TABLE OF CONTENTS

Chapter 1

Physics Building Blocks and their Reconstruction

If you can put your five fingers through it
it is a gate, if not a door.
–Stephen Dedalus, in James Joyce's
Ulysses

In order to convert the multitude of electrical signals read out by the subdetectors of ATLAS as a result of a sucessful trigger (c.f. Section **??**) into well-defined and meaningful representations of the underlying physics process that initiated them, at the level required for performing high-quality physics analysis, several steps of reconstruction and identification must take place. The physics analyses presented in the current work involve the use of leptons, jets, and the so-called missing transverse momentum, $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$. The methods used to deduce the presence of these objects within the ATLAS detector will be discussed in this chapter. Section 1.1 introduces the reconstruction of charged-particle tracks and *pp* interaction vertices within the ID, both of which are used as low-level seeds or inputs to the reconstruction of the high-level physics objects to be discussed in the subsequent sections. Section 1.2 goes on to discuss the reconstruction of the charged leptons relevant to the current work: electrons and muons. Sections 1.3 and 1.4 describe the reconstruction of jet objects and the identification of jets arising from the decay of heavy-flavor hadrons, respectively. Section 1.5 then goes on to describe the reconstruction of $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, which relies on an accurate description of leptons and jets. The methods used for reconstructing the leptons and jets are not one hundred percent accurate: detector information arising due to an electron may leave signatures similar to those of a jet, for example, and thus spoil their unambiguous description. Where relevant, in the following we will discuss the methods by which the reconstruction and identification of the physics objects is made more precise and how high levels of confidence about their actual presence within the detector are achieved. Section 1.6 will also introduce the notion of high-level object ambiguity resolution through the use of so-called *overlap removal* procedures.

1.1 Charged-Particle Tracks and Primary Vertices

The reconstruction of charged-particle tracks ('tracking') and primary interaction vertices ('vertexing') is based on information provided by the ID, primarily by the the pixel and SCT subdetectors [1, 2, 3, 4, 5]. Charged-particles produced in pp collisions will leave signals — *hits* — on the different layers of the ID. The aim of tracking is to translate these layer hits into *spacepoints* which are then combined to form a track following the particle's traversal trough the ID. Given its highly granular readout, the pixel detector provides three dimensional spacepoints from each layer hit while the back-to-back readout strips on each layer of the SCT must be combined, using the stereo-angle information from the second set of strips, to give three dimensional spacepoint information. The hit information provided by the TRT straws is two-dimensional in nature, providing only $r - \phi$ information in the barrel section and $\phi - z$ information in the end-caps.

Within the solenoidal magnetic field of the ID, charged-particle tracks follow helical trajectories in the plane transverse to the beam-pipe (xy-plane) and can be fully characterised by five *track (perigee) parameters*:

$$(d_0, z_0, \phi, \theta, q/p)\,, \tag{1.1}$$

where d_0 (z_0) is the transverse (longitudinal) impact parameter, ϕ and θ are the azimuthal and polar coordinate, respectively, of the track at the point at which d_0 and z_0 are defined, q/p is the ratio of the particle charge to the magnitude of its momentum. The charge of a track is determined by its curvature within the magnetic field. The track parameters are defined with respect to their associated primary vertex, whose reconstruction will be described shortly. An illustration describing the track parameters is provided by Figure 1.1.

The primary track reconstruction algorithm used in ATLAS follows an *inside-out* pattern recognition procedure and first starts with information provided by track *seeds*, composed of a few spacepoints, in the silicon detectors which then are extended outwards into the TRT [1]. The inside-out approach accounts for the majority of tracks reconstructed in ATLAS but it is complemented by an *outside-in* approach that starts with the TRT hits and moves inwards [1]. This latter approach is useful in recovering those tracks with ambiguous or missing inner-layer pixel hits; for example, in the case of photon conversions or long-lived neutral particle decays.

