

Double B -hadron Jet Tagging and
Identification of Gluon to $b\bar{b}$ jets with the
ATLAS Detector

Lic. María Laura González Silva

Tesis Doctoral en Ciencias Físicas
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Noviembre 2012



UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Física

**Double B -hadron Jet Tagging and Identification of
Gluon to $b\bar{b}$ jets with the ATLAS Detector**

Trabajo de Tesis para optar por el título de
Doctor de la Universidad de Buenos Aires en el área Ciencias Físicas

por **María Laura González Silva**

Director de Tesis: Dr. Ricardo Piegaia

Consejero de estudios: Dr. Daniel Deflorian

Lugar de Trabajo: Departamento de Física (CONICET-UBA)

Buenos Aires, 2012

AGRADECIMIENTOS

A Ricardo, a mi familia...

Abstract

El detector ATLAS, uno de los cuatro experimentos del acelerador LHC actualmente en construcción en el CERN, tiene como propósito principal el descubrimiento del bosón de Higgs, la partícula involucrada en el mecanismo de ruptura de la simetría electrodébil y el origen de la masa, así como la búsqueda de nueva física más allá Modelo Estándar hasta la escala de 1 TeV. En el colisionador chocarán haces de protones con una energía de centro de masa de 14 TeV y una frecuencia de interacción de 10^9 Hz.

This thesis describes a method that allows the identification of double B -hadron jets originating from gluon-splitting. The technique exploits the kinematic differences between the so called “merged” jets and single B -hadron jets using track-based jet shape and jet substructure variables combined in a multivariate likelihood analysis. The ability to reject b -jets from gluon splitting is important to reduce and to improve the estimation of the b -tag background in Standard Model analyses and in new physics searches involving b -jets in the final state. In the simulation, the algorithm rejects 95% (50%) of merged B -hadron jets while retaining 50% (90%) of the tagged b -jets, although the exact values depend on the jet p_T .

Contents

1	Introduction	4
1.1	Identification of b -jets from gluon splitting	4
1.2	Physics Motivation	6
2	The theory of the strong interactions	14
2.1	The Standar Model	14
2.2	Jet physics	14
2.2.1	Monte Carlo parton-shower programs	14
2.2.2	Jet algorithms	18
2.2.3	Jet substructure	18
3	The ATLAS detector at the LHC	19
3.1	The Large Hadron Collider	19
3.1.1	Luminosity and pile-up	23
3.2	The ATLAS Detector	26
3.2.1	Inner tracking system	28
3.2.2	The Calorimeter System	32
3.2.3	The Muon System	39
3.2.4	Forward detectors	40
3.2.5	Trigger and Data Adquisition	41

3.2.6	ATLAS Performance and Data quality	43
4	Jet reconstruction and <i>b</i>-Tagging	46
4.1	Jets reconstruction and calibration	46
4.2	<i>b</i> -jet Tagging	47

%

Chapter 1

Introduction

1.1 Identification of b -jets from gluon splitting

The ability to identify jets containing B -hadrons is important for the high- p_T physics program of the ATLAS experiment. Two robust b -tagging algorithms taking advantage of B -hadron pair via gluon splitting, as depicted in Fig. 1.1. We will henceforth call “merged” b -jets $b\bar{b}$ jets the b -tagged jets containing two B -hadrons. The ability to single out b -tagged jets from gluon splitting has several applications in different lines of analysis: measurement of QCD beauty production, $t\bar{t}$ and single top production, reduction of background in searches with b -quarks in the final state, and the study of substructure in fat jets, where $g \rightarrow b\bar{b}$ jets compete with boosted $Z \rightarrow b\bar{b}$ and $H \rightarrow b\bar{b}$ jets.

