge invariant. However, the non-zero vacuum expectation value of ϕ again generates masses for these particles:

$$\mathcal{L}_m^{\text{fermions}} = -\frac{v}{\sqrt{2}} \left(Y_{ij}^U \overline{U_{Li}} U_{Rj} + Y_{ij}^D \overline{D_{Li}} D_{Rj} + Y_{ij}^E \overline{E_{Li}} E_{Rj} + \text{h.c.} \right)$$

For any complex square matrix M, we can find unitary matrices A and B such that AMB is real and diagonal. Let V^{f_L} and V^{f_R} be unitary matrices that make $M^f = V^{f_L} Y^f V^{f_R \dagger}$ real and diagonal for $f \in \{U, D, E\}$. Then we have

$$\mathcal{L}_m^{\text{fermions}} = -m_j^U \overline{U_{Li}} (V^{U_L \dagger})_{ij} V_{jk}^{U_R} U_{Rk} - m_j^D \overline{D_{Li}} (V^{D_L \dagger})_{ij} V_{jk}^{D_R} D_{Rk} - m_j^E \overline{E_{Li}} (V^{E_L \dagger})_{ij} V_{jk}^{E_R} E_{Rk} + \text{h.c.}$$

where m_j^U, m_j^D, and m_j^E are the fermion masses of generation j. The up-type quark mass eigenstates are labeled up (u), charm (c), and top (t), and the down-type mass states are down (d), strange (s), and bottom (b). The charged lepton mass eignestates are the electron (e), muon (μ), and tau (τ). Since the neutrino fields do not have a right-handed chirality component, they do not have a Yukawa coupling like the other fermions, therefore neutrinos are massless in the SM. The dynamic component of the scalar field, H, is coupled to the fermions in the same way as the vacuum expectation value, so the coupling of the Higgs boson to each fermionic field is directly proportional to the fermion mass.

The couplings of the $W^\pm$ field to fermions are present in the terms

$$\mathcal{L}_{Wf} = -\frac{g}{\sqrt{2}} \left(\overline{N_{Li}} \not{W}^+ E_{Li} + \overline{U_{Li}} \not{W}^+ D_i + \text{h.c.} \right)$$

By inserting some of the unitary matrices described above and their inverses:

$$\mathcal{L}_{Wf} = -\frac{g}{\sqrt{2}} \left(\overline{N_{Li}} ({V^{E_L}}^\dagger)_{ij} \not{W}^+ V^{E_L}_{jk} E_{Lk} + \overline{U_{Li}} ({V^{U_L}}^\dagger)_{ij} \not{W}^+ V^{U_L}_{jk} D_{Lk} + \text{h.c.} \right)$$

We label the states $V^{E_L}_{ij} N_{Lj}$ that interact diagonally with the charged lepton mass eigenstates via W as ν_e, ν_μ, and ν_τ. These complete the elementary particle content of the Standard Model, which is summarized in Figure 2.2. The up-type quark mass eigenstates interact diagonally via W with $V^{U_L}_{jk} D_{Lk}$, which in general are not the mass eigenstates of the down-type quarks. The matrix that transforms between these bases, $V = V^{U_L} {V^{D_L}}^\dagger$, is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The off-diagonal entries of the CKM matrix, which have all been experimentally measured to have non-zero values, allow interactions between different generations of quarks.

The coupling strength of an interaction is not necessarily a constant, and in general it depends on the energy scale. In perturbative calculations, fields are renormalized so as to cancel out divergent terms, and this introduces an unphysical renormalization scale μ. The dependence of the coupling strength on this parameter is given by the β function:

$$\beta(\alpha) = \frac{d\alpha}{d \ln \mu}$$

where α is related to the coupling g by $\alpha = \frac{g^2}{4\pi}$. By taking μ to be close to the scale of the momentum transfer Q in a given process, we can find the effective strength of an interaction.

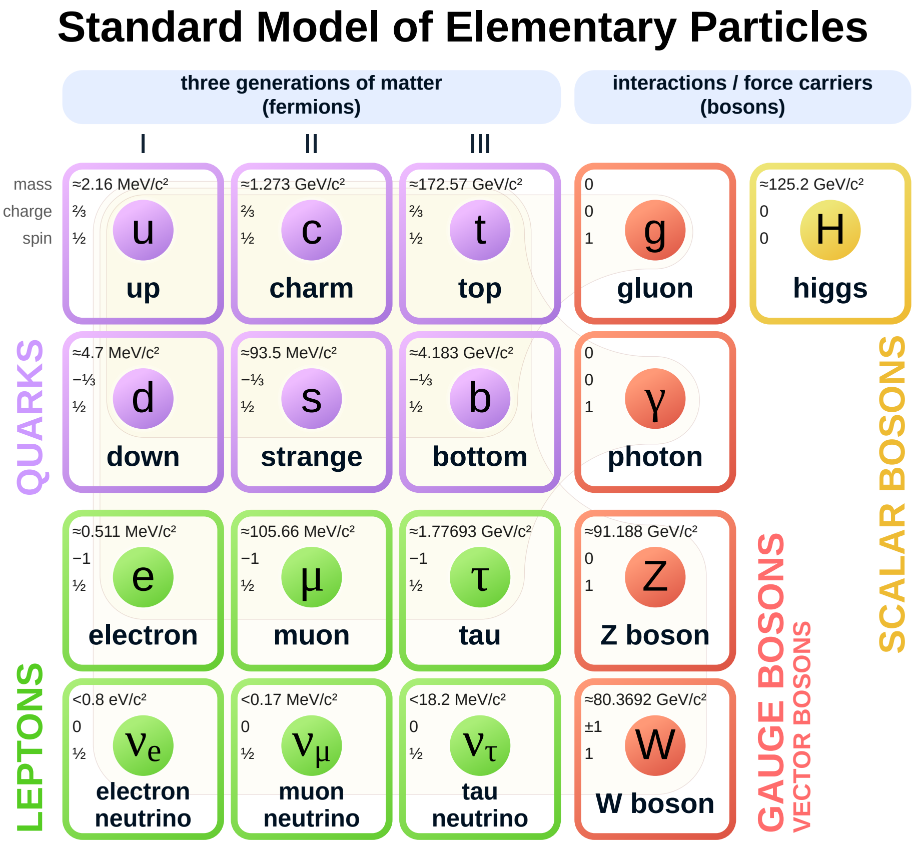


Figure 2.2: Summary of the elementary particles of the Standard Model (Wikimedia Commons, 2024).

For example, the β function of the strong interaction at lowest order is

$$\beta(\alpha_s) = -\frac{1}{2\pi^2}\left(11 - \frac{2n_f}{3}\right)\alpha_s^2 + O(\alpha_s^3)$$

where n_f is the number of quark flavors [21]. Since the observed number of flavors is six, the coefficient of α_s^2 is negative. This means that the strength of the strong interaction is higher at lower energies, as shown in Figure 2.3. Because of this, quarks are confined to composite states called hadrons. Hadrons are combinations of quarks that transform in the trivial representation of the $SU(3)$ symmetry of QCD. That is, hadrons have a net zero color charge. Individual quarks are never detected directly, so hadrons are the only observable states of quarks. However, at high energies, the coupling of quarks decreases and they behave almost as if they were free particles. This phenomenon is known as asymptotic freedom.

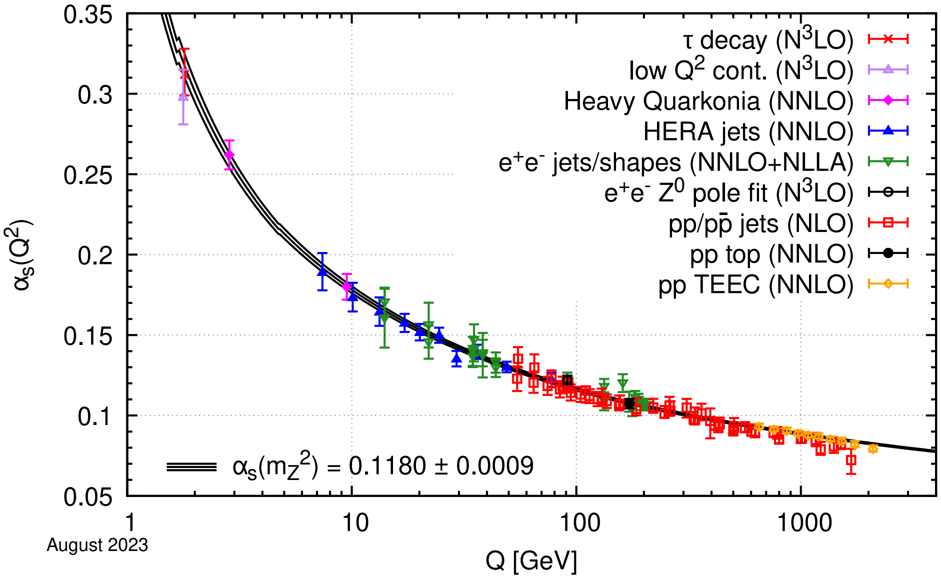


Figure 2.3: Running of the strong coupling α_s as a function of energy scale Q [2]. The colored bars indicate various determinations of the coupling at particular scales, and the band of black lines is calculated by appling the β function to the average of these results.

2.2.4 Limitations

As shown in Figure 2.4, the Standard Model has proven to be a theory capable of accurately predicting rates of processes across several orders of magnitude. However, it is not a complete theory of our universe. The most obvious omission is the absence of a particle-level description of gravity. The SM fields also cannot account for dark matter or the observed contribution of dark energy to the accelerated expansion of the universe, which together are responsible for the vast majority of the total energy in universe, as shown in Figure 2.5. Neutrinos have been found to have a small but non-zero mass, which cannot be directly incorporated into the SM Lagrangian with renormalizable terms. The SM also cannot explain the apparent asymmetry between the amount of matter and antimatter that currently exists. These problems suggest the existence of other fields and their associated particles.

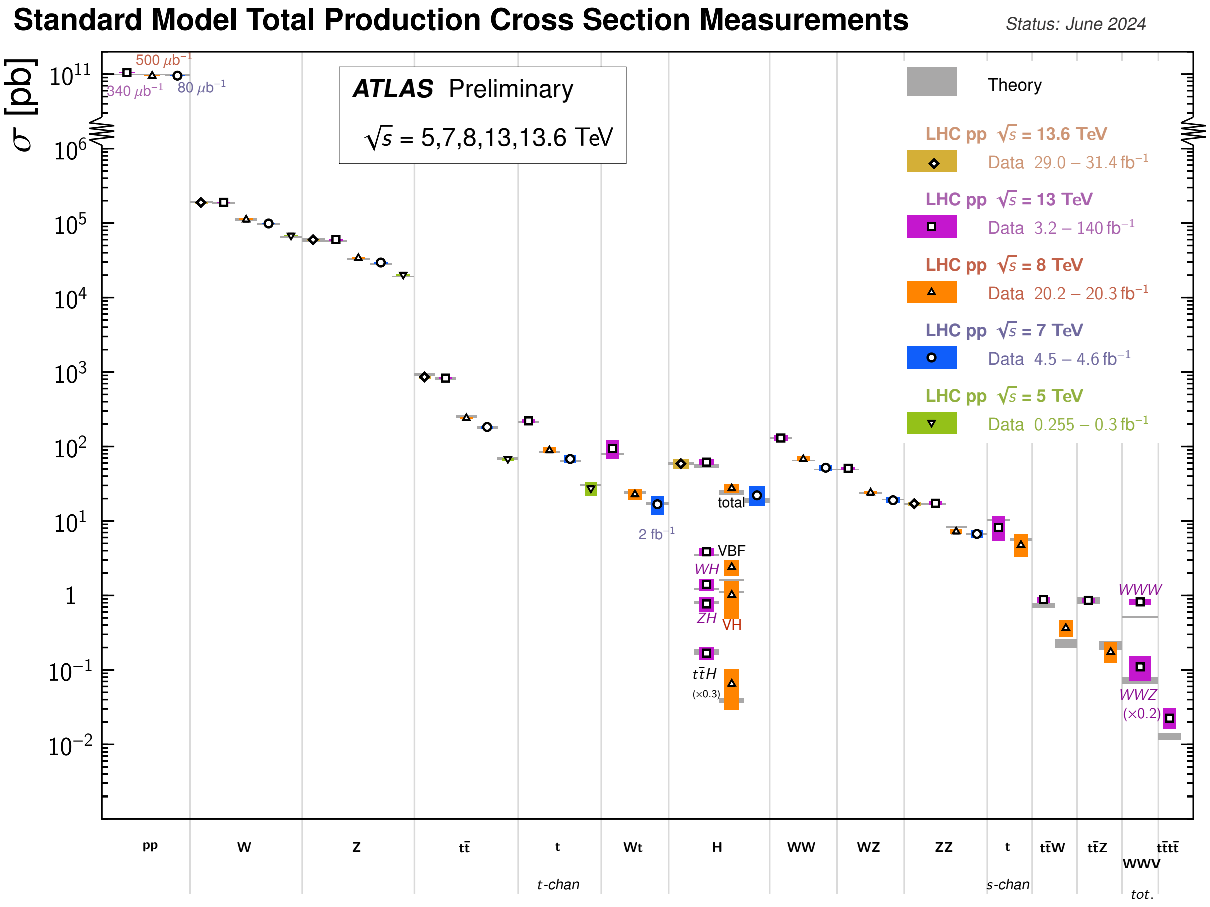


Figure 2.4: Theoretical Standard Model cross section predictions vs. measurements by the ATLAS experiment for the production of many different states [3].

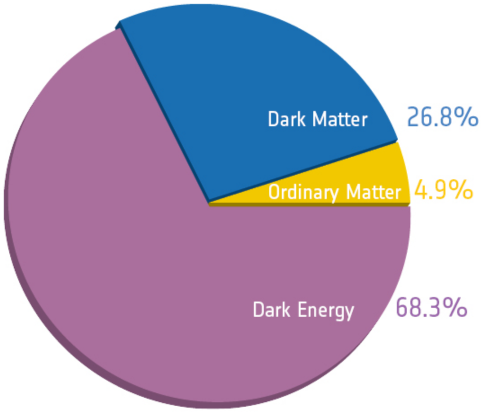


Figure 2.5: Total energy content of the universe by source, based on observations by the Planck space telescope (ESA and the Planck Collaboration, 2013). ”Ordinary matter” represents all energy attributable to the particles of the Standard Model.

2.3 Long-lived particles

This section will consider a class of extensions to the Standard Model that may help resolve the apparent inconsistencies between the SM and observations of our universe. These extensions involve particles with a particular range of lifetimes. Particles can broadly be classified as either stable or unstable. Unstable particles will spontaneously decay into lighter particles after a finite lifetime, while stable particles do not.

2.3.1 Particle lifetimes

Rates of any processes, including decays of unstable particles, are given approximately by the formula known as Fermi's golden rule [22]:

$$\Gamma_{fi} = 2\pi \int |T_{fi}|^2 \delta(E - E_i)\rho(E)\, dE$$

where T_{fi} is the transition matrix element between the initial and final states. The total decay rate is the sum of rates for each allowed decay mode:

$$\Gamma = \sum_i \Gamma_i$$

The branching ratio that describes how often a particle decays via mode i is given by

$$BR(i) = \frac{\Gamma_i}{\Gamma}$$

If we consider a very large, non-interacting ensemble of N identical unstable particles at rest, the rate of change in the number of remaining particles is

$$dN = -N\Gamma dt$$

After integrating and normalizing, we have

$$N(t) = Ne^{-\Gamma t}$$

This is simply N copies of the the probability density function of a decay at time t, which is an exponential distribution. That is, for the proper time t_{proper},

$$P(t_{\text{proper}}) = \Gamma e^{-\Gamma t_{\text{proper}}}$$

The mean lifetime of the particle τ in its rest frame is

$$\tau = \int_0^\infty tP(t)dt = \frac{1}{\Gamma}$$

Figure 2.6 shows the broad range of mean lifetimes of Standard Model particles. Particle species that can have a lifetime long enough that the displacement between production and decay is directly observable are called long-lived particles (LLPs). Examples of LLPs in the Standard Model include the μ lepton and the K_L^0 hadron.

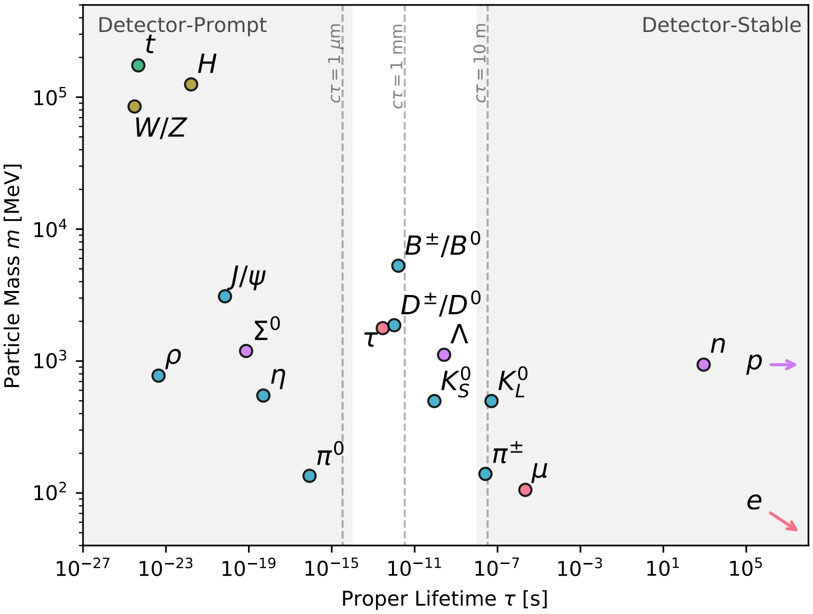


Figure 2.6: Mean lifetime vs. mass for various Standard Model particles [4]. The lighter middle region roughly represents the lifetime range in which current particle physics experiments could directly detect displaced decays.

2.3.2 Models

If there are any additional gauge symmetries that are spontaneously broken by a scalar field as in the electroweak sector, this scalar can mix with the Standard Model Higgs boson [23]. For example, a term such as $\kappa H^2 s^2$ in the Lagrangian, where s is the field for the new scalar particle, will cause mixing between H and s that determines the observable scalar mass eigenstates. If $m_s < m_H/2$, it is possible for the Standard Model Higgs boson to decay to two of these new scalars. If there are no lighter BSM particles that s can decay to, it will decay to SM fermions with the same branching ratios as the SM Higgs. The lifetime of s depends on κ and the potential of the new scalar field. For small mixing angles between H and s, the mean lifetime τ_s can result in macroscopic decay lengths. Figure 2.7 shows a Feynman diagram for such a hidden sector model.

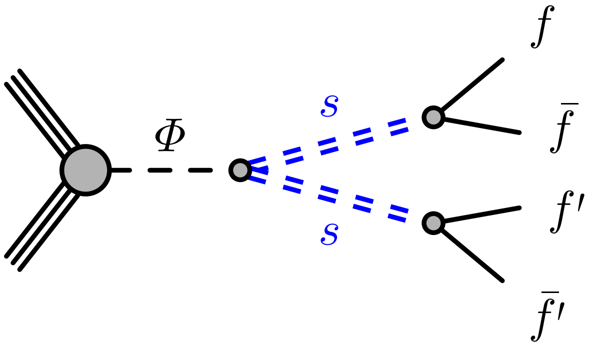


Figure 2.7: Feynman diagram for a generic hidden sector model with a scalar mediator Φ and a long-lived scalar s [5]. For the base model, Φ is the SM Higgs boson H.

While this is a very generic LLP model, there are many specific BSM theories that are motivated by limitations of the Standard Model and predict the existence of new LLPs [24]. Various theories of supersymmetry, dark matter, etc. include LLPs. Nearly any model that allows small couplings or includes approximate symmetries can lead to LLPs. A wide variety of lifetimes are possible, up to $\tau \approx 0.1$ s. This limit is set by observations of light nuclei abundances in the universe, which are sensitive to the delayed injection of energy that would be provided by LLP decays after protons and neutrons formed in the Big Bang [25].

Chapter 3

Experimental apparatus

The European Organization for Nuclear Physics, most often referred to as CERN (for the council that established it, Conseil Européen pour la Recherche Nucléaire), hosts the largest particle physics laboratory in the world. It was formed in 1954 and has since built and operated several particle accelerators and colliders. The CERN laboratory is located on the border between France and Switzerland, near Geneva. Three of the fundamental particles of the Standard Model—the W, Z, and Higgs bosons—were discovered in experiments at CERN. The current flagship project of CERN is the Large Hadron Collider (LHC).

3.1 The Large Hadron Collider

The LHC is the largest and highest-energy collider ever constructed. It is located in a tunnel originally created for the Large Electron–Positron collider (LEP), which operated from 1989 to 2000 [26]. The tunnel is 27 km long and 3.8 m in diameter, ranging in underground depth between 50 m and 175 m [7]. The path of the tunnel is approximately circular, consisting of curved arcs connecting several long straight sections. Figure 3.1 illustrates the location of LHC tunnel, which was engineered to fit between Lake Geneva and the Jura mountain range, with an inclination of 1.4% to keep the average depth at around 100 m. Figure 3.2 shows the layout of the LHC ring. Each of the long straight sections is numbered, starting with the southernmost, which is called point 1, and increment clockwise around the LHC. The beams are made to collide at points 1, 2, 5, and 8, and these locations house the four large experiments of the LHC. The remaining points are reserved for equipment used to prepare and manage the beams.

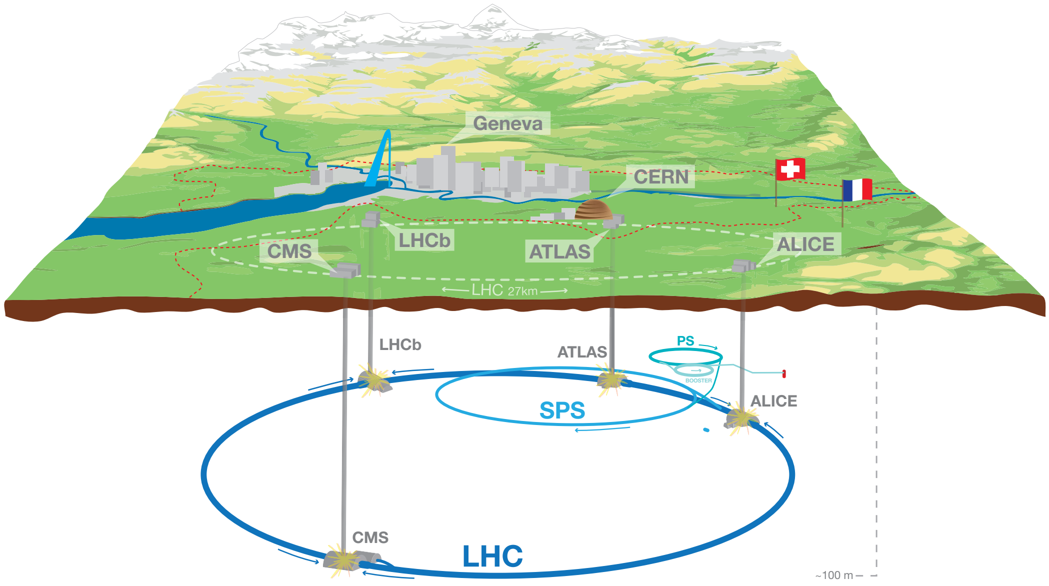


Figure 3.1: Illustration of the location of the Large Hadron Collider [6].

Particles are injected into the LHC from a series of smaller accelerators. Figure 3.3 shows the full system of particle accelerators at CERN. This chain of accelerators was incrementally built over the decades of CERN's history, with each new accelerator piggybacking on the infrastructure of the previous ones to reach higher and higher beam energies. For beams of protons, the most common mode of operation, a linear accelerator extracts protons from ionized hydrogen gas and injects the accelerated protons into the Proton Synchrotron Booster, which in turn injects into the Proton Synchrotron (PS) and then the Super Proton Synchrotron (SPS). Up to 2018, the LINAC2 was used as the initial linear accelerator. The proton source was then upgraded to use LINAC4, which has an increased beam energy and intensity. For beams of heavier ions such as lead nuclei, LINAC3 and the Low Energy Ion Ring (LEIR) are used to inject into the PS and from there to the SPS. In all modes of running, the final stage of injection is from the SPS into the LHC.

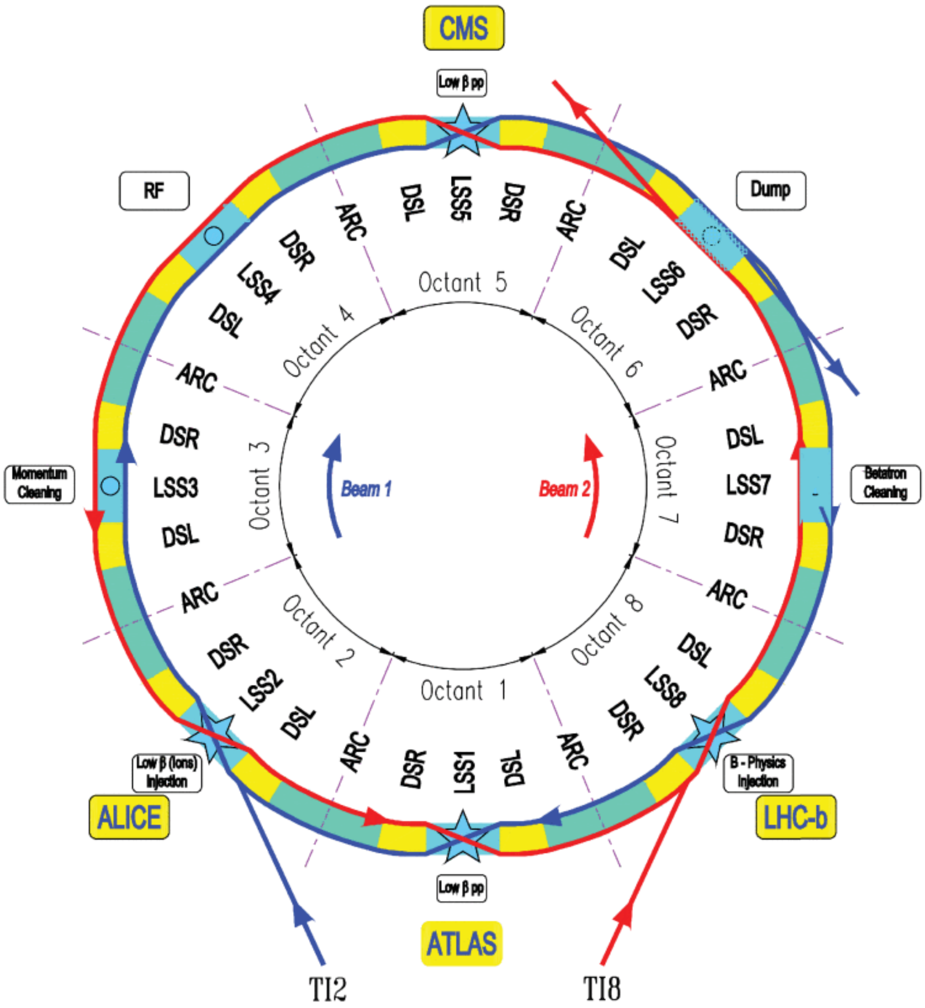


Figure 3.2: Layout of the arcs and insertion regions of the LHC (not to scale) [7].

Injection of particles occurs near points 2 and 8 for beam 1 and beam 2, respectively. These are at the ends of the transfer lines from the SPS, and a kicker magnet inserts the bunches into the LHC ring.

The LHC has two countercirculating beams in separate beam pipes. The beams are separated by 19 cm through most of the LHC, other than near the collision points. Each beam can be filled with up to 2808 bunches of particles. During proton operation, a nominal bunch consists of about 10^{11} protons. Bunches are spaced by 25 ns, and the bunch length is 1 ns. Traveling at a tiny fraction below the speed of light, the beam completes 11000 revolutions around the LHC per second.

The majority of the LHC ring consists of 1232 superconducting dipole magnets, which are responsible for bending the beams through the arcs. Each dipole magnet is 15 m long and has a mass of about 28

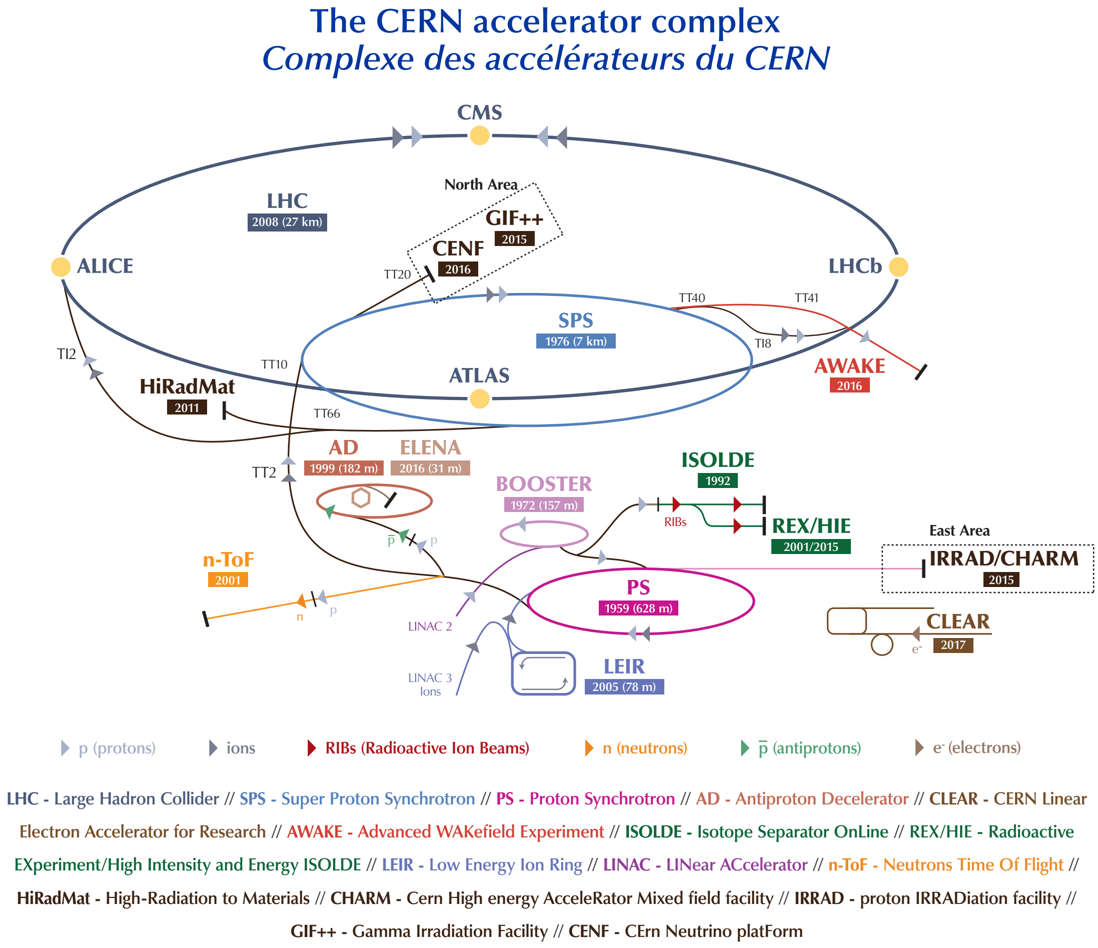


Figure 3.3: Diagram of the full CERN accelerator complex in 2018 [8]. LHC protons began in LINAC2 and passed through the PS Booster, the PS, and the SPS before being injected into the LHC ring. Heavy ions are accelerated by LINAC3, LEIR, and the SPS for injection into the LHC.

tons. Figure 3.4 shows some of the dipole magnets and a schematic of their cross section. In addition, there are quadrupole magnets for focusing the beams and a variety of types of magnets for beam cleaning. Most of the LHC magnets are made out of cables containing strands of niobium–titanium, an alloy that is superconducting below about 10 K, embedded in a copper matrix. These cables are held in place by a rigid non-magnetic collar, which is surrounded by a magnetic yoke. This full assembly is bathed in liquid helium, cooled to 1.9 K. At this temperature, helium is a superfluid. With essentially zero viscosity, the helium is extremely efficient in penetrating the magnets and transporting heat out of them.

Longitudinal acceleration of particles in the LHC occurs at point 4, where there are eight superconducting

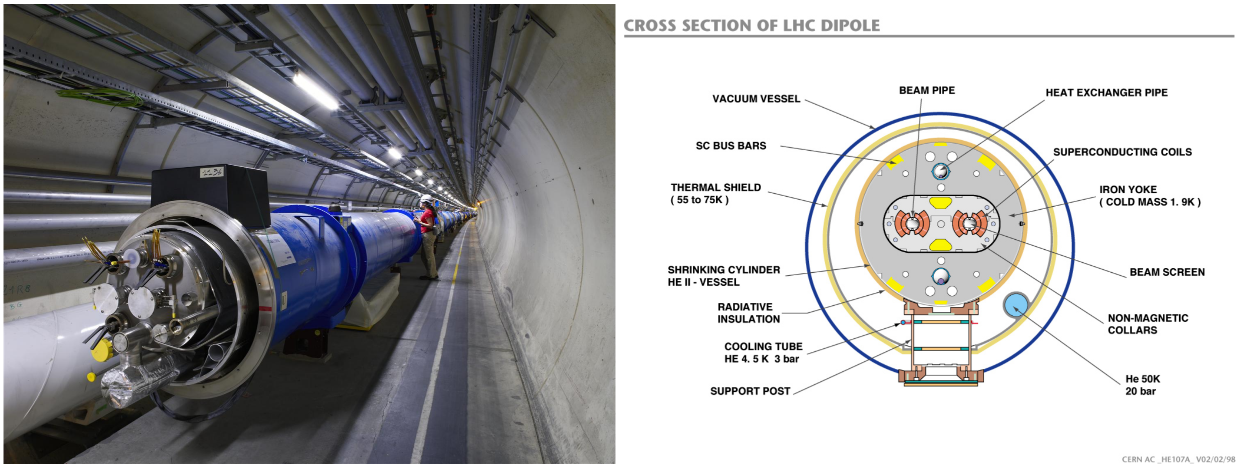


Figure 3.4: Left: View of dipole magnets in the LHC tunnel during installation [9]. Right: Cross sectional schematic of an LHC dipole magnet [10].

radio frequency (RF) cavities per beam. The cavities use niobium as the superconducting material, deposited on copper by sputtering. They produce electromagnetic waves timed in with the bunches so that they experience forward acceleration from the electric field such that the beam energy increases by 485 keV for every turn through the LHC ring. These cavities operate at 400 MHz and are powered via waveguides by a 300 kW klystron for each cavity. Klystrons use electron beams and resonance cavities to amplify RF waves.

There are four points around the LHC ring where the beams are made to collide. These are called interaction points (IPs). Each IP has an large-scale experiment surrounding the point of collision. These experiments are ATLAS (IP1), ALICE (IP2), CMS (IP5), and LHCb (IP8). ATLAS and CMS are general-purpose detectors with nearly full solid angle coverage designed to identify all SM particles other than neutrinos that are produced in the collisions. ALICE is specialized to study collisions of heavy ions and the very high multiplicity of charged particles that are produced in such events. LHCb is specifically designed to examine the decays of hadrons that include b quarks. Each IP has a set of separation/recombination dipole magnets that bend the beams so that they merge into a single beam pipe and cross and thus collide in these detectors. A set of quadrupole magnets on each side of the IP called an inner triplet focuses the beam to reach a minimum transverse size at the point of collision. This optimizes the instantaneous luminosity, which can reach up to about $2 \cdot 10^{34}$ cm^{-2} s^{-1} in ATLAS and CMS.

A facility at point 6 is used to remove the beams from the LHC ring. Ejection of the bunches requires a system of fast-pulsed magnets for each beam. The extraction magnet directs the beam out of the main ring and towards a beam dump area. Since this magnet requires some time to ramp up, there must always be a gap in the beam with no bunches. This abort gap is set to 3 µs, so the extraction magnet needs to be able to reach nominal field strength within that time. Then a set of 10 dilution magnets for each beam sprays the bunches in an “e”-shaped pattern so that the energy is distributed across a large enough area so as to not penetrate the full length of the beam dump target. The beam dump target is a cylinder of graphite, 8 m long and 0.7 m in diameter, encased in a steel jacket and surrounded by 900 tons of concrete shielding.

Nominal operation of the LHC during periods of high-energy collisions follows a cycle of several stages. The first step is injection of beam into the two rings of the LHC. This occurs at a beam energy of 450 GeV. A single pilot bunch of low intensity ($5 \cdot 10^9$ protons) is circulated in each beam pipe in order to establish that conditions are safe for full-intensity beams and to measure and corerct any beam parameters as necessary. Next, high intensity physics bunches are injected, around 10^{11} protons. Each ring can accomodate up to 2808 bunches.

Once injection has been completed, the LHC enters its ramp sequence. During ramping, the beam energy is increased by acceleration of the bunches in each pass through the RF cavities. Over approximately 20 minutes, the energy is increased from 450 GeV to the collision energy (3.5–6.8 TeV, depending on the year). During this time, the current in the magnets of the LHC is synchronously increased to maintain the trajectory of the protons as they increase in energy.

After the beams have reached their final energy, the emittance of beams at the interaction points is reduced by the inner triplet quadrupole magnets, referred to as squeezing the beams. This reduces the cross sectional area of the beam, so that the density of protons increases.