The collection of reconstructed tracks is used as input to the primary vertex reconstruction. Primary vertex reconstruction follows a so-called *adaptive vertex fitting* (AVF) [3, 5] procedure and occurs in two steps: primary vertex finding, in which tracks are associated to a particular vertex candidate, and vertex fitting, which involves the reconstruction of the actual vertex position and its errors. After the vertex fitting stage, the tracks associated with a given vertex are refit with the constraint of the vertex position and its errors. The track refitting can update the track parameters (Eqn. 1.1) associated with the tracks. Only vertices with at least two charged particle tracks with $p_T > 400\,\text{MeV}$ are considered.

In the high luminosity collisions at the LHC there will generally be multiple primary vertices

associated with each pp bunch crossing. A physics *event* in ATLAS, then, is chosen as the set of processes originating from the pp interaction associated with the *hardest* primary vertex — the *primary hard-scatter vertex* — taken as that primary vertex with the highest sum of squared p_T of tracks originating from that vertex. The subsequent event reconstruction takes place around the primary hard-scatter vertex and only those objects originating from it are taken as relevant when reconstructing the physics objects in the event. Any additional primary vertices are considered as *pileup vertices*.

The presence of so-called *secondary*, *tertiary*, and so on..., vertices are also important and will be described in Section 1.4.

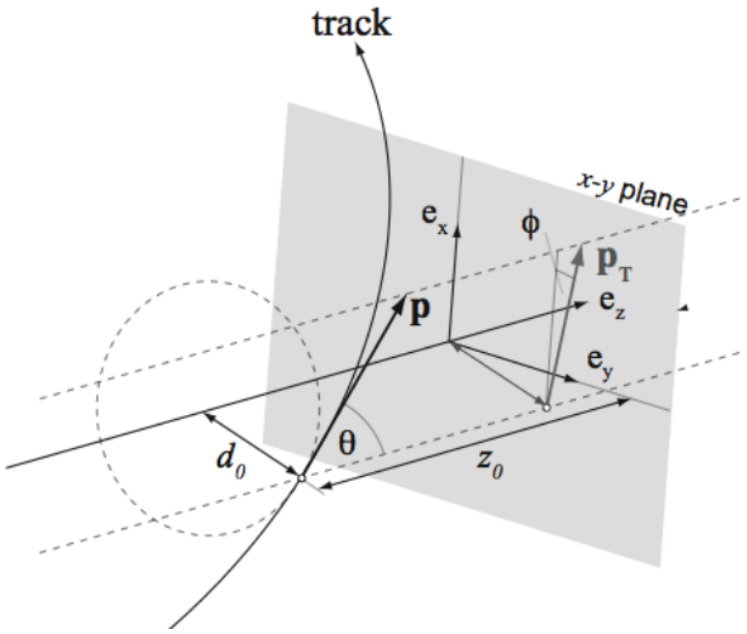

Figure 1.1: Illustration of the relationship between the track parameters and associated track. In this scenario, the hard scatter primary vertex is located at $(e_x, e_y, e_z) = (0, 0, 0)$, though this is not generally the case.