There are two possible strategies to attempt the identification of b -jets containing two B -hadrons. One of them relies on the direct reconstruction of the two B -decay secondary vertices [4]. This has the further advantage of allowing the measurement of the angular separation between the B -hadrons,

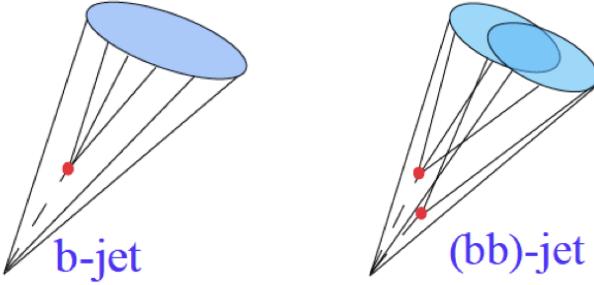


Figure 1.1: b -tagging algorithms select jets originating both from the fragmentation of a single b quark (“single” b -jets, left image) or from the splitting of a gluon into a pair of close-by $b\bar{b}$ quarks (“merged” b -jets, right image).

but suffers from the low efficiency of a double b -tag requirement plus additional reconstruction inefficiencies at small angular separation between the two B s. In this paper we develop an alternative method that does not rely on explicitly finding vertices, but exploits the substructure differences between single and merged b -jets, combining them in a multivariate analysis.

The note is organized as follows. In section 1.2 we review the physics cases where this tool finds natural applications. Sections ?? and 4.1 review the Data and Monte Carlo samples and their reconstruction and section ??, the criteria for selection of events, jets and tracks. The kinematic variables that differentiate between single and merged b -jets are discussed in section ?? and the validation of their MC distributions with QCD data in section ???. The construction of the multi-variate discriminator is presented in section ?? and the discussion of the systematic uncertainties in section ???. Section ?? investigates the performance of the tagger with other Monte Carlo generators and, finally, section ?? summarizes the results and discusses future improvements and new ideas.

1.2 Physics Motivation

Within the Standard Model (SM) a range of production channels exist for heavy-quark jets, pure QCD production or in association with heavy bosons (W, Z, H). Furthermore, b -quarks enter in many collider searches, notably because they are produced in the decays of various SM particles, top quarks and the Higgs boson (if light), and of numerous particles appearing in proposed extensions of the SM. The ability to distinguish genuine b -quark jets from those produced via gluon splitting is thus of wide application. Here we briefly discuss three cases, the measurement of QCD b quark production, the reduction of background in SM and BSM analyses with b quarks in the final state, and studies of jet substructure.

The measurement of the inclusive b -jet spectrum

The simplest and most fundamental measurement of heavy-quark jet production is the inclusive heavy-quark jet spectrum, which is dominated by pure QCD contributions. Studies of QCD bottom jets production are of intrinsic interest because of the correspondence between parton level production and the observed hadron level, and their potential to provide information on the b -quark parton distribution function, the only component of the proton structure thought to be generated entirely perturbatively from the DGLAP evolution of the other flavours. The theoretical calculation of the inclusive b -jet spectrum presents the striking feature of having rather important uncertainties ($\sim 50\%$), considerably larger than the corresponding ones for the normal (light) jet inclusive spectrum ($\sim 10\text{-}20\%$), see for example [5].

The origin of these uncertainties are reviewed in a recent paper by Banfi, Salam and Zanderighi [6], from which we have taken Fig. 1.2. Its top panel

shows the K -factor, the ratio of the next-to-leading order (NLO) to the leading order (LO) cross section, obtained with MCFM for the LHC design energy ($pp, \sqrt{s} = 14$ TeV). The fact that NLO terms are considerably larger than the LO ones indicates that the perturbative series is very poorly convergent, and implies that the NLO result cannot be an accurate approximation to the full result. It is for this reason that the scale dependence (middle panels) is large. The poor convergence of the perturbative series is related to the different channels for heavy quark production. At leading order only the so-called flavour creation channel (FCR) is present, $\ell\ell \rightarrow b$, where ℓ is a generic light parton (quark or gluon), see fig. 1.3. At NLO, two new channels open up, often referred to as flavour excitation (FEX) and gluon splitting (GSP). In the former, a gluon from one of the incoming hadrons splits collinearly into a $b\bar{b}$ -pair and one of those b -quarks enters the hard $b\ell \rightarrow b\ell$ scattering. In the gluon splitting process, the hard scattering LO diagram is of the form $\ell\ell \rightarrow \ell\ell$, and one of the final-state light partons (at NLO always a gluon) splits collinearly into a $b\bar{b}$ -pair that the clustering algorithm can classify within the same jet. A jet containing both b and \bar{b} is considered to be just a b -jet in standard definitions. The various channels can be conveniently separated with a parton shower Monte Carlo generator such as Herwig [8], where one can determine the underlying hard channel from the hard process in the event record. Their relative contributions to the total b -jet spectrum are shown in the bottom panel of fig. 1.2. One sees that the supposedly LO channel (FCR) is nearly always smaller than the two channels that enter only at NLO (FEX and GSP).