Up to this point in the LHC cycle, the beams are intentially separated from each other. Finally, the beams adjusted so that they intersect at the four interaction points, which initiates proton-proton collisions. After a short configuration and validation period to ensure the beams and collisions are within the desired parameters, stable beams are declared. This indicates that the experiments around the ring can safely enter their nominal configuration and begin taking collision data that is usable for physics analyses. During stable beams, the rate of collisions can be tuned by slight adjustments in the beam emittance or in the spatial offset between the centers of the two beams. The highest rate is obtained by minimizing the emittance and colliding the beams head on.

Each collision reduces the number of protons remaing in the beam, so the rate of collisions decreases over time. Eventually the rate decreases enough that it is more efficient to refill the LHC with fresh bunches. When this point is reached, or if any problem with the beam is detected, the beams are dumped. The LHC magnets are ramped down to the current needed at the injection energy, and then operation cycle repeats.

On a broader level, the LHC operates on a schedule of alternating periods of several years during which the accelerator is either nominally running most of the year or completely shut down. These running periods and long shutdowns are sequentially numbered. Run 1 occured from 2010 to 2012, followed by Long Shutdown 1 (LS1), which was during 2013–2014. The LHC operated at a beam energy of 3.5 TeV during 2010–2011 and 4 TeV in 2012. Run 2 spanned 2015 to 2018, during which the beam energy was increased to 6.5 TeV. LS2 lasted from 2019 to 2021, and Run 3 has been ongoing since 2022 and is currently scheduled to end in 2026. The beam energy in Run 3 is 6.8 TeV. Each long shutdown period is used to upgrade parts of the accelerator and the experiments at the interaction points. LS3 will include the most extensive upgrades, such that the program scheduled to start in 2029 (for Run 4 and beyond) is called the High-Luminosity LHC (HL-LHC), since the instantaneous luminosity delivered to the IPs will increase by a factor of 2–4 relative to Run 3. The HL-LHC is expected to be operational until at least the 2040s. The long-term LHC/HL-LHC running schedule is shown in Figure 3.5.

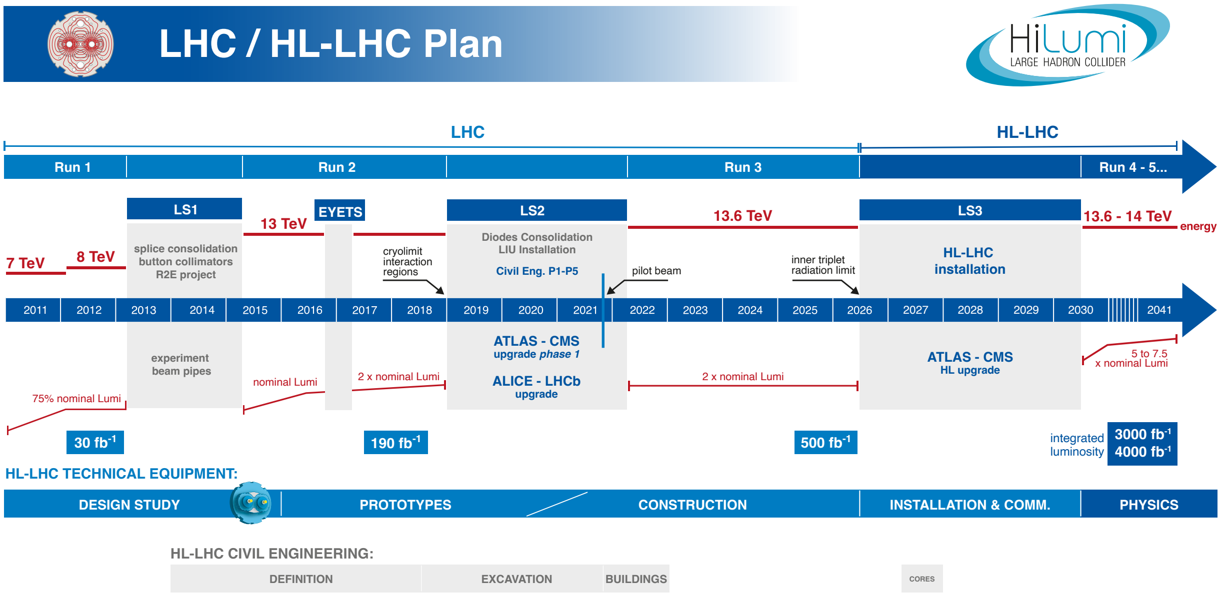


Figure 3.5: Schedule of LHC and HL-LHC operation as of January 2025.

3.2 The ATLAS detector

The first interaction point of the LHC houses the ATLAS detector. ATLAS is the largest of the LHC experiments and is designed as a general-purpose particle detector with nearly full 4π solid angle coverage around the point of collisions. It consists of an inner tracker surrounded by a solenoid magnet, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating toroidal magnets. The detector systems can generally can be classified as part of the "barrel", with layers of components in cylinders along the beam axis centered at the IP, or the "endcap", with components lying in transverse planes towards the ends of the detector in the beam axis directions. The overall layout is shown in Figure 3.6. ATLAS has contributed to many state-of-the-art results from the LHC, including the discovery of the Higgs boson in 2012. The ATLAS collaboration consists of about 5900 members across 183 institutes in 41 countries.

The ATLAS detector is located at point 1 of the LHC in a cavern that is 93 m underground. The cavern is 53 m long, 30 m wide, and 35 m tall. It is lined with reinforced concrete that is 1–2 m thick in the walls and ceiling and 5 m thick on the floor. It has two access shafts with diameters of 18 m and 13 m. The detector was constructed in place within the cavern by lowering components down these shafts. Figure 3.7

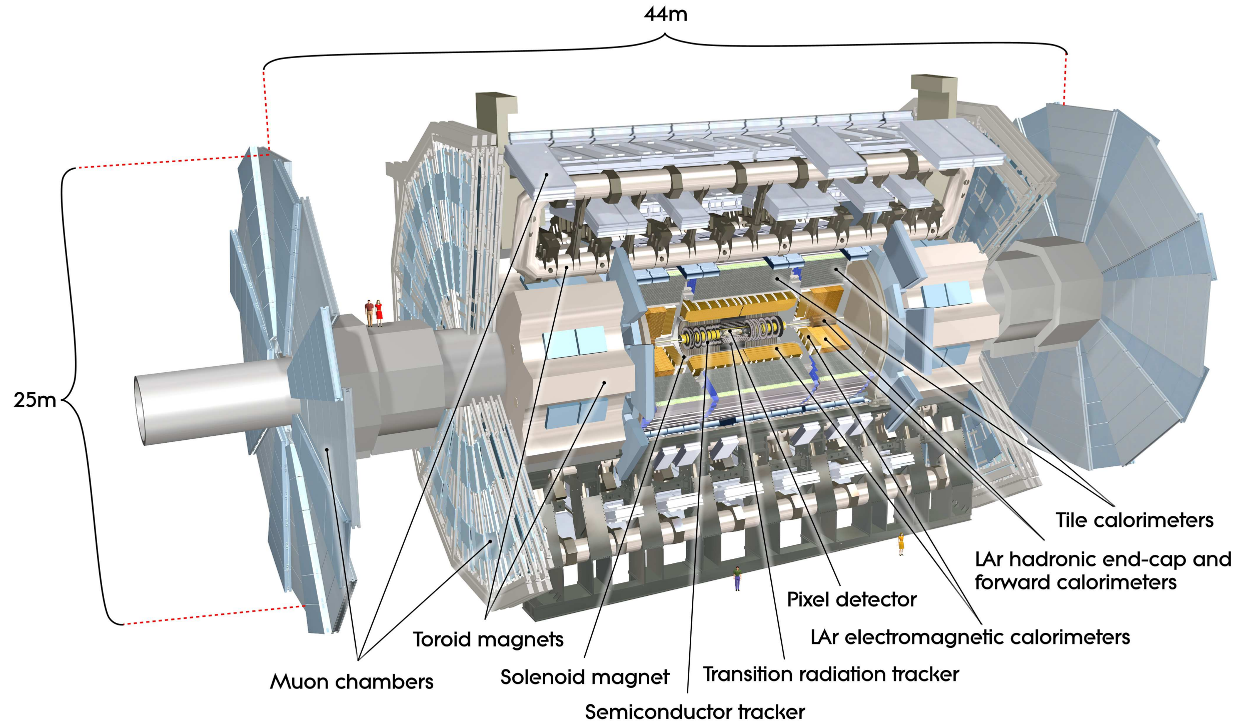


Figure 3.6: Cut-away view of the ATLAS detector.

shows the cavern just before installation of the detector. Work on the cavern was completed in 2003, and construction of the initial detector finished in 2008. Significant upgrades to detector components were made in LS1 and LS2, and the largest upgrade will occur during LS3.

ATLAS uses a right-handed coordinate system with its origin at the IP in the center of the detector and the z-axis along the beam pipe. The x-axis is horizontal and points from the IP towards the center of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, ϕ) are used in the xy plane, with ϕ being the azimuthal angle from the x-axis. Pseudorapidity η is defined in terms of the polar angle θ from the z-axis as $\eta = -\ln\tan(\theta/2)$ and is equivalent to rapidity $y = \frac{1}{2}\ln\left(\frac{E+p_z c}{E-p_z c}\right)$ in the ultrarelativistic limit $E \gg mc^2$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$.

3.2.1 Inner detector

The inner detector (ID) is composed of three subdetectors: the pixel detector, the semiconductor tracker, and the transition radiation tracker. Each of these employs a different type of hardware for tracking the trajectories of charged particles. Figure 3.8 shows an overview of the inner detector.

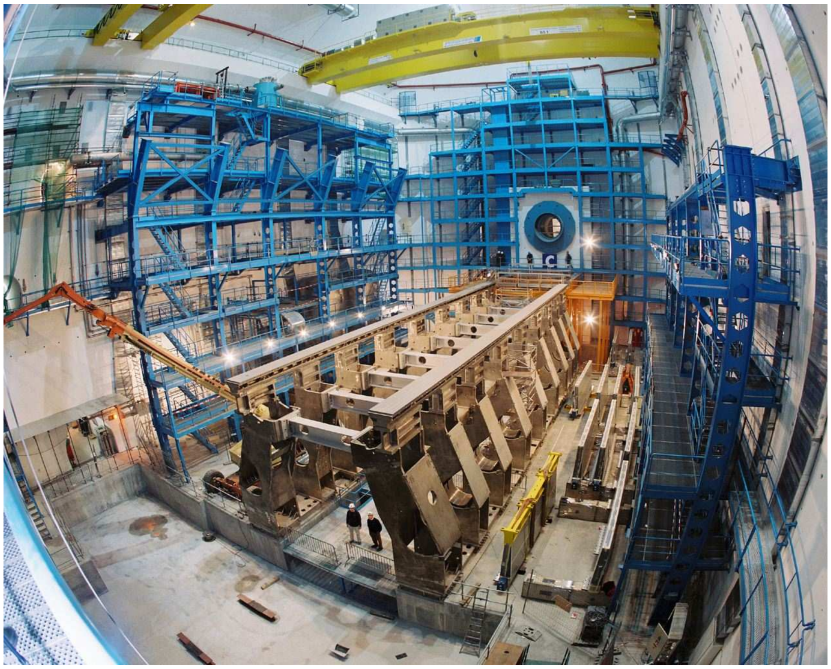


Figure 3.7: ATLAS cavern prior to installation of the detector.

The first two subdetectors operate based on p–n junctions in silicon with a reverse bias voltage applied. The positive carriers (holes) are attracted to the low-voltage end, and the negative carriers (electrons) are drawn to the high-voltage end. This leaves a large depletion region in between with a low density of charge carriers, and thus increases resistance and decreases current. When a charged particle from the collisions passes through the silicon and ionizes the atoms in it, electron–hole pairs are created and these new charge carriers drift to opposite sides of the material. This process induces a measurable current, which indicates a hit in the sensor.

The innermost part of the detector is made up of silicon pixel sensors. The pixels have the best resolution of all the ATLAS detectors. This resolution is needed for the high density of tracks in the volume closest to the point of collisions. The initial pixel system installed before Run 1 included three layers of sensors in both the barrel and endcap that are 50 µm × 400 µm. A fourth barrel layer, the insertable B layer (IBL) was inserted during LS1 between Run 1 and Run 2 and is the tracking layer closest to the beams at r = 33 mm. The IBL pixels are 50 µm × 250 µm. In total, there are 92 million pixels installed in ATLAS. Some of the pixel detector layers can be seen in Figure 3.9

The semiconductor tracker (SCT) is composed of silicon microstrip sensors. The principles of operation

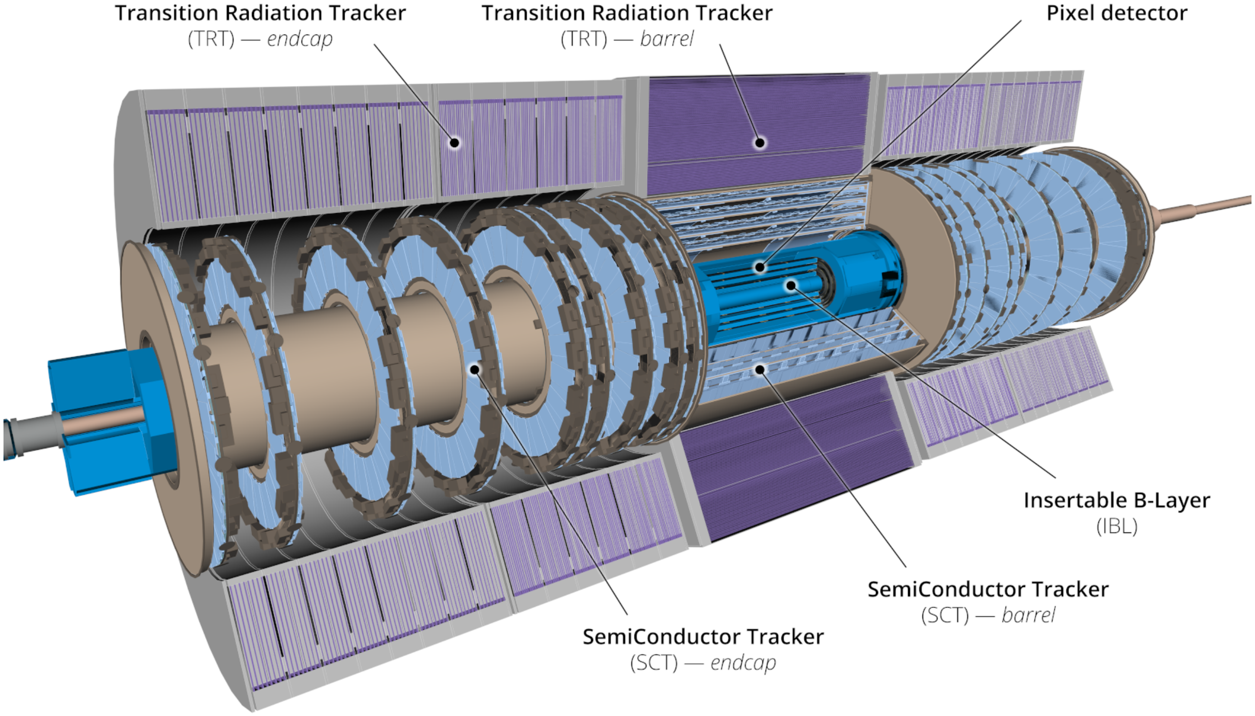


Figure 3.8: Cut-away view of the ATLAS inner detector.

are similar to that of the pixel sensors, but strip sensors are much longer in one dimension. Each readout channel corresponds to a strip that is 80 µm × 12 cm. The SCT modules include two layers of sensors that are rotated 40 mrad relative to each other in order to improve the resolution in the direction along the strips. There are a total of 6.3 million SCT channels. Example SCT components are shown in Figure 3.10.

The Transition Radiation Tracker (TRT) is based on drift tubes made of polyimide straws filled with a xenon-based gas mixture. In addition to providing tracking hits, the TRT is designed to be sensitive to transition radiation, which is emitted when a relativistic charged particle passes between two media with different dielectric constants. Measuring the transition radiation helps differentiate electron tracks from those of other types of particles, since the effect depends on the Lorentz factor $\gamma = E/m$. The straws are made from two layers of a 35 µm film composed of polyimide, aluminum, and graphite. They are 4 mm in diameter with a length of 144 cm in the barrel and 37 cm in the endcaps. In the center of each straw is a 31 µm diameter gold-plated tungsten wire, which is connected to the frontend electronics. The wires are grounded and the straw walls are kept at -1.5 kV. The straws are filled with a mixture of 70% Xe, 27% CO_2,

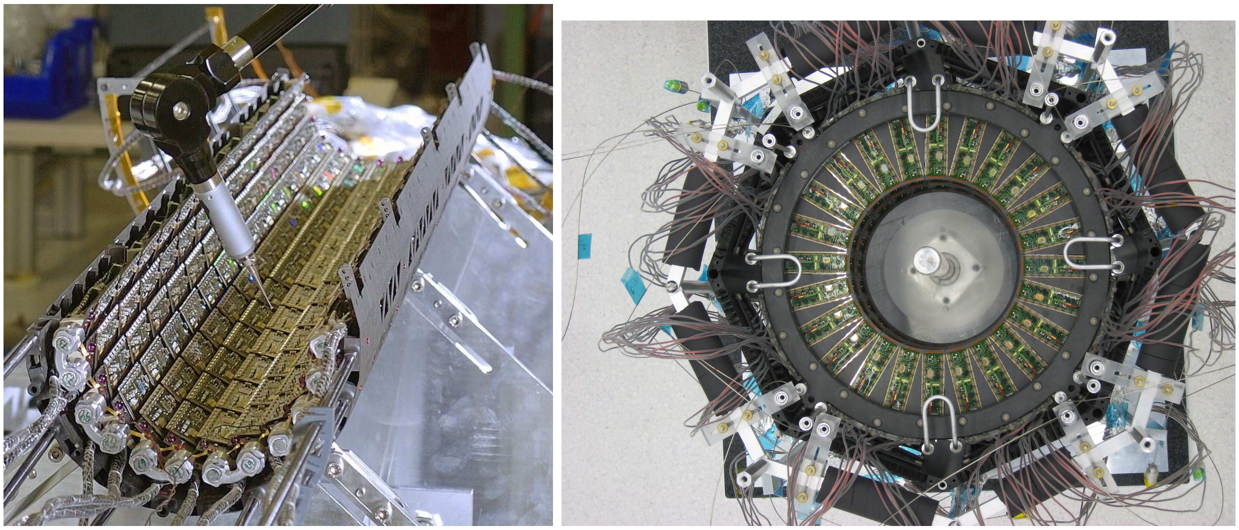


Figure 3.9: Half of a pixel barrel layer (left) and one completed pixel endcap (right). The layout of the individual modules can be seen wrapping around where the beam pipe would be. For the endcap layers, the other half of the pixel modules are on the back side of the disks.

and 3% O_2. Charged particles passing through the straw ionize the gas inside, and the liberated charges are collected on the wire and straws. The ionization charge is amplified by a factor around 2.5 $\times 10^4$ while drifting through the gas before the current is read out by the frontend. Transition radiation photons are absorbed by the gas mixture, which greatly increases the signal amplitude via increased ionization.

The barrel portion of the TRT is divided into three rings of 32 modules each. Each module has a carbon fiber shell and an array of straws embedded in a matrix of polypropylene fibers. There are a total of 73 cylindrical layers of straws in the barrel, and the straws are oriented parallel to the beam pipe. The TRT endcap consists of 20 wheels containing a total of 160 straw planes, with 768 radially-oriented straws in each plane. The space between endcap straw layers is filled with layers of polypropylene. The polypropylene material in both the barrel and the endcap increases the amount of transition radiation photons produced. Parts of the TRT can be seen in Figure 3.11.

Surrounding the inner detector systems (pixel, SCT, and TRT) is a solenoid that provides a 2 T axial magnetic field. This field bends the trajectories of charged particles, which allows for momentum measurements when the tracks are reconstructed, since particles of higher momentum have straighter tracks. This magnet is pictured in Figure 3.12. The solenoid is 5.8 m long, 2.5 m in diameter, and 0.1 m thick. It contains a coil of NbTi that is cooled to 4.5 K by liquid helium. The nominal current is 7.73 kA, which can be charged

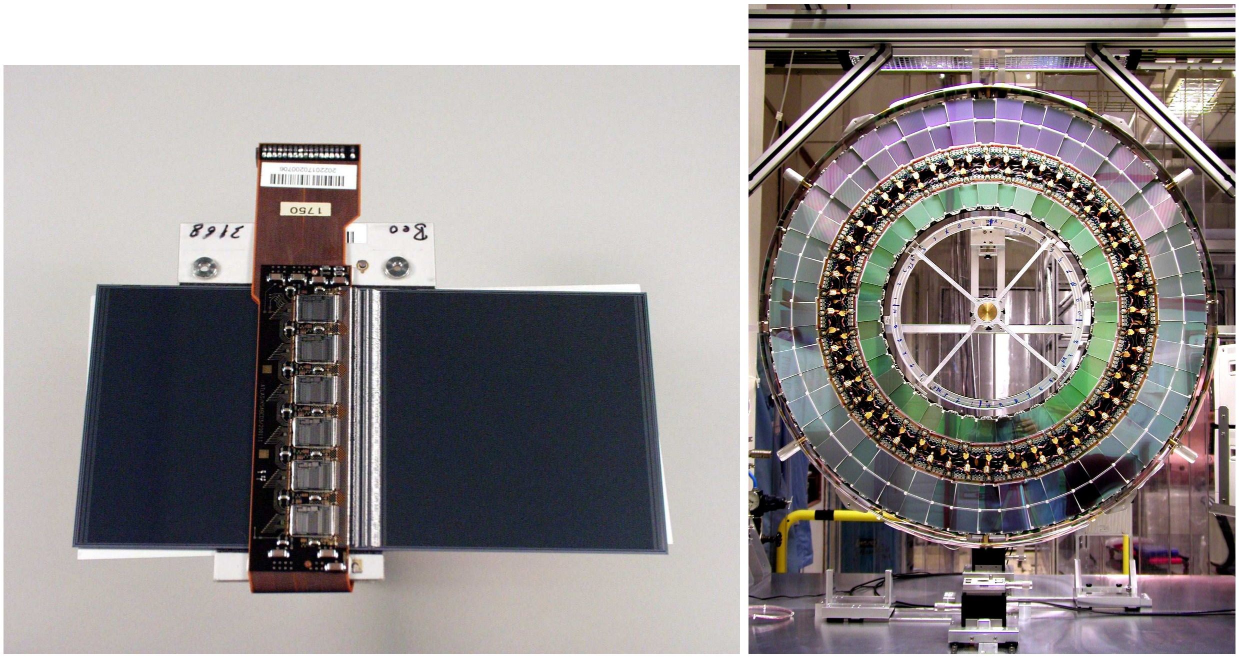


Figure 3.10: SCT barrel module (left) and an SCT endcap disk (right). The barrel modules are arranged in cylindrical layers around the beam pipe in a pattern similar to the pixel barrel modules. The endcap disk has a middle ring of modules on the other side, between the outer and inner rings shown here. For geometrical acceptance reasons, not all SCT endcap disks have all three rings of modules.

or discharged in about 30 minutes.

3.2.2 Calorimetry

The middle layer of the ATLAS detector, surrounding the inner detector, is the calorimeter system. An overview of this system is shown in Figure 3.13. The calorimeter is made up of dense material to absorb energy of particles passing through it interspersed with material that is sensitive to this energy deposition. Different parts of the calorimeter vary in which materials are used for these roles. Almost all Standard Model particles are stopped before reaching the outer edge of the calorimeter, and thus their initial energy can be determined by summing all of their associated energy deposits. Generally only muons and neutrinos are able to escape the calorimeter. The inner layers of the calorimeter, collectively called the electromagnetic calorimeter, are primarily designed to measure the energy of photons, electrons, and positrons. The outer layers are called the hadronic calorimeter and are designed to measure the energy of hadrons, which are more massive and do not lose much of their energy in the electromagnetic calorimeter.

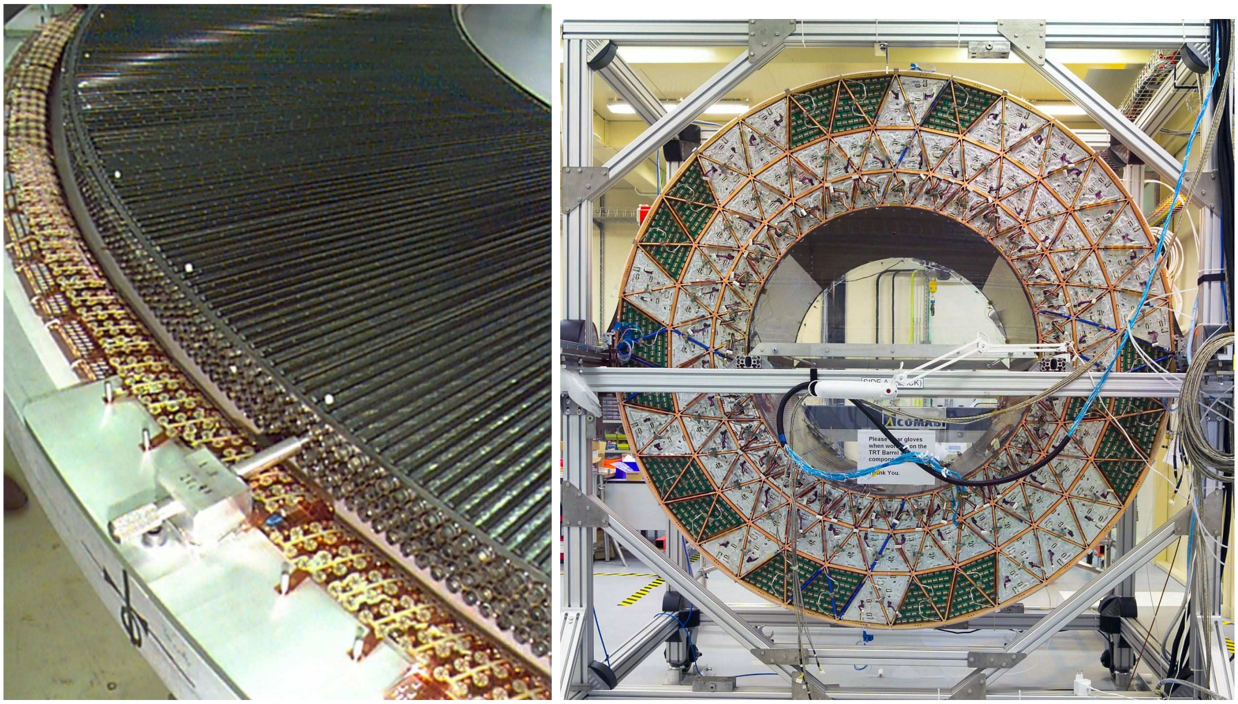


Figure 3.11: Left: A four-plane TRT endcap unit during assembly with the radial straw tubes visible. Right: End view of the TRT barrel.

The electromagnetic calorimeter uses liquid argon (LAr) as the active detector medium that is sensitive to energy deposits. In the electromagnetic barrel (EMB) and electromagnetic end cap (EMEC) calorimeters, the absorbers are lead plates. The EMB covers the region $|\eta| < 2.5$, and the EMEC covers $2.5 < |\eta| < 3.2$. For both, the modules have an accordion geometry as shown in Figures 3.14 and 3.15. This geometry provides full coverage in ϕ without any cracks. The thickness of the lead plates is 1–2 mm, and the gaps between them are 1–4 mm, depending on the location. Three layers of copper electrodes are positioned in the middle of these gaps, separated by insulating polyimide sheets. The outer two layers are maintained at 2000 V, and the inner layer provides readout of the charge collected from the liquid argon by capacitive coupling.

The barrel portion of the hadronic calorimeter uses steel plates as the absorber and scitillating plastic tiles as the active medium. This is called the tile calorimeter. The tile calorimeter is divided into three parts: a central barrel and two extended barrels. The central barrel is 5.8 m long and the extended barrels are each 2.6 m long. Along the whole length, the inner and outer radii are 2.28 m and 4.25 m, respectively. Each barrel consists of 64 wedge-shaped modules. The geometry of the tile calorimeter modules and overall layout can be seen in Figure 3.16. The absorber steel is made up of 5 mm thick master plates and 4 mm

Figure 3.12: Central solenoid magnet.

spacer plates that are glued together to form the pockets where the scintillators fit. The scintillating tiles are 3 mm thick and vary from 97 mm to 187 mm in the radial dimension and from 200 mm to 400 mm in the azimuthal dimension, depeding on the radial position. Wavelength-shifting fibers transmit the scintillation light through the steel girder at the outer radius where the readout electronics are located.

The hadronic endcap calorimeter (HEC) covers the region $1.5 < |\eta| < 3.2$. The absorbing material is copper, and liquid argon is used for the detecting medium. From the persepective of the IP, it is just behind the electromagnetic endcap calorimeter, with which it shares a cryostat. Each endcap of the HEC is composed of two wheels, both with a outer radius of 2 m and segmented into 32 modules. The inner radius increases from 370 mm at the front to 480 mm at the back. The modules consist of flat copper plates of 12.5–50 mm thickness with 8.5 mm gaps between them filled with liquid argon. Each LAr gap contains three evenly spaced electrode layers. The central layer is used for signal readout, and the electrode layer on either side is maintained at high voltage (1800 V). Figure 3.17 shows the structure of an HEC module and one of the endcaps after completing assembly.

The final component of the calorimeter system is the forward calorimeter (FCal), which covers the region $3.1 < |\eta| < 4.9$ and uses liquid argon as the active medium. The FCal units on each end share the cryostat of the electromagnetic and hadronic endcap calorimeters. The FCal begins at about 4.7 m from the IP and

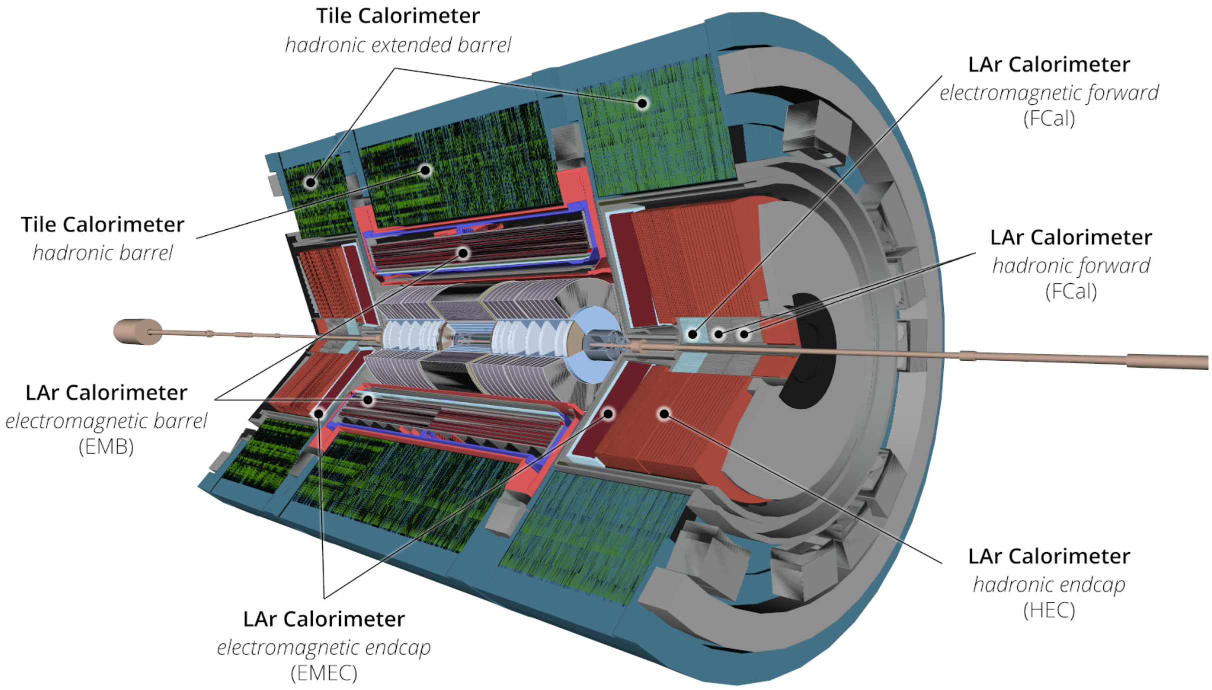


Figure 3.13: Overview of the calorimeter system.

is subject to the most intense particle fluxes of the calorimeter system due to being so close to the beamline. The FCal is segmented into three modules: an electromagnetic module (FCal1) and two hadronic modules (FCal2 and FCal3). Each module is 450 mm long in the z dimension.

FCal1 consists of solid copper rods inside of copper tubes, parallel to the beam pipe, that pass through holes drilled into stacked copper sheets. There is a 0.27 mm gap between the tubes and the rods that is filled with LAr. The rods and tubes form the electrode structures for high voltage and signal readout.

The hadronic modules (FCal2 and FCal3) use tungsten as the absorber material instead of copper. They have a similar electrode structure to the electromagnetic module, except that the rods are made out of tungsten. The LAr gap is 0.4 mm for FCal2 and 0.5 mm for FCal3. Instead of the copper sheet matrix, each tube is surrounded by tungsten slugs. There are only two copper sheets for each hadronic module, as end-plates on the front and back. Figure 3.18 shows a schematic of the hadronic FCal tube structure and a picture of the completed FCal assembly for one side.

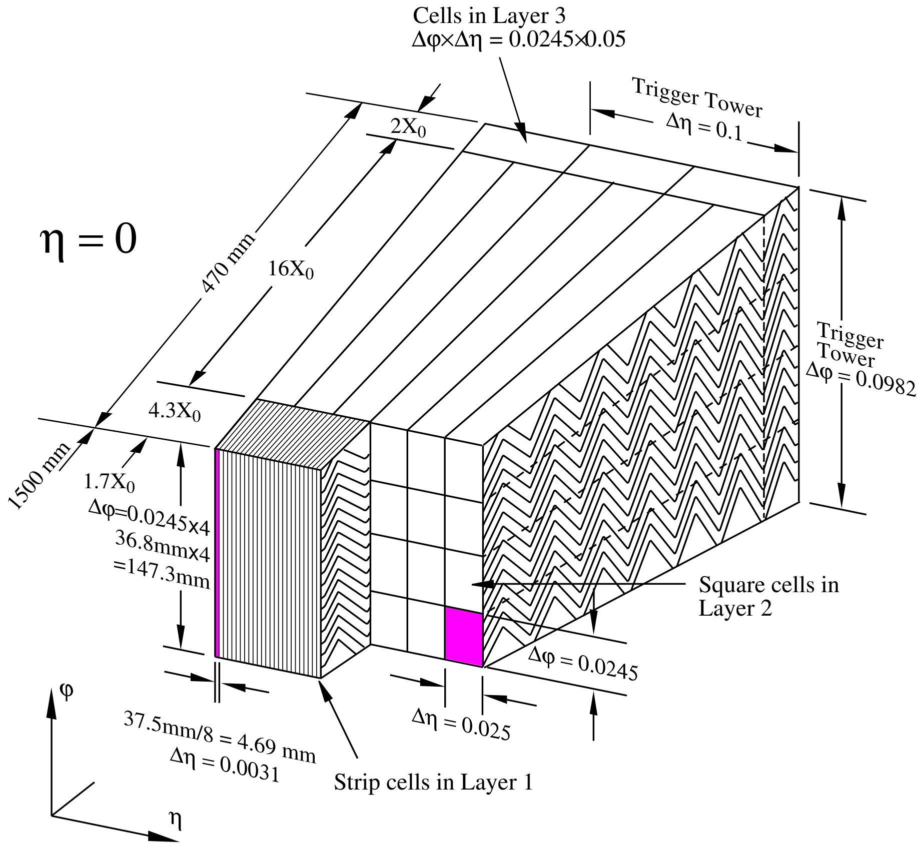


Figure 3.14: Schematic of the electromagnetic barrel calorimeter structure.

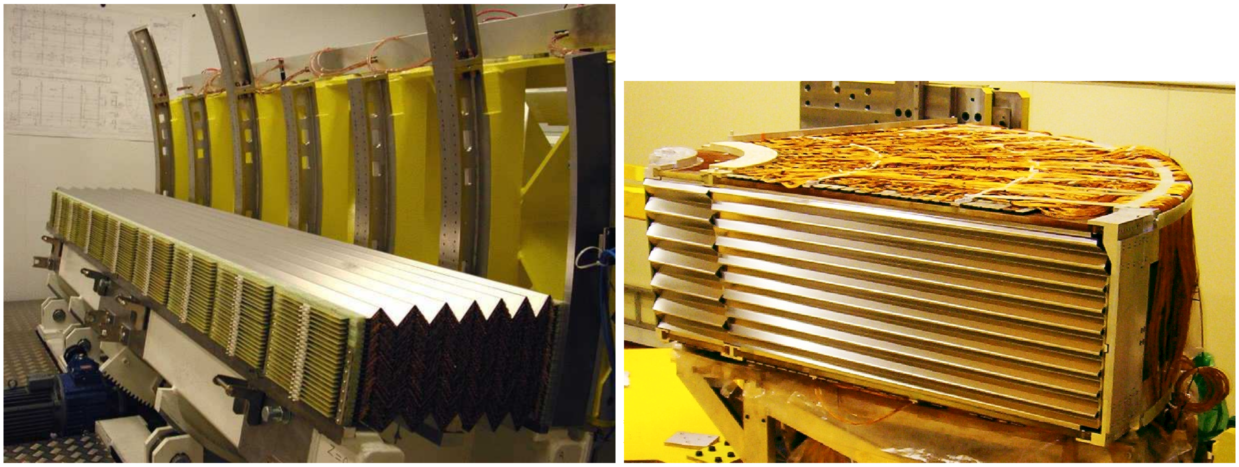


Figure 3.15: Electromagnetic barrel (left) and endcap (right) calorimeters during assembly.