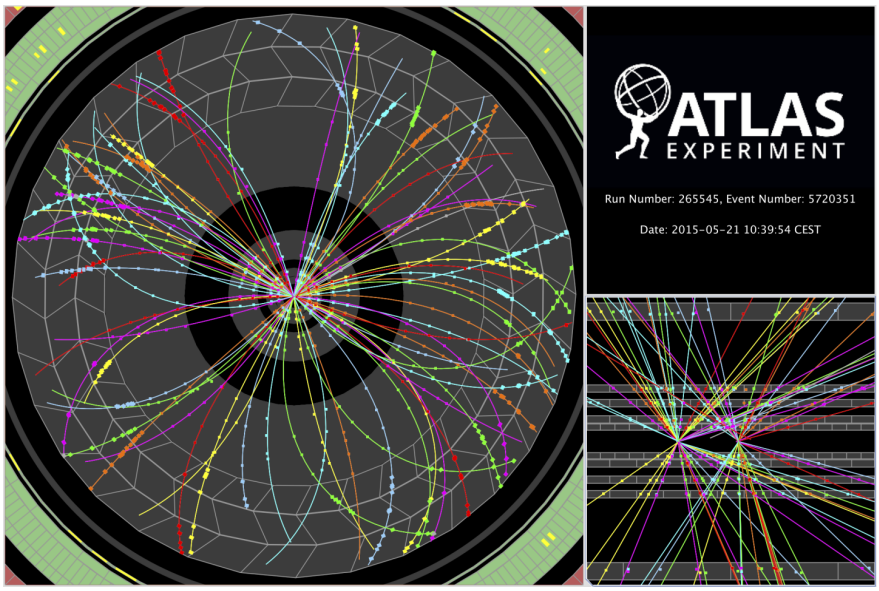

Figure 1.2: Event display of a low-pileup event recorded at the start of Run-II, in early 2015. *Left*: Transverse view of the ID. Seen in color are the reconstructed tracks traversing the inner layers of the pixel detector, SCT, and TRT. The colored dots are all reconstructed spacepoints used as input to the track fitting procedure. *Right, lower*: View in $r - z$ of the same pp bunch-crossing event as on the left. Two reconstructed primary vertices are clearly observed. On average, in Run-II there were roughly 30 primary vertices reconstructed per event, with up to ≈ 65 occuring at maximum.

1.2 Electrons and Muons

Electrons and muons, being charged particles, leave identifiable tracks within the ID. As a result, their reconstruction involves the use of the tracks and vertices described in the previous section, using them essentially as initial seeds for their complete reconstruction. Electron reconstruction, described in Section 1.2.1, complements the track information provided by the ID with calorimetric information provided by the EM calorimeter (Section **??**) and with knowledge about the pattern of transition radiation expected to occur in the TRT as a result of passing electrons. Muon reconstruction, described in Section 1.2.2, revolves around stitching together the tracks reconstructed in the ID with those tracks independently reconstructed in the MS layers at large radii.

1.2.1 Electrons

Electron Reconstruction

After 2016 they replaced sliding window algorithm with supercluster-based reco

The reconstruction of electron candidates is based on three components which characterise the signature of electrons: localised clusters of energy deposits found within the EM calorimeter, charged-particle tracks identified in the ID, close-matching (in (η, ϕ)) of the tracks to the clusters that form the final electron candidates [6]. It is generally possible to match multiple tracks to

the same electrogmagnetic cluster, all originating from the same primary electron produced in the hard-scatter. This is due to the fact that electrons lose significant amounts of energy to bremsstrahlung photons as they interact with and traverse the ID. These radiated photons can then undergo conversion to electron-positron pairs, which, too, can undergo further bremsstrahlung. The positrons, electrons, and photons are usually emitted in a very collimated fashion and thus deposit most of their energy in a localised fashion within the calorimter.

The search for localised energy deposits in the EM calorimeter is performed by following a sliding window algorithm over the individual cells whose dimensions are defined by the second sampling layer of the EM calorimeter (Figure **??**). Electron candidates are seeded by localised energy deposits whose summed transverse energy, across all layers of the EM calorimeter, is greater than 2.5 GeV [6]. These clusters act as seeds for the matching of reconstructed ID tracks. The reconstructed tracks are refit using a Gaussian Sum Filter (GSF) method [7] that accurately accounts for the bremsstrahlung energy losses characteristic of electron traversal and are then matched to the localised clusters using the cluster barycenter as the point of reference to match in $\eta - \phi$. If there is no GSF-track candidate matching to the EM calorimeter cluster seed, then the cluster is marked as an unconverted photon. The cluster is marked as a converted photon if a matched GSF-track candidate exists but is not associated with the primary hard-scatter vertex.