The largest residual uncertainties are associated with the channel with the most logarithms, gluon splitting. This channel however does not even correspond to one's physical idea of a b -jet, the one induced by a hard b -quark,

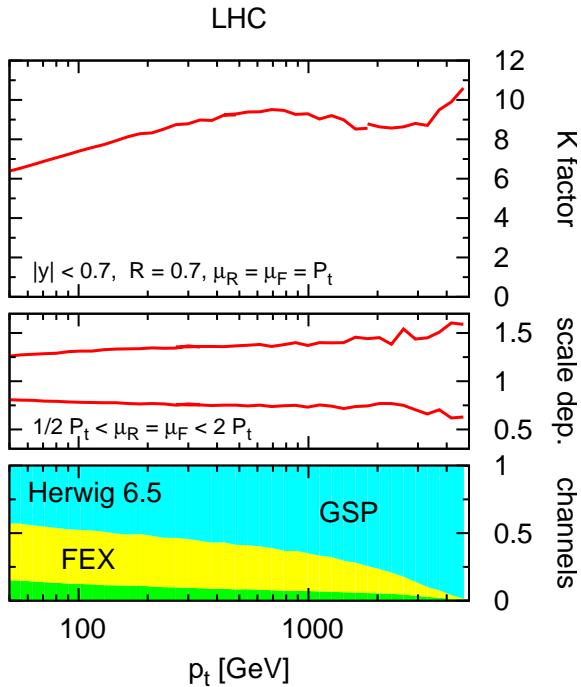


Figure 1.2: Top: K -factor for inclusive b -jet spectrum taken from [6], clustering particles into jets using the k_t jet-algorithm [7] with $R=0.7$, and selecting jets in the central rapidity region ($|y| < 0.7$). Middle: scale dependence obtained by simultaneously varying the renormalisation and factorisation scales by a factor two around p_T , the transverse momentum of the hardest jet in the event. Bottom: breakdown of the Herwig [8] inclusive b -jet spectrum into the three major underlying channels, flavor creation (FCR) flavor excitation (FEX) and gluon splitting (GSP).

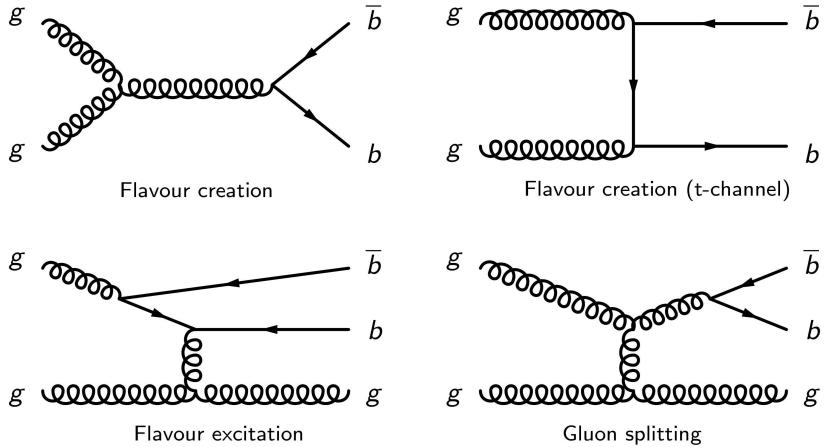


Figure 1.3: Typical Feynman diagrams for the 3 processes contributing up to next-to-leading order to QCD bottom production: (a-b) Flavor creation (FCR), (c) Flavor excitation (FEX), (d) Gluon splitting (GSP).

and it seems somehow unnatural to include it at all as part of one's b -jet spectrum. Reference [6] thus proposes a new approach to improving the accuracy of the prediction of the b -jet spectrum, where b -jets definition maintains the correspondence between partonic flavour and jet flavour. Specifically, a jet containing equal number of b -quarks and b -antiquarks is considered to be a light jet, so that jets identified as gluon splitting are removed from the b -jet spectrum.