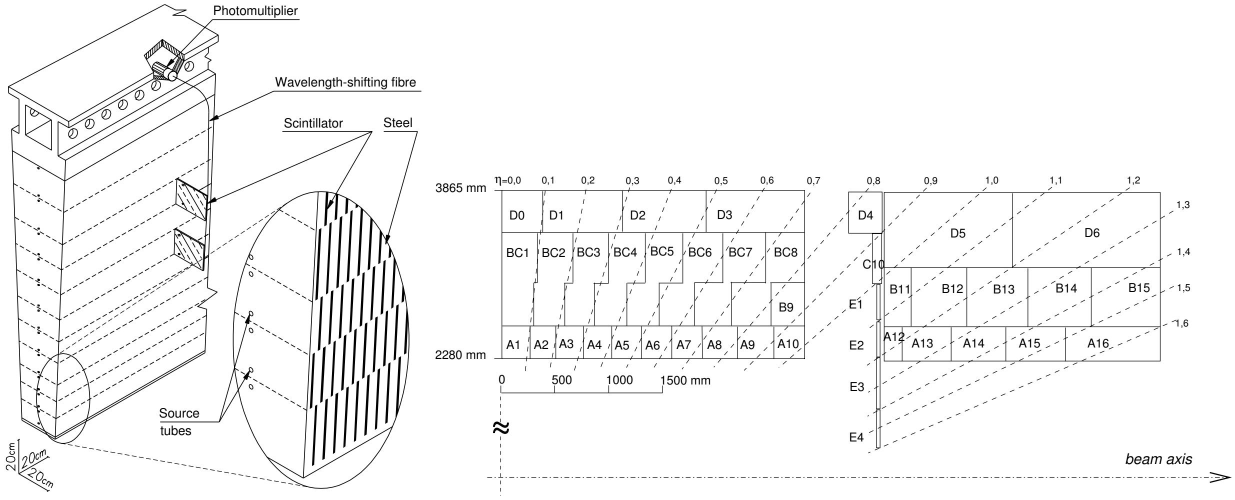


Figure 3.16: Left: Schematic of a tile calorimeter submodule. Right: Segmentation of the tile calorimeter readout for the central barrel and the extended barrel.

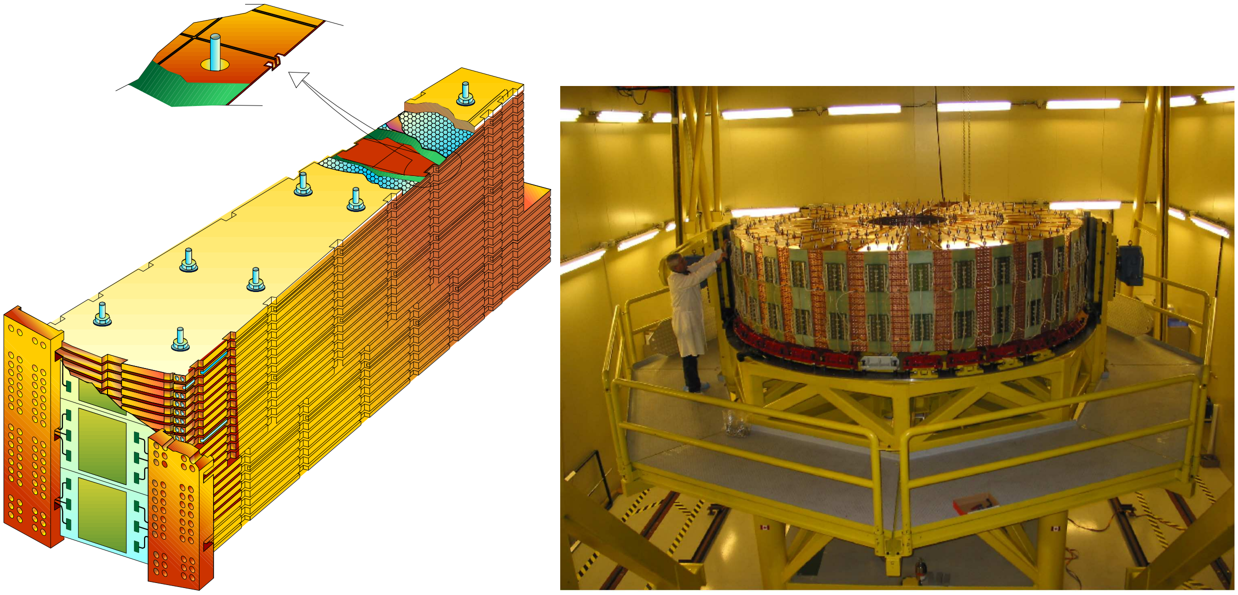


Figure 3.17: Illustration of a HEC module (left) and one completed endcap of the HEC (right).

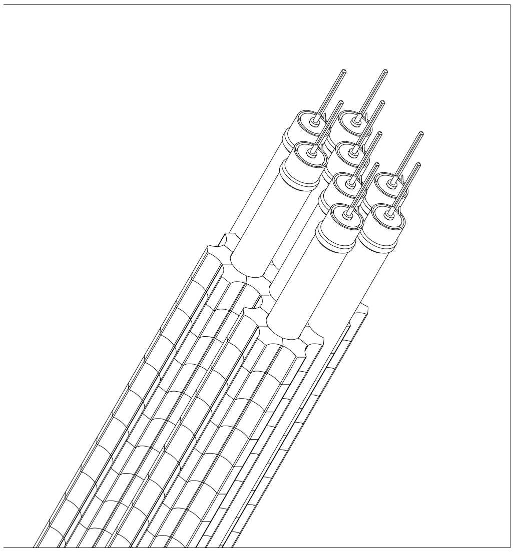



Figure 3.18: Left: Schematic detail of tube structure for the hadronic FCal modules, showing the tungsten slugs surrounding each tube. Right: Completed FCal assembly being insertedinto one of the endcap cryostats.

3.2.3 Muon spectrometer

The outermost part of the ATLAS is called the muon spectrometer (MS). The only detectable Standard Model particles that are regularly expected to escape the calorimeter are muons, so the MS is designed to detect these muons. The MS includes a toroid magnet system that bends the trajectories of the muons, which allows for determining their momenta. Like the inner detector and calorimeter, it is divided into barrel and endcap sections. Each endcap is composed of an inner or small wheel, a middle wheel, and an outer wheel. The detectors used in the MS can also be classified into two categories by their primary purpose: triggering readout of events or precision measurement in the bending direction of the magnets. An overview of the MS is shown in Figure 3.19.

This section describes the state of the MS as it was during Run 2 operation. The small wheels of the MS were replaced after Run 2 with new types of detectors, which are not included here.

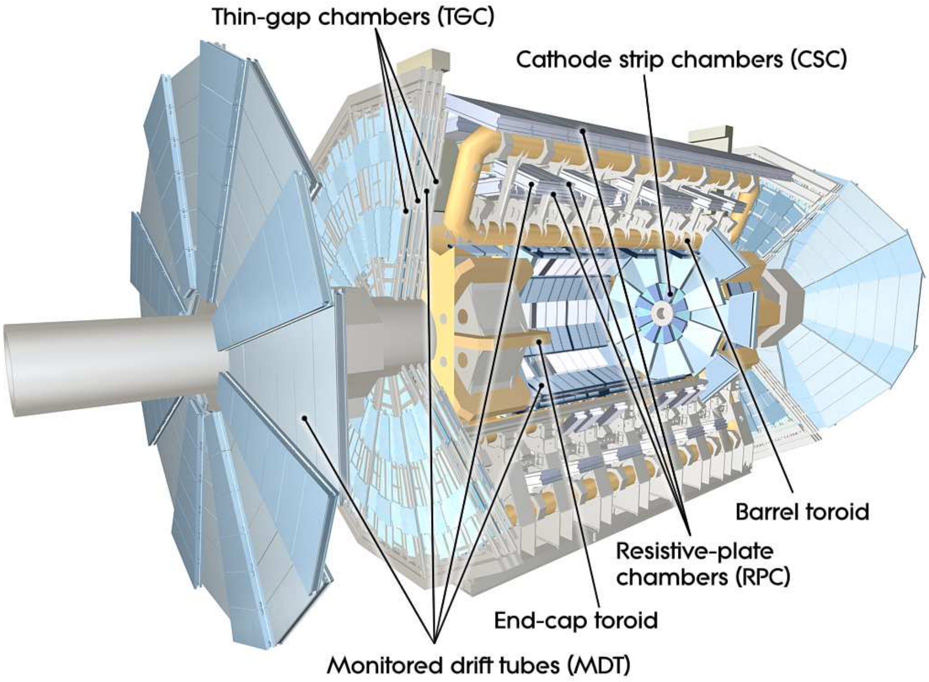


Figure 3.19: Muon spectrometer overview.

The toroid magnet system consists of a barrel toroid and two endcap toroids. All three of these contain eight toroidal coils, each lying in an $r - z$ plane, equally spaced around the azimuthal dimension. This produces a circumferential magenetic field. Like the solenoid, the toroid coils use NbTi as the superconducting material, cooled by liquid helium. The barrel toroid has inner and outer diameters of 9.4 m and 20.1 m, with a length of 25 m. Each barrel toroid coil is enclosed in a stainless steel vacuum vessel. The inner and outer diameters of the endcap toroids are 1.7 m and 10.7 m. The length of each endcap toroid is 5 m. Most of the endcap toroid structure, including the cryostat, is made out of aluminum alloy. The barrel and endcap toroids produce magnetic fields of approximately 0.5 T and 1 T for the muon detectors, respectively. The toroid magnets can be seen in Figure 3.20.

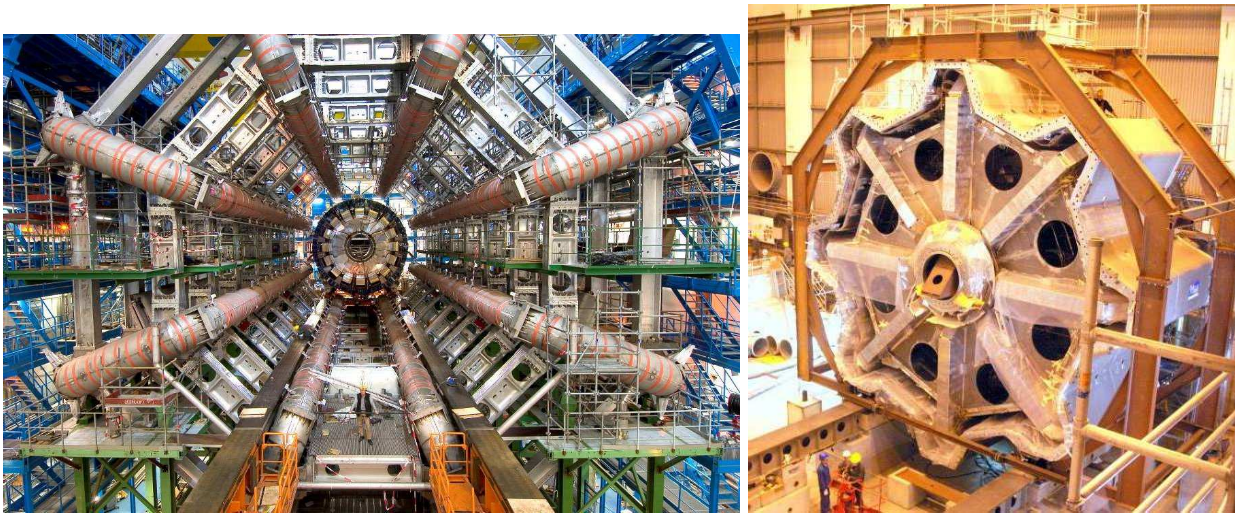


Figure 3.20: Left: Barrel toroid before installation of detectors. Right: An endcap toroid in its cryostat with the end open.

The detector types used for triggering readout based on muons are resistive plate chambers (RPCs) and thin gap chambers (TGCs). RPCs are used in the barrel ($|\eta| < 1.05$), and TGCs are used in the endcaps ($1.05 < |\eta| < 2.4$).

The RPCs contain resistive plates made of bakelite with a 2 mm gap filled with a gas mixture of $C_2H_2F_4$ (94.7%), isobutane (5%), and SF_6 (0.3%). A potential difference of 9.8 kV is maintained between the plates. The charges produced in the gas by an ionizing muon passing through are amplified and drift to the plates. The signal is read out by copper strips on the outer faces of the plates via capacitive coupling. The strips on each side of a gas gap are orthogonal in order to provide a measurement in both coordinates of the plane. An RPC unit consists of two of these gas gaps (and associated plates and strips) interleaved with a paper honeycomb material for spacing and supported by aluminum along the edges. A full RPC chamber is composed of two of these units held together by support pannels with a 65 mm overlap between the two units. There are three layers of RPC chambers in the barrel, and they are divided into 16 azimuthal sectors. There are small and large sectors, which alternate around the azimuthal direction in order to provide full angular coverage with no gaps. The RPC structure and layout are shown in Figure 3.21.

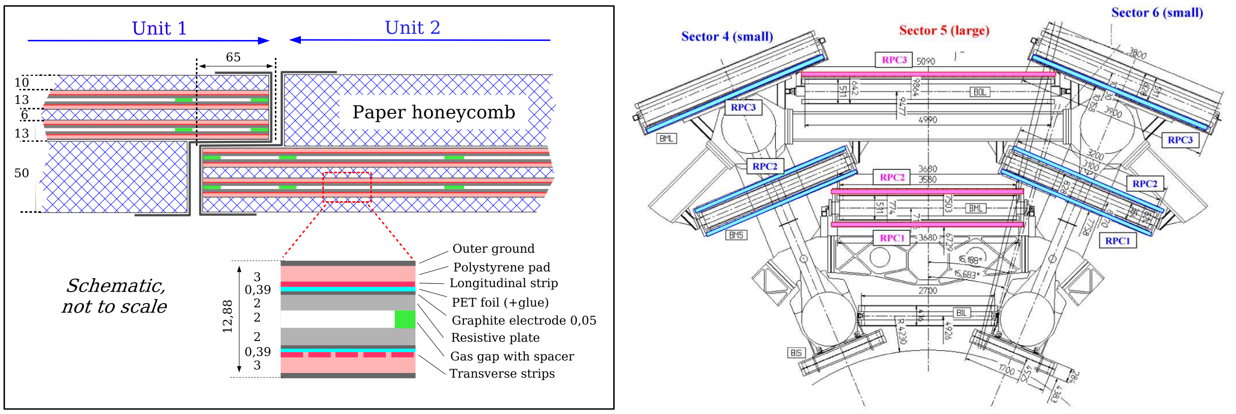


Figure 3.21: RPC chamber structure (left) and layout (right).

In the endcaps, TGCs are used instead of RPCs. TGCs contain 2.8 mm gas gaps of CO_2 and n-pentane between fiberglass laminate plates with wires in the center of the gas gap. The side of the laminate facing the gas gap is coated in graphite and grounded. A layer of copper is on the other side of the laminate, which can be either solid or segmented into strips. The wires are 1.8 mm apart from each other and kept at a potential of 2.9 kV. Each TGC unit contains either two or three of these wire chambers spaced by paper honeycomb material. Each unit contains two layers of strips. The signal readout of the amplified charges from gas ionization by muons comes from the wires and the strips. The TGC structure is shown in Figure 3.22.

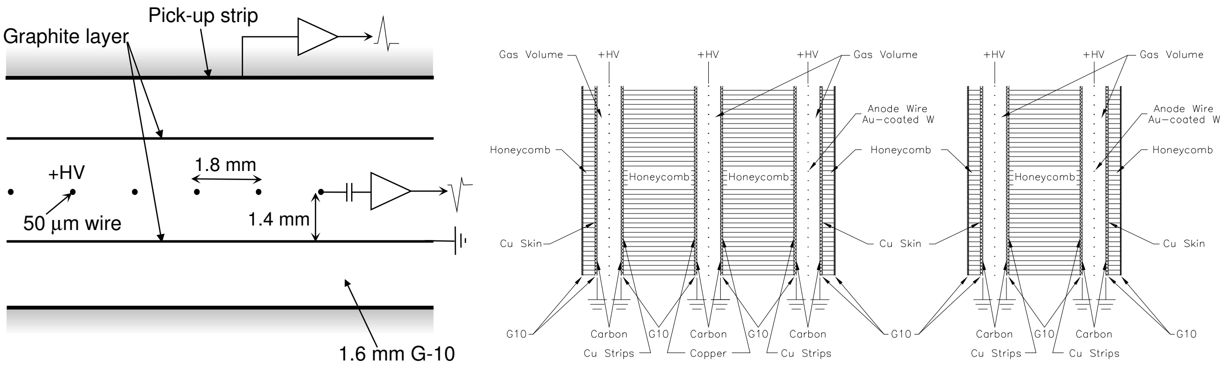


Figure 3.22: Left: TGC chamber structure. Right: Structure of TGC triplet and doublet units.

In all of the barrel and most of the endcap region, the type of detector used for precision muon momentum measurment is called a monitored drift tube chamber (MDT). MDTs are based on aluminum tubes filled with a gas mixture of argon (93%) and CO_2 (7%), pressurized to 300 kPa. The tubes have a diameter of 30 mm and contain a tungsten-rhenium wire running through in the center. The wires are held at a potential of 3.1 kV by a high voltage supply on one end, and the muon signal from gas ionization is read out on the opposite end. The tubes are grouped into multilayers made out of three or four layers of tubes glued together. Each MDT chamber consists of two multilayers separated by aluminum spacers. The spacer lengths vary from 6.5 mm to 320 mm, depending on location. Figure 3.23 illustrates the tube and chamber structures of the MDTs.

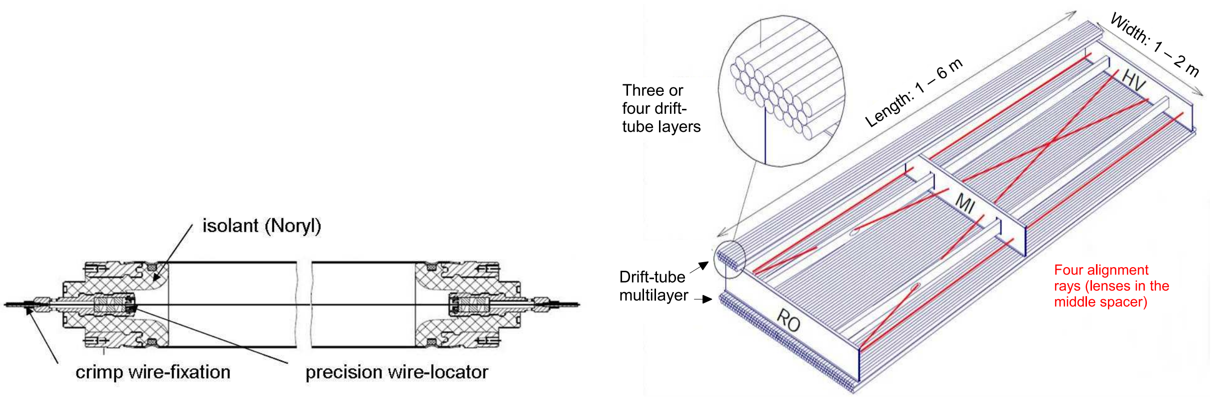


Figure 3.23: Left: Tube structure for the MDTs. Right: MDT Chamber layout.

In the most forward region ($2 < |\eta| < 2.7$) of the innermost (small) wheel of the endcap, cathode strip chambers (CSCs) are used instead of MDTs, since the particle flux here is beyond the safe limits for MDT operation. CSCs contain cells of a gas gap filled with argon (80%) and CO_2 (20%) between two layers of copper strips with wires running through the gap. The gold-plated tungsten-rhenium wires run radially, are separated from each other by 2.5 mm, and are kept at a potential of 1.9 kV. The strips on either side of a gas gap are orthogonal to each other, with one set of strips running paralel to the wires. The strips are 1.5–1.6 mm wide, and the gap between strips is 0.25 mm. Only every third strip is read out; the two strips between readout strips are floating and are capacatively coupled to the readout strips. The wires are not used for signal readout. Each chamber contains four wire planes separated by sheets of polyurethane foam, and there are two different sizes of chamber (referred to as large and small). In each endcap, there are two disks, each containing eight chambers of one size. Figure 3.24 shows the CSC structure and disk layout. The

CSCs are inclined so that a straight line from the IP to the center of each chamber is normal to the wire planes. This improves the spatial resolution by reducing the spread of the ionization in the wire plane.

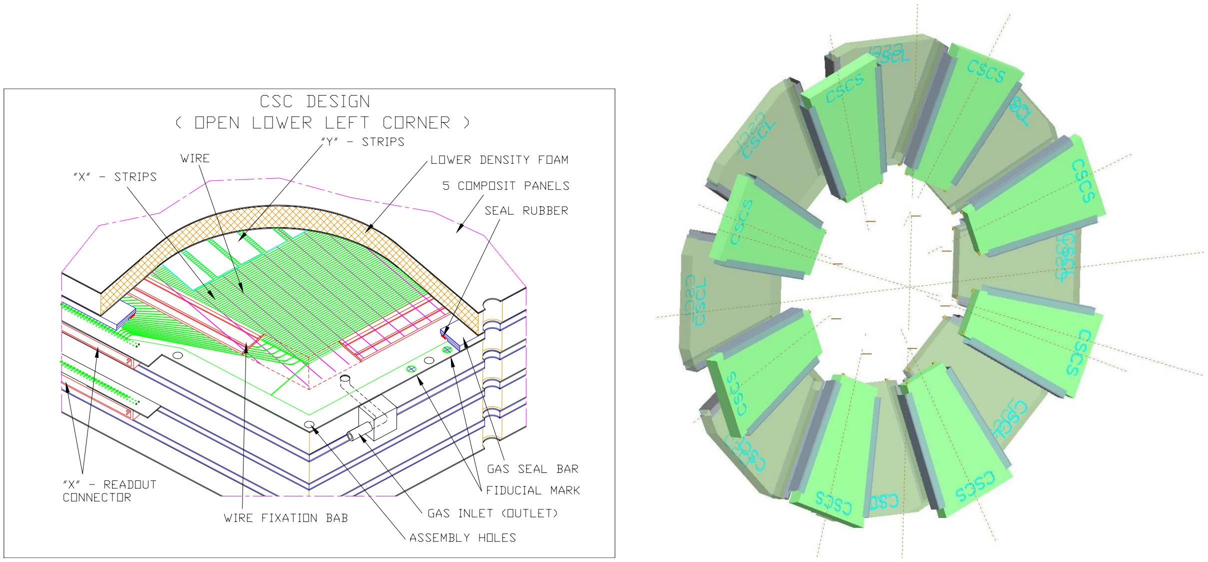


Figure 3.24: Left: Internal CSC structure. Right: Layout of large and small CSCs.

The full layout of the muon spectrometer is shown in Figure 3.25.

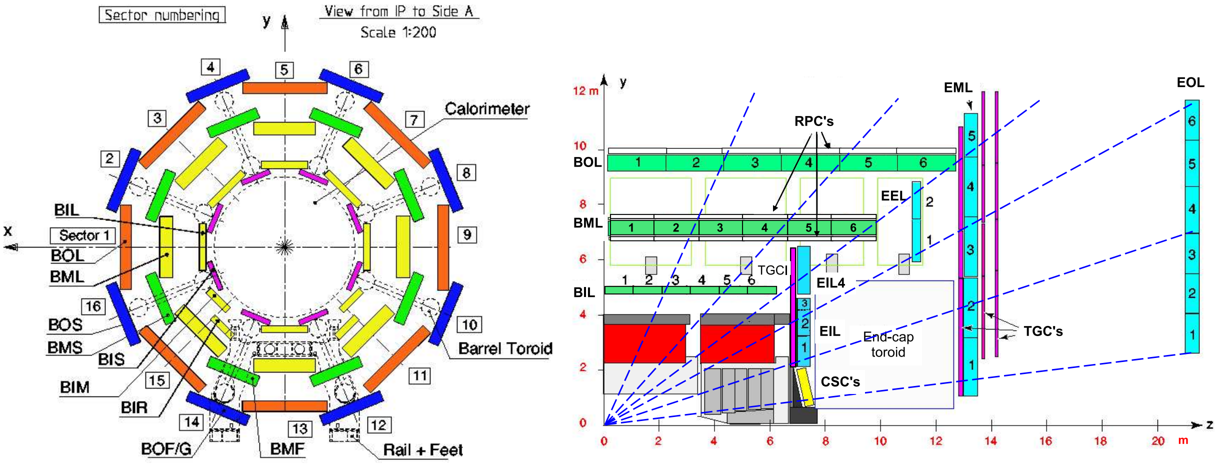


Figure 3.25: Muon detector layout. Left: Azimuthal RPC/MDT sectors in the barrel. Right: Detector layers arranged by r and η. RPC are in white, TGCs in purple, MDTs in green, and CSCs in yellow. Note that there is an extra endcap wheel (labeled EEL here) of MDTs between the inner and middle wheels to cover the part of the angular acceptance of the inner wheel that does not project onto the outer wheel.

3.2.4 Luminosity measurement

In addition to identifying individual high transverse momentum particles produced in collisions, it is also important to measure the total number of interactions at the IP in order to understand the luminosity that corresponds to the data collected. The primary detector used for measuring luminosity in ATLAS is called LUCID (LUminosity measurement using a Cherenkov Integrating Detector). The principle of operation for LUCID is to detect Cherenkov light via photomultipliers. Cherenkov light is produced when charged particles travel through a medium faster than the speed of light in that medium. Measuring the amount of this light with a photomultiplier allows counting the number of particles passing thourgh the medium, which is proportional to the number of interactions in each LHC bunch crossing.

The configuration of LUCID during Run 2 was called LUCID-2, since there was a significant upgrade to the Run 1 LUCID setup. LUCID-2 consists of of two modules, one on each side ($\pm z$) of ATLAS. Each module contains 20 photomultipliers with diameters of 10 mm arranged in five groups of four (Figure 3.26). Four of these groups are arranged at different positions in ϕ about 13 cm from the beam line at $|\eta| = 5.6$ and $|z| = 17$ m. The medium for production of Cherenkov radiation for these four groups is a quartz window at the end of the photomultiplier. One photomultiplier in each of these four groups is modified to have a reduced acceptance diameter of 7 mm by the addition of a thin aluminum layer on the window. The remaining fifth group of photomultipliers is housed in the shielding for the muon spectrometer endcap, with the Cherenkov medium being bundles of quartz fibers that run from near the beam line to the photomultipliers 1.5 m away. At the beginning of Run 2 in 2015, one of the photomultipliers in each of the four groups near the beam used a ^{207}Bi radioactive source for calibration, while the rest of these used an LED signal for calibration. It became clear during data taking in 2015 that the ^{207}Bi calibration was superior. This is because the radioactive sources lead to the production of light closer in wavelength and intensity to that produced by particles from LHC collisions, and the LED calibration components, such as optical fibers, are much more susceptible to radiation damage. Therefore the rest of the photomultipliers in these four groups were also equipped with ^{207}Bi sources at the start of either 2016 or 2017, and only photomultipliers with ^{207}Bi sources were used for the final measurements of Run 2 integrated luminosity.

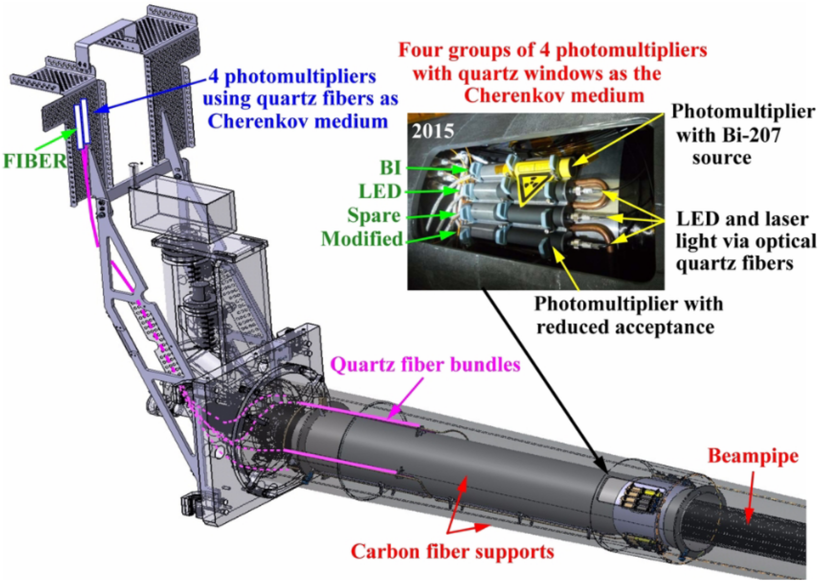


Figure 3.26: LUCID-2 configuration for one side of ATLAS in 2015. The LED calibration system was replaced with additional ^{207}Bi sources in 2016 and 2017. The support cylinder that four of the groups of photomultipliers are attached to has a diameter of 22 cm.

3.2.5 Trigger and data acquisition

While the rate of LHC bunch crossings can reach up to 40 MHz, technological and resource limitations require reducing the full readout rate to below a few kHz. The electronics and software that are used to decide whether or not data from a collision event is transfered to long-term storage is called the trigger system. It uses partially processed information from the event to make this decision while the data is held in temporary buffers. The vast majority of collisions involve relatively small momentum transfers that are not relevant the physics processes of interest at the LHC, so it is possible to greatly reduce the background rate while maintaining high signal efficiency.

The trigger system is divided into two stages: the Level-1 (L1) trigger and the High-Level Trigger (HLT).

The L1 trigger is hardware-based system that takes in information from the calorimeter and the muon spectrometer with reduced granularity [27]. Custom electronics are used to make a decision to accept or reject the event within 2.5 µs. This reduces the event rate from 40 MHz to below 100 kHz. The Level-1 Calorimeter (L1Calo) trigger identifies electron, photon, tau, and hadron/gluon candidates based on clusters of energy deposits in the calorimeter (Figure 3.27), while the Level-1 Muon (L1Muon) trigger identifies muon candidates based on hits in muon spectrometer detectors. The L1Muon trigger also uses tile calorimeter information to reject particles in the endcap not originating from the interaction point. There is also a Level-1 Topological (L1Topo) trigger that receives signals from both the L1Calo and L1Muon systems and can consider events with particular geometric or kinematic combinations between trigger objects. The final L1 trigger decision is determined by the central trigger processor (CTP), which receives inputs from all the other components of the L1 system and directly from some detectors, such as LUCID. Once an event is accepted by the L1 trigger, the front-end electronics read out the data for all detectors to readout drivers (RODs) for initial processing and then to the readout system (ROS) for buffering.

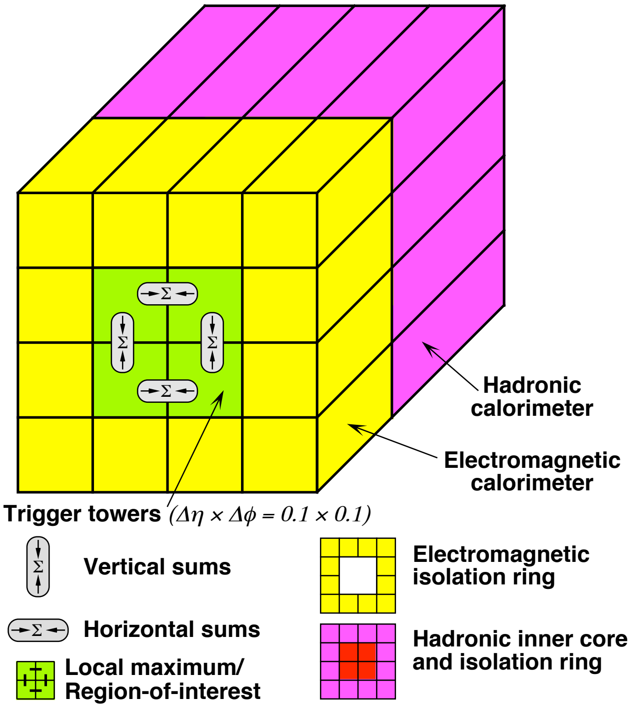


Figure 3.27: Schematic of trigger tower clusters used in the L1Calo trigger algorithms [11]. The transverse energy of a trigger tower is obtained by summing the energy deposited in each layer of the calorimeter in a $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ slice, and clusters are formed by adding the energy of adjacent trigger towers together. Electron/photon and tau triggers are based on 0.2×0.2 clusters, and jet triggers are based on 0.4×0.4 or 0.8×0.8 clusters.

The HLT system is software based and is designed to make event acceptance decisions within a few hundred milliseconds. The output event rate of the HLT for physics analysis was up to about 2 kHz in Run 2 depending on instantaneous luminosity, with an average rate of 1.2 kHz. The trigger algorithms are run on a dedicated computing farm of about 40,000 processing units at the ATLAS site. The HLT algorithms that are executed are based on the conditions that led to acceptance at L1 and request event features from the ROS as necessary to determine whether the final trigger requirements are satisfied or not. After an event is accepted by the HLT, the full event data is sent to the CERN Computing Center for permanent storage. The ATLAS detector records about 1 MB of data per event on average, leading to an average recording rate of 1.2 GB/s. An overview of the full ATLAS trigger and data acquisition system can be seen in Figure 3.28.

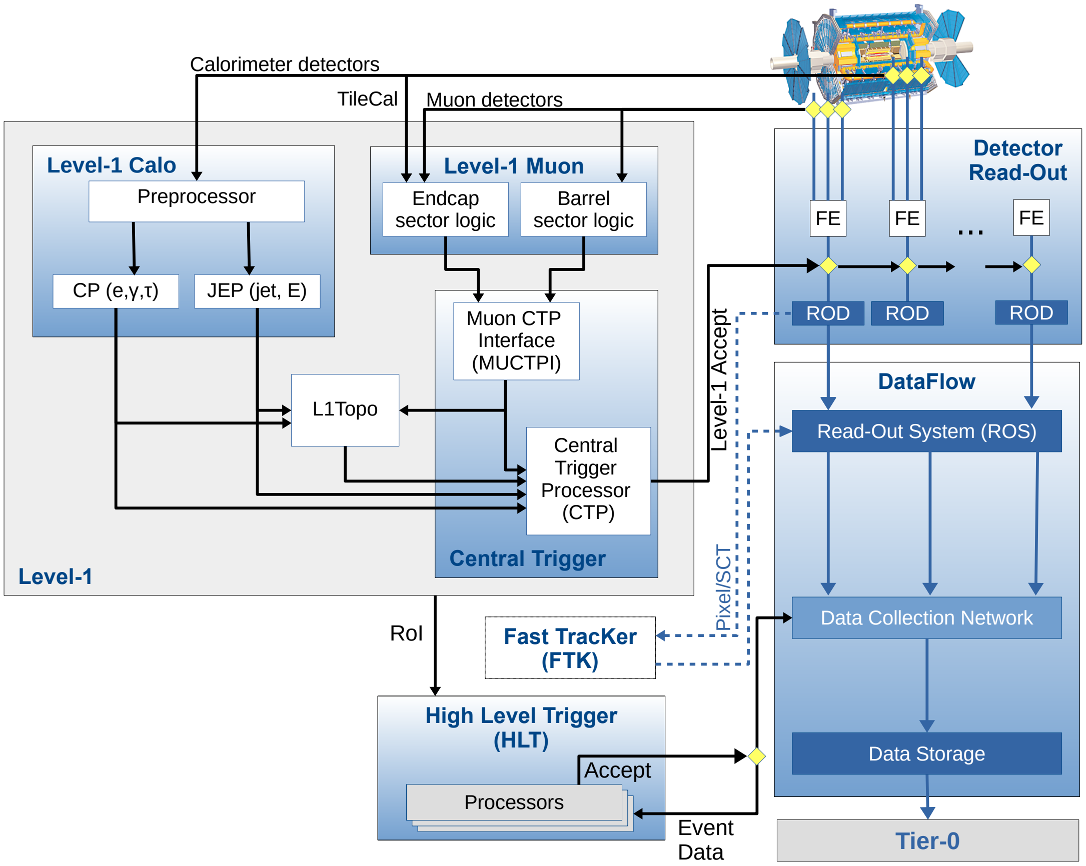


Figure 3.28: ATLAS trigger and data acquisition system in Run 2. The arrows illustrate data flow between components.

3.2.6 Physics object reconstruction

Most ATLAS physics analyses do not operate directly on the raw data read out by the subdetectors. Generally it is the types of particles created in collisions and their kinematic properties that are of interest. To this end, there are algorithms that have been standardized by the ATLAS collaboration that aim to reconstruct the relevant physics objects from low-level detector information. The most commonly used physics objects are described in the following sections.