Electron Identification

Once electron candidates are reconstructed, they are selected based on various levels of identification. A further set of identification criteria is required on top of the reconstruction so as to improve the selection of true electrons originating from the primary hard-scatter vertex — so-called *prompt* electrons — over *non-prompt* sources of reconstructed electrons such as those originating from photon conversions or the misidentifiation of charged pions that leave electron-like tracks in the ID. This identification criteria is based on the construction of a multivariate likelihood (LH) and is referred to as the *electron likelihood identification*. The inputs to the LH are listed in Table 1.1 and include measurements from the tracking system in the ID, calorimetric information, and quantities that combine the tracking and calorimetric information [6].

The electron LH is based on the products for the signal and background probability density functions (PDFs) associated with the set of inputs in Table 1.1:

$$L_{S\,(B)}(\mathbf{x}) = \prod_{i=1}^{n} P_{S\,(B),i}(x_i), \tag{1.2}$$

where $\mathbf{x}$ is the vector of quantities listed in Table 1.1 and the $P_{S\,(B),i}(x_i)$ are the values of the PDF for quantity i at value x_i for the signal (S) and background (B). The likelihoods are built using simulation and the signal is composed of samples of prompt electrons and the background is built from a combination of jets that mimic the signature of prompt electrons, electrons from photon conversions, and non-prompt electrons from the decay of hadrons containing heavy-flavours [6].

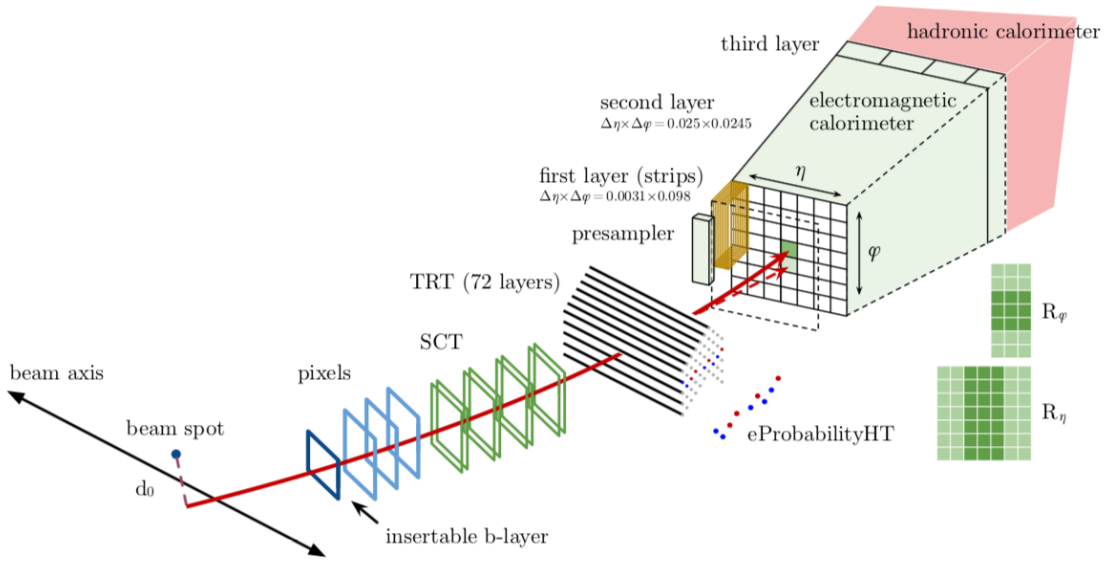

Figure 1.3

The final electron LH discriminant, shown in Figure 1.4, is based on a transformed version of the ratio,

$$d_L = \frac{L_S}{L_S + L_B}, \tag{1.3}$$

where the transformation acts to spread d_L to values not bounded by 0 and 1, motivated by the need to have well-defined working points based on selections on d_L.

There are four such fixed values of the final LH discriminant that are used to define four working points corresponding to increasing thresholds on the final LH discriminant: VeryLoose, Loose, Medium, and Tight. The efficiencies to identifiy prompt electron candidates are measured using samples of $Z \to ee$ and $J/\psi \to ee$ following a tag-and-probe approach. They are found, for electron candidates with $E_T > 40\,\text{GeV}$, to be 93%, 88%, and 80% for the Loose, Medium, and Tight working points, respectively [6].

