

UNIVERSITY OF OXFORD

LINCOLN COLLEGE

DOCTORATE OF PHILOSOPHY

PARTICLE PHYSICS

**ADVANCED MACHINE LEARNING APPLICATIONS
FOR HIGGS SEARCHES AND HEAVY FLAVOUR
QUARKS IDENTIFICATION AT ATLAS**

CANDIDATE

MAXENCE DRAGUET

SUPERVISOR

DANIELA BORTOLETTO

2020-2024



CONTENTS

1	The LHC & ATLAS Experiment	3
1.1	The Large Hadron Collider	4
1.2	The ATLAS Detector	6
1.2.1	The Inner Detector Tracker	9
1.2.2	Electronic and Hadronic Calorimeters	10
1.2.3	Muon Detection Systems	12
1.3	Operation and Reconstruction with the ATLAS Detector	13
1.3.1	Trigger System	13
1.3.2	Low-Level Signatures	14
1.3.3	Electrons	15
1.3.4	Muons	16
1.3.5	Jets	16
1.3.6	Taus	19
1.3.7	Missing Transverse Energy	19
	Bibliography	21
	Appendices	25

LIST OF ABBREVIATIONS

BSM	Beyond the Standard Model	MS	Muon Spectrometer
CERN	Centre Européen pour la Recherche Nucléaire	MVA	Multivariate Analysis
CPU	Core Processing Unit	PU	Pile-up
IBL	Insertable B-Layer	PV	Primary Vertex
ID	Inner Detector	QCD	Quantum Chromodynamics
IP	Impact Parameter	RNN	Recurrent Neural Network
JES	Jet Energy Scale	SCT	Semiconductor Tracker
JVT	Jet Vertex Tagger	SV	Secondary Vertex
LHC	Large Hadron Collider	TRT	Transition Radiation Tracker
		UE	Underlying Event

CHAPTER 1

THE LHC & ATLAS EXPERIMENT

Modern particle physics is at the edge of the technological reach of science. To discover the Higgs boson, an extremely complex infrastructure is required to probe physics at the required scale. The Large Hadron Collider (LHC) at Centre Européen pour la Recherche Nucléaire (CERN) is consequently the largest and most powerful particle accelerator and has held this title since its construction in 2008 [1]. It is easily ranks as one of the most complex machine ever created. Protons are accelerated to up to 99.999991% of the speed of light, in a giant 27 km long ring-shaped accelerator, buried 100 m below the surface of the French-Swiss border near Geneva. Superconducting magnets cooled down with liquid helium to 1.9 K steer this energetic beam thanks to powerful magnetic fields of 8.33 Tesla. The beams are made of bunches of particles, collided precisely at four specific interaction points where large detectors are built and operated by dedicated experiments: ATLAS [2], CMS [3], ALICE [4], and LHCb [5]. The first two are multipurpose experiments with overlapping physics programs, while ALICE and LHCb are respectively dedicated to heavy ion and heavy flavour physics. This chapter described the experimental setup of the LHC and the ATLAS experiment, focusing on proton-proton collisions and introducing relevant elements to the work presented in this thesis.

1.1 The Large Hadron Collider

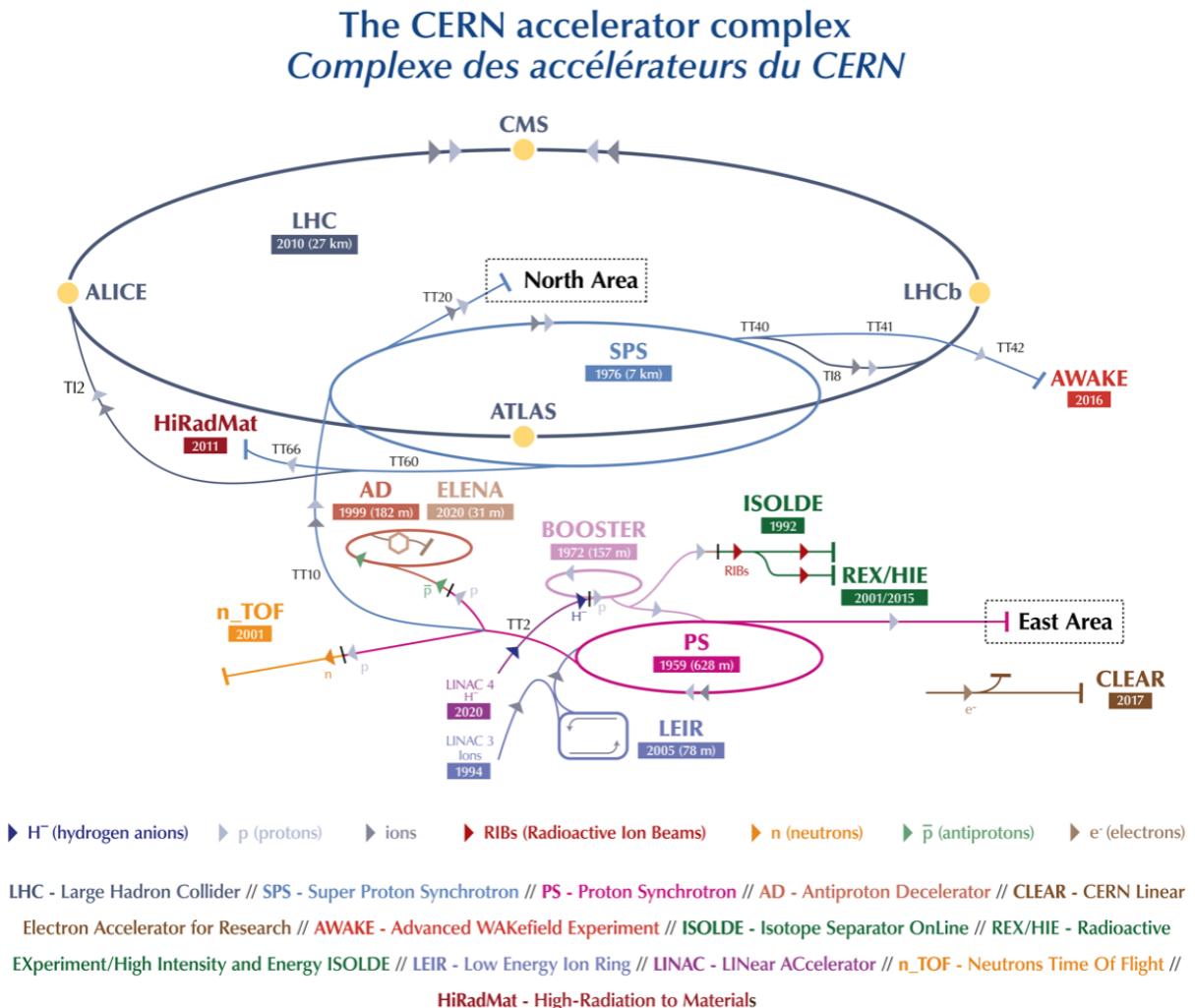


Figure 1.1: The accelerator complex of CERN during Run 3 [6]. The LHC is the top dark blue ring.

The last machine in a complex multi-stage accelerator complex of CERN displayed in Figure 1.1, the LHC is capable of frontally colliding proton or heavy ion beams. These beams are made of bunches of particles, collided precisely at four interaction points where large detectors are built and operated by dedicated experiments. The life of a proton beam starts innocuously ionised hydrogen H^- gas, passed to a linear accelerator called LINAC4¹ to reach energies of 160 MeV. After stripping the ionised hydrogen of its two electrons thus leaving a bare proton, the next acceleration stage is in the Proton Synchrotron Booster (BOOSTER) bringing the beam energy to 2 GeV. The protons are then passed to increasingly larger synchrotrons: the Proton Synchrotron (PS) to reach energies of 26 GeV, followed by the Super Proton Synchrotron (SPS) to reach energies of 450 GeV. The beam is finally injected into the LHC in two different beamlines

¹Before 2020, it was LINAC2.

circulating the proton in opposite directions. Superconducting dipole magnets reaching 8.33 T steer the highly-energetic beams, while complex geometries of magnets (quadrupoles, sextupoles, ...) are deployed to refine the bunch shape through focusing effects. Powerful radiofrequency cavities accelerate them to their final energy of 6.5 TeV, giving a total pp centre of mass energy of $\sqrt{s} = 13$ TeV in Run 2.