Tracks

Tracks represent the trajectories of charged particles. These are reconstructed from the signals produced in each layer of the inner detector via ionization in the detector media by the charged particle.

The first step in track reconstruction is to identify pixel and SCT clusters from the raw detector measurements [28]. A cluster is a sensor or a group of adjacent sensors in a layer that detected a charge above a set threshold. These clusters are used to form space-points, which represent a single three-dimensional measurement. Each pixel cluster equates to one space-point, while two clusters of SCT strips (one from each side of an SCT module) are required for one space-point.

Next, track seeds are obtained from sets of three space-points. The set of space-points provides a rough estimate of the impact parameter and momentum of a particle passing through them, which are used to make some intial quality cuts. A combinatorial Kalman filter is then used to build track candidates from these seeds by adding compatible space-points from other layers.

Track candidates at this stage can overlap in their space-points, so the next step is an ambiguity-solving algorithm. Each track candidate is assigned a score based on how well it matches the Kalman filter fit and whether there are holes (missing space-points in a layer). Several quality cuts are enforced for the tracks, including requirements on p_{T}, η, impact parameter, minimum number of clusters, and maximum number of holes. Clusters can be shared by no more than two tracks, and a track can have no more than two shared clusters. The candidates are processed in decresing order of score, and clusters are removed from a track candidate if they cause the current candidate or any previous candidate to fail the quality criteria.

The tracks that pass these selections are then extrapolated into the TRT to include any matching signals there. The extended track is rescored, and the TRT information is dropped from the track if it results in a worse score.

A final stage attempts to form additional tracks from unused segments in the TRT. These are extrapolated back to the SCT and pixel detectors, and any matching unused clusters are added. This improves tracking efficiency for particles coming from secondary vertices, such as electron/positron pairs produced by photons or decays of longer-lived hadrons.

Primary vertices

Primary vertices are the precise locations where two protons traveling in opposite directions interacted. There can be many primary vertices in a single bunch crossing, but generally there is no more than one that involves a large momentum transfer and is of interest for physics analyses. This is called the hard scatter, while the rest of the interactions are usually low momentum transfer (soft) QCD interactions known as pileup. Assigning tracks to the correct primary vertex is important to be able to properly reconstruct the full hard scatter event.

Primary vertex reconstruction uses an iterative fitting procedure starting from a seed position [29]. The x- and y-coordinates of the seed position are taken from the beam spot position. The beam spot is determined by the center a 3D Gaussian fit to the location of primary vertices reconstructed by this algorithm without the beam spot contraint in a separate stream of events not used for physics analyses. The z-coordinate is the mode of the z-coordinates of the set of tracks passing some basic quality criteria at their point of closest approach to the beam spot. Then the fit uses an iterative annealing procedure in which each track is weighted according to its compatibility with the current vertex position. In each iteration, a weighted χ^2 minimization fit leads to a new vertex position, and the temperature parameter of the annealing procedure is gradually reduced in predefined steps. After the last iteration, tracks incompatible with the vertex position by more than seven standard deviations are unassigned from the vertex. These unassigned tracks are then used to start the iterative fit again with a new primary vertex, and this continues until no more vertices can be found.

The primary vertex of the hard scatter interaction is taken to be the one with the highest sum of squared transverse momenta of tracks assigned to the vertex: $\sum p_{\mathrm{T}}^2$.

Jets

When elementary particles carrying QCD color charge (quarks and gluons) are created in the experiment, they are not detected directly. Other than the top quark, which almost instantly decays, these particles undergo a non-perturbative process called hadronization, in which a collimated shower of hadrons is produced. The net energy and momentum of this shower corresponds to that of the original particle. This shower is called a jet, and it is reconstructed from the cluster of detector signals produced by the hadrons and their decay products.

In the analysis described in this work, the type of jet objects used are reconstructed purely from calorimeter information. Jet reconstruction begins by forming clusters of topologically connected cell signals called topo-clusters [30]. A calorimeter cell is smallest readout unit in each part of the calorimeter that can provide an energy measurement. Several steps in the clustering algorithm use the cell significance:

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}$$

The cell signal $E_{\text{cell}}^{\text{EM}}$ is electromagnetic (EM) scale energy, which represents the energy loss of an electron or photon that would produce the observed signal in the cell, and $\sigma_{\text{noise,cell}}^{\text{EM}}$ is the expected noise in the cell. The noise includes both electronic detector noise and noise from pileup interactions.

Any cell with $|\zeta_{\text{cell}}^{\text{EM}}| > 4$ becomes a seed for a proto-cluster. Cells adjacent to the seed in all dimensions are added to the proto-cluster. If one of these adjacent cells has $|\zeta_{\text{cell}}^{\text{EM}}| > 2$, all cells adjacent to it are added as well, and this steps continues iteratively until no more cells can be added. If a cell with $|\zeta_{\text{cell}}^{\text{EM}}|$ is included in more than one proto-cluster, the proto-clusters are merged.

A cluster splitting algorithm is then applied. Any proto-clusters with two or more local maxima are split between these signal peaks. A local maximum is defined as a cell with $E_{\text{cell}}^{\text{EM}} > 500$ MeV, at least four neighboring cells, and no neighboring cells with a larger signal. The clusters after splitting are the final topo-clusters. Only topo-clusters with a total energy that is positive are used to reconstruct physics objects.

Once the topo-clusters have been formed, they are clustered into jets using the anti-k_t algorithm [31]. This algorithm defines a distance between cluster i and j:

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta R_{ij}^2}{R^2}$$

where k_{ti} is the transverse momentum of cluster i and R is a jet radius parameter. A value of $R = 0.4$ is used for standard jets in ATLAS. The distance from cluster i to the beam is defined as $d_{iB} = k_{ti}^{-2}$. The algoirthm procedes by finding all possible non-zero distances of these types and taking the smallest value. If this smallest distance is between two clusters, these are merged into one cluster. If the smallest distance is between a cluster and the beam, that cluster is labeled as a jet and removed from the pool of clusters. The distances are recalculated, and these steps are repeated until no clusters are left.

In order for the jet energy to accurately represent the total energy of the particles that created it, calibrations must be performed after reconstruction. This is done in several steps. First, corrections for pileup interactions are applied, which depend on the location of the jet and the amount of activity in the event [32]. Next, an absolute energy scale and η calibration is applied to account for non-linearities in the calorimeter response, energy losses not captured by the active material of the calorimeters, and biases in the η reconstruction. Then a global sequential calibration is used to reduce known dependencies of the response on the particle content of the jet, such as whether it was initiated by a quark or a gluon and the distribution of quark flavors that were produced. A final correction is applied only for jets in data (not in simulated events) called a residual in situ calibration, which accounts for mismodeling of the jet production and detector response.

Electrons and photons

Reconstruction of electrons and photons produced in collisions begins with the same topo-clusters described in the previous section [33]. The energy from hadronic calorimeter cells in the topo-clusters is removed to get the total electromagnetic (EM) energy of the cluster. Only topo-clusters with an EM energy greater than 400 MeV and more EM energy than hadronic energy are considered. These are referred to as EM topo-clusters.

The reconstruction proceeds by building superclusters, which are composed of a seed EM topo-cluster and potentially nearby clusters called satellite clusters. These satellite clusters are included to allow recovery of energy from radiated photons by electrons and from photons converting to e^+e^- pairs. Electron and photon superclusters are constructed independently, so there can be overlap between them. Electron candidates are only seeded by EM topo-clusters matching a track. Photon candidates do not have this requirement, but the seed cluster may match a track from a conversion vertex. Satellite clusters are added if they fall within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycenter. For electron candidates, satellite clusters are also added within 0.125 × 0.300 if they match the same track as the seed cluster. For photon candidate seed clusters matched to a conversion vertex, satellite clusters are also added if they match a track from the same conversion vertex. These rules are summarized in Figure 3.29.

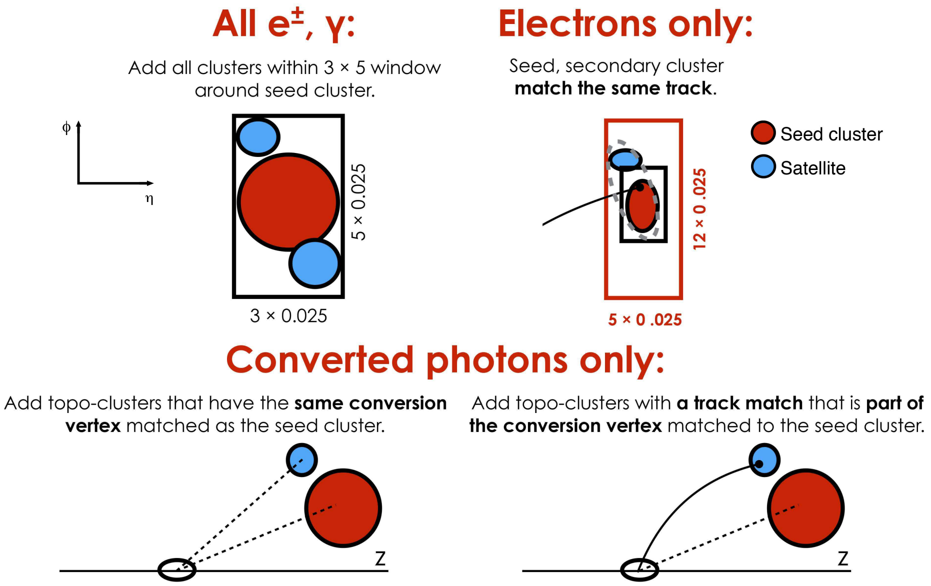


Figure 3.29: Diagram of the superclustering algorithm for electrons and photons. Seed clusters are shown in red, satellite clusters in blue.

If a seed cluster forms both an electron supercluster and a photon supercluster, an ambiguity resolution procedure is attempted. In the case that the supercluster can be clearly identified as originating from one type of particle, the candidate for the other particle is removed. Otherwise both candidates are kept, and final classification is deferred to individual analyses using these physics objects.

Muons

Muons are unique in the fact that they can regularly produce a signal in all of the major subdetector systems. Therefore information from the inner detector, calorimeter, and muon spectrometer can all potentially be used in their reconstruction. Several different muon reconstruction strategies are used in ATLAS [34]. The first step is the identification of short straight-line segments of hits in individual MS stations. These are used to reconstruct tracks within the MS via a procedure similar to that described in the earlier section on tracks, except using MS segments instead of ID space-points. Then there is a search for combined muons, which are identified by matching a track in the MS to a compatible track in the ID and doing a combined fit. Next, an inside-out algorithm is used to try to match ID tracks to loosely-aligned MS signals that were not sufficient to form an track. If an MS track cannot be matched to an ID track, a muon-spectrometer extrapolated muon is reconstructed by extrapolating to the beamline. Segment-tagged muons are identified by extrapolating ID tracks to at least one MS segment with a tight angular matching requirement. Finally, calorimeter-tagged muons are found by extrapolating ID tracks through the calorimeter to match to energy deposits consistent with a minimum-ionizing particle.

Missing E_T

Some particles created in collisions are not expected to be detected at all. For example, neutrinos have an extremely low probability of interacting with any part of the detector, so they essentially always escape detection. The presence of these particles can still be inferred by imposing momentum conservation in the transverse plane. Since the beams are traveling in a direction normal to this plane, the initial transverse momentum is zero. Thus the vector sum of all transverse momenta of particles created in a collision should also be zero. This can be divided into detectable and undetectable particles:

$$\sum_{\text{visible}} \boldsymbol{p}_{\mathbf{T},\boldsymbol{i}} + \sum_{\text{invisible}} \boldsymbol{p}_{\mathbf{T},\boldsymbol{i}} = 0$$

and therefore

$$\sum_{\text{invisible}} \boldsymbol{p}_{\mathbf{T},\boldsymbol{i}} = - \sum_{\text{visible}} \boldsymbol{p}_{\mathbf{T},\boldsymbol{i}}$$

The reconstructed quantity on the right side of this equation is called missing transverse momentum, and the magnitude is written as $E_{\mathrm{T}}^{\mathrm{miss}}$. In addition to non-interacting particles, a significant non-zero value of $E_{\mathrm{T}}^{\mathrm{miss}}$ could indicate problems with reconstruction of an event or the interference of particles from a non-collision background. $E_{\mathrm{T}}^{\mathrm{miss}}$ is determined from a sum across reconstructed electrons, muons, photons, and jets, as well as from tracks that are not associated with any other physics object candidate [35].

3.2.7 Simulation

In order to understand the data that is collected from a detector like ATLAS, it is important to have a model of how the detector is expected to respond to physics processes. This is done via Monte Carlo (MC) simulation. Every step from the proton–proton collisions to the signals that are read out by the subdetectors is simulated with as much accuracy as possible. This provides points of comparison for the real data, so that theory predictions can be confirmed or excluded, and experimental observations can be translated into measurements of fundamental parameters.

The first step of simulation is to simulate the initial proton–proton interaction that results in a hard scatter event. In addition to the creation of the final state particles that are relevant for a particular analysis, there are also soft QCD interactions of partons not participating in the hard scatter, known as the underlying event. Monte Carlo event generators are used to randomly sample the parton distribution functions to set the initial parton states and randomly sample the scattering cross sections to determine the outgoing partons and their momenta. The contributions of additional QCD interactions by the initial and final state particles are taken into account by a parton shower simulation. Then the process of hadronization is simulated using a phenomenological model.

Finally, the propagation of the particles through the detector is simulated, including the production of signals in the active components of the subdetectors. The signals are digitized in the same way as in the physical detector. These last steps are simulated with the GEANT4 toolkit, using a detailed model of the detector [36, 37].

As mentioned earlier in the section on track reconstruction, there are also generally many pileup interactions from other pairs of protons colliding in the same bunch crossing. The effect of these interactions is replicated by taking the hard scatter event and overlaying simulated minimum-bias events, which involve soft inelastic collisions.

After simulation of detector response, the same reconstruction algorithms described in the previous section are applied to the simulated detector readout for analysis. The same trigger algorithms are also run to determine if a simulated event would have passed the trigger.

Chapter 4

Analysis

This analysis searches for events in which a pair of neutral long-lived particles (LLPs) is produced, where one LLP decays within the calorimeter into a jet and the other decays before reaching the calorimeter to produce two separate additional jets. This three-jet approach was designed to improve sensitivity to shorter lifetimes and lower boosts in LLP models relative to the previous ATLAS calorimeter displaced jet search, which looked for two LLP jets [5].

Since the targeted decays occur past most or all of the tracking layers of the inner detector, the jets are not expected be associated with many tracks. In addition, the structure of the energy deposits in the calorimeter for these jets is different from those arising from promptly decaying particles and pure QCD events. The jets from these LLPs are narrower due to the shorter distance between the decay vertex and the calorimeter cells. For the LLP decaying within the calorimeter, little or no energy is deposited in the innermost layers of the calorimeter, which lie between the IP and the decay position.

4.1 Simulated samples

4.1.1 Signal samples

A simplified benchmark hidden sector model is used as the targeted signal for this analysis, where the SM and the hidden sector are connected via a scalar mediator Φ. While Φ could in principle be the Higgs boson, the mass range used for the analysis focuses on bosons with masses ranging from 60 GeV to 1000 GeV. In this model, the decay $\Phi \rightarrow ss \rightarrow f\bar{f}f'\bar{f}'$ is considered, where the two neutral scalars s decay to pairs of SM fermions via Yukawa coupling to SM fermions. Several sets of samples were generated with different assumptions for the masses of the LLPs and mediators. Some mass points were generated with two different assumptions about the mean LLP lifetime in order to validate extrapolation of the analysis results to shorter or longer lifetimes. Events were simulated using MadGraph5_aMC@NLO v2.6.2 [38] at leading order (LO) with the NNPDF2.3LO PDF set, and parton showering and hadronisation was modelled using Pythia 8.230 [39] with the A14 set of tuned parameters [40].

To model the effect of multiple *pp* interactions in the same or neighbouring bunches, simulated inclusive *pp* events were overlaid on each generated signal and background event. The multiple interactions were simulated with Pythia 8.186 using the A2 tune [41] and the MSTW2008LO PDF set [42].

The detector response to the simulated events was evaluated with the GEANT4-based detector simulation [36, 37]. A full simulation of all the detector components was used for all the samples. The standard ATLAS reconstruction software was used for both simulation and *pp* data.

4.1.2 Background simulation

The main background considered in this analysis is non-resonant QCD jet production known as multijet events. This background is simulated in 12 slices of transverse momentum (p_{T}) of the leading jet, and the samples are combined with the appropriate weights to account for the differences in production cross sections. They were generated at LO with Pythia 8.186 [43] using the A14 set of tuned parameters (tune) [40] for parton showering and hadronisation. The NNPDF2.3LO parton distribution function (PDF) set [44] was used. These multijet samples are used to train some of the machine learning algorithms in the analysis and evaluate associated systematic uncertainties.

4.2 Triggers

Two types of CalRatio triggers are used, which only differ in the L1 item used (L1_TAU or L1_LLP-NOMATCH). These are both used to seed the same CalRatio HLT algorithm. The triggers are designed to select narrow jets with little energy in the electromagnetic calorimeter and no ID tracks pointing towards the jet.

4.2.1 Level 1 signature-driven CalRatio triggers

L1_TAU60 and L1_TAU100: All the L1 seeds used in this analysis are based on L1 tau items rather than L1 jet items. A L1 tau item sums energy in a narrower region than a jet item: 0.2×0.2 in $\eta \times \phi$ versus the 0.8×0.8 that the L1 jet triggers use [45]. In the case of L1_TAU60 (L1_TAU100), the E_{T} of the item is required to be larger than 60 (100) GeV for the event to progress to the HLT. This high threshold is necessary to keep the trigger rate low enough so that it can be kept unprescaled at high luminosities. The L1_TAU60 trigger was active for the whole 2015 and 2016 data-taking runs and during part of 2017 and 2018, collecting 102.0 fb^{-1} of data. Due to the increasing instantaneous luminosity, starting in 2017 L1_TAU100 became the lowest L1 tau unprescaled trigger. It was active for the whole 2017–2018 data-taking runs, collecting 101.7 fb^{-1} of data.

L1_LLP-NOMATCH: This is a L1 topological trigger [46] which can make use of the combined information from the calorimeters and the MS at L1. In particular, L1_LLP-NOMATCH selects events containing a L1_TAU object isolated from EM energy, with most of its energy in the HCal. It begins with a L1_TAU30 item (with E_{T} threshold at 30 GeV). It is then required that no energy deposit in the ECal with $E_{\mathrm{T}} > 3$ GeV be found within a 0.2 cone around the most energetic L1_TAU object. In its 2016 version, the trigger additionally required that the second most energetic L1_TAU deposit (if present) was also isolated from EM energy. This condition was removed in 2018 to improve signal efficiency. By making use of the information from the ECal and HCal separately at L1, it was possible to keep a low E_{T} threshold version of this trigger unprescaled. This has particular advantages for the low-E_{T} signal models, where the L1_TAU60 (L1_TAU100) requirement of $E_{\mathrm{T}} > 60$ GeV ($E_{\mathrm{T}} > 100$ GeV) has a severely detrimental effect on the analysis efficiency. The L1_LLP-NOMATCH trigger was only introduced part-way through 2016 data-taking. This version collected 53.7 fb^{-1} of data in 2016 and 2017. The 2018 version was active during the whole 2018 data-taking period, collecting 58.5 fb^{-1} of data.

4.2.2 CalRatio trigger at HLT

The HLT chain for the CalRatio triggers is the same regardless of the L1 seed. In all cases, partial scan (PS) jet reconstruction is used. Partial scan jet reconstruction, as opposed to full scan jet reconstruction, uses only a section of the detector centered around the regions of interest (RoIs) seeding the trigger. The individual RoIs are combined into a super-RoI and jet reconstruction is performed with the anti-k_{T} algorithm using a cone of size $\Delta R < 0.4$.

Due to the unusual nature of the jets studied in this analysis, the standard jet quality requirements applied in most of ATLAS analyses are not applicable here. This standard cleaning is based on quantities that identify jets with a large part of their energy in a single calorimeter cell or sampling layer, or with a high portion of negative energy from fluctuations in the calorimeter cells response. Many of these characteristics are also features of the jets from long LLP decays and applying the standard cleaning would have a dramatic effect in the signal efficiency. Therefore a dedicated cleaning for jets created in the HCal was developed.

Reconstructed jets are required to pass the `cleanLLP` flag, which makes use of a number of parameters listed here:

- HECf: the fraction of jet energy in the Hadronic Calorimeter Endcaps (HEC);
- HECQ and AvgLArQF: quality parameters measuring the deviation of the pulse shape produced by the liquid Argon (LAr) cell from the predicted shape, and for which values near 0 indicate that the predicted and measured pulse shapes are very similar;
- negE: the negative energy in the jet;
- FracSamplingMax: the maximum fraction of energy deposited in a single sampling layer.

The first term of the `cleanLLP` flag contains the HEC spike-cleaning part of the standard jet cleaning requirements, $((\mathrm{HECf} > 0.5) \cap (|\mathrm{HECQ}| > 0.5) \cap (\mathrm{AvgLArQF}|/65535 > 0.8))$, which is very efficient to reject noise bursts in the LAr calorimeter. A second term with the requirement of $((|\mathrm{negE}| > 10\ \mathrm{GeV}) \cap (\mathrm{FracSamplingMax} > 0.85))$ is applied to further eliminate LAr noise. If the jet fulfills the requirements of either of these two terms, it is rejected and not considered for the next steps of the trigger chain. This flag was added to the trigger in 2016. Hence the 2015 version of the trigger does not include this flag online.

However, a slightly tighter version of this jet cleaning is applied offline, where the negative energy is required to be smaller than 4 GeV for the jet to be considered clean.

At the HLT, at least one jet passing the cleanLLP flag is required to satisfy $E_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.5$. The η requirement is a result of the coverage of the ID. To identify jets with a high fraction of their energy in the hadronic calorimeter, triggering jets are required to pass $\log_{10}(E_{\mathrm{H}}/E_{\mathrm{EM}}) > 1.2$. Jets satisfying these requirements are used to determine 0.8×0.8 square regions in $\eta \times \phi$ centered on the jet axis in which to perform tracking. Triggering jets are required to have no tracks with $p_{\mathrm{T}} > 2$ GeV within $\Delta R < 0.2$ of the jet axis.

Finally, jets satisfying all of the above criteria are required to pass a beam-induced background (BIB) suppression algorithm that relies on cell timing and position [47]. BIB is composed of muons created in the detector by beam-halo and beam-gas, travelling in-time with the beam and parallel to the beam pipe. Jets resulting from the deposition of energy in the HCal by a muon associated with BIB can fake signal jets and thus be accepted by the CalRatio trigger. It relies on the fact that calorimeter hits resulting from BIB will be aligned in a relatively horizontal line parallel to the beam pipe and will have a very specific time distribution, not consistent with the time taken for a particle to travel at the speed of light from the IP to the hit location. The algorithm rejects the event if at least four cells at the same ϕ but separated in $\Delta \mathrm{R}$ from the triggering jet, have timing consistent with that of a BIB muon from the current bunch-crossing. In this case, a BIB muon corresponds to a particle travelling at the speed of light, entering the calorimeter through one of the endcaps in time with the proton bunch but in a horizontal direction parallel to the beam pipe. In the 2016 data-taking, this algorithm was modified to deal with the introduction of partial scan jet reconstruction. It was observed that a large amount of BIB was not being identified by the algorithm in 2015, because only calorimeter cells falling inside the cone around the RoI defined for partial scan reconstruction were being considered by the algorithm. The routine was modified for all cells in the calorimeter to be inspected, leading to a better identification of BIB. A second modification with respect to 2015 data-taking is the requirement that the four (or more) BIB-cells are in the same calorimeter layer. This requirement avoids misidentifying cells which are not in a relatively horizontal line as coming from BIB.

4.2.3 Efficiency

The simulated trigger efficiency as a function of the LLP decay position is calculated as the number of LLPs decaying at a given position and firing the CalRatio trigger divided by the number of LLPs generated at that decay position. This efficiency strongly depends on the LLP decay position, as shown in simulation in Figure 4.1. In this figure the efficiency as a function of the LLP decay length in the x-y plane is shown for LLPs decaying in the barrel ($|\eta| < 1.4$). Only LLPs decaying in the HCal are considered in these plots. The left panel gives the efficiencies for one of the low mass samples, for the HLT CalRatio trigger seeded by the 2016 version of L1_LLP-NOMATCH as open markers, and to the 2018 version of L1_LLP-NOMATCH as filled markers, with an increase of $\sim 5\%$ in efficiency. The CalRatio trigger seeded by the L1_TAU60 trigger is shown as a filled area for comparison. At low masses, the high 60 GeV threshold gives a low trigger efficiency. The right panel of the figure is for one of the high mass samples. In this case, the efficiency increase between the 2016 and 2018 versions of L1_LLP-NOMATCH is much larger, about 35%. In cases where the LLP decays in the hadronic endcaps ($|\eta| > 1.4$), the efficiency is shown in Figure 4.2 as a function of the LLP decay length in the longitudinal direction.

The simulated trigger efficiency as a function of the truth LLP p_{T} is calculated as the number of LLPs firing the CalRatio trigger for a given p_{T} divided by the number of LLPs generated at that p_{T}. It is shown in Figures 4.3 for all the CalRatio triggers. The L1_TAU60-based trigger starts to be efficient for LLPs with $p_{\mathrm{T}} > 100$ GeV and reaches its plateau at $150 - 200$ GeV. The L1_LLP-NOMATCH-based trigger recovers a large portion of the LLPs with $p_{\mathrm{T}} < 100$ GeV, specially for low mass samples. The 2016 version of the L1_LLP-NOMATCH-based trigger showed a decrease in efficiency for high-p_{T} jets, due to the L1_LLP-NOMATCH requirement that there are no EM deposits at L1 matched to any of the two most energetic L1_TAU deposits (if there is more than one). Low-p_{T} LLPs decaying at the end of the ECal will deposit very little energy in the ECal and hence pass the L1_LLP-NOMATCH trigger. Higher-p_{T} jets will deposit larger amounts of energy in the ECal. Since the requirement is on both L1_TAU deposits, events where one of the LLPs decays in the HCal but the other decays in the ECal or before are much more likely to pass for low-p_{T} jets than for high-p_{T} ones. In the 2018 version, the requirement of the second L1_TAU deposit being isolated from an EM deposit was removed, hence a large part of the efficiency at high p_{T} was recovered.

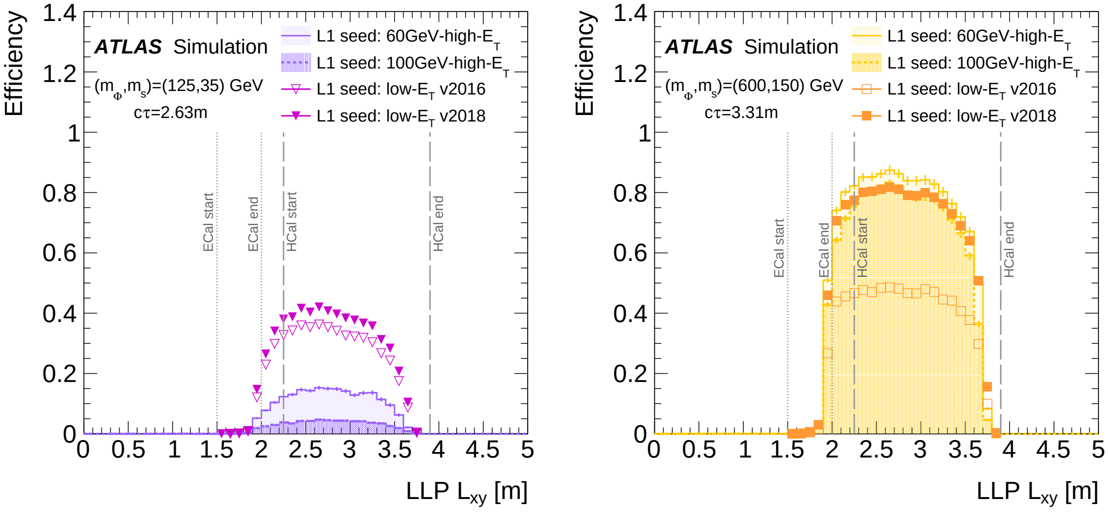


Figure 4.1: Simulated trigger efficiency as a function of the LLP decay decay position x-y plane for LLPs decaying in the barrel ($|\eta| < 1.4$). A comparison is shown for the three versions of the CalRatio triggers for one of the low-mass signal samples (left) and one of the high-mass signal samples (right). In each of the plots, the light filled area corresponds to the CalRatio trigger seeded by L1_TAU60, open markers correspond to the CalRatio trigger seeded by the 2016 version of L1_LLP-NOMATCH, and filled markers to the 2018 version of L1_LLP-NOMATCH.

4.3 Data samples

The main dataset used in the analysis represents a total integrated luminosity of 140 fb^{-1} collected with the ATLAS detector at $\sqrt{s} = 13$ TeV in 2015–2018, using the CalRatio trigger seeded by L1_TAU or L1_LLP-NOMATCH items. When running, all triggers ran unprescaled, although not all triggers ran for the full period. Only triggers that were unprescaled were used by this analysis, leading to slightly different trigger sets being used at different points in data taking. The lowest unprescaled L1_TAU trigger was L1_TAU60 during the 2015 and 2016 data-taking, and L1_TAU100 during 2017 and 2018 data-taking, where L1_TAU60 was disabled when the instantaneous luminosity was higher than 14 nb^{-1} s^{-1}. The L1_LLP-NOMATCH seed was originally defined for the 2016 data-taking period, where it ran unprescaled for approximately one third of the data-taking. This seed was modified at the end of 2017 and ran unprescaled in 2018. A logical OR of these triggers is taken for the analysis.

Additional background datasets are also collected to study non-collision backgrounds: cosmic rays and beam-induced background in particular. These background datasets are used for training machine-learning algorithms and for selection optimisation. They are collected either during empty bunches for cosmics studies

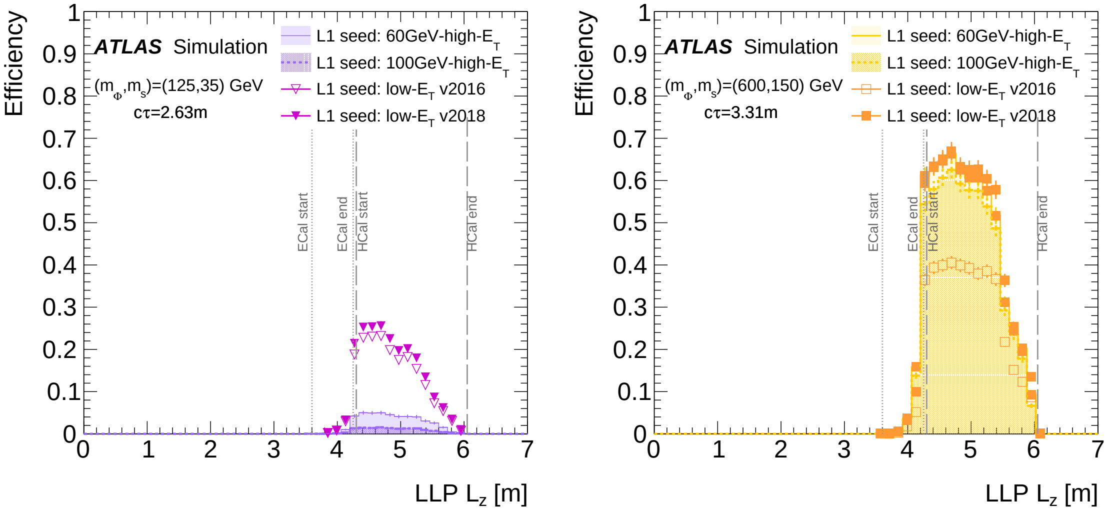


Figure 4.2: Simulated trigger efficiency as a function of the LLP decay decay position long the longitudinal direction for LLPs decaying in the endcap ($|\eta| > 1.4$). A comparison is shown for the three versions of the CalRatio triggers for one of the low-mass signal samples (left) and one of the high-mass signal samples (right). In each of the plots, the light filled area corresponds to the CalRatio trigger seeded by L1_TAU60, open markers correspond to the CalRatio trigger seeded by the 2016 version of L1_LLP-NOMATCH, and filled markers to the 2018 version of L1_LLP-NOMATCH.

or using a modified HLT trigger requirement for BIB studies.

4.4 Jet selection

This analysis uses jets that are reconstructed using the anti-k_t algorithm with a radius of 0.4 using calorimeter topoclusters as described in Section 3.2.6. The jets are first selected by applying quality cuts to candidate jets in a step called jet cleaning. Jets that pass this cleaning are then passed to a neural network that tags the jets with predicted likelihoods of whether the jet arose from BIB, QCD multijet activity, or an LLP decay.

4.4.1 Jet cleaning

The working point associated with the standard ATLAS jet cleaning excludes jets with low ECal activity [48]. Applying this cleaning would eliminate most of the signal jets for this analysis. The analysis instead uses a working point called `LooseBadLLP` that is identical to the standard one except for

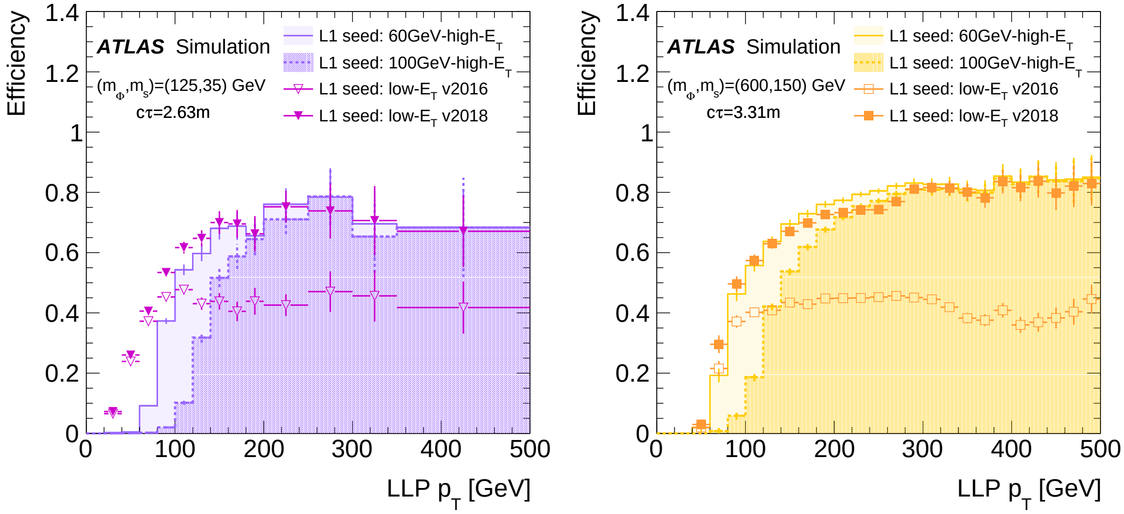


Figure 4.3: Simulated trigger efficiency as a function of the LLP p_T. A comparison is shown for the three versions of the CalRatio triggers for one of the low-mass signal samples (left) and one of the high-mass signal samples (right). In each of the plots, the light filled area corresponds to the CalRatio trigger seeded by L1_TAU60, open markers correspond to the CalRatio trigger seeded by the 2016 version of L1_LLP-NOMATCH, and filled markers to the 2018 version of L1_LLP-NOMATCH.

- the removal of the explicit rejection of low-EMF jets;

- the inclusion of a requirement filtering out jets with $((|\text{negE}| > 4 \text{ GeV}) \cap (\text{FracSamplingMax} > 0.85))$

The second condition rejects any jet with an absolute value of its negative energy higher than 4 GeV and more than 85% of its energy associated with a single calorimeter sampling layer. The negative energy of a jet is calculated as the sum of the energy in all cells with negative energy, which results from the subtraction of electronic and pileup noise.

Jets that fulfill the following critera are called *clean jets*:

- Jet passes *LooseBadLLP*;

- Jet $p_T > 40$ GeV;

- Jet $|\eta| < 2.5$.

CalRatio jet candidates are clean jets that are matched (within a cone of $\Delta R < 0.2$) to an HLT jet which fired one of the CalRatio triggers. A CalRatio jet candidate must also pass offline requirements consistent with the HLT CalRatio jet requirements: $\log_{10}(E_\mathrm{H}/E_\mathrm{EM}) > 1.2$ and $\Delta R_\mathrm{min}(\text{jet, track}) > 0.2$.

4.4.2 LLP jet tagging neural network

This analysis makes use of a neural network (NN) classifier trained to distinguish signal-like jets (arising from a long-lived particle decay) from beam-induced background and SM multijets. This displaced jet tagger was originally trained for an earlier CalRatio analysis [5]

The architecture chosen was a set of convolutional layers fed into a long short-term memory (LSTM) layer. An additional adversarial component was added to reduce the impact of the network exploiting mismodeling of some of the features. The inputs to the NN are low-level features of the jets from the tracker, calorimeters, and muon system. All jets used in the training are clean jets, as defined in Section 4.4.1. The NN was trained on SM multijet MC samples and the BIB dataset, as well as a combination of hidden sector signal samples. Two trainings were performed: one using signals with mediator masses below 200 GeV; the other using signals with mediator masses strictly above 200 GeV. These two versions of the NN are referred to as the *low-*E_T and *high-*E_T NNs respectively. Each NN outputs three scores for each jet, which relate to the probability that such a jet would be the result of an LLP decay, a beam-induced background hit, or a non-displaced jet. These scores are referred to as signal-score, BIB-score and multijet-score, and sum to unity. The per-jet NN is then used by the algorithms used for event selection.