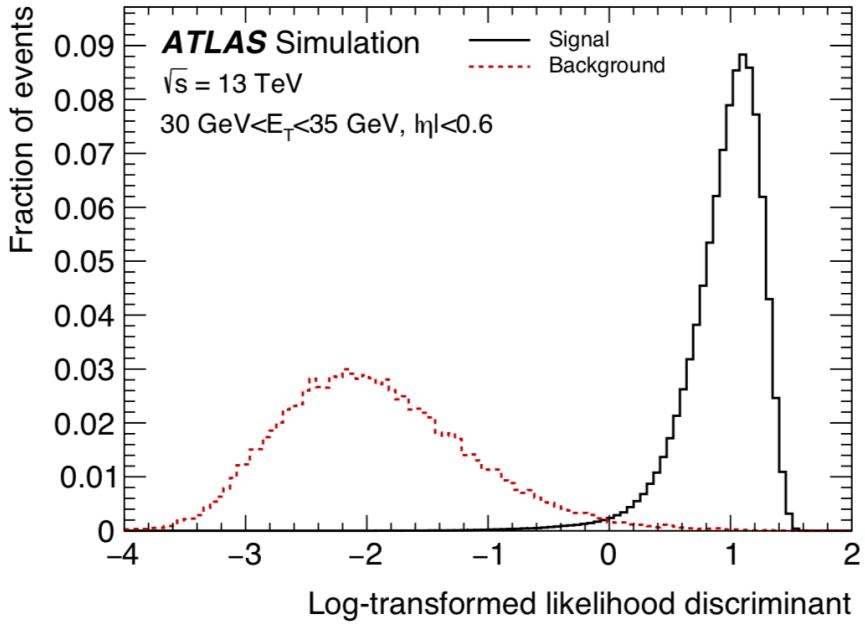

Figure 1.4: Transformed LH-based electron identification discrimiant for electron candidates with $30\,\text{GeV} < E_T < 35\,\text{GeV}$ and $|\eta| < 0.6$. From Ref. [6].

Table 1.1: From Ref. [6].

Input Type	**Name**	**Description**
Hadronic Leakage	R_{had1}	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster
	R_{had}	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster
Third layer of EM calorimeter	f_3	Ratio of the energy in the third layer to the total energy in the EM calorimeter. Only used for $E_T < 30\,\text{GeV}$ and $\lvert\eta\rvert \leq 2.37$.
Second layer of EM calorimter	$w_{\eta 2}$	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells centered at the electron cluster position.
	R_ϕ	Ratio of the energy in 3×3 cells over the energy in 3×7 cells centered at the electron cluster position.
	R_η	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centered at the electron cluster position.
First layer of EM calorimeter	w_{stot}	Shower width, $\sqrt{(\sum E_i (i - i_{\text{max}})^2)/(\sum E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η, and i_{max} is the index of the highest-energy strip. Used only for $E_T > 150\,\text{GeV}$.
	E_{ratio}	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies.
	f_1	Ratio of the energy in the first layer to the total energy in the EM calorimeter.
Track conditions	n_{Blayer}	Number of hits in the innermost pixel layer.
	n_{Pixel}	Number of hits in the pixel detector.
	n_{Si}	Total number of hits in the pixel and SCT detectors.
	d_0	Transverse impact parameter relative to the beam-spot.
	$\lvert d_0/\sigma(d_0)\rvert$	Significance of transverse impact parameter defined as the ratio of d_0 to its uncertainty.
	$\Delta p/p$	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee.
TRT	eProbabilityHT	Likelihood probability based on transition radiation in the TRT.
Track-cluster matching	$\Delta\eta_1$	$\Delta\eta$ between the cluster position in the first layer and the extrapolated track.
	$\Delta\phi_{\text{res}}$	$\Delta\phi$ between the cluster position in the second layer of the EM calorimeter and the momentum-rescaled track, extrapolated from the perigee, times the charge q.
	E/p	Ratio of the cluster energy to the track momentum. Used for $E_T > 150\,\text{GeV}$.