The operation of the LHC is split into dedicated *runs* of data taking, separated by *shutdowns* to maintain or upgrade the infrastructure. Some key metrics from these runs from the point of view of the ATLAS experiment are displayed in Table 1.1

	Year	\sqrt{s} [TeV]	$\langle\mu\rangle$	Luminosity \mathcal{L} [$\text{cm}^{-2}\text{s}^{-1}$]	$\int \mathcal{L}$ [fb^{-1}]
Run 1	2010 - 2012	7-8	18	0.8×10^{34}	26.4
Run 2	2015 - 2018	13	34	$1-2 \times 10^{34}$	140.1
Run 3	2022 - 2025	13.6	50	2×10^{33}	65

Table 1.1: Metrics on the accecelarator performance of the LHC in the different runs of data taking. The reported values correspond to those recorded by the ATLAS experiment [7, 8, 9]. Numbers for the ongoing Run 3 are preliminary, with the integrated luminosity listed considering events recorded until July 2023. The number of interactions per bunch crossing averaged over each run is displayed as $\langle\mu\rangle$.

Run 2 operated at a larger centre of mass energy (\sqrt{s}) and higher average instantaneous luminosity \mathcal{L} than Run 1, with Run 3 again pushing up the limits. The average instantaneous luminosity measures the quantity of data collected from the relation

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt} \quad (1.1)$$

relating the event rate of a particular process to its cross-section σ . The instantaneous luminosity is a machine parameter: it depends on the design and the operation of the accelerator. It is calculated from

$$\mathcal{L} = \frac{N_1 N_2 N_b f}{4\pi \sigma_x \sigma_y} \quad (1.2)$$

where N_1 and N_2 are the number of protons in each bunch, N_b the number of bunches, f is the collider revolution frequency, and σ_x and σ_y are the x and y geometrical extension of the beam density distributions. The integrated luminosity $\int \mathcal{L} dt$ is a direct measure of the quantity of data collected over a certain period, often expressed in units of inverse *barn* b^{-1} , where $\text{b} = 10^{-28} \text{ m}^2$. For Run 2, the total luminosity recorded corresponds to $140.1 \pm 1.2 \text{ fb}^{-1}$, with a small uncertainty of 0.83 % [8] thanks to a complex measurement involving luminosity-dedicated detectors such

as LUCID-2 [10]. Figure 1.2 shows the cumulative total integrated luminosity during Run 2, jointly with another important machine parameter: the average number of interactions per bunch crossing $\langle\mu\rangle$.

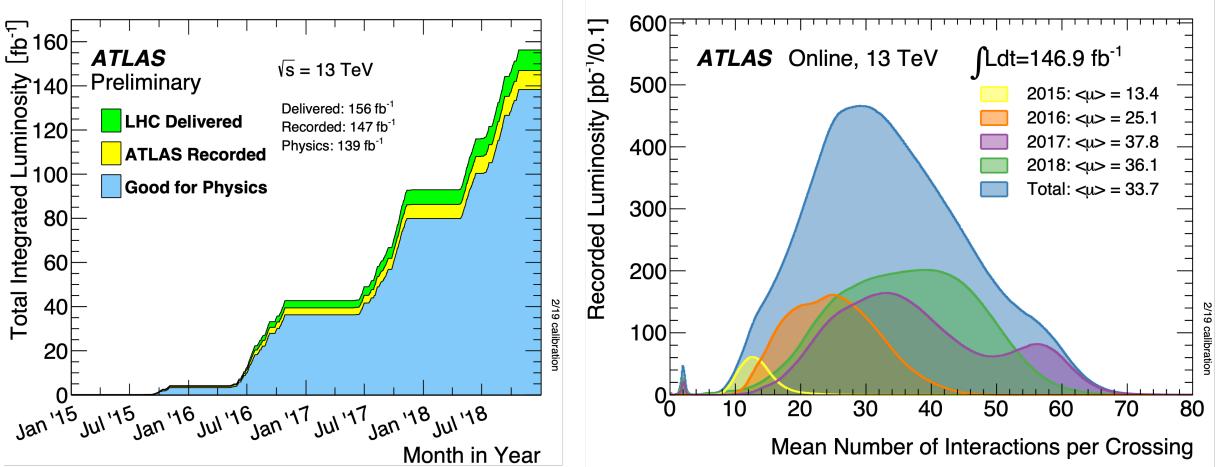


Figure 1.2: The ATLAS cumulative integrated luminosity delivered, recorded, and approved for physics (left) and the average pile-up distribution weighted by luminosity (right) during Run 2 [11]. The luminosities indicated correspond to an early calculation that was corrected in Ref. [8].

The main event during the collision of the proton bunches is the hard scatter where most of the energy transfer occurs. Other protons in the bunches do have softer interactions leading to background activity called *Pile-up (PU)*. Two types of pile-up are distinguished: *in-time PU* when the soft interaction is from protons in the same bunch as the hard scatter recorded, and *out-of-time PU* when the protons are from another bunch crossing. Bunches in the LHC are separated by a 25 ns delay, corresponding to a machine frequency of 40 MHz. To control the luminosity, the angle of attack of the beams can be tweaked so that their overlap on the impact point is modified. Having more frontal collisions leads to a larger overlap and a higher luminosity at the price of more PU.

1.2 The ATLAS Detector

The ATLAS Collaboration built and operates the eponymous cylindrically shaped multi-layered detector lying 100 m underground with a length of 45 long and a 26 m diameter [2], as presented in figure ???. The experiment is constructed to probe a broad range of physical phenomena, as required from the general purpose of the Collaboration. Aiming to be as hermetic as possible, the detector wraps around the interaction point with the barrel forming the central part of the cylinder and the end-caps closing the geometry at its extremity. Essential characteristics of the design had to be met to manage the extremely high event rate, requiring fast response,

radiation-hard sensors, and state-of-the-art readout electronics in combination with good spatial and temporal resolution to disentangle the effect of pile-up.