Rejection of background in SM analyses and beyond-SM searches

Efficient tagging of merged b -jets from gluon splitting can provide an important handle to understand, estimate and/or reject b -tag backgrounds to Standard Model and new physics searches at the LHC.

Standard Model physics analyses that rely on the presence of b quarks in the final state such as top quark physics, either in the $t\bar{t}$ or the single top

channels, and associated Higgs production: $WH \rightarrow \ell\nu b\bar{b}$ and $ZH \rightarrow \nu\nu b\bar{b}$, suffer from reducible backgrounds from QCD (that can produce b -jets from gluon splitting) and, most importantly, from the irreducible background due to W bosons produced in association with b quarks. Fig. 1.4 shows the two leading order processes that give rise to W bosons with at least one b -jet. In the first process, which can be thought of as a higher order correction to $W +$ jets production, the b quark pair is produced at small angles by gluon splitting and can often be reconstructed as a merged jet. NLO calculations

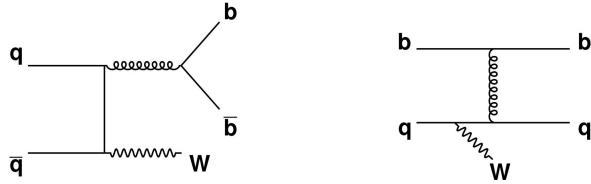


Figure 1.4: Leading order Feynman diagrams for W production in association with b quarks.

of the production of W bosons and two jets with at least one b quark at the LHC for jet $p_T > 25$ GeV, and $|\eta| < 2.5$ [9] indicate that the cross section for $W(b\bar{b})j$ is almost a factor of two higher than $Wb\bar{b}$, and about a third of Wbj , where $W(b\bar{b})j$ denotes the case in which the two b quarks are merged into the same jet.

New physics searches with b quarks in the final state can also greatly benefit from rejection of QCD and $W + b$ backgrounds with b -jets arising from gluon splitting. For example consider the search for supersymmetry in the $+ b$ -jets channel [10]. Within the framework of a generic R -parity conserving minimal supersymmetric extension of the SM The coloured superpartners of quarks and gluons, the squarks and gluinos, are expected to be copiously produced via the strong interaction at the LHC. The partners

of the right-handed and left-handed quarks, qR and qL, can mix to form two mass eigenstates, and these mixing effects being proportional to the corresponding fermion masses, they are expected to become most important for the third generation to yield sbottom (b_1) and stop (t_1) mass eigenstates significantly lighter than other squarks. Both sbottom and stop chain decay to b quarks and the lightest supersymmetric particle, producing the expected signal of + b -jets.

Jet substructure and boosted objects

At the LHC, many of the particles considered to be heavy at previous accelerators will be frequently produced with a transverse momentum greatly exceeding their rest mass. Good examples are the electro-weak gauge bosons W^\pm and Z^0 , the top quark, the Higgs boson or bosons and possibly other new particles in the same mass range. These boosted objects, produced either because they recoil against other energetic objects or because they arise from decays of even heavier BSM particles, can form upon decay a highly collimated topology too close to be resolved by a jet algorithm. For these cases, sophisticated tools have been developed in the last years [11, 12] to analyse the substructure of the ensuing jet and reveal its heavy-particle origin.

The study of $b\bar{b}$ jets from gluon splitting is an ideal testbed for studying jet substructure in data, as it provides a large supply of boosted, merged jets. Furthermore, understanding $g \rightarrow b\bar{b}$ jets is important as they are themselves the background to boosted object searches, like $Z \rightarrow b\bar{b}$ or $H \rightarrow b\bar{b}$. In particular, it has recently been suggested [13] that WH and ZH production become potential discovery and analysis channels by restricting ones attention to the $\sim 5\%$ of events in which the vector and Higgs bosons have large

transverse momentum, $p_{TH} > 200$ GeV. Understanding the much more common QCD events with merged $b\bar{b}$ jets will be essential before attempting to measure these rare final states.