For signal jets, a mix of all m_Φ and m_s samples is used. Signal events with odd event numbers were used in training, while those with even event numbers were used in the rest of the analysis chain. This is done so that limits/discovery are not done on the exact jets this network has trained on, which would bias the classification of these jets. Since this analysis looks for long-lived particles decaying into jets in the hadronic calorimeter or late in the electromagnetic calorimeter, a cut on decay radius of the truth particle associated with the jet is used to only include relevant jets in the training. For the QCD jet training sample, jets from the JZ2W, JZ3W and JZ4W samples are used, which correspond to dijet events with leading jets in the p_T range 60–800 GeV. The two highest p_T clean jets are used in the training. For BIB jet training data, the BIB dataset defined in Section 4.2 is used, with events that pass the HLT CalRatio triggers but that do not pass the HLT BIB-suppression algorithm. Then a ΔR matching procedure is used to find the offline jet matching the HLT-triggering jet. A trigger-matched jet that also meets the clean jet selection is labelled as a BIB jet for training purposes.

Signal and background jets can be characterized by all parts of the ATLAS detector and as such the jet-tagging NN should be able to take advantage of all this information. Therefore, information from objects in each subdetector are used to classify jets, as listed below:

- Spatial positions, momenta, impact parameter, and quality variables of tracks within $\Delta R < 0.2$ of the jet;
- Momenta, timing information, and spatial positions of topoclusters associated to the jet;
- Proportion of energy deposited in each layer of the electromagnetic and hadronic calorimeters for the jet;
- Spatial and timing information for muon segment hits within $\Delta\phi < 0.2$ of the jet;
- Momenta and spatial positions of the jet itself.

Preprocessing of input variables before training reduces biases between training events and to shape variables in a way that will speed up convergence during training. Once preprocessed, jets are input to the neural network to be tagged as signal, QCD, or BIB. The training was done using Keras with the TensorFlow backend.

A combination of two specialized network architectures were used for the jet tagger. The first is a 1D convolution neural network (CNN) which acts as a feature extractor. This convolution performs a series of operations on a 1D array of inputs. In this case the 1D arrays are the variables per subdetector, separately for the tracker, calorimeter, and muon segment variables. In this way, the network can utilize the correlations between the input variables in each jet cluster, track, and muon segment. The second step is the LSTM network, where the memory exploits relations between sequential inputs (for example, between the first two topoclusters).

The individual networks learning from the different subdetectors are concatenated with the jet inputs into one network which has a number of fully connected layers. The outputs of this final concatenated network are three predictions: the signal score, QCD score, and BIB score. The parameters of this architecture were determined using a hyperparameter scan.

Finally, an adversarial network is added to the output of the jet tagging NN. The adversarial network is set up such that it learns how to tell simulation and data jets apart in a control region separate from the signal selection used to train the jet tagging neural network. The adversarial network feeds back to the jet tagging neural network such that latter network does not learn from data/MC mismodeling. This greatly reduces the data/MC differences in the output of the jet tagging NN, as shown in Figures 4.4 and 4.5 for the BIB score. A diagram showing the various aspects of the network as well as the adversary is shown in Figure 4.6.

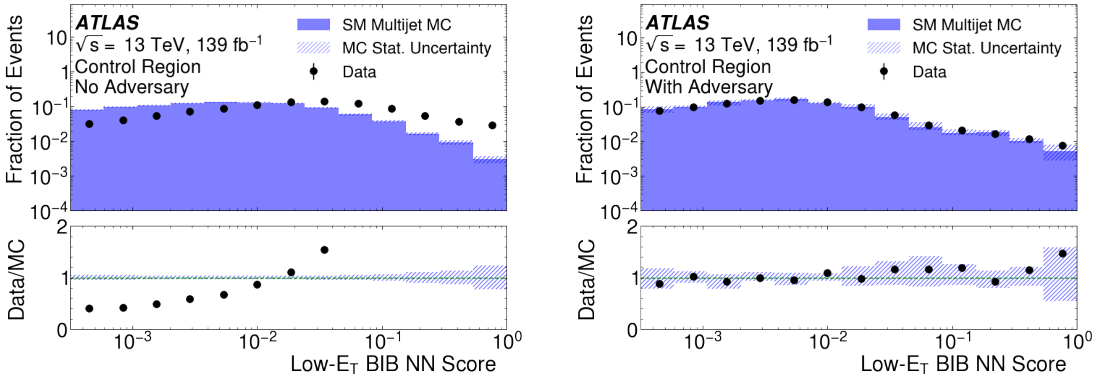


Figure 4.4: Low-E$_T$ per-jet NN BIB score, in the dijet control region, for an network with no adversary (left) and with an adversary (right).

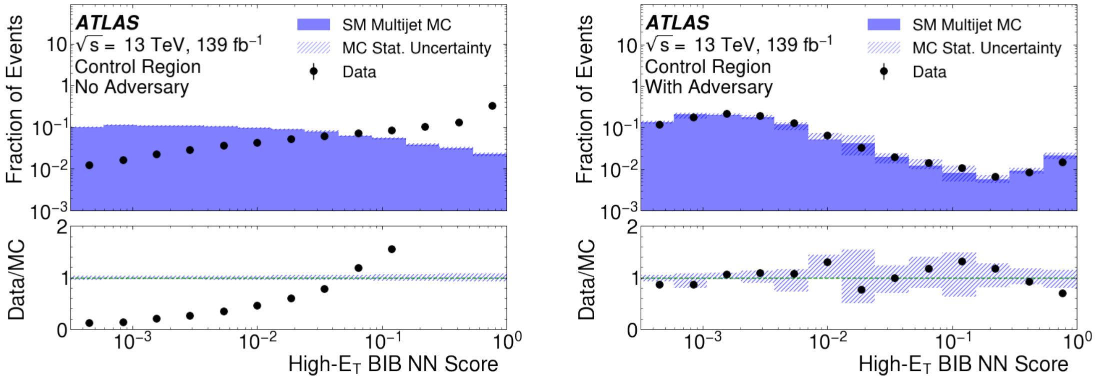


Figure 4.5: High-E$_T$ per-jet NN BIB score, in the dijet control region, for an network with no adversary (left) and with an adversary (right).

The training is divided into low-mass and high-mass trainings. Given that low-mass samples are harder

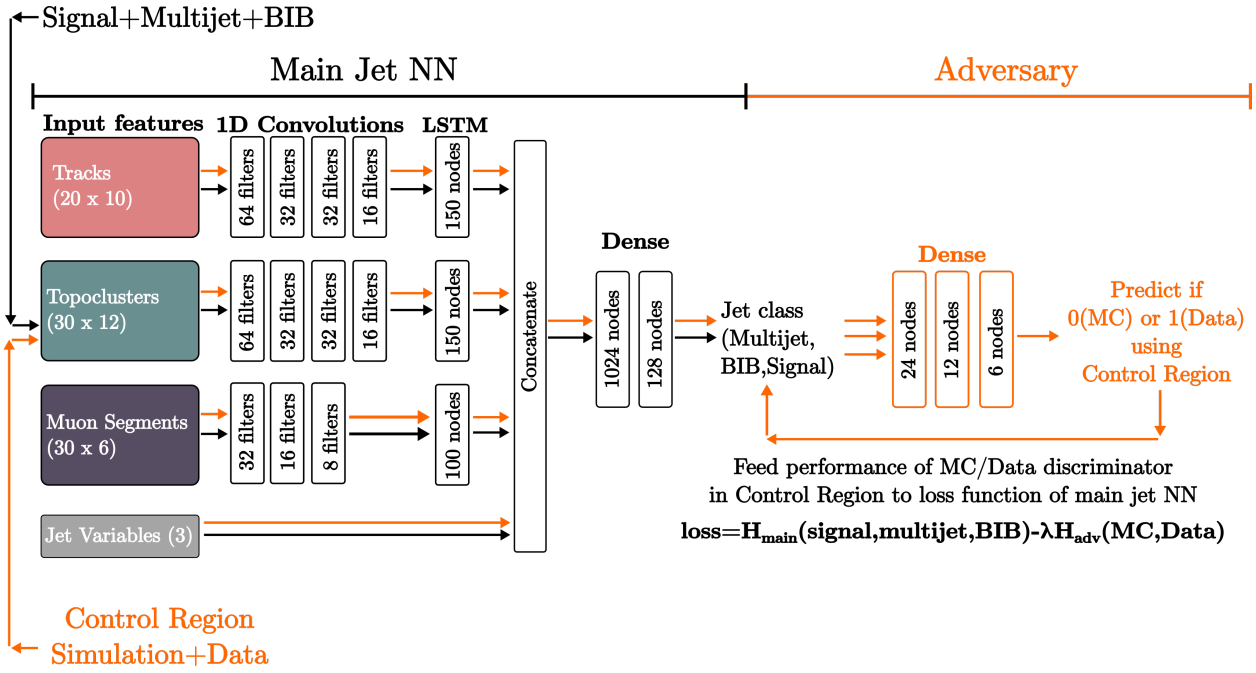


Figure 4.6: Diagram of how the adversary interacts with the main network. The main network, which discriminates between signal, QCD, and BIB, is shown with the inputs moving through the network as black arrows. The inputs to the adversary, which classifies jets as being from MC simulation or from data, are shown moving through the network as orange arrows. The whole network tries to minimize the loss function as shown here, which it does by minimizing the main network loss function, learning to discriminate signal from BIB and QCD, while minimizing the negative of the adversary loss function, learning not to discriminate between simulation and data.

to discriminate from BIB and QCD due to their low p_{T}, it was tested to see if separate or combined trainings would give better performance. It was found that separating them gave better performance for signal tagging efficiency than combined, and so the trainings were separated. Thus, low-mass training refers to the jet NN being trained on the $m_\Phi \in \{60, 125, 200\}$ GeV samples, while high-mass training refers to jet NN being trained on the $m_\Phi \in \{400, 600, 1000\}$ GeV samples.

Final training performance can be seen in Figures 4.7, 4.8 and 4.9, where a low score is indicative of the jet not being labelled as that class, whereas a score close to one shows high confidence in a jet belonging to that class. The training took about 50 epochs, or run-throughs of the whole input data, to converge. A dedicated procedure to assign a systematic for the modeling of the NN inputs was performed, as described in Section 4.8.

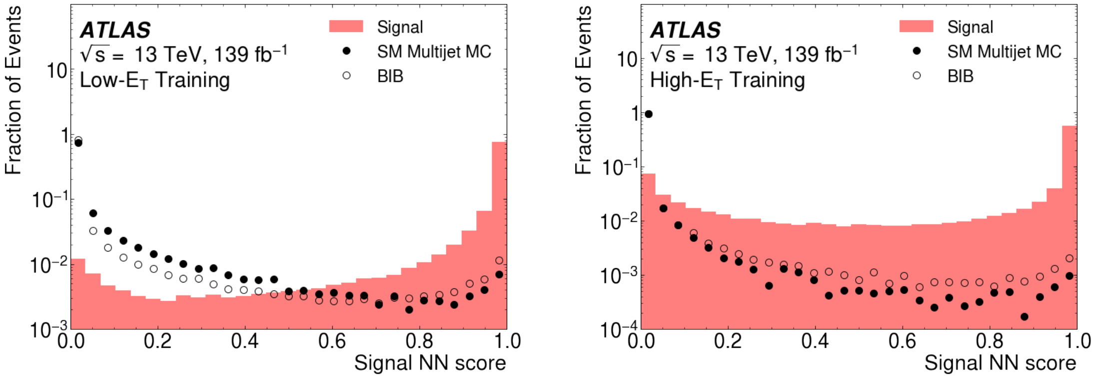


Figure 4.7: The output signal score (confidence of a jet being signal) shown for low mass (left) and high mass (right) training.

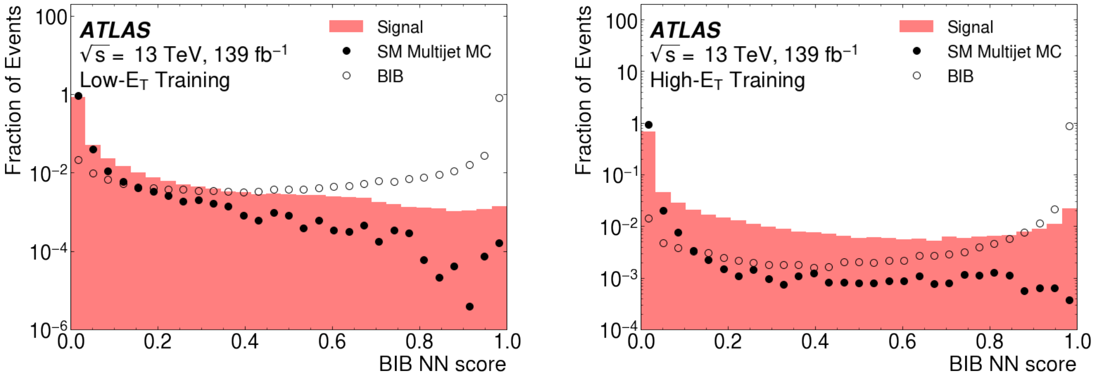


Figure 4.8: The output BIB score (confidence of a jet being BIB) shown for low mass (left) and high mass (right) selections.

4.5 Event selection

The steps in selecting events for the analysis are grouped into two phases: preselection and event cleaning. These phases are described in the following sections.

4.5.1 Preselection

The first step of preselection requires an event to have passed at least one of the CalRatio triggers described in Section 4.2. Events are then required to have at least three clean jets, one of which should be a CalRatio jet candidate as defined in Section 4.4.1. In a signal event, this is expected to be from an LLP that decayed

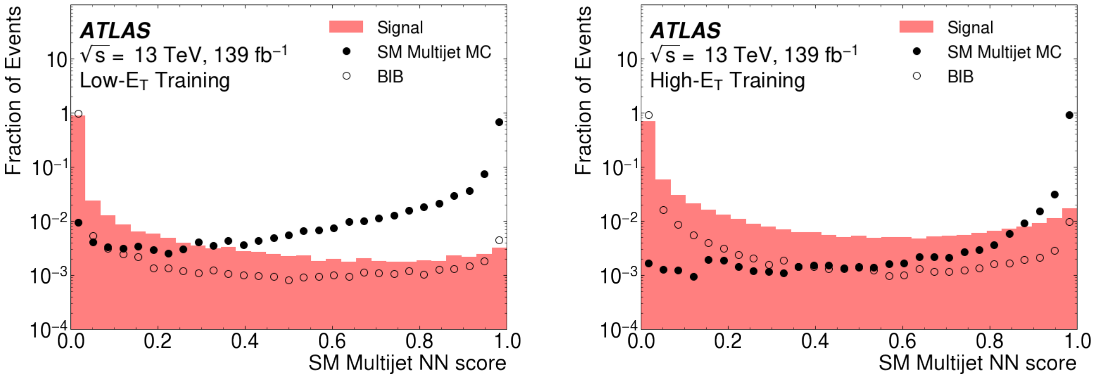


Figure 4.9: The output QCD score (confidence of a jet being QCD) shown for low mass (left) and high mass (right) selections.

in the hadronic calorimeter or near the outer edge of the electromagnetic calorimeter. The two other jets represent the resolved additional jets from a second LLP decay closer to the interaction point.

In order to select events with trackless jets, an additional event-level variable is defined: $\sum \Delta R_{\min}$ (jet, tracks). The value of $\sum \Delta R_{\min}$ (jet, tracks) is obtained by summing the $\Delta R_{\min}$ (jet, tracks) for each clean jet in the event with $p_{\mathrm{T}} > 50$ GeV. In regular QCD events, most jets have nearby tracks, and therefore $\Delta R_{\min}$(jet, tracks) for these jets will be small. Displaced jets are trackless, so the distance between the jet axis and the closest track is large. In events with three displaced jets, this sum of minimum distances to tracks will generally be much larger than in pure QCD events. The distribution of this variable is plotted in Figure 4.10 for both data and simulated signals. The $p_{\mathrm{T}} > 50$ GeV threshold used in this definition is different from the clean jet p_{T} threshold, but this cut was found to provide the best separation between signal and background for this variable. This quantity has been found to be stable up to pileup values of around 65 (approximately the maximum reached during Run 2), with a small dependence thereafter, as shown in Figure 4.11

4.5.2 Event cleaning

The jets with the three highest signal scores and the three highest BIB scores from the high-E_{T} and low-E_{T} NNs are considered signal jet candidates and BIB jet candidates. This leads to twelve total jet candidates, but there is no requirement that these all be distinct jets, so there is often significant or complete overlap between the different candidate types. These jet candidates must satisfy several conditions. First, they should all

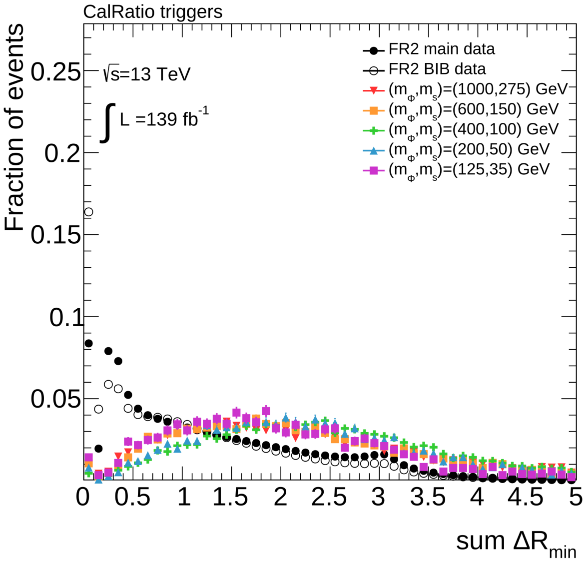


Figure 4.10: Distribution of $\sum \Delta R_{\mathrm{min}}$.

have a time (the energy-weighted average time of the jet's constituent energy deposits with respect to the collision time [49]) in the range $(-3, 15)$ ns to remove remaining BIB jets and jet candidates produced by detector noise while retaining signal jets originating from relatively slowly moving LLPs with timing up to 15 ns. Next, each of these jets should have $\log_{10}(E_{\mathrm{H}}/E_{\mathrm{EM}}) > -1.5$, where $\log_{10}(E_{\mathrm{H}}/E_{\mathrm{EM}})$ is the base-10 logarithm of their CalRatio value. This helps eliminate electromagnetic calorimeter noise and electrons or photons that have been misidentified as jets. Figure 4.12 shows the distribution of the signal jet candidate times and energy log ratios. The jets with the highest signal scores must also not be in the transition regions between the barrel and the endcap ($|\eta| \notin (1.45, 1.55)$), where there is a relative excess of jets due to the different distribution of detector material.

A neural network with a single output per event is used to remove the last remaining BIB events. This NN is trained to be decorrelated with another event-level variable via the ABCDisCo method [50] in order to be able to also use its output for background estimation as described in Section 4.6.

ABCDisCo neural network

The ABCDisCo NN uses jet-level and event-level variables, including the per-jet NN scores, and is trained on events from the BIB dataset and the HSS signal samples. Both the low-E_{T} and high-E_{T} versions of the

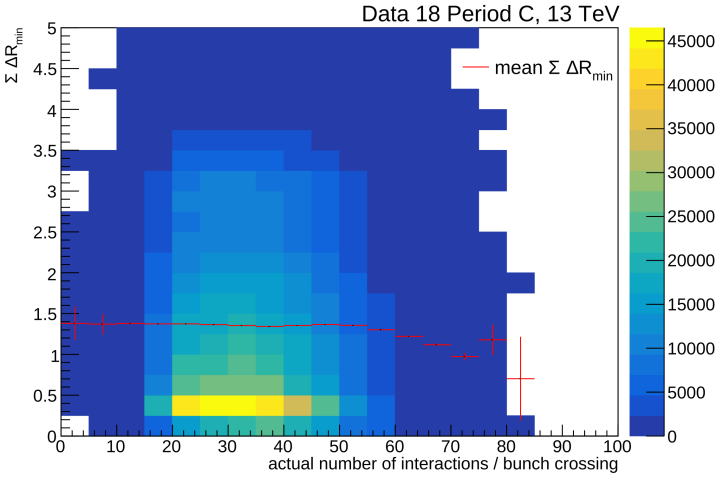


Figure 4.11: Stability of $\sum \Delta R_{\text{min}}$ (jet, tracks) against pileup.

per-jet NN outputs are used. Unlike the per-jet NN, only one version of the per-event NN is trained. The NN was set up with Keras and TensorFlow.

The NN is trained on events in all HSS signal samples and the BIB data sample. For all samples, events entering the training set are required to pass the preselection. Events in the signal samples are additionally required to pass cuts on the LLP decay locations. Each LLP must either decay with $L_{xy} < 3.8$ m if it passes through the barrel calorimeter ($|\eta| < 1.4$) or with $L_z < 6$ m otherwise (for the endcap calorimeter), where L_{xy} and L_z are the projected distance from the IP to the decay in the xy-plane and the z-coordinate of the decay, respectively. This ensures that both LLPs decayed in a location such that the decay products could produce at least one jet. For both the signal and BIB dataset events, only events with odd event numbers are used for training, and events with even event numbers are used as the validation set.

In addition to the per-jet NN outputs, inputs to the ABCDisCo NN include kinematic variables of each signal and BIB jet candidate and its associated tracks, angular distances between the jet candidates, and missing transverse momentum in the event. The inputs are linearly shifted and scaled via a normalization layer such that the training set has a mean of zero and unit variance. Three dense layers of 64 nodes each with rectified linear unit activation are used, with a batch normalization layer after the first dense layer. The output layer is a final dense layer with a single output with linear activation, which is interpreted as a logit (logarithm of the odds ratio: $\ln \frac{p}{1-p}$, where p is the estimated probability of an event arising from

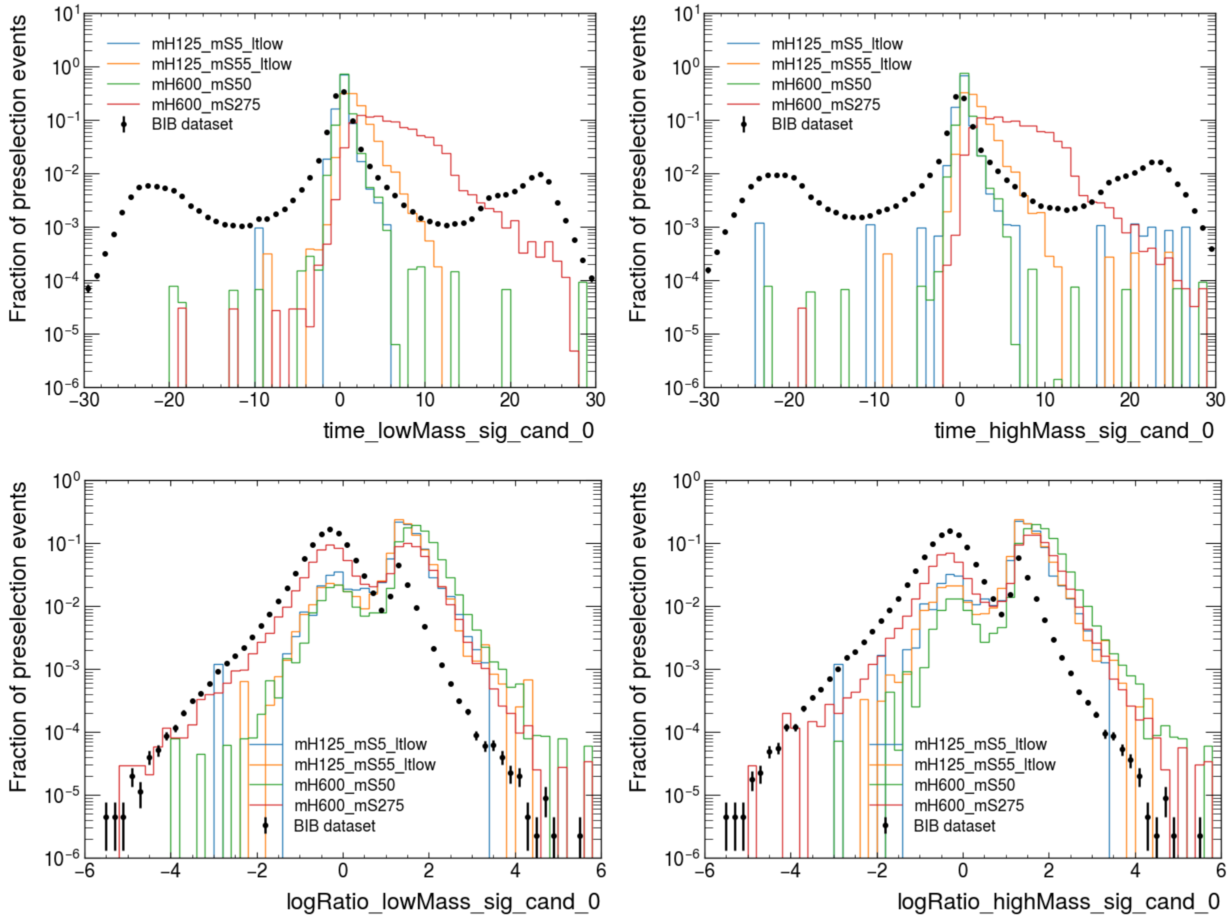


Figure 4.12: Top: Distribution of the time of the first signal jet candidates. The left plot is for the jet with the highest low-$\mathrm{E_T}$ NN score, and the right is for the jet with the highest high-$\mathrm{E_T}$ NN score. Bottom: Distribution of $\log_{10}(E_{\mathrm{H}}/E_{\mathrm{EM}})$ for the same jets.

signal based on the input values).

Following the ABCDisCo method, the NN is trained to discriminate signal from background while simultaneously decorrelating the NN output and another variable ($\sum \Delta R_{\mathrm{min}}$) by reducing their *dis*tance *co*rrelation. This allows using the NN output as one of the axes of a plane which can be used for the data-driven ABCD background estimation method.

The following loss function is used to train the ABCDisCo network:

$$\mathcal{L}_{\mathrm{ABCDisCo}} = \mathcal{L}_{\mathrm{classifier}} + \lambda \mathcal{R}^2 \left(\mathrm{NN}, \sum \Delta R_{\mathrm{min}} \right)$$

$\mathcal{L}_{\mathrm{classifier}}$ is the standard binary cross-entropy of the NN output. λ is a hyperparameter, which showed optimal performance at 0.03 for the final training. $\mathcal{R}^2(\text{NN}, \sum \Delta R_{\text{min}})$ is the squared distance correlation between the NN output and $\sum \Delta R_{\text{min}}$. Distance correlation is a measure of dependence between two variables that is sensitive to non-linear correlations, unlike the more commonly used Pearson correlation coefficient.

The training events are weighted such that all HSS signal models have the same total weight (in order to retain sensitivity to all models) and such that background and signal have the same total weight (in order to maintain overall class balance).

An Adam optimizer is used with the default parameters (including a learning rate of 10^{-3}). A batch size of 10,000 is used to reach a balance between the speed of training convergence and per-batch distance correlation estimation accuracy.

The network is trained for 200 epochs, and the weights are saved after each epoch. After training, a dedicated scan through each epoch is done to more carefully assess discrimination performance and correlation in various regions. This step is necessary as the per-batch loss calculation during training is not sufficient to ensure adequate overall decorrelation. After this study, the best candidate epoch is chosen for final validation and use in the analysis results. The best epoch was found to be epoch 68 out of 200.

Figure 4.13 shows the distribution of the NN output for the BIB dataset and some example signal models. BIB-like events are associated with lower values of the NN output. In order to use the ABCD background estimation method, rejection of BIB-like events is important to ensure that only one source of background (QCD multijet events) will be significant in the final event selection. This rejection is illustrated in Figures 4.14 and 4.15. Events in the BIB dataset with a NN output of 3 or higher do not exhibit the characteristic timing and angular distributions of BIB and for the purposes of this analysis are indistinguishable from pure QCD multijet events. Thus events are required to have a NN score of at least 3 to be considered in later steps of the analysis of this channel in order to effectively eliminate BIB as a background.

4.6 Background estimation

After applying all the selections detailed in the previous section, BIB events that are distinguishable from the QCD multijet background have been removed. Another potential source of background is cosmic rays.

The cosmic rays that pass through the ATLAS experiment are primarily muons that were produced in

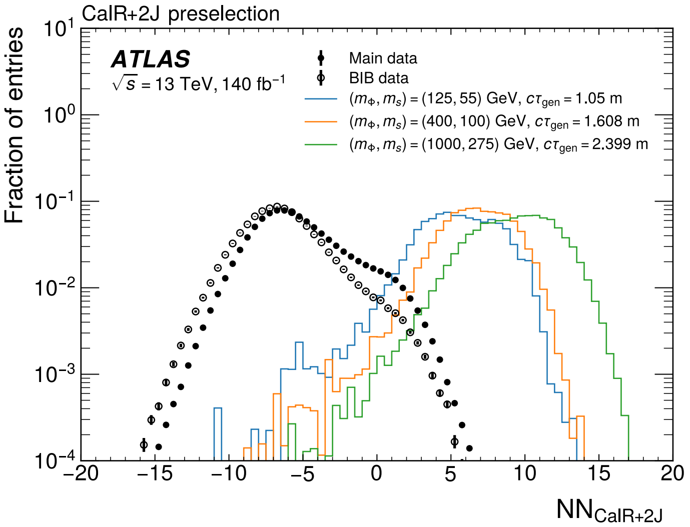


Figure 4.13: ABCDisCo NN output distribution

the upper atmosphere of the Earth. Since these originate from outside the detector, cosmic ray muons can deposit energy in the calorimeter with no tracks pointing back to the IP, similar to an LLP signal. The contribution of cosmic rays to the background was estimated in detail in the previous CalRatio analysis, which required two clean jets [5]. This was done by studying events in bunch crossings with no injected protons. After correcting for the smaller proportion of these empty bunch crossings and the lower number of tracks in these events, only about two cosmic ray events in total were expected to pass the final selections of the analysis, which is well within the statistical uncertainties of all regions. Since the analysis described here further reduces the background by requiring an additional clean jet, the cosmic ray contribution remains negligible, thus this background can be ignored.

Thus the analysis cuts effectively leave only one population of background events. The multijet background is not well modelled, so the data-driven ABCD method is used to estimate the background yield in the final signal regions. In the ABCD method, a plane is defined by two uncorrelated variables, and this plane is divided into four regions (A, B, C, and D) by a threshold value in each variable. Here, region A is identified as the signal region, and region D is diagonally opposite. Due to the fact that the variables are uncorrelated, the ratio of events from a single background source in region A to region B is the same as the ratio between region C and region D. Then the background yield in a region A (N_A) can be obtained from

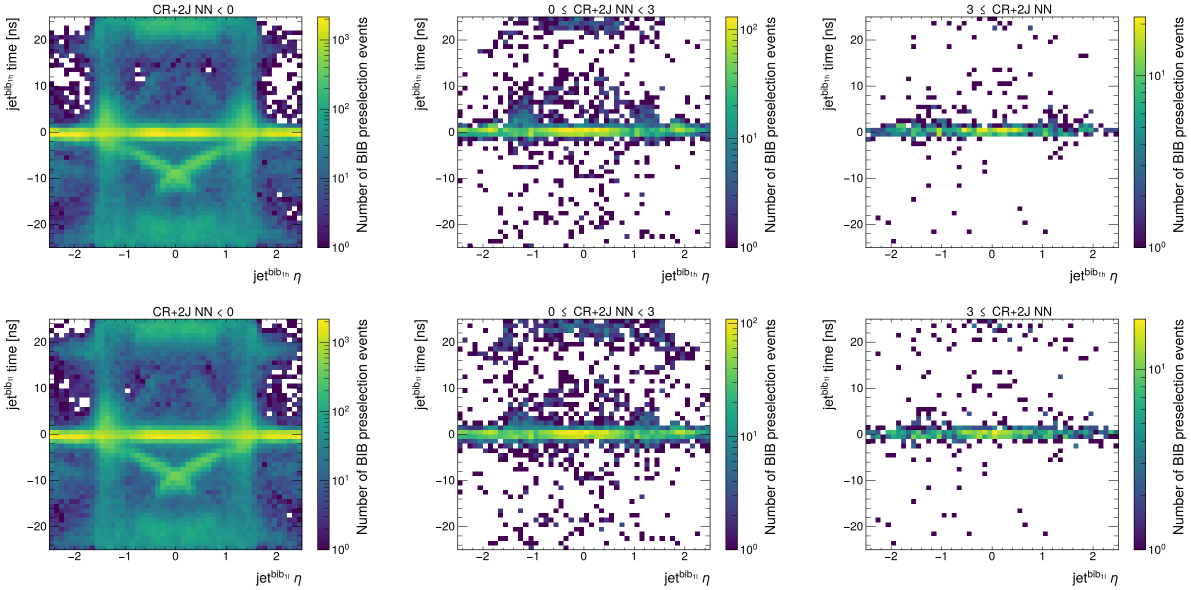


Figure 4.14: First BIB jet candidate time versus η from the BIB data sample in three different intervals of the NN output, increasing from left to right. The top row of plots shows the high-ET BIB jet candidate, and the bottom row shows the low-ET BIB jet candidate.

the yields in the three other regions (N_B, N_C, and N_D) by calculating $N_A = N_B \cdot N_C/N_D$.

Traditionally, the ABCD plane axes are chosen from the list of available variables that separate signal and background well. In cases where it is difficult to find two such variables that are sufficiently uncorrelated, another approach is to construct new discriminants that are uncorrelated. This analysis employed the single ABCDisCo method, where a neural network with a single output was trained to simultaneously discriminate signal from background and be decorrelated from another event-level variable (in this case, $\sum \Delta R_{\min}$). This network is the one used in the event cleaning and was described in detail in Section 4.5.2.

The boundaries between the ABCD regions were set such that the upper limit on the signal contamination of each of regions B, C, and D was less than 5% based on the upper limits on signal cross section times branching ratio from the previous CalRatio analysis. Under this constraint, the boundaries were optimized with respect to the expected upper limits across all HSS signal models by the procedure described later in Section 4.9. These boundaries are at 0.71 in $\sum \Delta R_{\min}$ and 7.61 for the NN output. Figure 4.16 shows the distribution of events for data and some example signal samples.

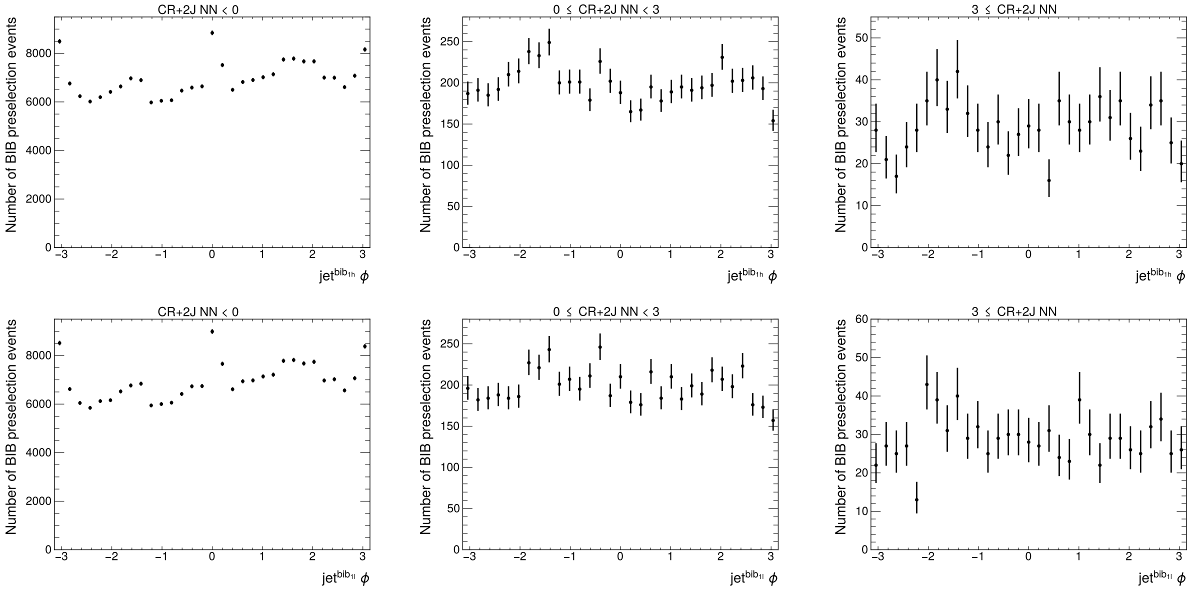


Figure 4.15: First BIB jet candidate ϕ from the BIB data sample in three different intervals of the NN output, increasing from left to right. The top row of plots shows the high-ET BIB jet candidate, and the bottom row shows the low-ET BIB jet candidate.

4.7 Lifetime extrapolation

Since each signal sample was generated for a particular LLP mean lifetime, it is necessary to extrapolate the signal efficiency to obtain results for other lifetimes. This is achieved by using a weighting method, which is applied separately to each signal sample. The weight assigned to an LLP with mean lifetime τ_{new} is calculated from a sample generated with lifetime τ_{gen} by:

$$w(t) = \frac{\tau_{\text{gen}}}{\exp(-t/\tau_{\text{gen}})} \cdot \frac{\exp(-t/\tau_{\text{new}})}{\tau_{\text{new}}}.$$

where t is the proper time of decay of the LLP. In the benchmark HSS model, the LLPs are pair produced, so each event is weighted by the product of the two individual LLP weights. The weighted sample is used to evaluate the signal efficiency for τ_{new}.