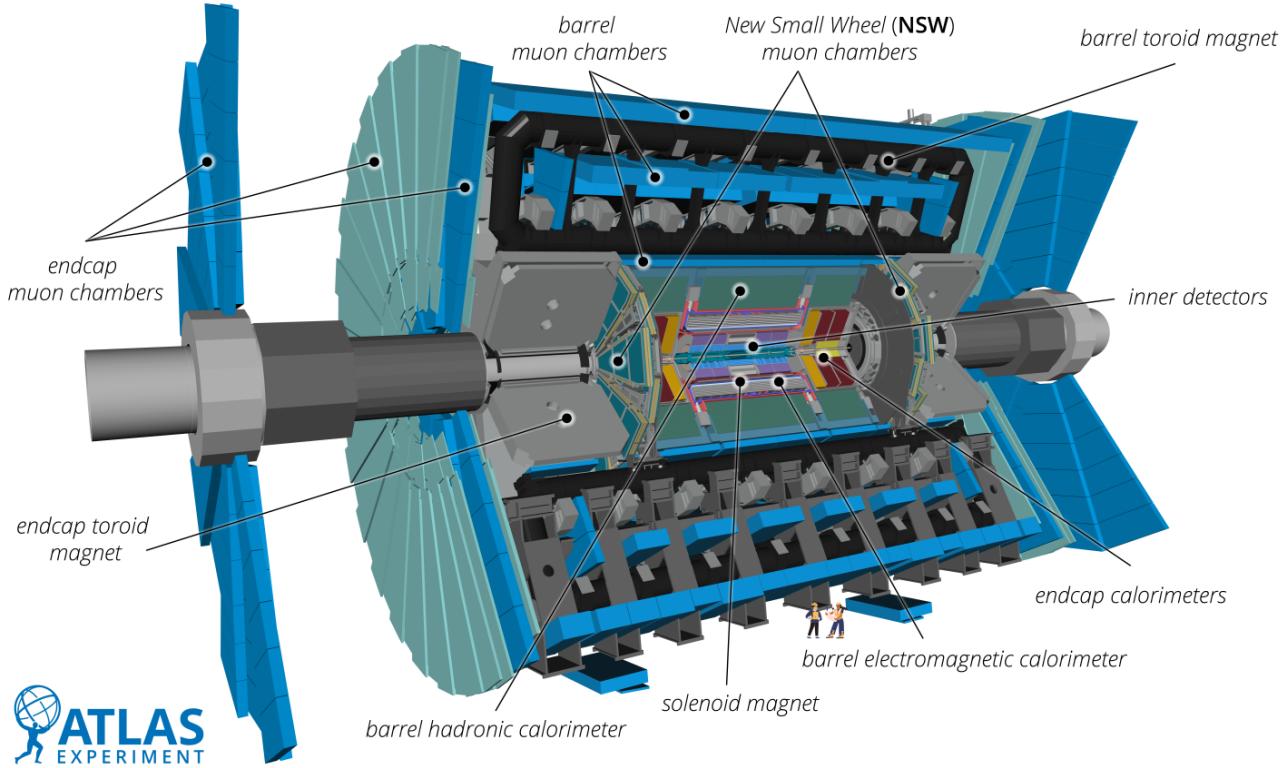


Figure 1.3: Cut-away view of the ATLAS detector [12].

The coordinate system adopted in ATLAS is described in Figure 1.4: the x -axis points to the centre of the LHC ring, the y -axis points upwards, and the z -axis in the direction of the beamline (anti-clockwise from the top). The azimuthal angle ϕ is defined in the transverse plane $x - y$ and the polar angle θ is measured upwards from the beam-axis (along y). Therefore, the transverse momentum p_T of a particle is obtained from its momentum vector $\mathbf{p} = (p_x, p_y, p_z)$ as: $p_T = \sqrt{p_x^2 + p_y^2}$. This is of particular interest as the momentum's longitudinal component p_z is not measurable due to the openings for the beamline and the partonic interaction carrying only a fraction of the original proton momentum. Only the transverse momentum can be correctly assessed, as the initial parton is mostly boosted in the direction of the beamline. The rapidity y of a particle, playing a crucial role in the special relativity, is expressed as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (1.3)$$

with E and p_L the particle's energy and longitudinal momentum. In the ultrarelativistic limit, when $p \gg m$, the rest mass is negligible and $E \approx p$. In such cases, the rapidity y is well

approximated by the experimentally reconstructable pseudo-rapidity η that is expressed as

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (1.4)$$

Like the rapidity, $\Delta\eta$ is an invariant under Lorentz boosts along the longitudinal z -axis. It is often combined with the azimuthal angular aperture $\Delta\phi$ to define the angular separation ΔR between two objects as

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = \sqrt{\Delta(\phi_2 - \phi_1)^2 + \Delta(\eta_2 - \eta_1)^2}. \quad (1.5)$$

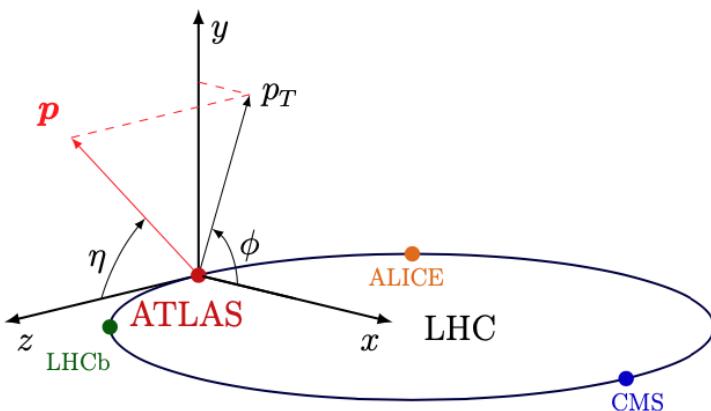


Figure 1.4: The ATLAS coordinate system, from [13].

As depicted in Figure 1.3, ATLAS combines different systems into one precision machine. These subdetectors are able to measure information in the range $|\eta| < 2.5$, with some specialised subdetectors extending further. An essential component in the detector system are its two types of superconducting magnets. Four central solenoid magnets wrap the point of impact and generate a powerful 2 T magnetic field within the inner detectors along the z -axis, while toroidal magnets are placed externally on the barrel and the endcaps muon systems to generate a 3.5 T magnetic field deflecting these leptons in the η -direction. A q -charged particles of momentum p is deflected by a magnetic field B due to the Lorentz force, leading to a connexion between the radius of curvature R of the trajectory and the momentum p such as:

$$p_{\perp} = 0.3 qBR [\text{GeV}/c] \quad (1.6)$$

where p_{\perp} is the magnitude of the momentum perpendicular to the magnetic field \mathbf{B} , and q is expressed in unit of proton charge. Therefore, from the measurement of the curvature, experimentalists are able to infer the component of the momentum transversial to their generated B .

Higher magnetic fields lead to larger curvature simplifying the measurement of R and improving the resolution of p_{\perp} .

The rest of this chapter reviews the different subdetectors of ATLAS and introduces some common reconstruction methods that are relevant to the work presented in this thesis.

1.2.1 The Inner Detector Tracker

The detector placed closest to the point of interaction is the Inner Detector (ID) [14]. This is a tracker covering the range $|\eta| < 2.5$ in a radius of 3 cm to 1 m, designed to record hits in silicon semiconductors or straw-tubes when charged particles fly through so that their trajectory or *track* can be reconstructed from the collected hits. The powerful 2 T magnetic field of the central solenoids enable this detector to measure both charge and momentum. The ID combines three subsystems, as represented in Figure 1.5.

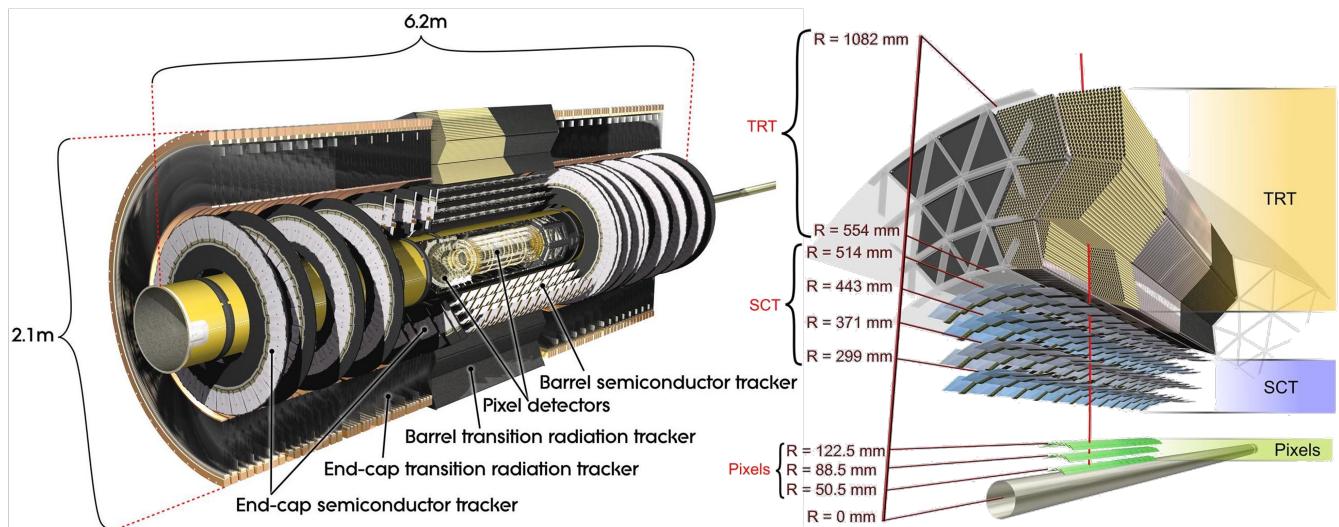


Figure 1.5: The Inner Detector of ATLAS [12].