%

Chapter 2

The theory of the strong interactions

2.1 The Standard Model

2.2 Jet physics

2.2.1 Monte Carlo parton-shower programs

We have used a perturbative language to describe the short-distance interactions of quarks, leptons and gauge bosons. For leptons and colourless bosons this language is sufficient. However, for quarks and gluons it must be complemented with the structure of incoming hadrons, and a picture for the hadronization process, wherein the coloured partons are transformed into jets of colourless hadrons, photons and leptons. The hadronization can be further subdivided into fragmentation and decays. This process is still not yet understood from first principles, but has to be based on models. In the data, fluctuations arise from the quantum mechanics of the underlying

theory. In generators, Monte Carlo techniques are used to select all relevant variables according to the desired probability distributions, and thereby ensure (quasi-)randomness in the final events.

QCD perturbation theory, formulated in terms of quarks and gluons, is valid at short distances. At long distances, QCD becomes strongly interacting and perturbation theory breaks down. In this confinement regime, the coloured partons are transformed into colourless hadrons, a process called either hadronization or fragmentation.

Different classes of techniques can be used to make QCD predictions at hadron colliders, and in particular at the LHC. The so called Matrix Element Monte Carlos, including LO, NLO, and NNLO calculations. These “fixed-order predictions” involve the first couple of terms in the QCD perturbative expansion for a given cross section; as one includes further orders in the expansion one can reasonably hope to see systematic improvement in the accuracy of one’s predictions.

The Monte Carlo parton shower programs, which simulate the full hadronic final state, are crucial in evaluating detector acceptances and response. Knowing QCD predictions is crucial in the design of methods to search for new physics, as well as for extracting meaning from the data.

(This) procedure for generating a parton shower event is present in PYTHIA (pt-ordered and HERWIG (angular ordering), the ones we use in this thesis and others like SHERPA. (REFERENCIAS!).

Hadronization cannot be calculated from first principles, but has to be modelled. The two most commonly used model classes are the string and cluster ones.

Real events consist not of partons but of hadrons. Since we have no idea how to calculate the transition between partons and hadrons, Monte Carlo

event generators resort to hadronization models. One widely-used model involves stretching a colour string across quarks and gluons, and breaking it up into hadrons [REFERENCIAS]. This is the Lund model in Monte Carlo program PYTHIA. Another model breaks each gluon into a q q pair and then groups quarks and anti-quarks into colourless clusters, which then give the hadrons. This cluster type hadronization is implemented in the HERWIG event generator [REFERENCIAS] and recent versions of SHERPA. Hadronization models involve a number of non-perturbative parameters. The parton-shower itself involves the non-perturbative cutoff Q_0 . These different parameters are usually tuned to data from the LEP experiments. In addition to the hard interaction that is generated by the Monte Carlo simulation, it is also necessary to account for the interactions between the incoming proton remnants. This is usually modelled through multiple extra 2 2 scattering, occurring at a scale of a few GeV, and known as multiple parton interactions. This modelling of the underlying event is crucial in order to give an accurate reproduction of the (quite noisy) energy flow that accompanies hard scatterings in hadron-collider events.

Pythia

PYTHIA is a general-purpose event generator. It has been used extensively for e^+e^- , ep and $pp/p\bar{p}$, e.g at LEP, HERA and the Tevatron, and during the last 20 years has probably been the most used generators for LHC physics studies.

PYTHIA contains an extensive list of hardcoded subprocesses, over 200, that can be switched on individually. These are mainly 2→1 and 2→2, some 2→3, but no multiplicities higher than that. Consecutive resonance decays may of course lead to more final-state particles, as will parton showers.

In PYTHIA generator showers are ordered in transverse momentum (REFERENCIA) both for ISR and for FSR. Also MPIs are ordered in p_T . The initial- and final-state algorithms are partly based on a dipole-type approach to recoils.