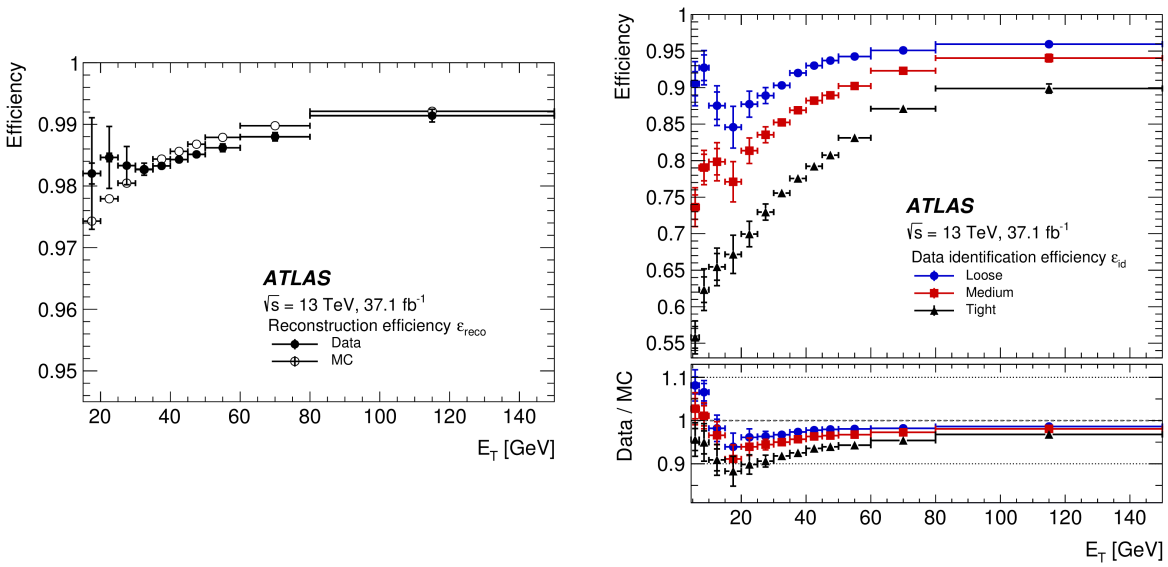

Figure 1.5: 2015-2016 *Left*: Electron candidate reconstruction efficiency, measured in simulation and in data, as a function of the candidate E_T. *Right*: Electron identification efficiency measured in data, as a function of electron E_T, for the three standard LH identification working points LOOSE (blue), MEDIUM (red), and TIGHT (black). The lower panel shows the ratio of the efficiency measured in data over that measured in simulation.

1.2.2 Muons

Muon Reconstruction

The reconstruction of muon candidates is performed by combining the tracking capabilities of the ID and the MS [8]. Muon reconstruction first starts with the independent reconstruction of charged-particle tracks within the ID and the MS. The independently formed tracks are subsequently combined to form a complete track representing the traversal of a muon through the full detector. The muon track in the ID is reconstructed like any other charged-particle track (Section 1.1).

Muon reconstruction within the MS starts with a pattern finding phase, looking for hit patterns each of the muon chambers to then form track segments. The track segments between different MS layers are then fit together to form muon track candidates. At least two matching track segments are required in order to build a muon track candidate, except in the transition region between the barrel and end-cap where single track-segment candidates can be used. Once a muon track candidate is formed from the combined segments, a global χ^2 fit is performed to improve the association of hits to each muon candidate. The χ^2 is repeated several times, removing outlying hits as necessary, until a threshold is met for all associated hits.