First, the high-granularity *Pixel Detector* covers the innermost region with three barrel and three endcap layers, for a total of 80 million sensitive semiconductor-based pixels [14, 15]. During Run 2, an additional *Insertable B-Layer (IBL)* was added at 33 mm from the centre, with 12 million pixels [16]. This detector gives robust and precise tracking performance and plays a major role in flavour tagging, as described in Chapter ???. Pixels are $50 \mu\text{m} \times 400 \mu\text{m}$ in the $R\phi \times z$, with a smaller $50 \mu\text{m} \times 250 \mu\text{m}$ for the IBL. The current resolution delivered is of $10 \mu\text{m}$ ($67 \mu\text{m}$) in the transverse $R\phi$ plane (z -direction) [17, 18].

The *Semiconductor Tracker (SCT)* is the next detector, constructed by arranging pairs of

silicon microstrips layers into modules assembled into 4 concentric barrel layers and 9 disks in each endcap [19, 20]. The resolution is typically of $17\ \mu\text{m}$ in $R\phi$ and $580\ \mu\text{m}$ in z [21].

The final system is the *Transition Radiation Tracker (TRT)*, a gas-based straw-tube tracker aiding track reconstruction by delivering numerous hits [22]. Approximately 300 000 drift tubes of a 4 mm diameter filled with a mélange of argon and xenon are arranged along the beamline in the barrel and radially in the endcaps. Each tube has a conducting wire at its centre and its surface is electrically charged, so that a charged particle flying through ionises the gas, leading to a measurable discharge. Polyethylene is placed between the tubes to encourage the emission of transition radiation from relativistic particles proportionally to their Lorentz boost $\gamma E/m$. Consequently, the TRT is used both for tracking and electron and pion identification, by reconstructing the mass of the charged particles from the amount of γ -radiation. For tracking, the position resolution provided is of $130\ \mu\text{m}$ in the $R\phi$ plane for the barrel and the $z - \phi$ plane in the endcaps [23]. Altogether, the inverse momentum resolution of the ATLAS ID is

$$\sigma(1/p_T) = 0.36 \oplus \frac{13}{p_T \sin \theta} \text{TeV}^{-1} \quad (1.7)$$

where \oplus is the sum in quadrature [2]. This corresponds to a relative error of 0.01% for a track with $p_T \sim 500\ \text{MeV}$, and 4% at a $p_T \sim 100\ \text{GeV}$.

1.2.2 Electronic and Hadronic Calorimeters

Covering the $|\eta| < 4.9$ region, calorimeters collect the energy of all interacting particles, neutral and charged, except for the muons. They are separated into an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL) covering jointly the $|\eta| < 3.2$ region and a forward calorimeter for the $3.2 < |\eta| < 4.9$ region, each interlaying specific layers of active and passive materials, as displayed in Figure 1.6 [`\cite {TheATLASCollaboration_2008}`]. The passive material typically has a large atomic number to induce cascade of particles called *shower*. The active material, typically liquid argon (LAr) collects the energy from these showers through ionisation or scintillation light.

The ECAL is designed to collect the energy of electrons and photons, as well as contributing to the measurement of the energy of jets. The active material is LAr and absorbing plates of lead are used as passive material, to encourage bremsstrehlung $e \rightarrow e\gamma$ and pair production from photons $\gamma \rightarrow e^+e^-$. The ECAL have a depth of $22\ X_0$, where the units of *radiation length*

X_0 track the distance for an electron to have only $1/e$ its original energy. The energy resolution is parametrised into three terms for the sampling, noise, and a constant contribution, summed in quadrature as [24]

$$\frac{\sigma_E^{\text{ECAL}}}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{0.12[\text{GeV}]}{E} \oplus 0.3\%, \quad (1.8)$$

giving an energy resolution between $\sim 0.5\%$ for 10 GeV electrons, and $\sim 0.6\%$ for 60 GeV photons.

The HCAL is designed to capture the energy of hadronic shower, with LAr as active material for the endcap and forward calorimeters and scintillating plastic tiles for the barrel. As passive material, the endcaps use copper plates, the forward calorimeters use copper and tungsten, and the tile calorimeter in the barrel uses steel. The depth of the hadronic calorimeter is $\sim 10X$, where X is the nuclear interaction length tracking the average distance before a hadron interacts with a nucleus. The calorimeters collect the majority of the energy, with an HCAL resolution expressed as [24]

$$\frac{\sigma_E^{\text{HCAL}}}{E} = \frac{52\%}{\sqrt{E}} \oplus \frac{1.6[\text{GeV}]}{E} \oplus 3\%. \quad (1.9)$$

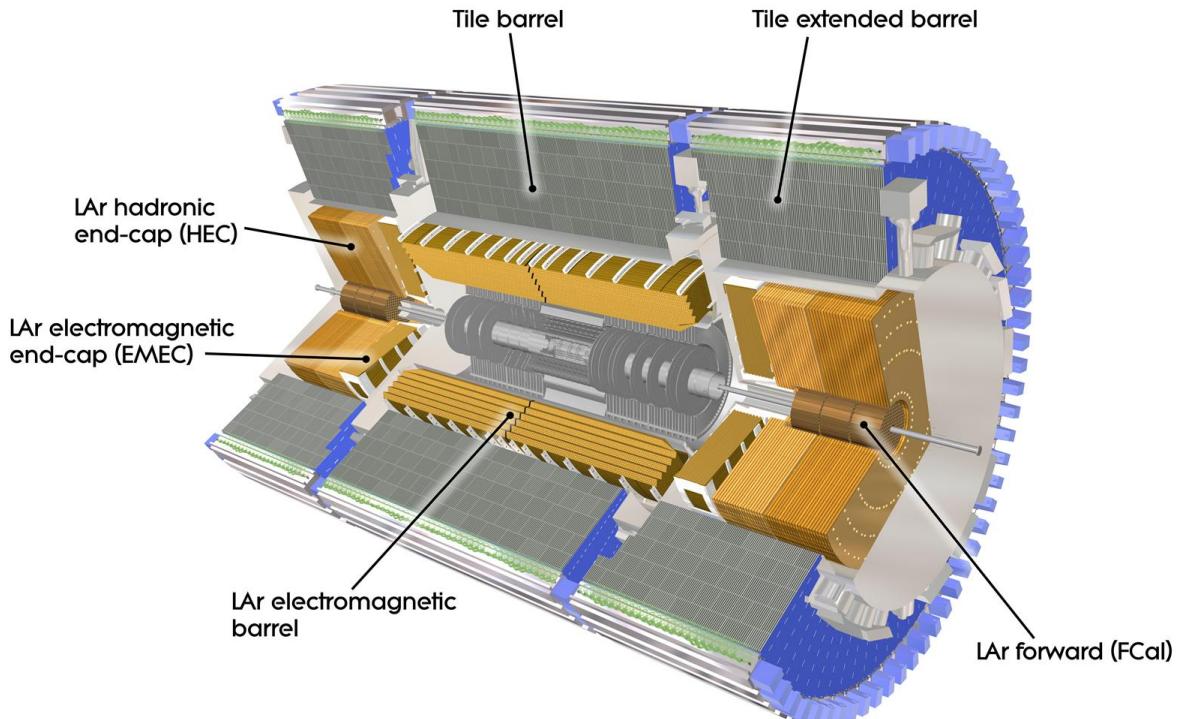


Figure 1.6: The calorimeter systems of ATLAS [12].

1.2.3 Muon Detection Systems

Muons require dedicated detection systems to be efficiently and precisely reconstructed in ATLAS. Their high-mass coupled with long lifetime at the energy studied mean they barely interact with the previously described system, only leaving a few hits in the inner detector and flying largely undisturbed through the calorimeters. For this reason, the outmost subdetectors of ATLAS are specially designed to interact with muons. The Muon Spectrometer (MS), shown in Figure 1.7, is a dedicated muon tracking system which also provides an efficient triggering hardware, as described later in this chapter.