MPI modelling has traditionally been a hallmark of PYTHIA.

Hadronization is based solely on the Lund string fragmentation framework.

Samples of dijet events from proton-proton collision processes were simulated with PYTHIA 6.423 [11] using a 2 → 2 matrix element at leading order in the strong coupling to model the hard subprocess, pT -ordered parton showers to model additional radiation [12], underlying event and multiple parton interactions [13], and fragmentation and hadronisation based on the Lund string model [14]. The ATLAS AMBT2 tune of the soft model parameters was used [15].

Herwig++

Herwig++ is based on the event generator HERWIG (Hadron Emission Reactions With Interfering Gluons), which was first published in 1986 and was developed throughout the era of LEP. HERWIG was written in Fortran, and the new generator, Herwig++ developed in C++. The version used in this thesis is versin 2.4.2 released in 2009. Some distinctive features of Herwig++ are:

- Angular ordered parton showers.
- Cluster hadronization.
- Hard and soft multiple partonic interactions to model the underlying event and soft inclusive interactions (REFERENCIA)

A more detailed descripcion can be found elsewhere (REFERENCIAS).

2.2.2 Jet algorithms

2.2.3 Jet substructure

Chapter 3

The ATLAS detector at the LHC

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [14] is a proton-proton (pp) synchrotron located in the previous Large Electron Positron (LEP) collider tunnel at CERN Laboratory, just outside the city of Geneva (Switzerland), approximately 100 m underground. It is designed to collide bunches of up to $\sim 10^{11}$ protons every 25 ns at a center-of-mass energy of 14 TeV (seven times the 2 TeV reached by the Tevatron accelerator at Fermilab Laboratory, in Chicago).

The experiments analyzing the collisions produced by the LHC are distributed around the 27 km ring at the various interaction points. The ATLAS experiment is located at Point 1, which is closest to the main CERN site. Point 5 houses the other general purpose detector, CMS. ALICE and LHCb experiments are located at Point 2 and Point 8, respectively. The former is designed to investigate heavy ion collisions; the latter, to investigate rare decays of b-mesons. The layout of these four experiments along the LHC

ring is shown in Fig. 3.1.