This lifetime extrapolation method can be validated by comparing two samples that were generated for different LLP mean lifetime assumptions. The validation is performed by applying the efficiency extrapolation to the samples for each lifetime assumption and comparing the extrapolated efficiency curves. The results are shown in Figure 4.17. The figures show the extrapolated efficiencies for each sample. The blue and orange curves show the results for the low-lifetime and high-lifetime samples respectively. The vertical

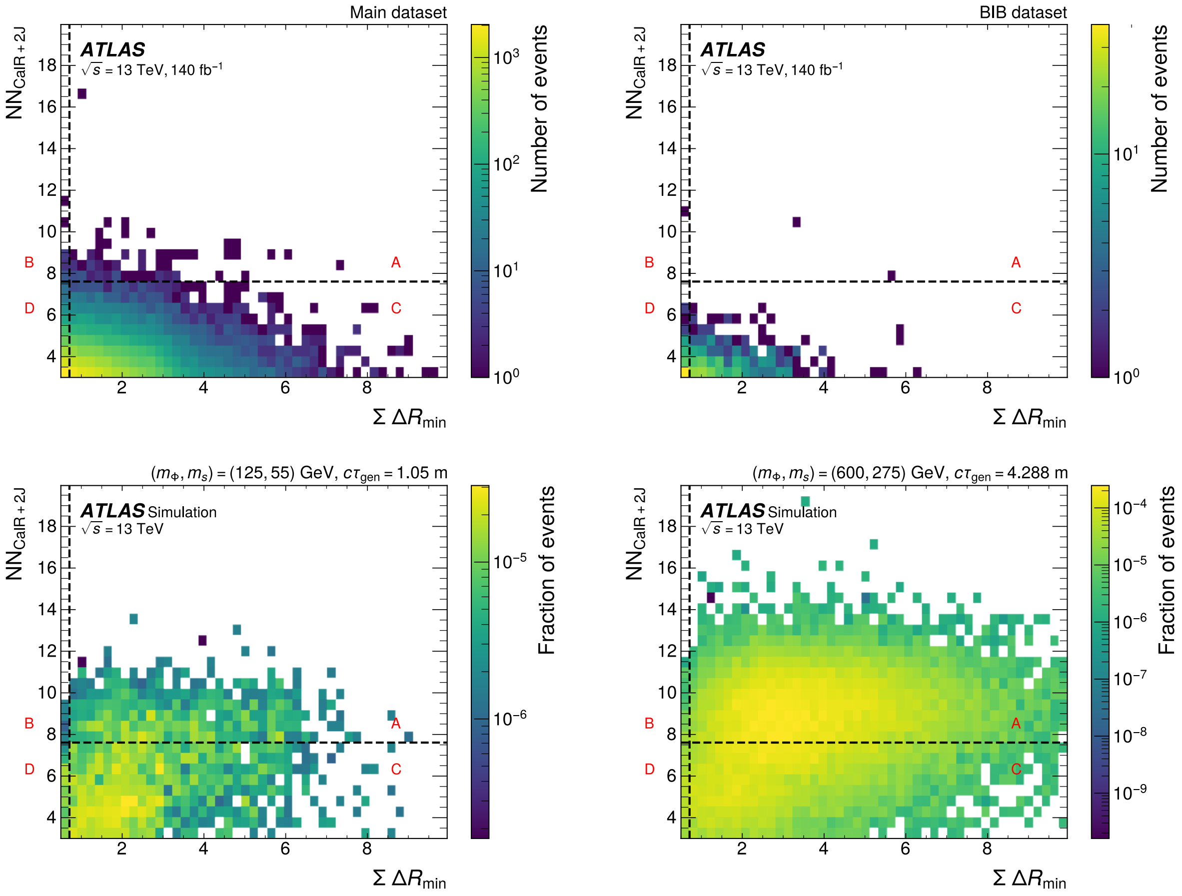


Figure 4.16: Distributions of events in the ABCD plane. The upper plots show data from the main dataset (left) and the BIB dataset (right). The lower plots show the distribution for two different signal samples.

dashed lines of each color show the original mean lifetime used in the simulation for the two cases. For the high-mass samples, where the selection efficiency is highest, the extrapolations agree almost perfectly. For the lower-mass samples, where the selection efficiencies are small, the extrapolations still agree within statistical uncertainties across almost all of the range.

4.8 Systematic uncertainties

In addition to the statistical uncertainty of estimating the event rate in each ABCD region with a finite dataset, various systematic uncertainties affect the analysis results. The derivation of these uncertainties is described in this section.

Since the background estimation is data driven, it is not directly affected by theoretical or experimental uncertainties. However, residual correlations between the two ABCD plane variables can bias the prediction for region A. Validation tests for this method were performed with alternative ABCD regions.

Regions B, C, and D were used to set up the validation regions $\mathrm{VR}_{\mathrm{BD}}$ and $\mathrm{VR}_{\mathrm{CD}}$. $\mathrm{VR}_{\mathrm{BD}}$ consists of the nominal B and D regions combined and is divided into four alternative ABCD regions (A′, B′, C′, and D′). The ABCD estimation procedure was tested with these regions and with variations on the cut values in each ABCD plane variable. The same procedure was performed for $\mathrm{VR}_{\mathrm{CD}}$, composed of the nominal C and D regions. These validation regions represent the largest rectangular regions that be tested without using events from the signal-rich region A. The results can be seen in Figure 4.18.

The difference between the number of observed and expected events in region A′ of $\mathrm{VR}_{\mathrm{CD}}$ is slightly larger than one standard deviation as calculated by the statistical uncertainties. The left plot in Figure 4.19 shows the ratio of observed to expected events in this region as a function of the $\sum \Delta R_{\mathrm{min}}$ cut with only statistical uncertainties included. In this range of cut values, the observed number of events is consistently a bit higher than the expected number. The mean of these ratio points is 1.05, and a flat line at this value fits the points well. Therefore a systematic uncertainty on the ABCD estimate of 5% is inferred. The right side of Figure 4.19 shows the same ratio plot with this 5% uncertainty added.

All the remaining systematic uncertainties affect the expected number of signal events in each region given a particular integrated luminosity and signal cross section times branching ratio.

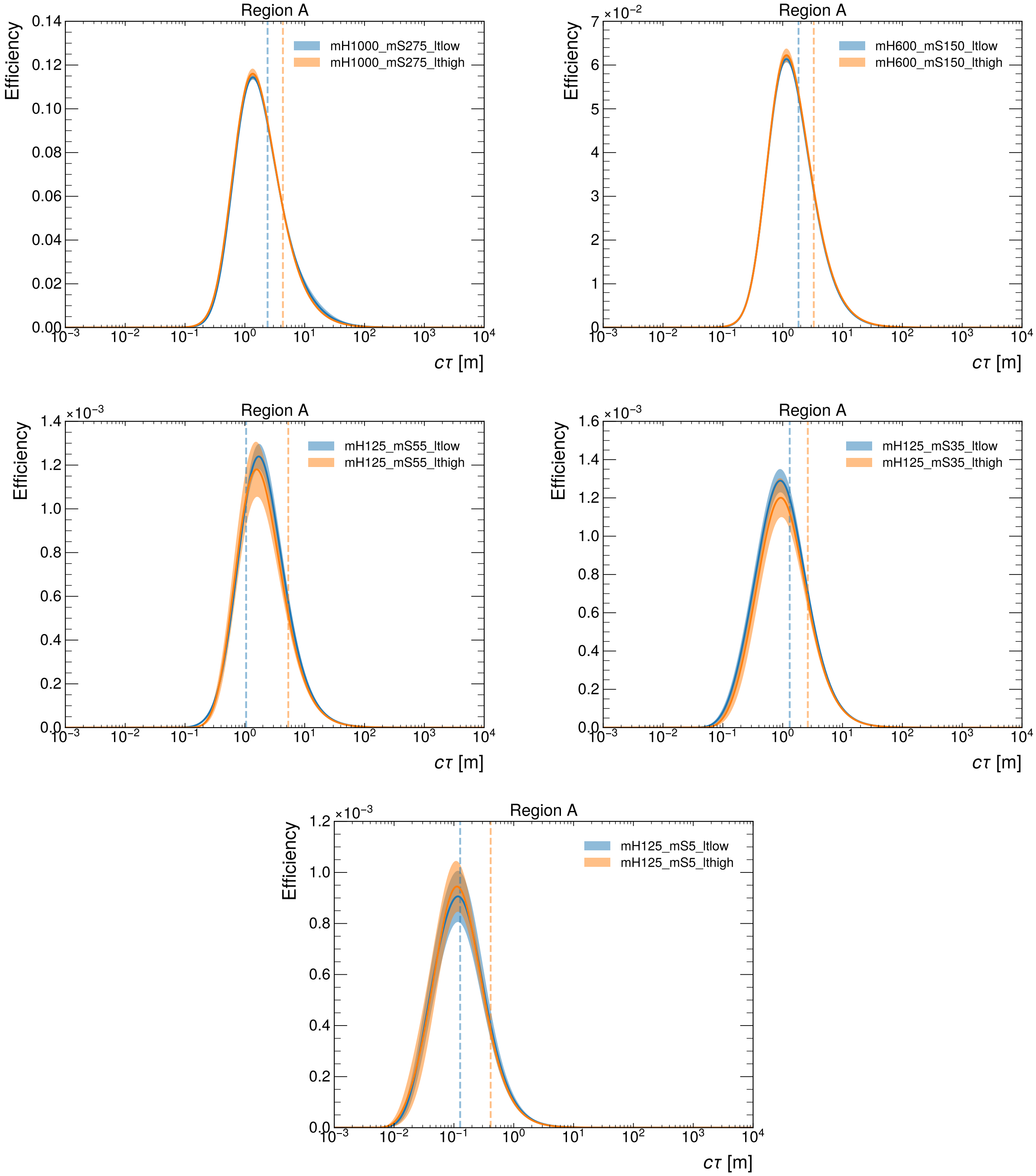


Figure 4.17: The extrapolated signal efficiency in region A as a function of mean LLP lifetime. The bands represent statistical uncertainties only. The blue and orange curves are obtained from the same model, where events were generated using two different mean lifetime assumptions. The dashed vertical lines represent the originally generated mean lifetimes. The lower-lifetime sample is shown in blue, while orange represents the higher-lifetime sample.

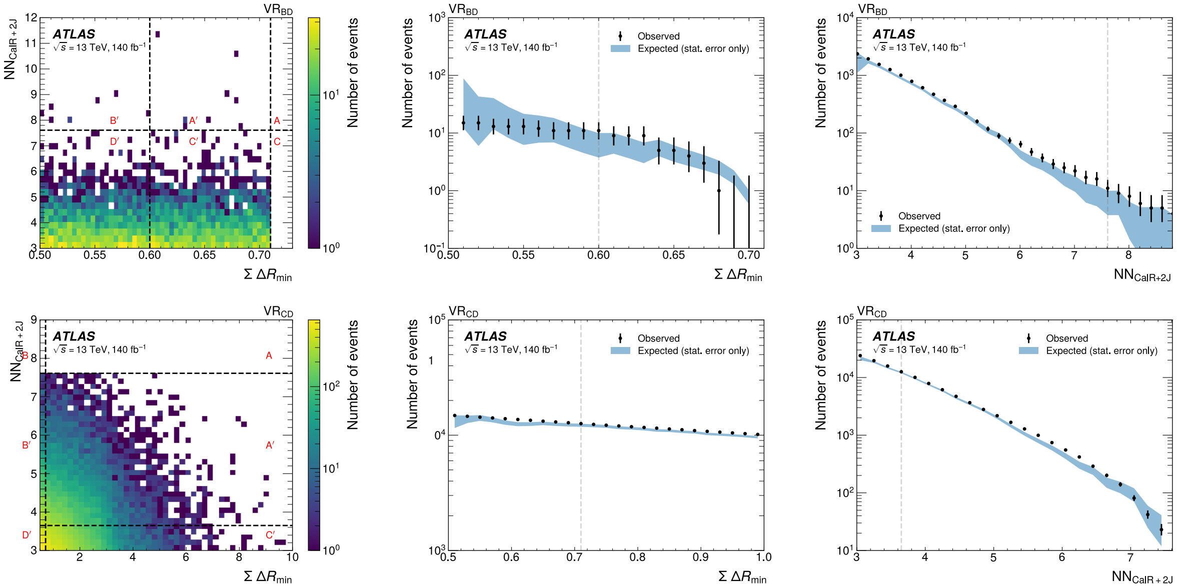


Figure 4.18: CR+2J validation regions. The left plots demonstrate how VR$_{\mathrm{BD}}$ and VR$_{\mathrm{CD}}$ are defined and show the distribution of events within them. VR$_{\mathrm{BD}}$ only includes events in the main dataset with NN values less than 7.61, and VR$_{\mathrm{CD}}$ only includes main dataset events with $\sum \Delta R_{\mathrm{min}}$values less than 0.71. The plots on the right compare the observed and expected number of events in A′ in VR$_{\mathrm{BD}}$ and VR$_{\mathrm{CD}}$ by applying the standard ABCD procedure while varying cut values between the subregions. The vertical dashed lines indicate the nominal boundaries between the VR subregions, shown in the 2D plots and used when varying the cut in the other variable. These were set to be the boundaries between the nominal ABCD regions where possible and to be the median values in the respective validation region otherwise.

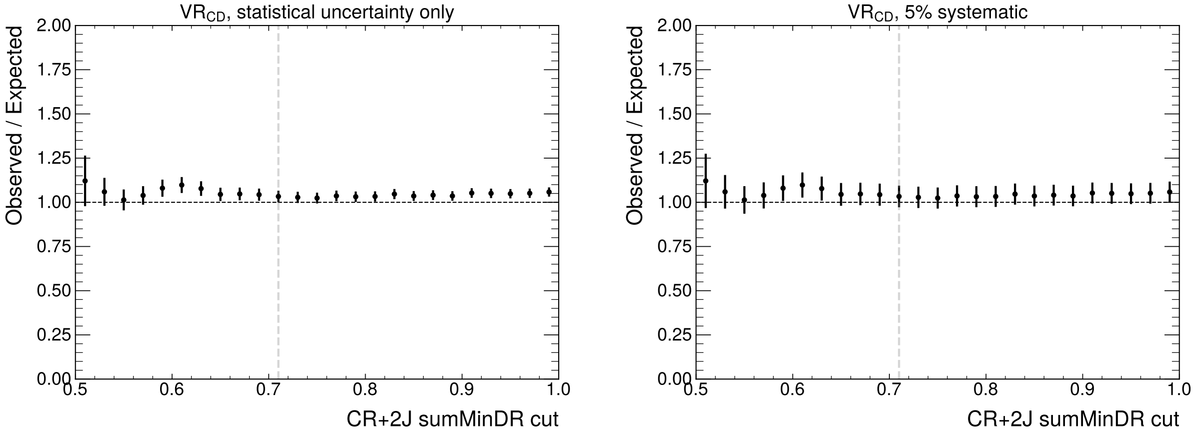


Figure 4.19: Ratio of observed and expected events in region A′ of VR$_{\mathrm{CD}}$. The left plot includes only statistical uncertainties, while the right includes an additional 5% non-closure systematic.

The uncertainty on the integrated luminosity of the data collected by ATLAS during LHC Run 2 is 0.83%, obtained using the LUCID detector for the primary luminosity measurements [51]. This impacts the results via the total number of expected signal events at a given signal cross section times branching ratio.

The rest of the sources of uncertainty affect the estimation of the fraction of signal events that fall within each ABCD region.

An uncertainty is assigned on the reweighting of events in simulation to match the observered distribution of pileup in data. A higher number of pileup interactions leads to more tracks in the event, which decreases the signal efficiency of analysis selections, since the isolation of jets from tracks is used in several steps of the analysis. The uncertainty in the pileup reweighting of the reconstructed events in the MC simulation is estimated by comparing the distribution of the number of primary vertices in the MC simulation with the distribution in data as a function of the instantaneous luminosity. Differences between these distributions are reduced by scaling the mean number of interactions per bunch crossing in the MC simulation and the $\pm 1\sigma$ uncertainties are assigned to these scaling factors [52, 53]. The effect on the signal event yields varies between 0.3% and 7% depending on the signal model.

Signal efficiency is also affected by uncertainty in the energy of the reconstructed jet, both in absolute scale and resolution. These uncertainties are calculated via the standard procedure prescribed by the ATLAS collaboration [54]. The uncertainties for the signal models studied range from 0.1% to 15%. Since the displaced jets used in this analysis are non-standard due to their unusual electromagnetic energy fraction (EMF), the jet energy uncertainties were rederived as a function of the EMF and pseudorapidity. These additional uncertainties are found to have an effect of up to 4% on the signals considered in this analysis. They are conservatively added in quadrature with the standard jet energy uncertainties.

Since LLPs are used for triggering events in this analysis, an uncertainty is assigned on the signal trigger efficiency by studying the modeling of the key variables in the trigger: the jet energy, the CalRatio, and the p_{T} of the jets' tracks. The data–MC agreement of these variables between HLT-reconstructed quantities and offline-reconstructed quantities is considered. A tag-and-probe technique is applied to a pure sample of multijet events obtained using standard jet triggers in both data and MC simulation [5]. Scale factors that represent the degree of mismodeling in each variable are derived and then applied in an emulation of the CalRatio triggers. The change in yield relative to the nominal unscaled trigger emulation after the full analysis selection is taken as the size of the systematic uncertainty, which is between 1% and 6% depending on the signal model.

A systematic uncertainty on the machine learning algorithms used in the analysis is also included. Specifically, this uncertainty accounts for potential mismodeling of input variables used in the per-jet NN tagger. In the dijet control region, the distribution of the inputs of the per-jet NN were studied. In most cases the agreement was found to be good, but for variables with residual differences, the mismodeling was encapsulated in transfer factors required to reach compatibility between data and MC. In an ensemble of pseudo-experiments, the per-jet NN variables are modulated by random Gaussian variations with a mean of zero and a width determined by the transfer factors determined above. The signal efficiencies are recalculated in each pseudo-experiment with the modulated variables. The spread in the signal efficiencies across the ensemble of pseudo-experiments gives the uncertainty associated with this effect. This effect is of the order of 5% or smaller depending on the exact model.

Finally, an uncertainty due to the NLO p_{T} reweighting of the signal samples was obtained by comparing the next-to-next-to-leading order (NNLO) predictions for the p_{T} spectrum of a 125 GeV mediator mass with those at NLO accuracy. The difference between these is used as an additional event weight for each signal sample, and the difference in the signal efficiency relative to the nominal NLO weighting is taken as the uncertainty. This results in an additional signal efficiency uncertainty of 1% to 6% for most samples.

4.9 Statistical interpretation

The data-driven background estimation procedure described in Section 4.6 is used in a simultaneous fit to assess the compatibility of the data with the background prediction and the presence of signal events. An overall profile likelihood function is constructed from the product of the Poisson probabilities of observing the number of events N_X^{obs}, given an expectation N_X^{exp} in each region X, where $X =$ A, B, C or D. The value of N_X^{exp} in each region is the sum of the expected background yield N_X^{bkg} and the expected signal yield N_X^{sig}, given by the number of simulated signal events entering region X multiplied by the signal strength μ (the parameter of interest). In the fit, the expected background yields are constrained to obey the ABCD relation

$N_A^{bkg} = (N_B^{bkg} \cdot N_C^{bkg})/N_D^{bkg}$. Since the Poisson constraints only apply to N_X^{obs} relative to N_X^{exp}, it follows that the background prediction may change dynamically in the fit as a function of the signal strength μ. Additional Gaussian constraint terms in the likelihood are included for nuisance parameters corresponding to the background esimation unceratinty and the total signal uncertainty. The likelihood was implemented using the `pyhf` framework [55, 56, 57].

If the signal strength parameter is fixed to zero, the only remaining components of the expected yields which can be modified relate to the background. Ignoring the observed yield in region A, the expected background yields in B, C, and D will exactly match the observed yields in those regions, while the yield in A will reflect the naive ABCD relation. This is called the *a priori* background estimate. It corresponds to the case where the observed value in A is not known (i.e., before the data has been unblinded). Including the observed yield in region A in the fit may modify the expected background yields in the other regions– they will move away from observed yields in each region (within the allowed statistical uncertainties) such that the ABCD relation on expected yields can be maintained, even if the observed yields do not exactly satisfy this relation. This is called the *a posteriori* background estimate. It corresponds to the case after the observed yield in region A is known (after unblinding). The introduction of a signal component will dynamically modify the allowed background prediction, for example by allowing a potential excess in region A to be explained by a non-zero signal strength.

In the absence of an excess, upper limits on LLP production cross-section times branching ratio are extracted. The CL_s method is used for this purpose, using the test statistic $\tilde{q}$ as implemented in `pyhf` [58, 59].

Chapter 5

Results

The full cutflow for data and some example signal samples can be seen in Table 5.1. A summary of the expected and observed number of events in each region of the ABCD plane can be found in Table 5.2.

As no significant excess is observed, upper limits on signal cross section times branching ratio are set. The lifetime extrapolation procedure described in Section 4.7 is used to set limits across a broad range of mean LLP lifetimes. Figure 5.1 shows the extrapolated HSS signal efficiencies, and the expected and observed upper limits are shown in Figure 5.2.

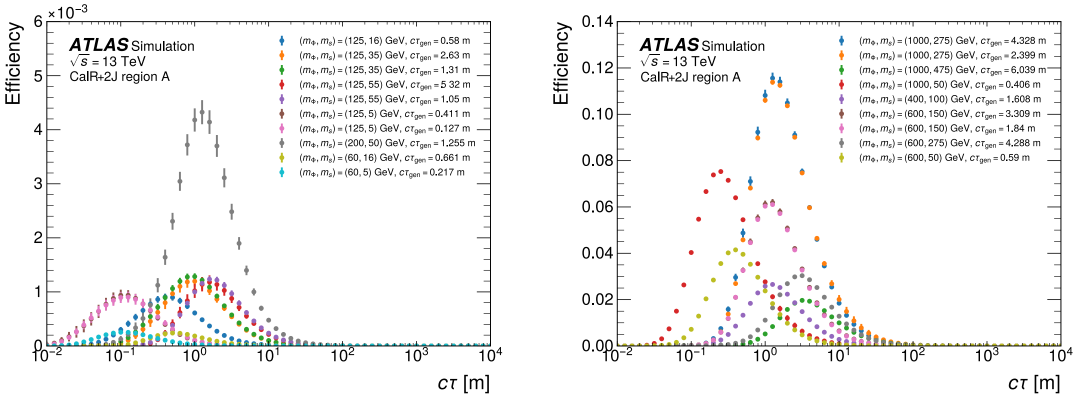


Figure 5.1: Signal efficiencies for HSS signals. Extrapolation of efficiencies to new mean LLP lifetimes is done via the event reweighting procedure.

		Main dataset	BIB dataset	$m_\Phi = 125$ GeV $m_s = 55$ GeV $c\tau_{\text{gen}} = 0.127$ m	$m_\Phi = 600$ GeV $m_s = 275$ GeV $c\tau_{\text{gen}} = 4.288$ m
Preselection	CalRatio triggers			1.4%	11%
	≥ 3 clean jets			0.60%	7.7%
	$\sum \Delta R_{\text{min}} > 0.5$	5,738,136	446,794	0.59%	7.6%
Event cleaning	Trigger matching	2,068,592	154,986	0.53%	6.5%
	-3 ns $< t < 15$ ns	2,609,223	99,398	0.51%	5.8%
	$\log_{10}(E_{\text{H}}/E_{\text{EM}}) > -1.5$	2,289,758	89,380	0.46%	5.1%
	$\lvert\eta\rvert \notin (1.45, 1.55)$	2,068,592	80,555	0.41%	4.5%
	$\text{NN}_{\text{CalR+2J}} \geq 3$	30,097	408	0.35%	4.4%
Region A	$\sum \Delta R_{\text{min}} \geq 0.71$, $\text{NN}_{\text{CalR+2J}} \geq 7.61$	92	2	0.10%	2.8%
Region B	$\sum \Delta R_{\text{min}} < 0.71$, $\text{NN}_{\text{CalR+2J}} \geq 7.61$	18	1	0.00%	0.01%
Region C	$\sum \Delta R_{\text{min}} \geq 0.71$, $\text{NN}_{\text{CalR+2J}} < 7.61$	25213	328	0.24%	1.6%
Region D	$\sum \Delta R_{\text{min}} < 0.71$, $\text{NN}_{\text{CalR+2J}} < 7.61$	4774	77	0.01%	0.04%

Table 5.1: Sequential impact of each requirement on the number of events passing the selection. For the example signal models considered, the columns represent the cumulative fraction of events passing each step.

Main selections	B	C	D	Estim. A (*a priori*)	A	Estim. A (*a posteriori*)
CR + 2J selection	18	25213	4774	95 ± 23	92	93 ± 10

Table 5.2: Application of the likelihood-based ABCD method to the final selections. The columns A, B, C and D contain the number of observed events in each region. "Estim. A" refers to the estimated contribution of background in the search region A assuming no signal, as calculated using the full statistical model. As described in Section 4.9, the a priori estimate refers to the pre-unblinding case and the a posteriori estimate refers to the post-unblinding case, including the observed data in region A into the background-only global fit. Only statistical uncertainties are included in the quoted error.

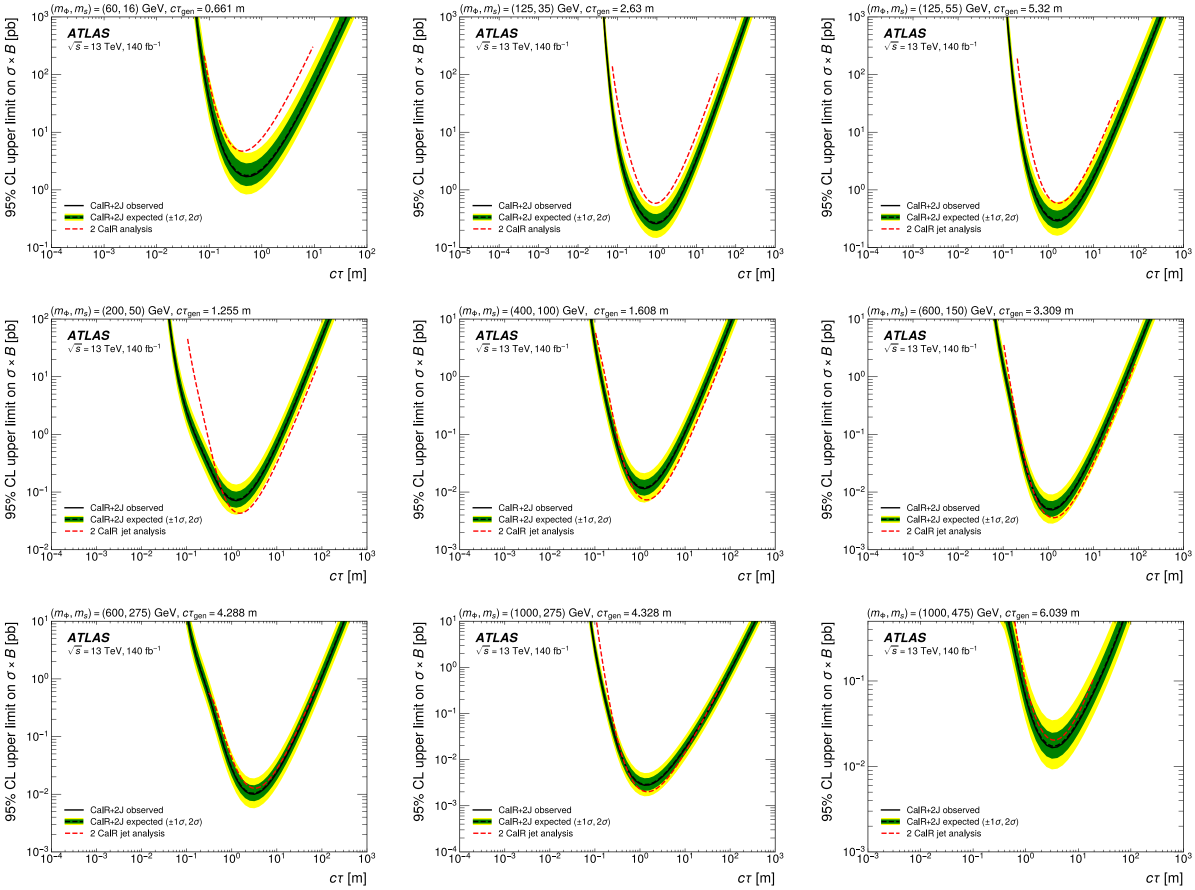


Figure 5.2: Expected and observed upper limits for HSS signal models. The “2 CalRatio jet analysis” lines refer to the search for two CalRatio jets performed in [5].

Chapter 6

Conclusions

A search for neutral long-lived particles decaying in the ATLAS detector was presented. The targeted event topology was that of three displaced jets being produced by two long-lived particles resulting from the decay of a heavy mediator particle. These jets were identified by characteristics of LLPs decaying in or near the calorimeter, such as isolation from particle tracks and relatively little energy being deposited in the innermost layers of the calorimeter. The results complement those from other ATLAS searches using different detector signatures, as shown in Figure 6.1. In particular, this analysis improves the ATLAS upper limit on LLP signals in an intermediate lifetime region between searches focused on displaced vertices in the inner detector and in the muon spectrometer.

One limitation for the ABCDisCo neural network used in this analysis was the limited statistics available for an independent sample of background events. The BIB sample used for training contains fewer events than the main dataset and especially fewer of the very rare QCD events that make up the most signal-like background. Since this analysis was particularly sensitive to low-boost LLPs, the jets for many of the signal samples studied tended to have late timing. This is both because the LLPs are traveling at a speed appreciably slower than c and because the wide opening angle of the displaced decay can lead to a significantly increased path length to the calorimeter. Compared to an ultrarelativistic particle traveling directly from the IP, there can be a delay of several nanoseconds before the decay products are detected in the calorimeter. This is already exploited by the ABCDisCo neural network, but it could also be used at

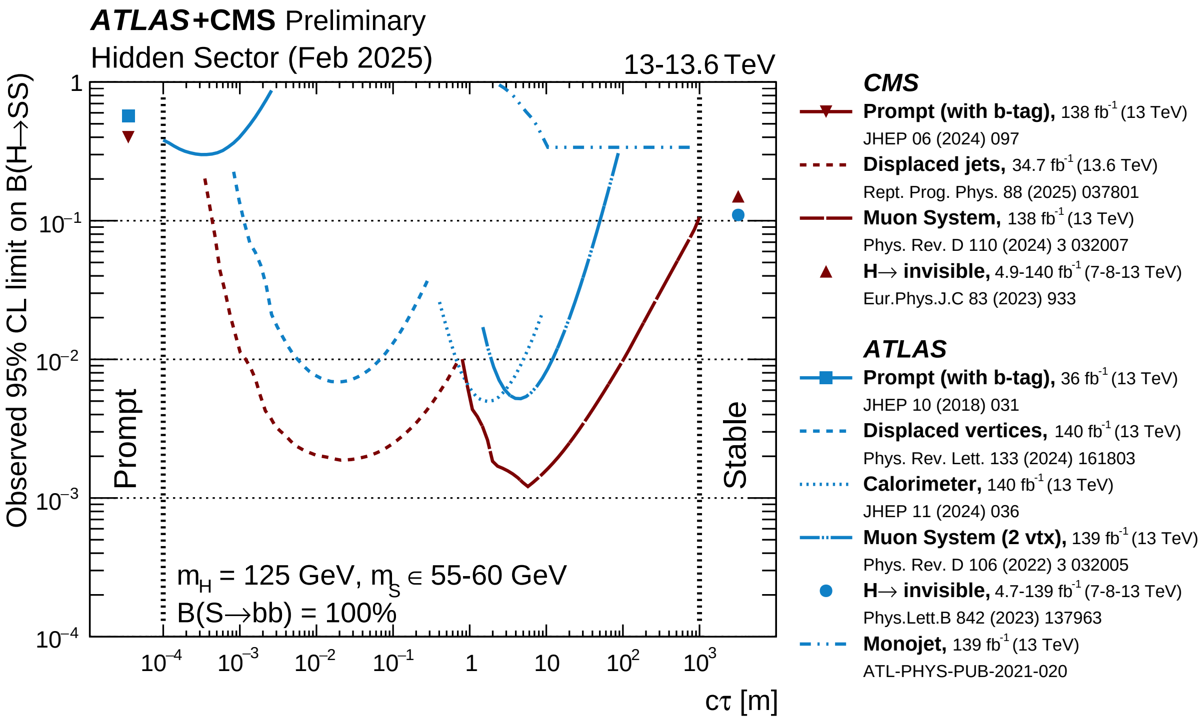


Figure 6.1: 95% confidence level upper limits on the SM Higgs boson decaying to a pair of neutral scalars as a function of $c\tau$ for a hidden sector model where m_H = 125 GeV and m_s is in the range of 55 to 60 GeV. In this plot, the long-lived scalars are then assumed to decay to b-quarks with 100% branching fraction. This results of this analysis are repesented by the line labeled “Calorimeter” in the legend.

the trigger level to improve efficiency to these signals.

As this analysis is primarily limited by statistical uncertainty, the expected limit would improve with additional data. The statistical significance of a number of signal events s with a number of background events b is approximately $s/\sqrt{b}$, which is proportional to $\sqrt{\mathcal{L}}$, where $\mathcal{L}$ is the total integrated luminosity. A caveat is that the analysis is sensitive to the number of pileup interactions, which decrease the signal efficiency of the trigger and analysis selections. Since pileup will increase significantly after the High-Luminosity LHC upgrade, investigating strategies to mitigate this effect will be important.

Another option to improve or expand sensitivity is to search for LLPs produced in association with Standard Model particles. For example, this was done for combinations with a W or Z boson decaying to leptons, setting upper limits on new signal models. This allows using a trigger based on the associated object, which can help avoid the issue of the large QCD background that limits low-p_{T} searches using the CalRatio triggers.

Looking further ahead, many dedicated LLP experiments have been proposed, both at current and future colliders. Now that interest in LLP searches has increased significantly since the start of the LHC, the implications of these types of analyses should ideally inform the designs even of general-purpose detectors at future colliders. Such foresight could improve sensitivity to LLP signals beyond a simple scaling of energy or luminosity.

Appendix A

MATHUSLA

Searches for neutral LLPs at collider experiments are limited by both geometric acceptance and high Standard Model backgrounds. These LLPs can only be detected if their decay length is short enough to be within the experiments, and this signal can be difficult to distinguish from very rare QCD events. For ultra long-lived partilces (ULLPs) with lifetimes such that $\beta\gamma c\tau \gtrsim 100$ m, most particles will escape the detectors at the LHC interaction points before decaying. Their presence could be hinted at by missing transverse momentum, but this is not discernable from the case of stable neutral particles, such as neutrinos.

A.1 Proposed experiment

MATHUSLA (MAssive Timing Hodoscope for Ultra Stable neutraL pArticles) is a proposed experiment dedicated to searching for ULLPs produced at the HL-LHC [60]. As shown in Figure A.1, the proposed site lies adjacent to the surface infrastructure for the CMS experiment at Point 5 of the LHC. The front of the detector would be displaced 81 m vertically and 70 m horizontally from the IP (Figure A.2) [61]. The more than 100 m of rock between completely shields MATHUSLA from all Standard Model particles produced in the LHC other than muons and neutrinos. The floor and wall facing the IP are instrumented with veto layers that detect incoming muons from the LHC. The ceiling and the remaining walls feature several tracking layers, which are used to reconstruct tracks and vertices from LLPs decaying inside the detector volume. Both the veto layers and the tracking layers utilize plastic scintillators and silicon photomultipliers to detect hits.

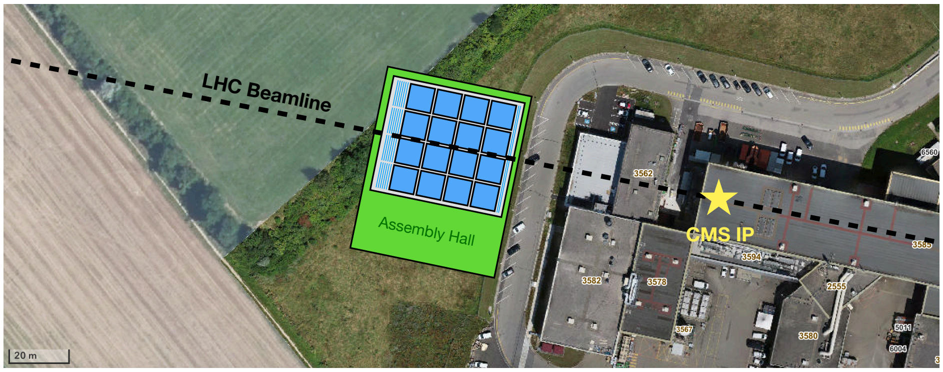


Figure A.1: Location of proposed MATHUSLA detector at the CMS site.