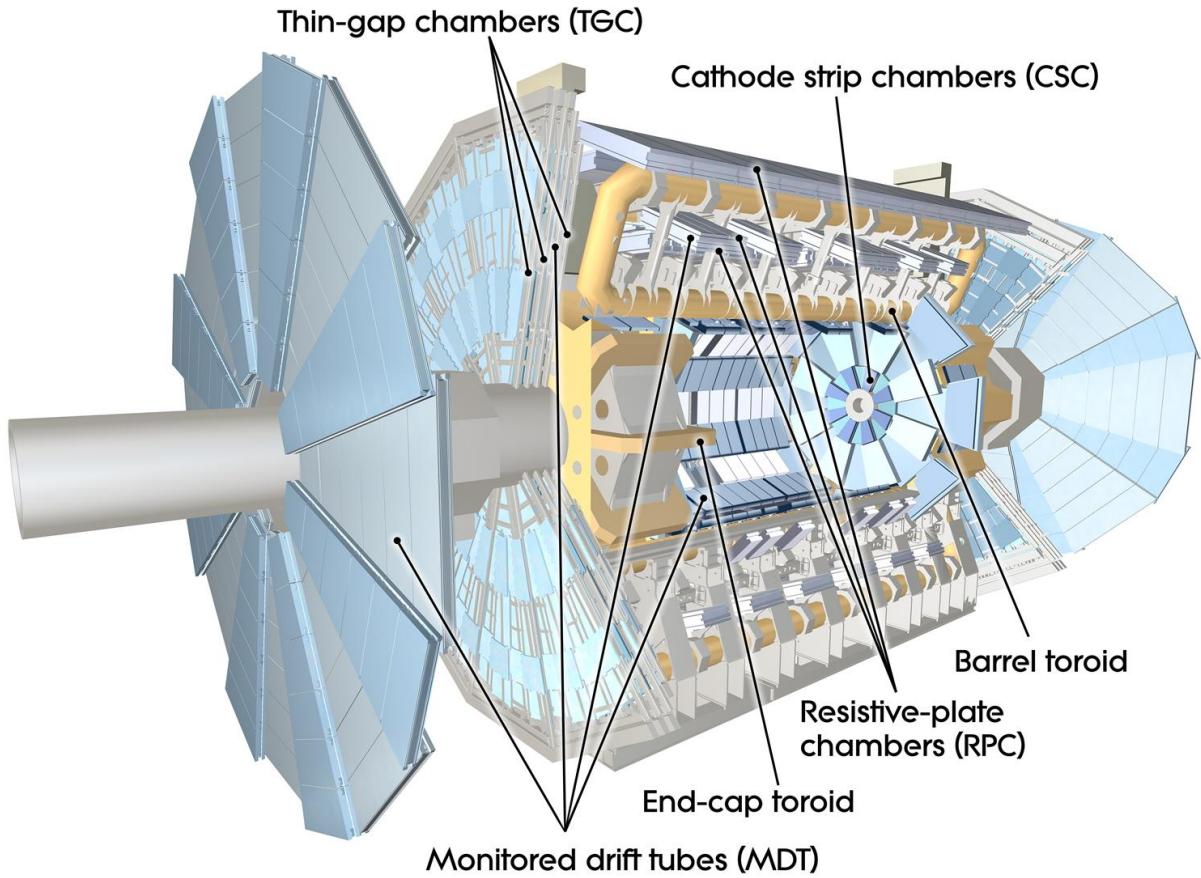


Figure 1.7: The muon detectors of ATLAS [12].

The muon tracker is composed of drift tubes split between a barrel region for $|\eta| < 1.2$ and the endcaps $1.2 < |\eta| < 2.7$, with cathore strip chambers in the inner-layers of the endcaps to resist the high activity. The trigger system relies on resitive-plate chambers in the barrel and thin gap chamber in the end-caps. For muon momentum and charge measurements from the tracks, powerful superconducting toroidal magnets are used to deflect muons in the MS. The resolution on p_T is measured to be of $\sim 1.7\%$ (2.3%) for muons from J/ψ decays in the central (forward) region [25].

1.3 Operation and Reconstruction with the ATLAS Detector

For physics-quality data taking, the different subdetectors of ATLAS must be performing according to specifications. In operation, the event rate produced by the LHC in the heart of the ATLAS detector is of 40 MHz, from the 25 ns bunch-crossing. This unfortunately leads to a data generation rate that is too high for the electronics and computing resource available, requiring the Collaboration to design specific approaches to reduce the rate to a manageable level [26]. This is the task of the trigger system, which is described first in the section. Events that pass the trigger thresholds are stored and must be further processed to reconstruct the physics processes from the low-level measurements performed by the different subdetectors: this is the task of reconstruction, the last subject described in this chapter only for objects types relevant to the presented work. This latter step is performed thanks to the extensive ATLAS software [27, 28], exploiting the specific signatures of the different detected particles as schematised in Figure 1.8.

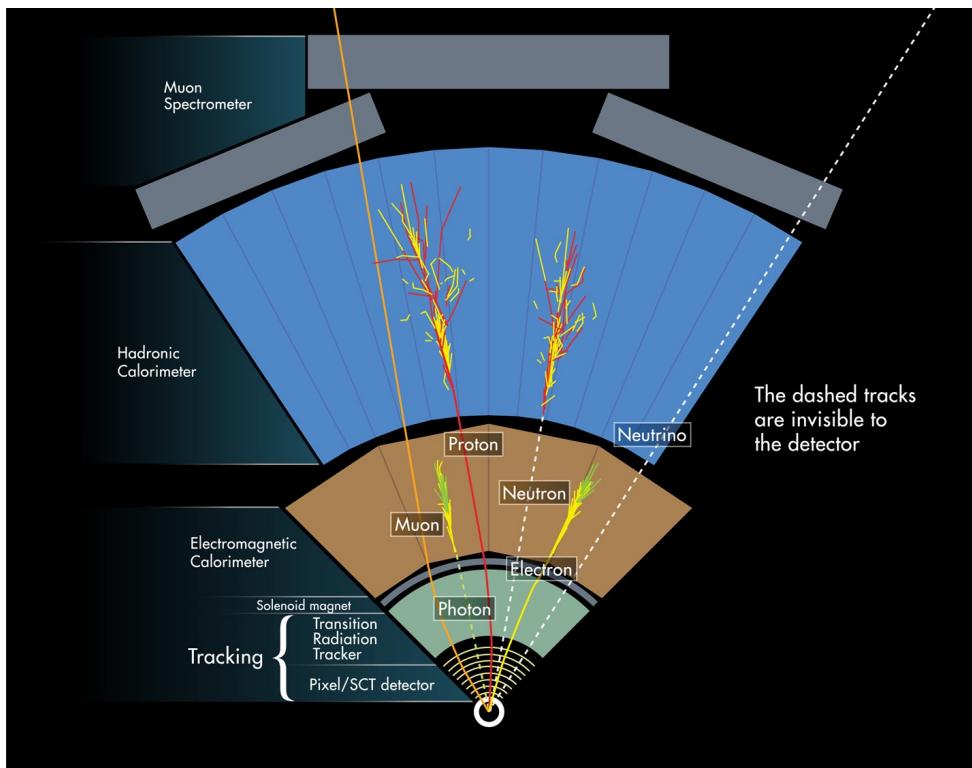


Figure 1.8: Schematics of different particles signatures in the ATLAS detector [29].

1.3.1 Trigger System

The ATLAS trigger system relies on a hierarchical approach to progressively reduce the data rate and select events deemed interesting for physics. Firstly, the *low-level* Level-1 (L1) trigger is built on fast electronic hardware accessing only coarse information to reduce the rate to 100 kHz

in $\sim 2.5 \mu\text{s}$. This is followed by the High Level Trigger (HLT) that runs on a farm of 40,000 Core Processing Units (CPUs) to implement a finer software-based selection, bringing the rate down to a 1.2 kHz or 1.2 GB/s for data storage [30]. In this process, gradually more complex information is accessed by dedicated readout and measurement systems. Some commonly used triggers are based on signatures of electrons, muons, missing transverse energy, and b -jets. Different trigger menus are designed by the Collaboration, with dedicated data taking period for each setup. Analyses are then able to select data from the whole collected data based on the specific signatures they search.