There are several algorithms used to combine the muon track candidates in the ID and MS, each using different sets of information related to the ID, MS, and calorimeters. At the time of the current work, there are four standard combination algorithms used each based on the subdetectors

used in their construction:

- **Combined Muon (CB)** This type of muon is formed with a global refit using all muon track candidate hits in the ID and the MS. Hits may be added or removed from the MS track candidate during the refit. Muons are reconstructed following an outside-in pattern recognition algorithm, in which the muon is first reconstructed in the MS and extrapolated inwards to the ID hits. A complementary, albeit non-standard, inside-out algorithm also exists.

- **Segment-tagged Muon (ST)** An ID track is classified as a muon if, once extrapolated to the MS, it is associated with at least one local track segment in the MDT or CSC chambers. Segment-tagged muons are used when a muon candidate crosses only one layer of the MS chambers, either because of their low p_T or because they fall into un-instrumented regions of the MS.

- **Calorimeter-tagged Muon (CT)** An ID track is classified as a muon if it is matched to an energy deposit in the caloriemter that is compatible with a minimum ionising particle. Calorimter-tagged muons have the lowest purity, but recover acceptance in regions of the MS that are only partially instrumented to allow for cabling and services to the calorimeter and ID systems, particularly in the region $|\eta| < 0.1$.

- **Extrapolated Muon (ME)** This type of muon is based only on the track candidates formed in the MS and a loose requirement that the track candidate be pointing back towards the IP. Extrapolated muons are mainly used to extend the acceptance of muon reconstruction into the region $2.5 < |\eta| < 2.7$ that is not covered by the ID acceptance.

Muon Identification

Muon identification refers to the act of applying aditional quality criteria on the reconstructed muon candidates in order to mainly suppress contamination from background sources that mimic muon signatures, such as pion and kaon decays, while ensuring high rates for the acceptance of prompt muons. There are three standard muon identification working points in ATLAS, each a subset of the previous one, and are referred to as the Loose, Medium, and Tight muon identification working points. Medium muons are the default in ATLAS analyses and can only be composed of CB and ME muons. As all muons used in the present thesis are Medium muons, only this identification working point will be described in detail.

Reconstructed muon candidates originating from non-prompt sources such as in-flight decays of charged hadrons in the ID, are often characterised by the presence of a 'kink' in the reconstructed muon's track. It is therefore expected that the independent momentum measurements made in the ID and MS may be incompatible for non-prompt sources of muon candidates. The muon identification criteria, then, make use of quantities that relate the ID and MS muon track candidates. These quantities are described in Table 1.2. Medium muons have a rather loose selection on the

compatibility between the ID and MS momentum measurements and, with respect to those quantities in Table 1.2, are only required to have a q/p significance less than 7. On top of requirements on those quantities described in Table 1.2, the muon identification working points place additional requirements on the number and type of hits in the ID and MS. All identification working points require, in the ID, that there be at least 1 hit in the pixel subdetector, at least 5 hits in the SCT subdetector, less than 3 silicon holes,[1] and at least 10% of the TRT hits originally assigned to the muon track candidate exist after the combined reconstruction. MEDIUM muons further require that the CB muons have at least 3 hits in at least two MDT layers, except in $|\eta| < 0.1$ where tracks with at least one MDT layer but no more than one MDT hole are allowed. The ME Medium muons are required to have at least 3 MDT or CSC layer hits, and are employed only in $2.5 < |\eta| < 2.7$. reference n MDT/CSC hits figure?

Quantity Name	**Measurement**	**Description**
q/p significance	$\lvert(q/p)^{\mathrm{ID}} - (q/p)^{\mathrm{MS}}\rvert / \sqrt{\sigma^{\mathrm{MS}}_{p_T} + \sigma^{\mathrm{ID}}_{p_T}}$	Absolute value of the difference between the ratio of the charge and momentum of the muon candidates measured in the ID and MS, divided by the quadrature sum of the corresponding uncertainties.
ρ'	$\lvert p_T^{\mathrm{MS}} - p_T^{\mathrm{ID}}\rvert / p_T^{\mathrm{Combined}}$	Absolute value of the difference between the transverse momentum measurements in the ID and the MS, divided by that of the combined muon candidate.
χ^2_{norm}	–	Normalized χ^2 of the combined muon track fit

Table 1.2

[1] A missing hit is considered a 'hole' only if it falls between hits successfully assigned to a given track.