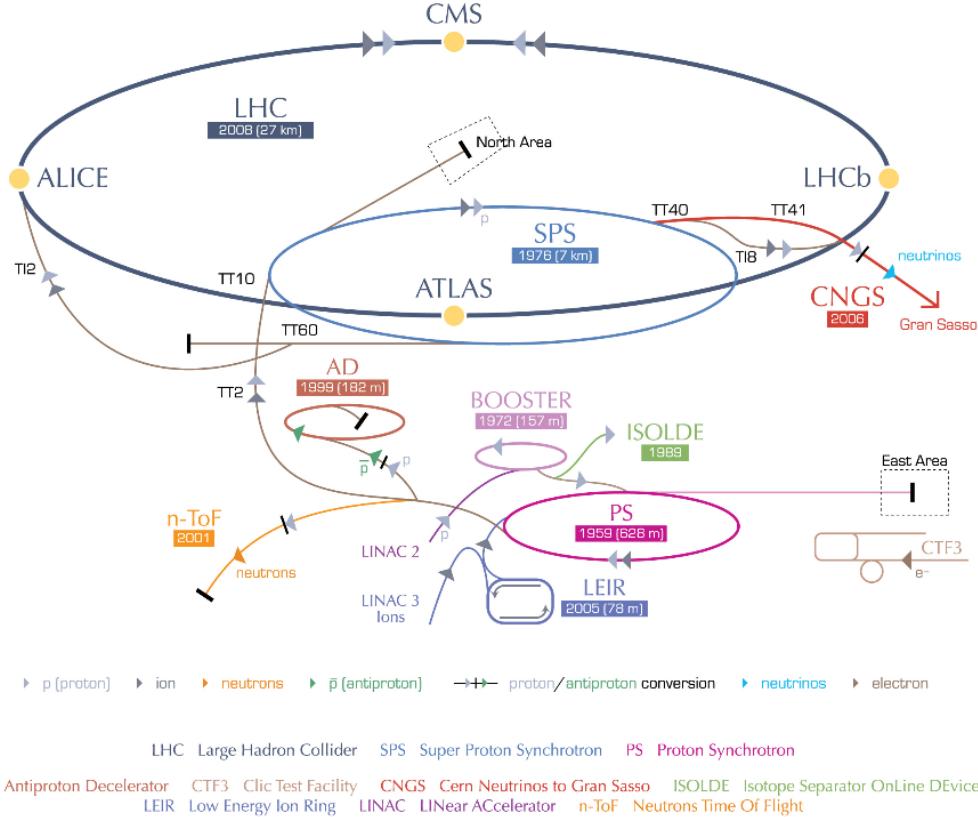


Figure 3.1: The CERN accelerator complex, showing the injection system, along with each components date of construction, and the placement of the four main experiments.

Proton beams are formed, before insertion into the main LHC ring, using a succession of smaller machines with increasingly higher energies, as shown in Fig. 3.1. The chain begins as protons are injected into the PS Booster (PSB) at an energy of 50 MeV from Linac2. The booster accelerates them to 1.4 GeV. The beam is then fed to the Proton Synchrotron (PS) where it is accelerated to 25 GeV. At desin strength, the bunch structure, known as a

bunch train, contains 72 bunches of protons upon entry to the Super Proton Synchrotron (SPS). The SPS accumulates up to four fills of 72 bunches from the PS and accelerates them to 450 GeV, with a bunch spacing of ~ 25 ns. They are finally transferred to the LHC (both in a clockwise and an anticlockwise direction) where they are accelerated for 20 minutes to their nominal energy of 7 TeV. Beams will circulate for many hours inside the LHC beam pipes under normal operating conditions.

The bunch structure is a direct consequence of a radio frequency (RF) acceleration scheme used to attain the desired high proton beam energy. In RF acceleration, particles travel through a series of time-varying electrical fields and they can only be accelerated when the RF field has the correct orientation when particles pass through an accelerating cavity, which happens at well specified moments during an RF cycle. The result of a sequence of RF accelerations is several bunches of protons. It is important to note that when we speak about “beams” we refer to many bunches of protons separated by some uniform distance. Increasing the number of bunches is one of the ways to increase luminosity in a machine (more about luminosity in subsection 3.1.1). At desinged beam intensitiy, when the bunches cross, there will be a maximum of about 20 collisions.

A larged magnetic field is needed to guide and maintain the beam particles in their circular orbit. The needed field is achieved using superconducting electromagnets built from NbTi coils that operate in a superconducting state, efficiently conducting electricity without resistance or loss of energy. The currents through the coils produce magnetic fiedls perpendicular to the direction of motion of the protons that deflect the protons into their orbits. The whole magnetic system comprises 1232 dipole magnets of 15 m length which are used to bend the beams, and 392 quadrupole magnets, each 57 m

long, to focus the beams. At a peak beam energy of 7 TeV, the dipoles need to produce an 8.33 T magnetic field, requiring a current of ~ 12 kA. In order to deliver the current densities and magnetic field required for 7 TeV proton beams, the magnets are kept at 1.9 K by circulating superfluid helium.

The first pp collisions produced by the LHC occurred on November 23 2009, at the SPS extraction energy of 450 GeV per beam. Very quickly after, on December 8, ATLAS and CMS detectors started recording data at energy of 2.36 TeV. By this time the LHC became the highest energy accelerator in the world. During this period, bunch intensities were limited by machine-protection considerations to 1.5×10^{10} protons.

In February 2010, the LHC was commissioned once more with 450 GeV beams, and a series of tests were performed to ensure that the magnet systems could operate safely at the currents necessary to control 3.5 TeV beams. This was followed by the very first collisions at 7 TeV center-of-mass energy on March 30. During the 2010 run the beam parameters were tuned (the beam widths squeezed and the number of protons per bunch and the number of bunches in each beam increased) in order to increase the beam intensity. In particular, as the intensity of the beams increased, the mean number of interactions per bunch crossing increased.