1.3.2 Low-Level Signatures

Low-level signatures are used in higher-level reconstruction processes to identify physics objects, such as electrons and jets. Three types are described here: the trajectory of charged particles called *tracks*, the construction of vertices, and the formation of calorimeter clusters.

Tracks are the reconstructed trajectory of charged particles through the detector from collected *hits*: localised energy deposits. With denser PU activity, the number of hits in a single event becomes significant turning track reconstruction into a challenging computational problem [31, 32]. The trajectories are curved thanks to the previously described superconducting magnets. From a set of hits, tracks are fitted inside-out [32]: clusters of three hits in the Pixel or SCT detectors are first identified as *seeds*, with additional hits associated by a combinatorial Kalman Filter [33] based on compatibility criteria on the initial track. Hits can initially be shared by several tracks, with the ambiguity resolved later when the reconstructed tracks are ranked by quality and χ^2 -fits are performed to quantify the best possible association, favouring high p_T tracks. The process is then extended to the TRT from the outside-in, and followed by additional quality criteria such as requiring tracks to have a $p_T > 500 \text{ MeV}$ in $|\eta| < 2.5$, a minimum of 7 hits in the Pixex and SCT, at most one hole, and at most two shared hits. Tracks are parametrised by the longitudinal (along z) and transversal (in the $x - y$ plane) Impact Parameter (IP), respectively z_0 and d_0 : these measure the distance from the primary vertex to the perigee of the track².

Tracks emanate from the decay of particles or radiated through physics processes. If a reconstructed charged particle is produced in the main event, its trajectory leads back to this special location called the *pv* [34]. If the particle is produced in another decay, the point of emanation

²The point of closest approach of the track with respect to the Primary Vertex (PV).

can sometimes be distinguished and is labelled *Secondary Vertex (SV)*. Identifying vertices is crucial for the physics programme of the Collaboration. The primary vertex is identified from a seed vertex first identified from the set of all well reconstructed tracks. The vertex position is then iteratively refined by removing tracks incompatible with the reconstructed vertex and refitting, until some quality criteria are met. Discarded tracks are then used to identify secondary or tertiary vertices. The primary vertex is the one with the largest sum of squared track- p_T .

The calorimeters of atlas are composed of many granular layers. So-called clusters are identified by grouping clusters with certain energy deposit measured, either from the *sliding-window* or *topo-clusters* algorithm [35]. The former one generates fixed-size rectangles by translating the window to maxime the transverse energy E_T measured. The latter clusters neighbouring cells based on a signal to noise criterion. The advantage of the sliding-window is that it is easier to calibrate and is hence used in electron, photon, and hadronic τ 's reconstruction. Topo-clusters are robust against noise and therefore used for jet and missing transverse energy reconstruction.

1.3.3 Electrons

Electrons leave a rich signature in the ATLAS detector, in the ID and ECAL. For the central region $|\eta| < 2.5$, electrons are identified and reconstructed with both subdetectors, while the forward region $2.5 < |\eta| < 4.9$ is only covered by the calorimeters with the shape of the shower used to discriminate electrons. Here, only centrally produced *prompt*-electrons are described, where *non-prompt*-electrons are not produced from the main physics process but through later decays or interactions with the detector itself. Photons, pions, and sometimes jets can be mistaken as electrons, with identification and isolation criteria derived to provide high purity and efficiency samples to analyses. Reconstruction relies on calorimeter clusters and track information. When relevant, tracks are matched to the clusters based on the expected energy loss. The track is also extrapolated to insure compatibility with the cluster barycentre. The process is re-ran with more stringent conditions after refitting the tracks matched to clusters, so that energy loss is taken into account. A prompt-electron is required to have a track matched to the PV. The absence of a precision hits or a matched track leads to consider the calorimeter clusters as a photon deposit.

To further discriminate prompt-electrons from non-prompt electrons and photons, which act as a background due to photon conversion to e^+e^- , a likelihood-based identification algorithm built on a Multivariate Analysis (MVA) is deployed [36]. Parameters used include the number of

hits in each tracker layer, the IPs, a calorimeter cluster parameters. Several working points that are progressively more selected are defined on the discriminant, from *Very Loose*, *Loose*, *Medium*, to *Tight*. Prompt-electron candidates are required to be isolated from other tracks and other energy deposits, with specific criteria that are either ID- or calorimeters-based. In the former case, the sum of tracks p_T in a ΔR cone around the electron is used, while the latter analyses the sum of energy deposit in a calorimeter cone around the electron cluster. As further described in Chapter ??, the efficiency of the electron reconstruction, including identification and isolation, are computed by comparing the measured and simulated measurements of the $Z \rightarrow e^+e^-$ and $J\psi \rightarrow e^+e^-$. Photons are reconstructed similarly to electrons, with no tracks being matched for unconverted photons. Converted photons are allowed hits in the ID.

1.3.4 Muons

The MS is the main detector to reconstruct muons, with other subdetectors playing a smaller role: muons can indeed leave a track in the ID and some energy deposits in the calorimeters. These fragmented signatures in the subdetectors are combined to identify muon candidates. In the MS, tracks are constructed from a fit of the successive hits in the different chambers. *Combined muons* are defined by matching a track in the MS to a track in the ID, with additional information from the calorimeters. Prompt-muons are rejected from background-produced muons (such as in the decay of a b -hadron) by specific criteria targeting discrepancies in the p_T between the MS and ID. Increasingly selective working points are defined to identify muons as *Loose*, *Medium*, *Tight*, *High- p_T* . Isolation requirements are applied similarly to the electron case, either track- or calorimeter-based in a ΔR cone around the candidate muons. The calibration of muons is performed similarly to the electrons, on a $Z \rightarrow \mu^+\mu^-$ and $J\psi \rightarrow \mu^+\mu^-$ samples.

1.3.5 Jets

Quarks and gluons are the most commonly produced particles in a hadron collider setting. As described in Chapter ??, these particles carry colour-charges and undergo hadronisation when produced to neutralise their free colour. This complex process leads to a unique signature in the detector: a spray of particles emitted within the original parton direction called a *jet*. Electrically charged and neutral particles are contained within jets, with most of their energy deposited in the hadron calorimeters. These aggregated objects are therefore constructed by applying a clustering algorithms on tracks and/or calorimeter clusters, depending on the jet definition. The most notorious clustering method is the anti- k_T algorithm, thanks to the robustness of the defined jet

to collinear splitting or additional soft emissions [37]. The algorithm starts by considering high momentum object, to which softer objects are added thus growing the clusters. Two measures of distances d are used as parameters when considering whether to cluster object i and j :

$$d_{ij} = \min \left(\frac{1}{k_{Ti}^2}, \frac{1}{k_{Tj}^2} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad (1.10)$$

and

$$d_{iB} = \frac{1}{k_{Ti}^2}. \quad (1.11)$$

The first distance, d_{ij} , combines the angular aperture ΔR_{ij} between i and j ³ with the transverse momentum k_T of the two objects and a fixed *radius* parameter R . This distance defines a jet radius limiting the size of the associated cone. It is compared to the second distance, d_{iB} , assessing the size of the already formed cluster i . If $d_{ij} < d_{iB}$, j is clustered with i , otherwise i is identified as a jet and removed from consideration. The algorithm proceeds after updating the distances, until all constituents are assigned a cluster. Typical radius values for ATLAS are $R = 0.4$ and $R = 1.0$, defining respectively small- R and large- R jets. The former is commonly used for quark and gluon jets, while the latter is employed by analysis to identify boosted heavy objects decay, such as W or Higgs bosons.