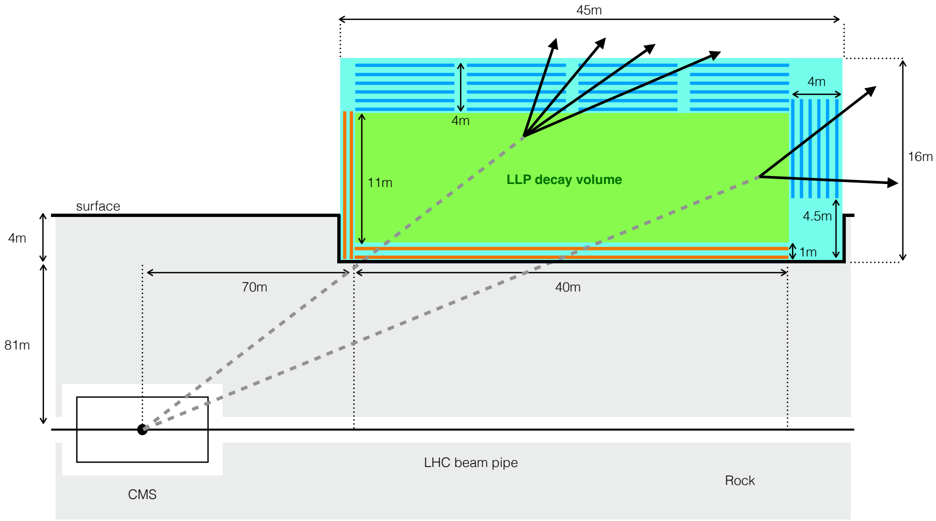


Figure A.2: Diagram of MATHUSLA location and potential LLP decays.

The primary disadvantage of a detector at the surface is the much higher flux of cosmic ray particles. The downward cosmic ray tracks can be identified by the timing of the hits in the tracking layers. However, these downward-going particles can interact with material in or near the detector to produce upward-going particles.

Figure A.3 shows the expected sensitivity of MATHUSLA compared to some other experiments for Higgs boson decays to LLPs. These results are based on 40 m × 40 m horizontal area of the detector and incorporate the signal efficiency after applying selections that eliminate the simulated backgrounds of cosmic rays, LHC muons, and neutrinos.

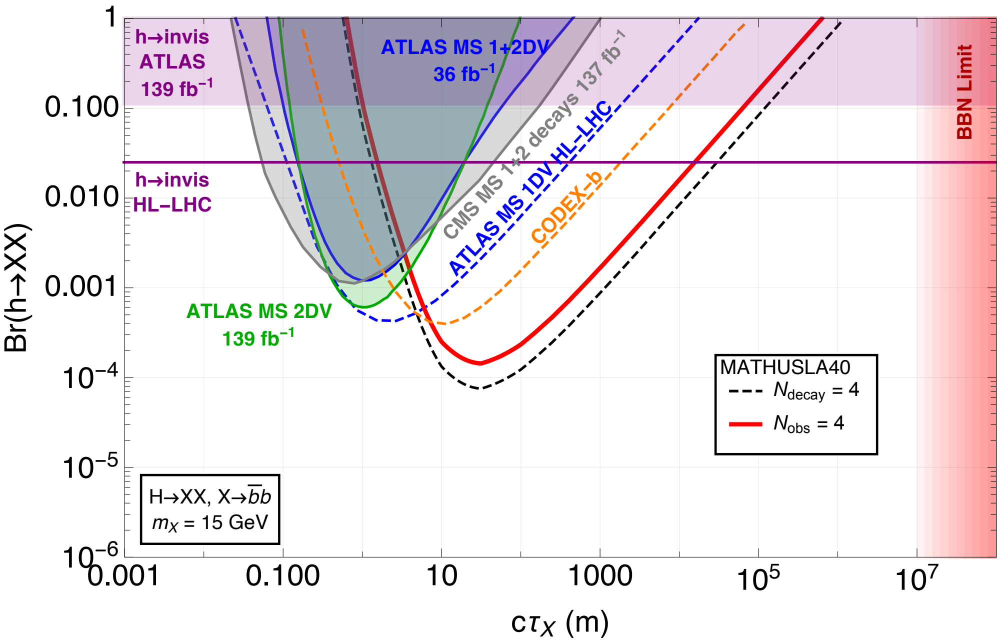


Figure A.3: Sensitivity of MATHUSLA and other experiments to hadronically decaying LLPs produced in exotic Higgs boson decays.

In order to help validate design assumptions and quantify backgrounds at the surface above LHC, a test stand was built and operated above ATLAS at Point 1 during collisions in 2018 [62]. The rest of this appendix describes this test stand in detail.

A.2 Test stand

The MATHUSLA test stand comprised two planes of scintillation counters, one at the top and one at the bottom, with six layers of RPCs between them that were grouped into three double-layers. Figure A.4(a) shows the basic design of the test stand. The overall structure was 6.8 m tall with a distance of 6.5 m between the lowermost and uppermost scintillators. The three RPC double-layers were located at approximately 2 m, 3.7 m, and 5.5 m above the lowermost scintillators. The structure had a base of 2.98 × 2.91 m^2 with an active area of approximately 2.5 × 2.5 m^2. Figure A.4(b) shows the test stand in the SX1 building at CERN, 80 m above the ATLAS IP.

The scintillator planes were used to form the trigger for the test stand, while the RPC layers were used to measure spatial and time coordinates for tracking.

(a)

(b)

Figure A.4: (a) 3D model of the MATHUSLA test stand. (b) Photo of the final assembled structure installed above the ATLAS IP. The green dots identify the two scintillator layers used for triggering, while the red dots mark the three RPC double-layers used for tracking.

A.2.1 Scintillation counters

The scintillators used in the MATHUSLA test stand are spare scintillation counters from the forward muon trigger system of the DØ detector at the Tevatron at Fermilab [12]. The scintillator tiles are made of 12.7 mm-thick BICRON 404A plastic, which has a light emission peak at 420 nm and an attenuation length of 1.7 m. Each tile has two wavelength-shifting (WLS) bars with an absorption peak matching the emission peak of the scintillator. As shown in Figure A.5, the WLS bars are located on two edges of each scintillator and double as light guides. The bars are made of SOFZ-105, based on PMMA (polymethylmethacrylate) plastic, and contain the wavelength-shifting fluorescent dopant Kumarin 30. One end of each bar directs the light signal into a 25 mm-diameter MELZ 115M photomultiplier tube (PMT). The sensitivity peak of the PMTs matches the 480 nm emission peak of the WLS bars. At the end opposite to the PMT there is a mylar tape reflector. The scintillator tile and WLS bars are wrapped in TYVEK type 1056D and photographic paper to ensure light-tightness and are encased in an aluminum outer shell. Each PMT is connected to a high-gain base and is surrounded by a magnetic shielding tube.

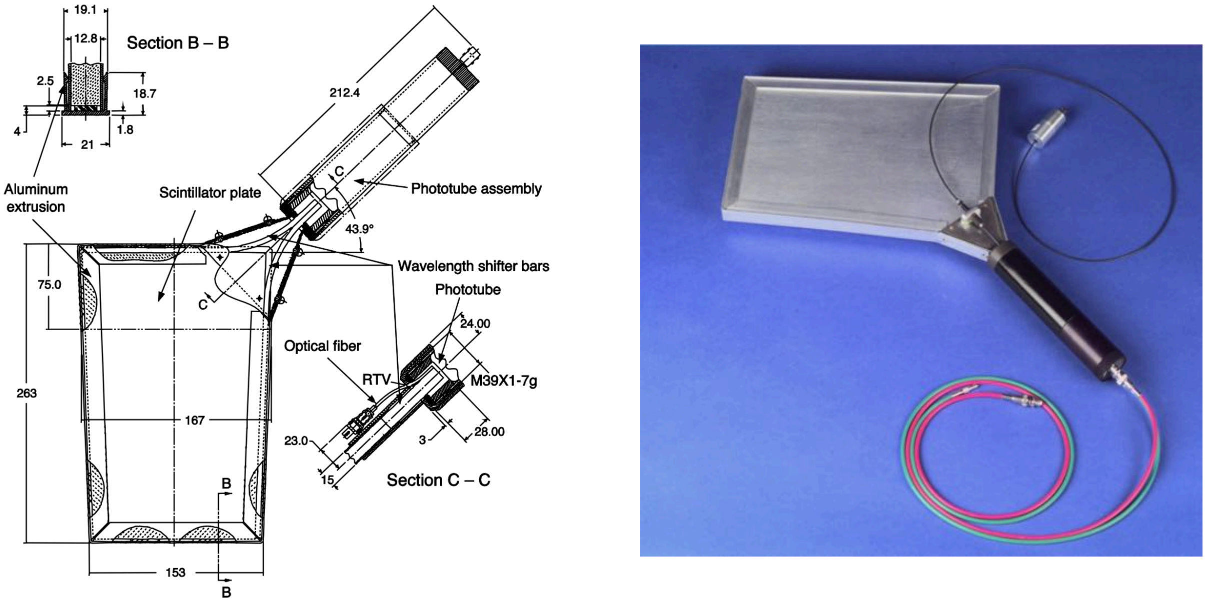


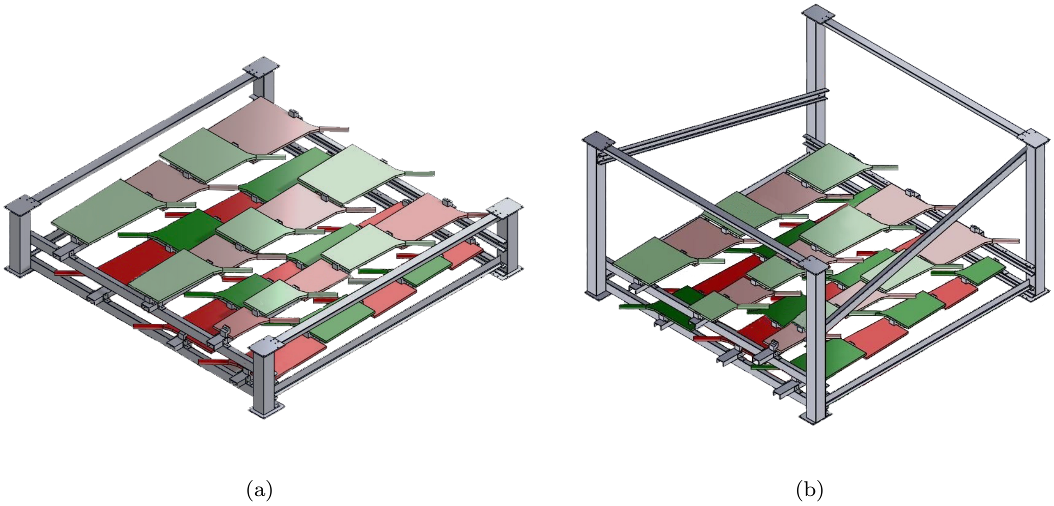
Figure A.5: Schematic [12] and photo of a DØ forward muon scintillation trigger counter.

Scintillator plane assembly

The counters were arranged into two approximately square planes with an area of 2.5 × 2.5 m^2 each. The counters were placed in rows at different heights as shown in Figure A.6 to allow for overlap in order to avoid gaps in the area coverage. The top and bottom scintillator planes were composed of 28 and 31 counters, respectively. The size of the smallest counters is approximately 22 × 37 cm^2 while the largest counters measure 63 × 69 cm^2.

A.2.2 Resistive plate chambers (RPCs)

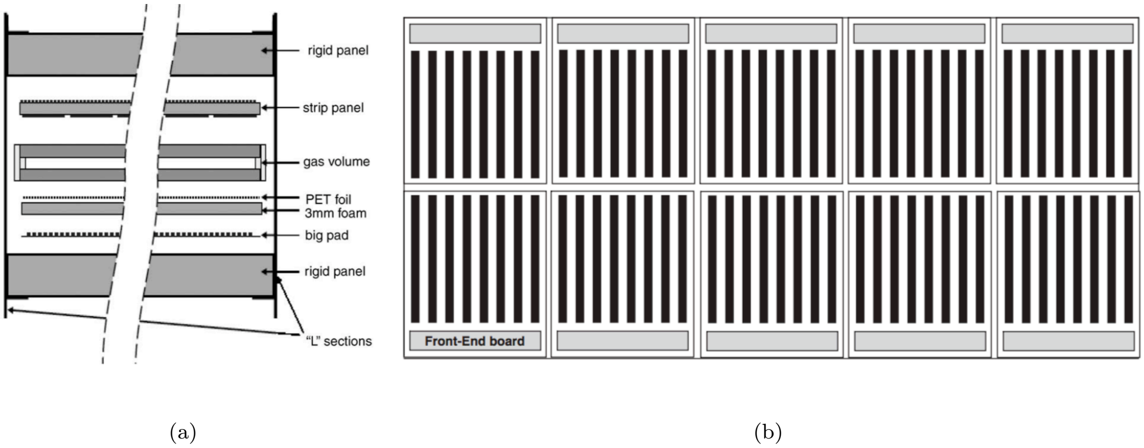
Spare RPCs from the ARGO-YBJ experiment [63] were used for the MATHUSLA test stand. Each chamber consists of a 2 mm-thick gas gap with a sensitive area of 2.70 × 1.23 m^2 and a readout strip panel, both assembled inside a 47 mm-thick Faraday cage that also serves as a mechanical support for the chamber. The readout panel can pick up the signals generated inside the gas gap by means of 80 copper strips of 6.76 × 62.35 cm^2. Figure A.7 shows a cross-section view of an ARGO-YBJ chamber and a sketch of the strip panel used for readout. The front-end boards are soldered at the end of the strips and embedded in



(a) (b)

Figure A.6: Layout of the top (a) and bottom (b) scintillator planes, each of which has a 2.5 × 2.5 m^2 active area.

the Faraday cage. Eight contiguous strips form a pad with a size of 55.68 × 62.35 cm^2. The pad signal is the logical OR of the eight strips and is used for timing.



(a) (b)

Figure A.7: (a) Schematic of the cross-section of an ARGO-YBJ chamber [13]. (b) Strip panel used for readout. The total size of each chamber is 285 × 126 × 4.7 cm^3.

For the MATHUSLA test stand, it was required that the RPCs be able to track both LHC muons and cosmic rays. The ARGO-YBJ chambers were originally designed for counting cosmic ray shower particles and not for tracking. Consequently the strip size was not optimized for spatial resolution. Additionally, in ARGO-YBJ the chambers operated in streamer mode with a gas mixture of 75% tetrafluoroethane, 15%

argon, and 10% isobutane [13]. RPCs operating in streamer mode have been used for tracking in the past [64]. In the test stand, the RPCs were operated in streamer mode using the standard ATLAS RPC gas mixture (94.7% tetrafluoroethane, 5% isobutane, and 0.3% sulfur hexafluoride [65]) with an addition of 15% of argon.

RPC layer assembly

The RPCs were arranged into six layers, each consisting of two chambers placed side by side. The RPCs in a layer were vertically offset by 10 cm to allow for overlap in order to avoid gaps in coverage. The layers were grouped into three double-layers, which were each composed of two layers horizontally rotated by 90° relative to each other. The rotated strips in each double-layer provided a measurement of two orthogonal spatial coordinates in a horizontal plane. The double-layers were also rotated slightly relative to each other. An example of one of the double-layers is shown in Figure A.8.

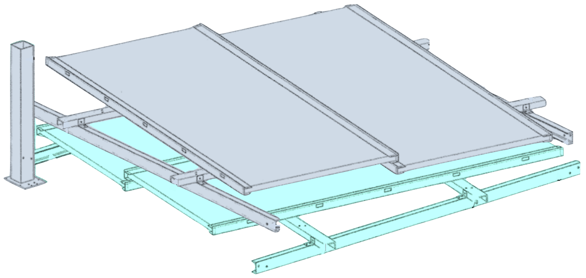


Figure A.8: Layout of one of the RPC double layers illustrating the support structure that ensures overlap of individual chambers to avoid gaps in coverage.

A.2.3 Electronics, trigger, and data acquisition

The analog signal of each scintillator was split into two paths. One was sent to an analog-to-digital converter (ADC) to measure the charge of each pulse. The other was sent into a discriminator, providing a logic signal. This logic signal was input to a multi-hit time-to-digital converter (TDC) to measure signal arrival times.

RPC data were acquired on receiver cards housed in a Local Station [66]. Each card recorded the address of the strips that were hit as well as the time from the corresponding pads.

The trigger for the MATHUSLA test stand was based on the top and bottom scintillator layer signals. The top (bottom) layer signal is defined as the logical OR of all the scintillation counters in the top (bottom) scintillator plane. There were two primary triggers, corresponding to *upward*-going particles and *downward*-going particles.

The logical AND of the top and bottom layer signals with the relative timing delay expected for upward-going (downward-going) particles traveling near the speed of light provided the upward (downward) trigger. Additional triggers considering only single-layer information were used for crosschecks and scintillator efficiency estimates.

When a trigger was received, the digitized charge and timing information for the scintillators was stored in the buffers of the ADCs and TDC, and the Local Station transferred the RPC data to an ARGO Memory Board. These modules were read out after each event by a PC via a VME controller board. The PC saved the complete raw data to files on disk for offline analysis.

A.2.4 Coordinate system

The origin of the test stand coordinate system is defined to be the center of the entire detector, at the midpoint of the full height, width, and length of the overall support structure. The x-axis is parallel to the ground and is aligned with the counterclockwise direction of the LHC (approximately east). The y-axis is also parallel to the ground and points away from the center of the LHC (approximately south). The horizontal edges of the support structure are aligned with these axes. The z-axis is directly downward. In this coordinate system, the ATLAS IP is at $(x, y, z) = (2.4 \text{ m}, 0.0 \text{ m}, 83.0 \text{ m})$.

The strips of each RPC layer in the test stand were approximately aligned with either the x- or y-axis. Each RPC double-layer consisted of one x-measuring layer and one y-measuring layer. Based on the RPC strip width (Section A.2.2), the resolution of each spatial coordinate measurement is approximately 2 cm.

Tracks in the test stand are spatially parameterized by two angles. *Zenith angle* refers to the smallest angle between the track and the z-axis (either the $+z$ or $-z$ direction) and ranges from $0°$ to $90°$. The *azimuthal angle* is the 2D polar angle of the projection of the $+z$ direction of the track onto the (x, y)-plane and ranges from -180° to 180°. The angles as defined here rely only on spatial information and do not depend on whether a track is upward-going or downward-going.

A.2.5 Timing calibration

The use of timing information is crucial in reliably reconstructing good tracks and distinguishing upward-going tracks from downward-going tracks. As such, it is imperative to ensure that the timing is consistent across all detector elements in the test stand by applying appropriate timing calibrations. Two types of timing calibration are applied: one that addresses characteristic delays between detector elements and one that addresses time slewing. After applying these calibrations, the timing resolution of the scintillation counters is better than 3 ns and a typical RPC pad has a resolution better than 4 ns [67].

Characteristic delays

There is a delay from the time when a particle hits a detector to the time when the hit is recorded. This delay depends on characteristics such as high voltage settings, drift time within the detector, and cable lengths. Hence, each of the 59 scintillators and 120 RPC pads is characterised by its own typical delay. To use timing information appropriately, these delays are calibrated.

The calibration is performed with downward-going cosmic rays. In each event and for each possible pair of detectors that were hit, the time difference is recorded, minus the expected time of flight between them for a particle traveling at the speed of light. If all delays were the same for all detectors, the mean of this distribution would be centered at zero. In reality, the mean is the difference between the characteristic offsets of the detectors considered. A Gaussian is fit to the distribution of the timing difference for each pair of detectors. The offsets are determined by performing a least-squares fit of the means of all the Gaussians. After all timing calibrations have been applied, a similar fit is performed for the widths of all the Gaussians in order to calculate the timing uncertainty for each scintillator and RPC pad.

Time slewing correction for scintillators

The time of a hit in a scintillator is determined by the instant when the voltage of the signal pulse passes a discriminator threshold. This introduces a time slewing effect: larger pulses cross the discriminator threshold earlier than smaller pulses. The following calibration procedure ensures consistent timing information regardless of the pulse size.

For each scintillator, the time difference between hits in the given scintillator and all other detectors (corrected for time of flight) is plotted against the integrated charge of the corresponding pulse in the scintillator. A power function is fitted to the distribution and the parameters of that function are used to provide a time correction based on pulse charge.

A.2.6 Track reconstruction

Tracks are reconstructed by a least-squares fit using spatial and timing information from the RPC and scintillator hits. The algorithm starts by fitting all RPC and scintillator hits in an event to a straight line consistent with the speed of light. Tracks are identified as either upward-going or downward-going by the direction which results in the smallest χ^2 value. In order to remove detector noise and to separate hits produced by multiple particles, an iterative process is run in which the hit with the largest χ^2 contribution is removed and the track is refitted to the remaining hits. This process proceeds until all hit residuals are smaller than a given threshold, resulting in the final reconstructed track. The entire process is then repeated with all the discarded hits to form additional tracks, and this continues until no more tracks can be formed. Good tracks for analysis are required to contain at least one hit in the top scintillator layer, at least one hit in the bottom scintillator layer, and hits in at least four different RPC layers. Only upward-going (downward-going) tracks in events passing the upward (downward) trigger and failing the downward (upward) trigger are considered in the data analysis.

In general, events recorded by the test stand are very clean: 97% of the good reconstructed tracks are in events where a maximum of 2 RPC hits and a maximum of 2 scintillator hits were discarded by the tracking algorithm. More than 70% of the good tracks are in completely clean events where no hit was discarded. Only 0.1% of events with at least one good track contain more than one good track.

Figure A.9 shows an example of a downward-going track (top) and an upward-going track (bottom) from test stand data. The left panels of the figure are event displays of the example tracks, showing the scintillation counters and RPC pads corresponding to track hits in green. The red line represents the fitted track. The right panels of the figure show the time and z-coordinates of the hits forming the track.

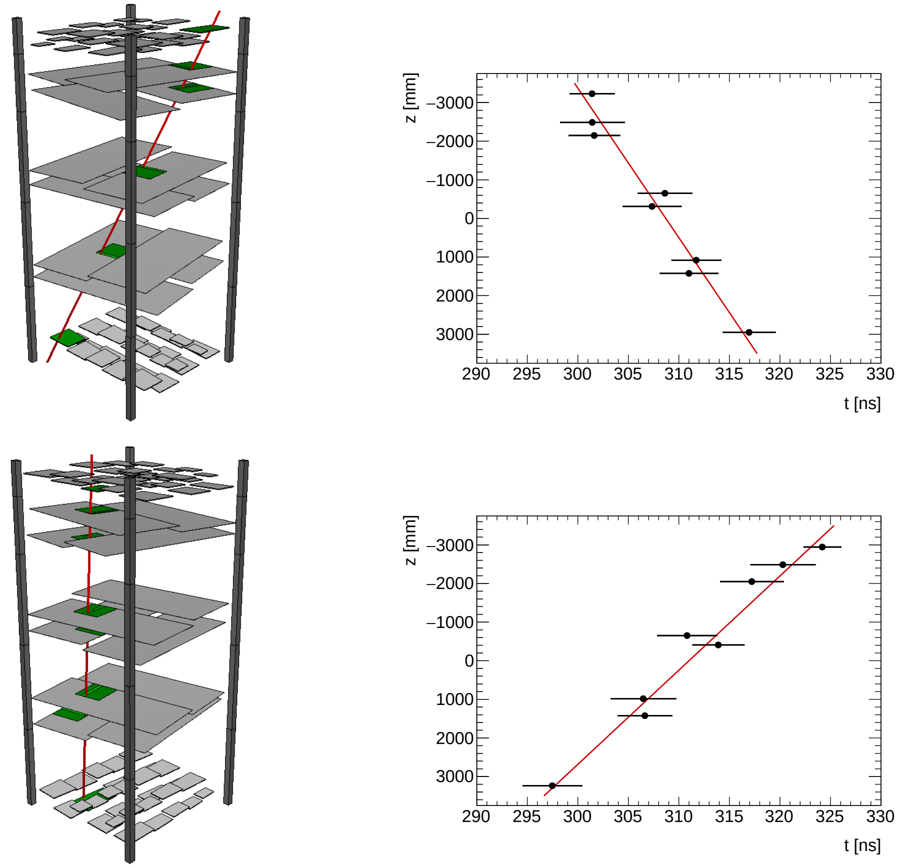


Figure A.9: Distribution of hits for examples of a downward-going track (top) and an upward-going track (bottom) from data. Left: Event display of the scintillation counters and RPC pads comprising the track hits. Right: Plot of the z-coordinate versus time for each hit. The red line represents the fitted track in each case.

A.2.7 Detector efficiency

The efficiencies of all scintillation counters and RPC strips were estimated using the data collected by the test stand. RPC efficiencies were calculated using data from the primary triggers, using information from both the top and the bottom scintillator layers. To avoid a trigger bias in the calculation of the scintillators' efficiency, data selected by the single-layer triggers were used. For this procedure, tracks were reconstructed as in Section A.2.6 but with less restrictive hit requirements to allow for potentially missing hits. If a track intercepts a given scintillator or RPC, this is considered to be an expected hit for the corresponding detector. If a hit was indeed recorded in event data from the intercepted detector, this is additionally considered to be a good hit. The efficiency of each detector is the ratio of the number of good hits to the number of expected hits. Purely geometric effects induced by this procedure on the calculated efficiencies were corrected for by performing the same procedure on simulated events.

All detectors used in the test stand were obtained after extensive use in previous experiments, so they spanned a large range of efficiencies. The individual scintillator efficiencies ranged from 60% to 98%. The RPCs had efficiencies ranging from 55% to 85% except for two of them that contained dead modules and so had very low total efficiency.

A.2.8 Simulation

The geometry and material of the ATLAS cavern and the test stand and its surroundings as shown in Figure A.10 (to scale) were modeled with GEANT4 10.6 [68]. Starting from the ATLAS IP (green star in the diagram), the material of the ATLAS detector (blue box in the diagram), equivalent to approximately 11 nuclear interaction lengths, is simulated by introducing a 1.85 m-thick cylindrical shell of iron. The rock (gray hatched area in the diagram) surrounding the ATLAS cavern was approximated by 45.30 m of sandstone, 18.25 m of marl, and 36.45 m of an equal mixture of sandstone and marl, as determined from a geological survey [69, 70]. Both the scintillation counters and the RPCs, as well as parts of the supporting structure, are included in the simulation. Air, comprised primarily of a standard admixture of nitrogen and oxygen, fills the gaps. The position of the test stand (red crosshatched area in the diagram) relative to the IP was determined by a combination of direct measurements of the test stand inside the SX1 building and engineering drawings of SX1 and the ATLAS cavern.

The energy deposited by an ionizing particle passing through a detector element in GEANT4 is saved as several small deposits within a few nanoseconds. These energy deposits in scintillators and RPC gas were integrated to form candidate detector hits. Each candidate hit is assigned an efficiency, corresponding to the one calculated in Section A.2.7 for the scintillator or RPC that was hit.

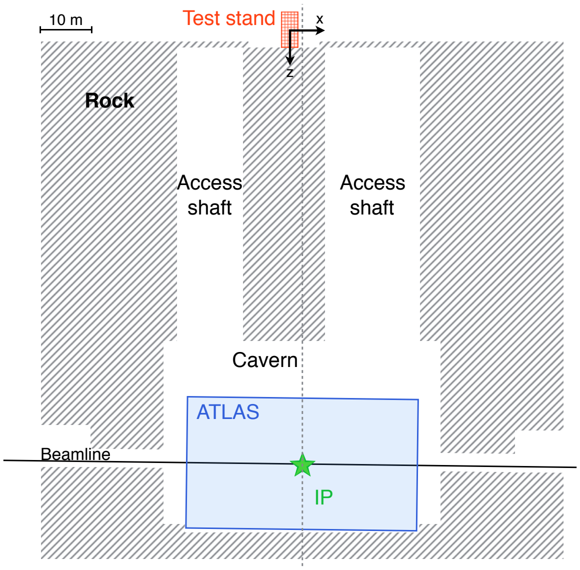


Figure A.10: Diagram with the geometry and material of the ATLAS cavern, the test stand, and its surroundings.

Cosmic rays

Cosmic rays are the dominant source of energetic charged particles at Earth's surface and represent the vast majority of tracks in the test stand. Since these particles originally arise from primary cosmic rays entering the atmosphere, they are initially directed downwards toward the surface. Downward cosmic rays can inelastically scatter or decay in test stand material or in the concrete floor and generate additional charged particles. These processes can produce particles traveling in any direction, including upwards. This upward contribution, referred to as *cosmic ray inelastic backscattering*, can activate the upward trigger of the test stand and produce upward-going tracks.

Downward cosmic ray particles, including muons, electrons, positrons, protons, neutrons, and photons, were generated by sampling energies and zenith angles from distributions predicted by PARMA4.0 [71, 72, 73], an analytical model for estimating cosmic ray fluxes on Earth. The particles were simulated in the GEANT4 model of the test stand and detector hits were recorded. In order to study both downward tracks from incoming cosmic ray particles and upward tracks produced by secondary particles with reasonable statistics, two separate sets of simulations were run. For the first set, initial cosmic ray particles were uniformly spatially distributed in GEANT4 just above the top scintillator plane of the test stand within a horizontal square area of 4.6 × 4.6 m^2. To study upward tracks generated by secondary particles, the initial cosmic ray particles were uniformly distributed 0.6 m above the bottom of the test stand within a horizontal square area of 6.2 × 6.2 m^2. This height was chosen so as to be just above the raised concrete floor adjacent to the test stand. The horizontal area chosen for the initial distribution of particles in the first (second) set of simulations ensured that any point on a top (bottom) scintillator received 95% or more of the total cosmic ray flux.

LHC pp collisions

The dominant LHC pp collision processes that can produce particles reaching the test stand are the production of W, Z, $c\bar{c}$, $b\bar{b}$, and $t\bar{t}$. Muons from the immediate or sequential decays of these particles are the main source of hits from pp collisions in the test stand.

Direct simulation of the above processes at $\sqrt{s} = 13$ TeV by Pythia 8.2 [39] was used to estimate the acceptance, which is defined as the number of events in which Geant4 records a sufficient collection of candidate hits to reconstruct a good upward track (as defined in Section A.2.6) divided by the total number of events generated. From the acceptance, a raw track rate can be computed that does not account for the efficiency of the detector elements.

The processes for W, Z, and $t\bar{t}$ were normalized to their measured cross-sections at 13 TeV [74, 75]. The processes $c\bar{c}$ and $b\bar{b}$ were simulated with a minimum p_{T} threshold on the leading outgoing parton of 25 GeV. Lower p_{T} thresholds give a very small contribution to the test stand rate because the minimum p_{T} a muon must have in order to reach the surface is approximately 30 GeV. Both ATLAS and CMS have measurements of the inclusive b-jet cross-section at 7 TeV [76, 77], but measurements at 13 TeV were not available. Although neither experiment measures the cross-section in a region of phase space directly relevant to the test stand, in the most relevant regions of phase space Pythia over-predicts the data by approximately 25%. Thus a 25% systematic uncertainty was assigned to the cross-section for $b\bar{b}$ and $c\bar{c}$.

There are considerable uncertainties in the material composition between the test stand and the IP. What is simulated is an approximation of a nearby geological survey, but the magnitude of local variations in the material is unknown. Furthermore, a small change in the total material can lead to a large change in the rate of upward-going muons because the initial momentum spectrum decreases steeply with increasing momentum. For instance, varying the depth of the rock between the IP and the test stand by ±1 m (a 2% change in the total rock) results in a 5–10% change in the rate of muons, depending on the process. The rate of muons from W bosons in particular is sensitive to the material description because the momentum distribution for these muons peaks around 40 GeV, which is in a range for which the survival probability to reach the surface is changing quickly. The material uncertainty leads to a 5–10% systematic uncertainty on the predicted rate of muons.

Normalizing the rates to an instantaneous luminosity of 10^{34} cm^{-2} s^{-1} results in an expected total raw track rate of 12.3 ± 1.4 per hour, including systematic uncertainties, before applying detector inefficiencies. These results are summarized in Table A.1.

Table A.1: Summary of cross-sections, acceptances, and expected rates at the test stand from various LHC pp collision processes at $\sqrt{s} = 13$ TeV. The rates are normalized to an instantaneous luminosity of 10^{34} cm^{-2} s^{-1} and do not include test stand detector inefficiencies.

Process	Cross-section [nb]	Acceptance $\left(\times 10^{-6}\right)$	Raw track rate [hr^{-1}]
$W \to \mu\nu$	20.6 ± 0.7	6.10 ± 0.17	4.5 ± 0.4
$W \to \tau\nu$		0.57 ± 0.02	0.42 ± 0.04
$Z \to \mu\mu$	1.98 ± 0.06	17.3 ± 0.4	1.23 ± 0.11
$Z \to \tau\tau$		1.59 ± 0.04	0.11 ± 0.01
$c\bar{c}$	5600 ± 1400	0.0052 ± 0.0007	1.1 ± 0.3
$b\bar{b}$	5300 ± 1300	0.0257 ± 0.0016	4.9 ± 1.3
$t\bar{t}$	0.75 ± 0.09	4.64 ± 0.12	0.12 ± 0.02
Total	—	—	12.3 ± 1.4

A.2.9 Results

The track rates and distributions from data are compared to expectations from simulation after applying detector efficiencies. Data are separated into periods when no beam was present in the LHC (runs with no beam) and periods when there were beams circulating in the LHC (runs with beam). Only good tracks as defined in Section A.2.6 are included in these studies.

Figure A.11 shows the zenith angle (left) and azimuthal angle (right) distributions of downward tracks in all data and the expected distributions from the downward cosmic ray simulation, normalized to data. The good agreement between these sets of events confirms that the downward-going tracks are properly reconstructed.

The number of downward cosmic ray tracks increases with zenith angle from 0° to 10° due to the increasing solid angle. The distribution peaks at about 10° and decreases at higher zenith angles as the geometric acceptance of the test stand diminishes. The fluctuations in the number of tracks as a function of azimuthal angle also reflects the geometric acceptance of the test stand.

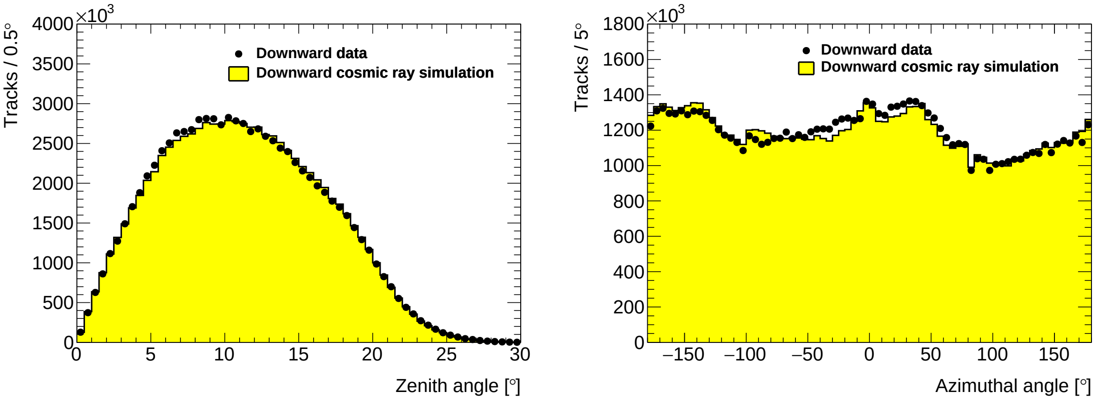


Figure A.11: Distribution of reconstructed downward-going tracks as a function of the zenith angle (left) and the azimuthal angle (right). Data events are shown as black markers. The yellow area corresponds to the downward cosmic ray simulation, normalized to data.

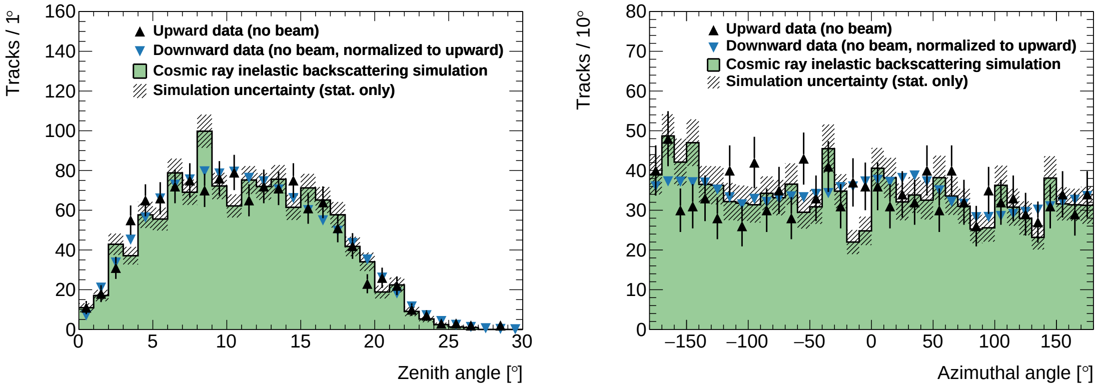
In the case of upward-going tracks, there are two components to take into account. The first component comes from cosmic ray inelastic backscattering. Upward particles generated by cosmic rays can be produced by cosmic muon decays or by interactions with material in the test stand or in the floor of the SX1 building. These upward particles are emitted across the entire geometric acceptance of the test stand, as seen in Figure A.12, where upward tracks in data with no beam are shown as upward-pointing black triangles. The angular distributions for downward tracks from the same dataset, normalized to the number of upward tracks, are shown as downward-pointing blue triangles. These upward and downward data distributions are consistent within uncertainties. The final angular distributions of these tracks are dominated by the narrow geometric acceptance of the test stand and are insensitive to the initial angular distributions of the interacting particles.