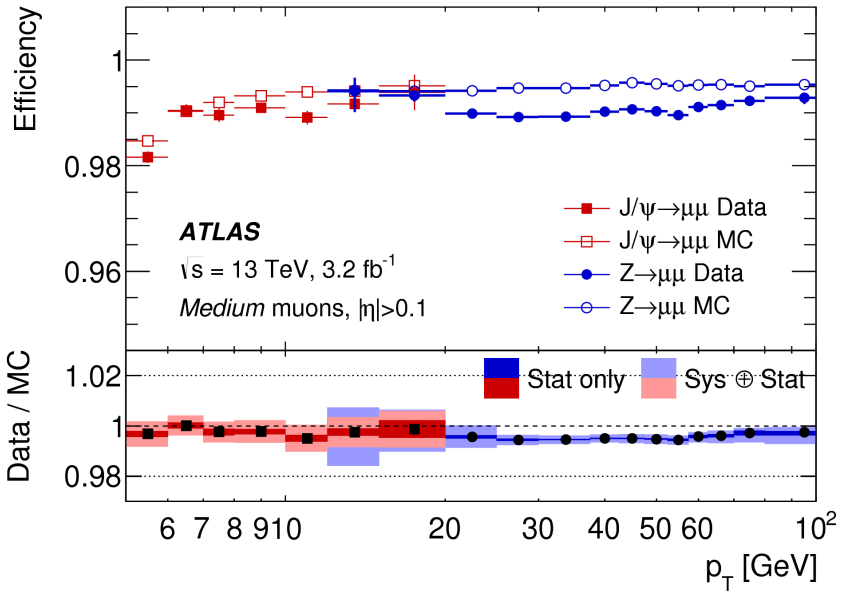

Figure 1.6: From Ref. [8].

1.3 Jets

1.4 Flavor Tagging of Jets

1.5 The Missing Transverse Momentum

1.6 Object-level Ambiguity Resolution

Bibliography

[1] T Cornelissen et al. *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*. Tech. rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002. Geneva: CERN, Mar. 2007. URL: https://cds.cern.ch/record/1020106 (cit. on p. 2).

[2] *The Optimization of ATLAS Track Reconstruction in Dense Environments*. Tech. rep. ATL-PHYS-PUB-2015-006. Geneva: CERN, Mar. 2015. URL: https://cds.cern.ch/record/2002609 (cit. on p. 2).

[3] ATLAS Collaboration. "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC". In: *Eur. Phys. J.* C77.5 (2017), p. 332. DOI: 10.1140/epjc/s10052-017-4887-5. arXiv: 1611.10235 [physics.ins-det] (cit. on p. 2).

[4] *Performance of primary vertex reconstruction in proton-proton collisions at* $\sqrt{s}$ *=7 TeV in the ATLAS experiment*. Tech. rep. ATLAS-CONF-2010-069. Geneva: CERN, July 2010. URL: https://cds.cern.ch/record/1281344 (cit. on p. 2).

[5] G Piacquadio, K Prokofiev, and A Wildauer. "Primary vertex reconstruction in the ATLAS experiment at LHC". In: *Journal of Physics: Conference Series* 119.3 (July 2008), p. 032033. DOI: 10.1088/1742-6596/119/3/032033. URL: https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032033 (cit. on p. 2).

[6] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data". In: (2019). arXiv: 1908.00005 [hep-ex] (cit. on pp. 4, 5, 6, 7, 8).

[7] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, May 2012. URL: https://cds.cern.ch/record/1449796 (cit. on p. 5).

[8] Georges Aad et al. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex] (cit. on pp. 9, 12).