Jets can be constructed from tracks, calorimeter clusters, or both. In ATLAS, several types of jets are deployed, depending on the desired properties and the objects they are meant to represent. The following types are all reconstructed with the anti- k_T algorithm, clustering different objects:

- PFlow jets are reconstructed from particle-flow objects [38] with a radius $R = 0.4$. These objects combine tracking information from the ID to the calorimeter clusters, leading to a better energy resolution at low p_T and lower PU contamination after calibration [39].
- EMTopo jets are constructed from denoised topological calorimeters clusters called *topoclusters*, based on the per cell energy significance $S_{\text{cell}} = E_{\text{cell}}/\sigma_{\text{cell}}$, where E_{cell} is the energy and σ_{cell} the expected noise level in the cell [40]. The topoclusters are then used with the anti- k_T method with a small 0.4 or large 1.0 radius.
- Large- R jets are built from topological calorimeter clusters with a radius $R = 1.0$. These jets are applied trimming to remove the contribution from soft contamination, which is mainly due

³ $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$.

to PU and Underlying Event (UE) activity, leading to an improved mass resolution [41]

- Track-jets are constructed are based on the anti- k_T clustering algorithm with a variable radius R depending on jet p_T , such that the wide cone used at low p_T ($R \approx 0.4$) becomes narrower at high p_T ($R \approx 0.02$). They are typically matched as sub-jets of a large- R jet, to deploy the flavour tagging techniques described in Chapter ??.

PFlow and variable-radius jets are used to train the algorithms of Chapter ??, while EMTopo, large- R , and track jets are used in the analysis of Chapter ???. Jets can be assigned a flavour based on the presence of an original parton within a $\Delta R = 0.3$ cone around the jet axis. Experimentally, the flavour is often determined based on the hadrons associated to the jet, as described in detail in Chapter ??.

Jets benefit from extensive calibration to correct the reconstructed properties such as the mass, the energy, and the jet axis direction. In particular, corrections to account for PU activity and out-of-cone emissions and deposits are considered. Detector effects are also taken into account, such as differences between the electromagnetic and hadronic calorimeters and leakage out of the active regions. The so-called *Jet Energy Scale (JES)* calibrations corrects the jets in successive steps [42]:

- *Origin*: the jet axis, initially constructed from the centre of ATLAS, is corrected to point from the PV, and the p_T is updated.
- *Pile-up*: both in-time and out-of-time pile up leave additional energy deposit in the calorimeters. This is subtracted from the jet, from first an overall estimation based on the average pile-up and then from the actual number of interactions and vertices in the event.
- *Absolute*: absolute energy corrections dependent of the energy E and η are derived to match the data energy scale to the particle-level energy scale of dedicated simulated samples.
- *Eta inter-calibration*: the detector is not homogenous and the forward regions measurements are typically less accurate. Corrections to these forward jets are applied based on central jets ($|\eta| < 1.4$).
- *Global sequential calibration*: energy leakage in the calorimeters is accounted for with a set of four-momentum Corrections based on five different observables tracking the shape and the energy of the jet.

- *In-situ calibration*: corrects any possible differences due to an incorrect description of the detector in the simulations by performing a fit to data a dedicated measurement of a well reconstructed object. Events from the following processes are used in increasing p_T scales: $Z + \text{jet}$ events with the Z leptonically decaying, $\gamma + \text{jet}$, and Quantum Chromodynamics (QCD) multi-jet events at high momentum.

The JES is parametrised by p_T , and dedicated uncertainties are derived for analyses to use. The *Jet Energy Resolution (JER)* is then defined as σ_{pT}/p_T , and also calibrated with dedicated uncertainties derived from a fit to di-jet events [42]. Despite the JES correction procedure, PU jets can be still significant and a dedicated Jet Vertex Tagger (JVT) is used to reject this background [43]. This implements a 2D likelihood method built from tracking variables from which different selection criteria are derived as working points with specific PU jet rejections.

1.3.6 Taus

Taus make the heaviest generation of the charged lepton, with a mass of 1.8 GeV slightly higher than that of c -quarks [44]. Their lifetime is so short they mostly decay within the beampipe. They leptonically decay to neutrinos and a e or a μ 35% of the time, while the rest occurs hadronically. The leptonic decays are hard to disentangle from prompt electrons and muons. Hadronic decays however leave a signature reconstructed as a small- R jet in ATLAS. These hadronic taus are identified by a Recurrent Neural Networks (RNNs), to disentangle them from PU and QCD jets [45]. This discriminant is used with specific working points derived to define different efficiencies.

1.3.7 Missing Transverse Energy

Some physics objects do not interact with the ATLAS detector, such as neutrinos and hypothetical Beyond the Standard Model (BSM) particles. Consequently, their presence is not detectable but can be inferred thanks to the negligible initial p_T of the two interacting partons in the proton-proton collision. Requiring the transverse energy and momentum to be balance to 0, the so-called *Missing Transverse Energy* E_T^{miss} or *MET* is calculated as the negative vectorial sum of all physics object transverse momentum

$$\mathbf{E}_T^{\text{miss}} = - \sum_{\text{hard}} \mathbf{p}_T - \sum_{\text{soft}} \mathbf{p}_T, \quad (1.12)$$

where the sum is decomposed into a *hard* term englobing all physics objects (photons, electrons, muons, taus, and jets) and a soft term including good-quality ID tracks associated to the PV but not to a physics object [46]. The performance of the reconstruction is measured by comparing simulations to data, with scale and resolution derived with uncertainties to be used by physics analyses.

BIBLIOGRAPHY

- [1] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008), S08001. DOI: [10.1088/1748-0221/3/08/S08001](https://doi.org/10.1088/1748-0221/3/08/S08001). URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08001>.
- [2] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08003. DOI: [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003). URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08003>.
- [3] CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08004. DOI: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004). URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08004>.
- [4] ALICE Collaboration. “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08002. DOI: [10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002). URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08002>.
- [5] LHCb Collaboration. “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08005. DOI: [10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005). URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08005>.
- [6] *Accelerator Complex of CERN*. <https://home.web.cern.ch/science/accelerators>. Accessed: 2024-03-21.
- [7] ATLAS Collaboration. “Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC”. In: *The European Physical Journal C* 73.8 (2013), p. 2518. DOI: [10.1140/epjc/s10052-013-2518-3](https://doi.org/10.1140/epjc/s10052-013-2518-3). URL: <https://doi.org/10.1140/epjc/s10052-013-2518-3>.
- [8] ATLAS Collaboration. “Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC”. In: *Eur. Phys. J. C* 83.10 (2023), p. 982. DOI: [10.1140/epjc/s10052-023-11747-w](https://doi.org/10.1140/epjc/s10052-023-11747-w). arXiv: [2212.09379 \[hep-ex\]](https://arxiv.org/abs/2212.09379).
- [9] ATLAS Collaboration. *Preliminary analysis of the luminosity calibration of the ATLAS 13.6 TeV data recorded in 2022*. Tech. rep. Geneva: CERN, 2023. URL: <https://cds.cern.ch/record/2853525>.
- [10] G. Avoni et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *Journal of Instrumentation* 13.07 (2018), P07017. DOI: [10.1088/1748-0221/13/07/P07017](https://doi.org/10.1088/1748-0221/13/07/P07017). URL: <https://dx.doi.org/10.1088/1748-0221/13/07/P07017>.
- [11] *Public ATLAS Luminosity Results for Run-2 of the LHC*. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>. Accessed: 2024-03-22.
- [12] *ATLAS Schematics*. <https://atlas.cern/Resources/Schematics>. Accessed: 2024-03-22.