From these runs without beam, a ratio that relates the rate of upward inelastic backscattering to the rate of incident downward cosmic rays is obtained:

$$R_{\mathrm{up-to-down}} = \frac{(\text{Number of upward tracks})_{\text{data, no beam}}}{(\text{Number of downward tracks})_{\text{data, no beam}}} = (7.0 \pm 0.2) \times 10^{-5},$$

where purely statistical uncertainties are shown.

Figure A.12 also compares the data distributions to the results of the cosmic ray inelastic backscattering simulation. The statistical uncertainty of the simulation is shown by the hatched area. The predicted rate and angular distributions from this simulation are compatible within uncertainties with the observed data.



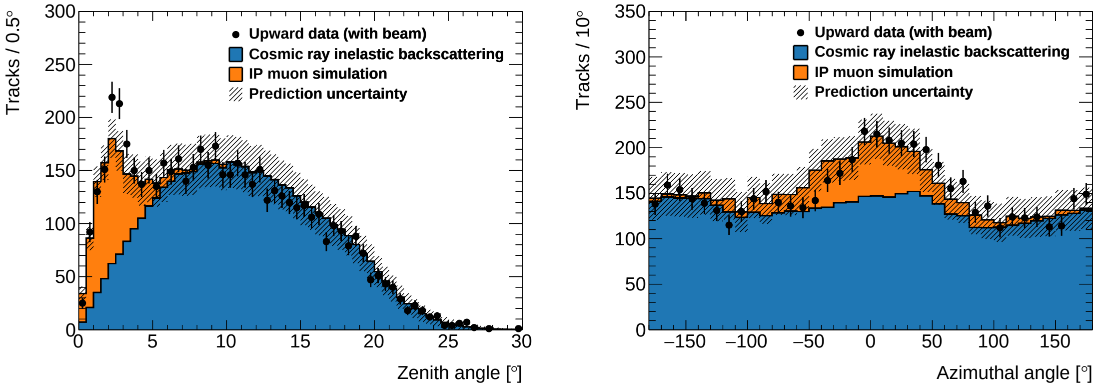
(a)

(b)

Figure A.12: Distribution of reconstructed upward-going tracks as a function of the zenith angle (left) and the azimuthal angle (right). Data events with no beam are shown as upward-pointing black triangles. The green area corresponds to the cosmic ray inelastic backscattering simulation, with the statistical uncertainty shown by the hatched area. This distribution is normalized by multiplying it by the ratio of downward tracks in simulation to downward tracks in data without beam. Downward-going tracks from data events with no beam, normalized using the $R_{\mathrm{up-to-down}}$ factor, are included in the plot as downward-pointing blue triangles.

The second component of upward tracks comes from muons created in LHC *pp* collisions. Since the test stand operated almost directly 80 m above the ATLAS IP, tracks from these muons are expected to be concentrated at small zenith angles.

Figure A.13 shows the angular distributions of upward tracks in data with beam (black markers). The blue area is the prediction for tracks from cosmic ray inelastic backscattering that is derived from the distribution of downward tracks reconstructed in data with beam, normalized by the ratio $R_{\mathrm{up-to-down}}$ defined above. Given that both angular distributions are identical, downward tracks are used in this normalization to avoid large statistical fluctuations from the smaller upward tracks dataset. The orange area in this figure shows the expected tracks from the simulation of muons produced in LHC collisions. The predicted track distributions from this simulation, after accounting for detector efficiencies, are normalized by the total integrated luminosity reported by ATLAS during the test stand runs with beam. The hatched area in this plot shows the combination of uncertainties of the cosmic ray inelastic backscattering prediction and the IP muon simulation.



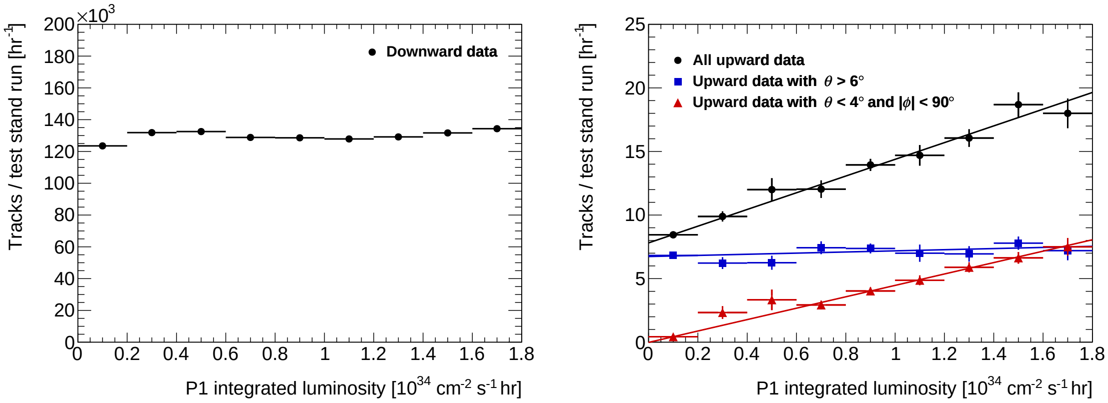
(a) (b)

Figure A.13: Distribution of reconstructed upward-going tracks as a function of the zenith angle (left) and the azimuthal angle (right). Data events in runs with beam are shown as black markers. Overlaid is a simulation of particles coming from the ATLAS IP in orange. The blue distribution corresponds to cosmic ray inelastic backscattering, showing the downward-going tracks from data runs with beam, normalized using the $R_{\mathrm{up-to-down}}$ factor. The hatched area shows the combination of the uncertainties of the IP muon simulation and the cosmic ray inelastic backscattering prediction.

The number of tracks per hour was studied as a function of luminosity. Figure A.14(a) shows that the rate of downward tracks is independent of the luminosity, as expected from cosmic rays. The small fluctuations in the number of tracks for different luminosity points are caused by fluctuations in RPC efficiency. Figure A.14(b) shows the rates for upward tracks. Black circles show the rate for all upward tracks. Blue squares represent the tracks with a zenith angle (θ) greater than 6° and correspond to the majority of the inelastic backscattering from cosmic rays. As expected, this distribution is independent of the luminosity at the LHC. Tracks with a zenith angle less than 4° and an absolute value of azimuthal angle (ϕ) less than 90° correspond mainly to particles coming from LHC *pp* collisions. These are shown by the red triangles, where a clear trend is observed with the rate increasing linearly as the luminosity increases. This confirms that the peak of upward tracks at small zenith angles is correlated with beam activity. A linear fit to these points was performed, providing the following result, in which the quoted uncertainties are purely statistical.

$$\text{Upward tracks } (\theta < 4^\circ,\, |\phi| < 90^\circ) = (4.48 \pm 0.16) \times \left(\frac{\text{Integrated luminosity}}{10^{34}\ \text{cm}^{-2}\ \text{s}^{-1}\ \text{hr}}\right) + (-0.02 \pm 0.03)$$

The positive slope and intercept near zero are strong evidence that tracks in this solid angle selection are predominantly coming from LHC particles.



(a) (b)

Figure A.14: Distribution of the number of reconstructed tracks as a function of the ATLAS integrated luminosity during each one-hour test stand run. Left: Downward tracks. Right: Upward tracks (black circles), including tracks with a zenith angle (θ) $> 6^\circ$ (blue squares) and tracks with a zenith angle $< 4^\circ$ and absolute value of azimuthal angle (ϕ) $< 90^\circ$ (red triangles).

In order to compare the rate of upward-going tracks between simulation and the data, tracks with a zenith angle less than 10° are selected. This selection is different than what is used in the fit estimate and is deliberately loose in order to avoid uncertainties from potential mismodeling of the zenith angle distributions that might bias the result. After accounting for detector inefficiency, the predicted rate of IP muon tracks from all sources with a reconstructed zenith angle less than 10° is 4.8 ± 0.5 per (10^{34} cm^{-2} s^{-1} hr). A rate of 5.7 ± 0.7 IP muon tracks per (10^{34} cm^{-2} s^{-1} hr) is measured after subtracting the expected number of tracks from cosmic ray inelastic backscattering.

A.2.10 Conclusions

The test stand results confirmed that downward-going cosmic ray particles could be clearly separated from upward-going particles based on the timing of hits in a track, and the rate of these downward tracks agreed with the expected flux within estimated uncertainties. The significant rate of upward tracks measured in the test stand, present even during periods of no collisions, was not expected beforehand, but extensive simulation studies of cosmic ray inelastic backscattering were able to reproduce the observed rates and angular distributions. After accounting for this background, the excess events observed during high-intensity *pp* collisions also agreed well with simulations of LHC muons scaled by luminosity. These results greatly informed the subsequent designs and simulations for the full-scale MATHUSLA experiment.

Appendix B

IRIS-HEP

The luminosity that will be provided by the HL-LHC presents great opportunities for many physics searches and measurements, but it also poses considerable challenges–not just to the detectors themselves, but on computing hardware, software, and data storage as well. For example, Figure B.1 shows computing resource needs for ATLAS projected into the HL-LHC era. The needs are expected outpace the rate of improvements in the underlying technologies without aggressive development and adoption of new more efficient data workflows. These advances will be necessary in order to fully take advantage of the HL-LHC data to look for new physics. This is the goal of the Institute for Research and Innovation in Software for High Energy Physics (IRIS-HEP). The rest of this appendix describes some of the projects being developed within IRIS-HEP.

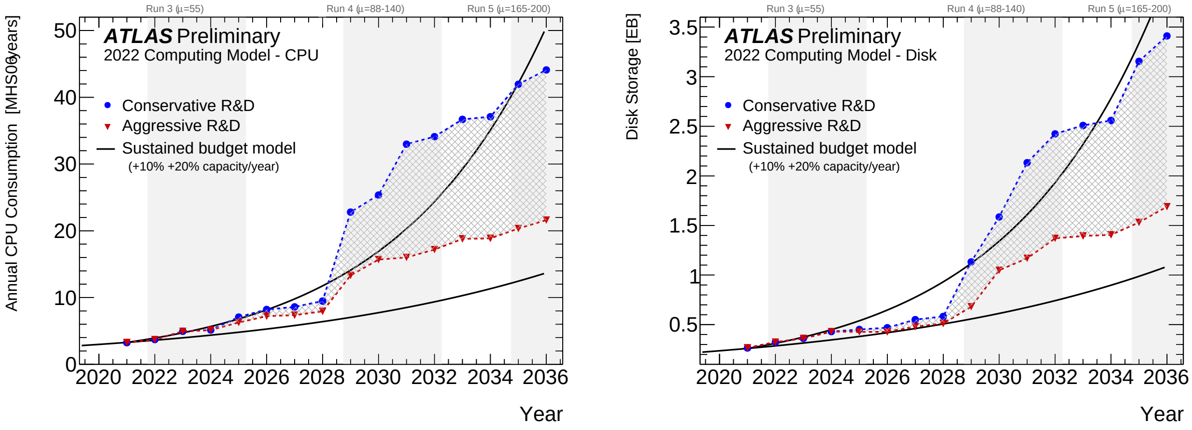


Figure B.1: Projected computing resource needs for ATLAS for CPU (left) and disk (right).

B.1 Analysis description languages

Analyses of high energy physics (HEP) data typically consist of executing a carefully designed algorithm on millions of collision events. The algorithm is specific to the analysis but identical for all events in a given data-taking period. The per-event results are aggregated over the entire dataset, such as in the form of histograms or a table recording the number of events passing certain selections. One of the most common ways to write analysis code is essentially a for-loop over events, generally using C++ and the ROOT software framework [78, 79]. This procedure tightly couples the analysis code to both the file format of the data and the steps of program execution.

There have been several recent efforts to provide more declarative interfaces for data analysis in order to simplify analysis code and to allow taking advantage of optimization via parallelization and vectorized operations. Within the ROOT package, the `RDataFrame` class aims to provide these features by drawing on the concept of data frames present in other platforms like R, Apache Spark, and the Python package pandas [80]. Coffea is a CMS-focused project designed to handle cuts and histograms similarly to `RDataFrame`, but entirely in Python, utilizing the Uproot package for reading and writing ROOT files [81].

Many of the issues faced in HEP data analysis are similar to those mitigated by work in database management systems. Two important aspects of database management are data query languages and data independence [82]. Query languages such as SQL are naturally declarative, as they only specify relationships between data fields. Data independence is the concept that the interface for accessing data should not depend on how and where the data is actually stored. There is not yet a widely adopted declarative interface to write HEP analyses across a variety of underlying storage formats. Declarative languages that focus on the physics content of analysis rather than the specifics of data formats or execution steps are called Analysis Description Languages (ADLs).

B.1.1 ADL benchmarks

In order to facilitate comparisons between different ADLs, a list of functionality benchmarks were compiled that represent common types of tasks that are required in analyses. These representative tasks were inspired by conversations within the HEP Software Foundation Data Analysis Working Group. While none are extensive enough to be a complete analysis, these typical simple data transformations can serve as building blocks for an analysis of realistic complexity. The definitions of the benchmark tasks and a table with links to implementations of them in various languages are maintained in a dedicated repository [83]. This community is able to add new implementations and propose additional benchmarks by opening pull requests and GitHub issues. Implementations can be validated by performing the tasks on a standardized CMS open data file also linked from the repository and comparing to the expected results.

B.1.2 FuncADL

FuncADL (Functional ADL) is a project that aims to provide a query interface for analysis using Python as a host language [84, 85, 86]. The fundamental base class is `EventDataset`, which generically represents a dataset of collision events, regardless of the storage format or location. All queries are constructed by applying operations on an instance of `EventDataset`. The query operations were inspired by LINQ (Language INtegrated Query), a feature built into the C# language [87]. This follows a functional programming paradigm, in which selections and transformations are defined by passing functions as arguments to these operators. Query operators always act on a sequence of objects, which may be the sequence of events or a sequence that exists within each event, such as a jet collection. The most important query operations can be categorized as projection, filtering, or aggregation operators.

Projection operators apply a transformation to each element of a sequence. This transformation is based on the properties of each element. For example, this can be used to select only the properties of an event relevant for a particular analysis and to disregard the rest of the information. `EventDataset` itself is effectively a trivial projection operator passing all event information through. There are two non-trivial projection operators: `Select` and `SelectMany`. Both of these take a lambda function as an argument that specifies the transformation to apply to each sequence element. The result of `Select` is simply the sequence of results of this transformation. The `SelectMany` projection operator expects this transformation to produce a sequence for each element, and the operator concatenates all of these sequences. This allows flattening collections across all events, such as all jets in a dataset.

Filtering operations filter out elements of a sequence based on a certain condition. The filtering operator is `Where`. `Where` expects a lambda function argument that results in a boolean value. An element of the input sequence only passes through the filter if applying the lambda function to it evaluates to `True`. Otherwise the sequence element is dropped.

Aggregation operators map an input sequence to a single value. These act on each element sequentially and reduce the dimensionality. The generic form is `Aggregate` operation, which takes a lambda function and a seed value as arguments. The lambda function is executed on the first element of the sequence with the seed value as a second argument. The resulting value becomes the seed value for applying the function to the second element, which provides the seed value for the third element, and so on. There are several common aggregation operations that are specializations of the `Aggregate` operator which are implemented directly in the language for convenience. These include the operators `Count`, `Max`, `Min`, and `Sum`. `Count` outputs the number of elements in the input sequence. `Max`, `Min`, and `Sum` expect the input sequence to be numerical values and output the maximum element, the minimum element, and the sum over all elements, respectively.

Only an example subset of operators have been enumerated here. There are other operations supported for more specialized usage. In addition, several native Python types and operations retain identical functionality within FuncADL. Lists, tuples, and dictionaries are allowed within queries. All arithmetical, boolean, and bitwise Python operators are supported as well. One important special operation supported in FuncADL is `Zip`. `Zip` takes several sequences of the same length and turns them into one sequence by bundling together the elements at a given position in each input sequence. The functionality is similar to the Python builtin class `zip` [88]. However, rather than only yielding tuples, FuncADL's `Zip` operator also allows zipping dictionaries together to output a sequence of dictionaries. This makes it possible to create new ad hoc data structures inside a query.

Useful queries are generally composed of several operations. These can be applied in succession on the same sequence, or they can be nested such that some of the query operations are run on sequences within each event. The level of nested queries can continue to any depth of object structure present in the data.

Query examples

The FuncADL code for some of the ADL benchmark tasks are shown in Table B.1 as example queries. All eight benchmarks have been implemented, but only the first four are shown here for brevity. Only the FuncADL queries are shown here, but the resulting histograms made from the returned arrays can be seen in the GitHub repository [89].

#	ADL benchmark task and corresponding FuncADL query
1	Plot the missing E_{T} of all events. `ds.Select(lambda event: event.MET_pt)`
2	Plot p_{T} of all jets in all events. `ds.SelectMany(lambda event: event.Jet_pt)`
3	Plot p_{T} of jets with $\lvert\eta\rvert < 1$.

```
ds.SelectMany(lambda event: Zip({'pT': event.Jet_pt,
                                 'eta': event.Jet_eta})\
                            .Where(lambda jet: abs(jet.eta) < 1)\
                            .Select(lambda jet: jet.pT))
```

#	ADL benchmark task and corresponding FuncADL query
4	Plot the missing E_{T} of events that have at least two jets with $p_{\mathrm{T}} > 40$ GeV.

```
ds.Where(lambda event: event.Jet_pt\
                       .Where(lambda pT: pT > 40)\
                       .Count() >= 2)\
.Select(lambda event: event.MET_pt)
```

Table B.1: Implemented benchmark tasks. In each case, `ds` is the `EventDataset` object that contains the events. Only the FuncADL query is provided rather than the plot, which is a histogram of the elements in the result. These plots can be seen in the GitHub repository.

Code generation and execution

The key to achieving data independence with a query language is having backends that can handle the details of how to actually execute a query. For FuncADL, this means taking a representation of the user's query and producing a script or executable that will apply the required transformations to the data. This is the code generation step.

The FuncADL query is internally represented by an abstract syntax tree (AST). An AST is composed of nodes corresponding to syntactic elements of Python and FuncADL, such as a lambda function or a `Select` operation. Each node contains links to child nodes, corresponding to the components of that syntactic element (for example, the arguments of a lambda function and the body of the function). The AST is built using the `ast` module included in Python [90]. The backend implementations of FuncADL take the AST of a query and traverse the tree with an `ast.NodeTransformer` (or an `ast.NodeVisitor`). For each node, these `NodeTransformer`s build a representation of the corresponding operation in the appropriate framework for the particular backend. This process translates a query AST into standalone code that will perform the retrieval and transformations specified by the query. The output is the generated code text, which can be compiled or executed as needed.

One of the available backend implementations is for the xAOD format. This is the format used by ATLAS for data files produced by the collaboration for general use by analysis teams. The xAOD format is a specially structured ROOT file produced by the Athena software framework [91]. Because of the data structures used, it is generally necessary to run a collaboration-specific framework like AnalysisBase in order to usefully read and extract information from the files. Thus a FuncADL backend for xAOD files was written that generates C++ code utilizing the ATLAS xAOD libraries. This allows accessing xAOD data structures from within the FuncADL query language, such as leptons and jet collections.

Another backend available for FuncADL is based on Uproot and Awkward. Uproot is a Python package completely independent from any ROOT code that is able to read ROOT files [92]. Awkward is a related Python package that provides an interface for accessing and manipulating arrays with potentially jagged dimensions [93]. These packages are designed to be lightweight and fast, benefiting from compiled vectorized optimizations. Uproot is capable of reading any ROOT `TTree`s that contain standard data types like strings, integers, floating point values, and variable-length vectors. This makes it ideal for operating on ROOT files with a simple structure like NanoAOD (CMS) or DAOD_PHYSLITE (ATLAS) [94, 95]. The Uproot FuncADL backend translates a query AST into generated Python source code for a function that can be evaluated on a data file and returns the selected and transformed values as an Awkward array. This backend was used to generate the code for executing the example queries in the previous section.

New backends for FuncADL can be implemented at any time. All that is required is to write a new `NodeTransformer` that can take in a FuncADL query AST and produce a script or executable which applies the necessary transformations to a data file of the appropriate format. It is possible to have multiple backends that can operate on the same format. For example, a prototype for an `RDataFrame`-based backend was created for an early version of FuncADL. The Uproot and `RDataFrame` backends are both able to run on flat ntuples, so this provides the option for alternative backends and performance benchmark comparisons between them.

The final piece of implementing FuncADL is tying the frontend query language to the backends via an interface that is able to execute a query and return the result to the user. There are a few different options for this, depending on where the query should actually be executed. The main options available are running via a suite of services called ServiceX or running the full transformation locally in the same process that is submitting the query [96]. The choice is determined by which subclass of `EventDataset` is used as the base for the FuncADL query. In both cases, the communication between the frontend parsing the user's query and the backend generating and executing code is done via a format called Qastle [97].

Query AST Language Expressions (Qastle) is a way of specifying data queries detached from any host language. Qastle is essentially a plain text format consisting of LISP-like expressions corresponding to the nodes of a query AST [98]. This format was created to provide a human-readable interface between the frontend and the backends. The Qastle format also removes extraneous information from the generic Python AST structure that is not relevant to FuncADL queries. The FuncADL frontend translates a user query into Qastle and passes it to the executor associated with the particular `EventDataset` subclass used.

The query built by combining an `EventDataset` instance and operations acting on it is only evaluated on demand. That is, the query is only executed when the `.value()` method is called. Therefore the construction of the query can be done in multiple lines and can be separated from the evaluation step.

The modular design of the frontend and backends of FuncADL allows for the possibility of adding further methods of execution in the future. It is possible to drop in a replacement for the FuncADL frontend as well. Any package that is capable of producing the Qastle format can be used in conjunction with the backends. This means that it is possible to run queries with the FuncADL backends that were not even written in the FuncADL query language. This strategy has been implemented by the package `tcut_to_qastle`, which translates ROOT `TCut`-formatted strings into Qastle, and these `TCut` queries can then be executed via ServiceX [99].

B.2 ServiceX

ServiceX is a full-featured data delivery service [96]. ServiceX itself consists of several services that perform dataset resolution, code generation, and data transformation, as shown in Figure B.2. It is possible to run ServiceX locally, although it is intended to be run on a cluster as a highly scalable platform. A ServiceX instance allows for centrally handling caching and extraction of selections and can be shared by many analysis teams. This reduces the need for duplicating data and streamlines the early selection steps that generally run on data files that are too large to easily process on a single PC. Query execution via ServiceX is possible through the classes `ServiceXSourceXAOD` and `ServiceXSourceUpROOT` for the xAOD and Uproot backends, respectively. The FuncADL frontend package produces a Qastle-formatted query and sends this to ServiceX. The FuncADL backends drive the code generation and data transformation steps. The result is then sent back to the analysis user.

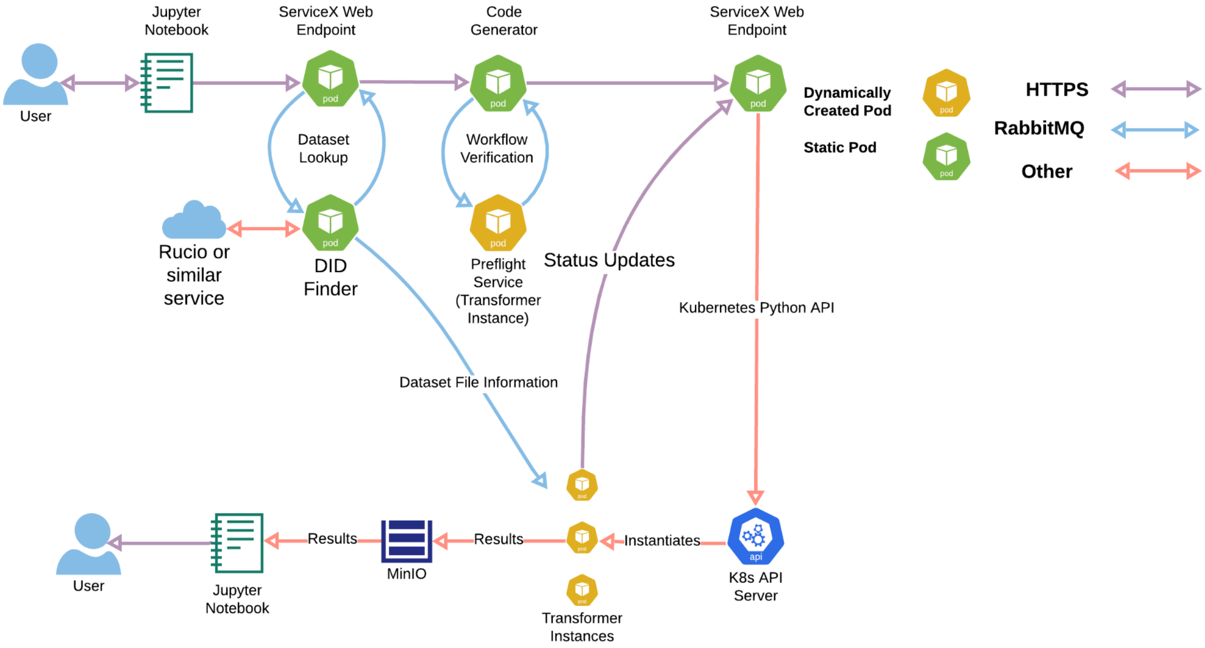


Figure B.2: Diagram of the components of ServiceX and how they interact. The user only needs to interact with the Jupyter Notebook interface.

Bibliography

[1] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 2013.

[2] S. Navas et al. Review of Particle Physics. *Phys. Rev. D*, 110:030001, Aug 2024.

[3] ATLAS Collaboration. Standard Model Summary Plots June 2024. Technical report, CERN, Geneva, 2024. ATL-PHYS-PUB-2024-011.

[4] Lawrence Lee et al. Collider searches for long-lived particles beyond the Standard Model. *Progress in Particle and Nuclear Physics*, 106:210–255, 2019.

[5] ATLAS Collaboration. Search for neutral long-lived particles in pp collisions at $\sqrt{s} = 13$ TeV that decay into displaced hadronic jets in the ATLAS calorimeter. *JHEP*, 06:005, 2022.

[6] CERN Service graphique. Overall view of the LHC. Vue d'ensemble du LHC. 2014. General Photo.

[7] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[8] Esma Mobs. The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018. 2018. General Photo.

[9] Maximilien Brice and Claudia Marcelloni. A worker inside the LHC tunnel. Vue du tunnel du LHC. 2006.

[10] Jean-Luc Caron. Cross section of LHC dipole.. Dipole LHC: coupe transversale. AC Collection. Legacy of AC. Pictures from 1992 to 2002., 1998.

[11] ATLAS Collaboration. Performance of the ATLAS trigger system in 2015. *The European Physical Journal C*, 77(5):317, May 2017.

[12] V.M. Abazov et al. The muon system of the Run II DØ detector. *Nucl. Instr. and Meth. A*, 552:372–398, 2005.

[13] Argo-YBJ Collaboration. Layout and performance of RPCs used in the Argo-YBJ experiment. *Nucl. Instr. and Meth. A*, 562:92–96, 2006.

[14] X. Fan, T. G. Myers, B. A. D. Sukra, and G. Gabrielse. Measurement of the electron magnetic moment. *Phys. Rev. Lett.*, 130:071801, Feb 2023.

[15] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012.

[16] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, 2012.

[17] Steven Weinberg. *The Quantum Theory of Fields.* Cambridge University Press, 1995.

[18] Howard Georgi. *Lie Algebras in Particle Physics.* CRC Press, second edition, 2019.

[19] Brian Hall. *Lie Groups, Lie Algebras, and Representations.* Springer Cham, second edition, 2015.

[20] Yuval Grossman and Yossi Nir. *The Standard Model: From Fundamental Symmetries to Experimental Tests.* Princeton University Press, 2023.

[21] Mark Srednicki. *Quantum Field Theory.* Cambridge University Press, 2007.

[22] J. J. Sakurai. *Modern Quantum Mechanics.* Addison-Wesley, revised edition, 1994.

[23] David Curtin et al. Exotic decays of the 125 GeV Higgs boson. *Phys. Rev. D*, 90:075004, Oct 2014.

[24] David Curtin et al. Long-lived particles at the energy frontier: the MATHUSLA physics case. *Reports on Progress in Physics*, 82(11):116201, oct 2019.

[25] Masahiro Kawasaki, Kazunori Kohri, and Takeo Moroi. Big-bang nucleosynthesis and hadronic decay of long-lived massive particles. *Phys. Rev. D*, 71:083502, Apr 2005.

[26] CERN. *LEP design report.* Report. CERN, Geneva, 1984. Copies shelved as reports in LEP, PS and SPS libraries.

[27] ATLAS Collaboration. Operation of the ATLAS trigger system in Run 2. *Journal of Instrumentation*, 15(10):P10004, oct 2020.

[28] ATLAS Collaboration. Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2. *The European Physical Journal C*, 77(10):673, Oct 2017.

[29] ATLAS Collaboration. Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC. *The European Physical Journal C*, 77(5):332, May 2017.

[30] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *The European Physical Journal C*, 77(7):490, Jul 2017.

[31] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-kt jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, apr 2008.

[32] ATLAS Collaboration. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} =$ 13 TeV with the ATLAS detector. *The European Physical Journal C*, 81(8):689, Aug 2021.

[33] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006, dec 2019.

[34] ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $sqrts =$ 13 TeV. *The European Physical Journal C*, 81(7):578, Jul 2021.

[35] ATLAS Collaboration. The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using 140 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions. 2 2024.

[36] ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J. C*, 70:823, 2010.

[37] S. Agostinelli et al. GEANT4 - a simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[38] 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*, 07:079, 2014.

[39] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159, 2015.

[40] ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. ATL-PHYS-PUB-2014-021, 2014.

[41] ATLAS Collaboration. Summary of ATLAS Pythia 8 tunes. ATL-PHYS-PUB-2012-003, 2012.

[42] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63:189, 2009.

[43] T. Sjöstrand, S. Mrenna, and P. Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[44] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys. B*, 867:244, 2013.

[45] ATLAS Collaboration. The ATLAS Tau Trigger in Run 2. ATLAS-CONF-2017-061, 2017.

[46] T. Berger-Hryn'ova, R. Caputo, K. Nagano, C. Ohm. Physics Uses and Hardware Constraints of the L1 Topological Trigger.

[47] ATLAS Collaboration. Search for long-lived neutral particles decaying in the hadronic calorimeter of ATLAS at $\sqrt{s} = 13$ TeV in 3.2 fb^{-1} of data. ATLAS-CONF-2016-103, 2016.

[48] ATLAS Collaboration. Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector. ATLAS-CONF-2015-029, 2015.

[49] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C*, 77:490, 2017.

[50] Gregor Kasieczka, Benjamin Nachman, Matthew D. Schwartz, and David Shih. Automating the ABCD method with machine learning. *Phys. Rev. D*, 103(3):035021, 2021.

[51] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ tev using the atlas detector at the lhc, 2022.

[52] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ using the ATLAS detector at the LHC. ATLAS-CONF-2019-021, 2019.

[53] G. Avoni et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *JINST*, 13(07):P07017, 2018.

[54] ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 96(7):072002, 2017.

[55] Lukas Heinrich, Matthew Feickert, and Giordon Stark. pyhf: v0.6.3. https://github.com/scikit-hep/pyhf/releases/tag/v0.6.3.

[56] Lukas Heinrich, Matthew Feickert, Giordon Stark, and Kyle Cranmer. pyhf: pure-python implementation of histfactory statistical models. *Journal of Open Source Software*, 6(58):2823, 2021.

[57] Mason Proffitt. abcd-pyhf, 2024. https://github.com/masonproffitt/abcd-pyhf.

[58] Alexander L. Read. Presentation of search results: the CL_S technique. *J. Phys. G*, 28:2693, 2002.

[59] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011.

[60] John Paul Chou, David Curtin, and H.J. Lubatti. New detectors to explore the lifetime frontier. *Physics Letters B*, 767:29–36, 2017.

[61] MATHUSLA Collaboration. Conceptual design report for the mathusla long-lived particle detector near cms, 2025.

[62] MATHUSLA Collaboration. The mathusla test stand. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 985:164661, 2021.

[63] ARGO-YBJ Collaboration. The ARGO-YBJ experiment in Tibet. *Nucl. Instr. and Meth. A*, 588:7–13, 2008.

[64] M. Abbrescia et al. Resistive plate chambers performances at cosmic rays fluxes. *Nucl. Instr. and Meth. A*, 359:603–609, 1995.

[65] G. L. Alberghi. Performance of the ATLAS RPC first level Muon trigger during the Run-2 data taking. *JINST*, 14:C06007, 2019.

[66] ARGO-YBJ Collaboration. Local Station: the data read-out basic unit for the ARGO-YBJ experiment. *Nucl. Instr. and Meth. A*, 518:549–553, 2004.

[67] ARGO-YBJ Collaboration. Temperature effect on RPC performance in the ARGO-YBJ experiment. *Nucl. Instr. and Meth. A*, 608:246–250, 2009.

[68] GEANT4 Collaboration, S. Agostinelli, et al. GEANT4 – a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[69] Géotechnique Appliquée Pierre & Claude Deriaz & Cie Sa - Genève. Sondage SLHC 20 lot 1. LHC-DZ-1-3999.120, June 1996.

[70] Géotechnique Appliquée Pierre & Claude Deriaz & Cie Sa - Genève. Sondage SLHC 21 lot 1. LHC-DZ-1-3999.121, June 1996.

[71] Tatsuhiko Sato. Analytical Model for Estimating Terrestrial Cosmic Ray Fluxes Nearly Anytime and Anywhere in the World: Extension of PARMA/EXPACS. *PLOS ONE*, 10(12):1–33, 2015.

[72] Tatsuhiko Sato. Analytical Model for Estimating the Zenith Angle Dependence of Terrestrial Cosmic Ray Fluxes. *PLOS ONE*, 11(8):1–22, 2016.

[73] EXPACS Homepage in English. `https://phits.jaea.go.jp/expacs/`, May 2019.

[74] 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*, 759:601, 2016.

[75] CMS Collaboration. Measurement of the Top Quark Pair Production Cross Section in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 116:052002, 2016.

[76] ATLAS Collaboration. Measurement of the inclusive and dijet cross-sections of b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Euro. Phys. J. C*, 71, 2011.

[77] CMS Collaboration. Inclusive b-jet production in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. *JHEP*, 04:084, 2012.

[78] TTree Class Reference. `https://root.cern/doc/master/classTTree.html`.

[79] AnaAlgorithm Introduction. `https://atlassoftwaredocs.web.cern.ch/ABtutorial/alg_basic_intro/`.

[80] Danilo Piparo et al. RDataFrame: Easy Parallel ROOT Analysis at 100 Threads. *EPJ Web of Conferences*, 214:06029, 01 2019.

[81] Lindsey Gray et al. CoffeaTeam/coffea: Version 0.7.1, February 2021.

[82] Serge Abiteboul, Richard Hull, and Victor Vianu. *Foundations of Databases: The Logical Level.* Addison-Wesley Longman Publishing Co., Inc., USA, 1st edition, 1995.

[83] adl_benchmarks_index. `https://github.com/iris-hep/adl-benchmarks-index/`.

[84] func_adl. `https://github.com/iris-hep/func_adl`.

[85] Proffitt, Mason and Watts, Gordon. FuncADL: Functional Analysis Description Language. *EPJ Web Conf.*, 251:03068, 2021.

[86] C Huh, M Proffitt, H B Prosper, S Sekmen, B Sen, G Unel, and G Watts. Declarative interfaces for HEP data analysis: FuncADL and ADL/CutLang. *Journal of Physics: Conference Series*, 2438(1):012075, feb 2023.

[87] Language Integrated Query (LINQ). `https://docs.microsoft.com/en-us/dotnet/csharp/programming-guide/concepts/linq/`.

[88] Built-in Functions. `https://docs.python.org/3/library/functions.html`.

[89] func_adl_benchmarks. `https://github.com/masonproffitt/func_adl_benchmarks/blob/master/swan_notebook.ipynb`.

[90] ast — Abstract Syntax Trees. `https://docs.python.org/3/library/ast.html`.

[91] ATLAS Collaboration. Athena, April 2019.

[92] Jim Pivarski et al. scikit-hep/uproot4: 4.0.4, February 2021.

[93] Jim Pivarski et al. scikit-hep/awkward-1.0: 1.1.2, February 2021.

[94] Rizzi, Andrea, Petrucciani, Giovanni, and Peruzzi, Marco. A further reduction in CMS event data for analysis: the NANOAOD format. *EPJ Web Conf.*, 214:06021, 2019.

[95] Elmsheuser, Johannes et al. Evolution of the ATLAS analysis model for Run-3 and prospects for HL-LHC. *EPJ Web Conf.*, 245:06014, 2020.

[96] ServiceX - Data Delivery for the HEP Community. `https://github.com/ssl-hep/ServiceX/`.

[97] qastle. `https://github.com/iris-hep/qastle`.

[98] Common Lisp HyperSpec. `http://www.lispworks.com/documentation/lw50/CLHS/Front/index.htm`.

[99] TCutToQastleWrapper. `https://github.com/ssl-hep/TCutToQastleWrapper`.