- [13] Giles Chatham Strong. “On the impact of selected modern deep-learning techniques to the performance and celerity of classification models in an experimental high-energy physics use case”. In: *Mach. Learn. Sci. Tech.* 1 (2020), p. 045006. DOI: [10.1088/2632-2153/ab983a](https://doi.org/10.1088/2632-2153/ab983a). arXiv: [2002.01427 \[physics.data-an\]](https://arxiv.org/abs/2002.01427).
- [14] ATLAS Collaboration. *ATLAS inner detector: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331063>.
- [15] Karolos Potamianos. “The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider”. In: *PoS EPS-HEP2015* (2015), p. 261. arXiv: [1608.07850 \[physics.ins-det\]](https://arxiv.org/abs/1608.07850).
- [16] M Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. 2010. URL: <https://cds.cern.ch/record/1291633>.
- [17] H. Pernegger. “The Pixel Detector of the ATLAS experiment for LHC Run-2”. In: *Journal of Instrumentation* 10.06 (2015), p. C06012. DOI: [10.1088/1748-0221/10/06/C06012](https://doi.org/10.1088/1748-0221/10/06/C06012). URL: <https://dx.doi.org/10.1088/1748-0221/10/06/C06012>.
- [18] ATLAS Collaboration. *IBL Efficiency and Single Point Resolution in Collision Events*. Tech. rep. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2203893>.
- [19] ATLAS Collaboration. “The silicon microstrip sensors of the ATLAS semiconductor tracker”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 578.1 (2007), pp. 98–118. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2007.04.157>. URL: <https://www.sciencedirect.com/science/article/pii/S0168900207007644>.
- [20] ATLAS Collaboration. *Technical Design Report for the ATLAS Inner Tracker Strip Detector*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2257755>.
- [21] ATLAS Collaboration. “Operation and performance of the ATLAS semiconductor tracker”. In: *Journal of Instrumentation* 9.08 (2014), P08009. DOI: [10.1088/1748-0221/9/08/P08009](https://doi.org/10.1088/1748-0221/9/08/P08009). URL: <https://dx.doi.org/10.1088/1748-0221/9/08/P08009>.
- [22] ATLAS TRT collaboration. “The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance”. In: *Journal of Instrumentation* 3.02 (2008), P02013. DOI: [10.1088/1748-0221/3/02/P02013](https://doi.org/10.1088/1748-0221/3/02/P02013). URL: <https://dx.doi.org/10.1088/1748-0221/3/02/P02013>.
- [23] A Vogel. *ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates*. Tech. rep. Geneva: CERN, 2013. URL: <https://cds.cern.ch/record/1537991>.
- [24] Francesca Cavallari. “Performance of calorimeters at the LHC”. In: *Journal of Physics: Conference Series* 293.1 (Apr. 2011), p. 012001. DOI: [10.1088/1742-6596/293/1/012001](https://doi.org/10.1088/1742-6596/293/1/012001). URL: <https://dx.doi.org/10.1088/1742-6596/293/1/012001>.
- [25] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}=13$ TeV”. In: *The European Physical Journal C* 76.5 (2016), p. 292. DOI: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y). URL: <https://doi.org/10.1140/epjc/s10052-016-4120-y>.
- [26] M. zur Nedden. “The LHC Run 2 ATLAS trigger system: design, performance and plans”. In: *Journal of Instrumentation* 12.03 (2017), p. C03024. DOI: [10.1088/1748-0221/12/03/C03024](https://doi.org/10.1088/1748-0221/12/03/C03024). URL: <https://dx.doi.org/10.1088/1748-0221/12/03/C03024>.
- [27] ATLAS Collaboration. *The ATLAS Collaboration Software and Firmware*. Tech. rep. Geneva: CERN, 2021. URL: <https://cds.cern.ch/record/2767187>.
- [28] ATLAS Collaboration. *ATLAS Computing Acknowledgements*. Tech. rep. Geneva: CERN, 2020. URL: <https://cds.cern.ch/record/2717821>.
- [29] Joao Pequenao and Paul Schaffner. “How ATLAS detects particles: diagram of particle paths in the detector”. 2013. URL: <https://cds.cern.ch/record/1505342>.

- [30] ATLAS Collaboration. “Operation of the ATLAS trigger system in Run 2”. In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10004. DOI: [10.1088/1748-0221/15/10/P10004](https://doi.org/10.1088/1748-0221/15/10/P10004). URL: <https://dx.doi.org/10.1088/1748-0221/15/10/P10004>.
- [31] ATLAS Collaboration. *The Optimization of ATLAS Track Reconstruction in Dense Environments*. Tech. rep. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2002609>.
- [32] ATLAS Collaboration. “Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2”. In: *The European Physical Journal C* 77.10 (2017), p. 673. DOI: [10.1140/epjc/s10052-017-5225-7](https://doi.org/10.1140/epjc/s10052-017-5225-7). URL: <https://doi.org/10.1140/epjc/s10052-017-5225-7>.
- [33] R. E. Kalman. “A New Approach to Linear Filtering and Prediction Problems”. In: *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45. ISSN: 0021-9223. DOI: [10.1115/1.3662552](https://doi.org/10.1115/1.3662552). eprint: https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/82/1/35/5518977/35_1.pdf. URL: <https://doi.org/10.1115/1.3662552>.
- [34] ATLAS Collaboration. “Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC”. In: *Eur. Phys. J. C* 77.5 (2017), p. 332. DOI: [10.1140/epjc/s10052-017-4887-5](https://doi.org/10.1140/epjc/s10052-017-4887-5). arXiv: [1611.10235 \[physics.ins-det\]](https://arxiv.org/abs/1611.10235).
- [35] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. Geneva: CERN, 2008. URL: <https://cds.cern.ch/record/1099735>.
- [36] ATLAS Collaboration. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J. C* 79.8 (2019), p. 639. arXiv: [1902.04655](https://arxiv.org/abs/1902.04655). URL: <https://cds.cern.ch/record/2657964%7D>.
- [37] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti- k_t jet clustering algorithm”. In: *JHEP* 04 (2008), p. 063. DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063). arXiv: [0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189).
- [38] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *The European Physical Journal C* 77.7 (2017), p. 466. DOI: [10.1140/epjc/s10052-017-5031-2](https://doi.org/10.1140/epjc/s10052-017-5031-2). URL: <https://doi.org/10.1140/epjc/s10052-017-5031-2>.
- [39] ATLAS Collaboration. “Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Rev. D* 96 (7 Oct. 2017), p. 072002. DOI: [10.1103/PhysRevD.96.072002](https://doi.org/10.1103/PhysRevD.96.072002). URL: <https://link.aps.org/doi/10.1103/PhysRevD.96.072002>.
- [40] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *The European Physical Journal C* 77.7 (2017), p. 490. DOI: [10.1140/epjc/s10052-017-5004-5](https://doi.org/10.1140/epjc/s10052-017-5004-5). URL: <https://doi.org/10.1140/epjc/s10052-017-5004-5>.
- [41] ATLAS Collaboration. “Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector”. In: *Journal of High Energy Physics* 2013.9 (2013), p. 76. DOI: [10.1007/JHEP09\(2013\)076](https://doi.org/10.1007/JHEP09(2013)076). URL: [https://doi.org/10.1007/JHEP09\(2013\)076](https://doi.org/10.1007/JHEP09(2013)076).
- [42] ATLAS Collaboration. “Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *The European Physical Journal C* 81.8 (2021), p. 689. URL: <https://doi.org/10.1140/epjc/s10052-021-09402-3>.
- [43] ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. Geneva: CERN, 2014. URL: <https://cds.cern.ch/record/1700870%7D>.

- [44] Particle Data Group et al. “Review of Particle Physics”. In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020), p. 083C01. ISSN: 2050-3911. DOI: [10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104). eprint: <https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf>. URL: <https://doi.org/10.1093/ptep/ptaa104>.
- [45] ATLAS Collaboration. *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*. Tech. rep. Geneva: CERN, 2019. URL: [%5Curl%7Bhttps://cds.cern.ch/record/2688062%7D](https://cds.cern.ch/record/2688062).
- [46] ATLAS Collaboration. “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 78.11 (2018), p. 903. DOI: [10.1140/epjc/s10052-018-6288-9](https://doi.org/10.1140/epjc/s10052-018-6288-9). URL: <https://doi.org/10.1140/epjc/s10052-018-6288-9>.

Appendices