Finally, the data samples analysed in this thesis correspond to proton-proton collisions at $\sqrt{s} = 7$ TeV delivered by the LHC and recorded by ATLAS between May and November 2011, with the LHC running with 50 ns bunch spacing. Table ?? summarizes the basic beam parameters expected for design energy and luminosity and the beam parameters as of May 2011. The LHC performance steadily improved during 2011. The average number of interactions per bunch crossing throughout the data-taking period considered rapidly increased approximately from ~ 3 to 8 until (northern hemisphere)

summer 2011, with a global average for this period of ≈ 6 . Starting in August 2011 and lasting through the end of the proton run, this number ranged from approximately 5 to 17, with an average of about 12. This evolution is illustrated in Fig. 3.2, which shows the maximum mean number of collisions per beam crossing versus day in 2011.

Parameter	2011 runs	Design
Center-of-mass energy [TeV]	7	14
Instantaneous luminosity [$\text{cm}^{-2}\text{s}^{-1}$]	3.65×10^{33} (year peak)	10^{34}
Bunches per beam	38 (May)	2808
Protons per bunch	0.8×10^{11} (May)	1.5×10^{11}
Mean interactions per crossing	6 to 12 (year average)	23

Table 3.1: Summary of beam conditions during the 2011 7 TeV runs and those foreseen at design energy and luminosity.

3.1.1 Luminosity and pile-up

The rate of events produced by the colliding beams depends on the luminosity of the collisions, which is a measure of the number of events per second per unit cross section, typically measured in units cm^2s^{-1} . The number of events of a particular process, then, is given by the product of the integrated luminosity, $\int dt L$, and the cross section of the process, σ_{event} . The integrated luminosities are typically quoted in units of inverse picobarns, $\text{pb}^{-1} = 10^{-36}\text{cm}^2$. In order to measure processes with very little cross sections a very high luminosity is required.

The delivered luminosity can be written as [15]:

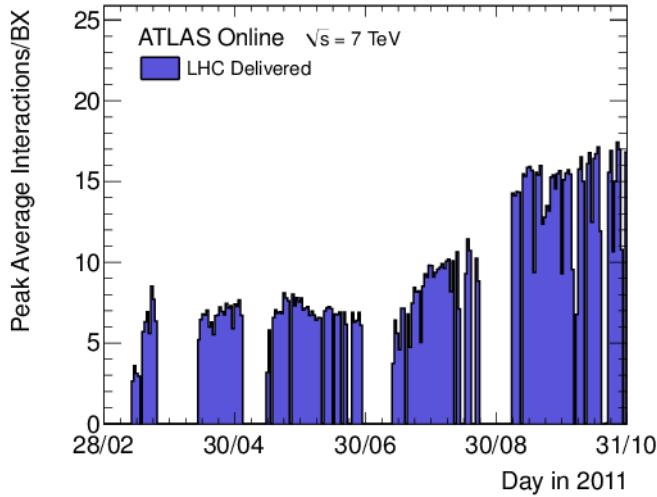


Figure 3.2: The maximum mean number of events per beam crossing versus day in 2011.

$$L = \frac{n_b f_r n_1 n_2}{2\pi \Sigma_x \Sigma_y} \quad (3.1)$$

where n_b is the number of colliding bunch pairs, n_1 and n_2 are the bunch populations (protons per bunch) in beam 1 and beam 2 respectively (together forming the bunch charged product), f_r is the machine revolution frequency, and Σ_x and Σ_y are the width and the height of the proton beams.

The number of protons per bunch, the number of bunches per beam, and the revolution frequency are all set by the beam operators. The widths of the proton beams are measured in a process known as a Van der Meer (*vdm*) scan [16]. In a *vdm* scan, the beams are separated by steps of a known distance. The collision rate is measured as a function of this separation, and the width of a gaussian fit to the distributions yields the width of the beams in the direction of the separation.

The total integrated luminosities provided by the LHC and recorded by

ATLAS in 2011 are shown in Figure 3.3. These events form the dataset analyzed in this thesis. By means of the beam-separation or vM scans, as well as other techniques to measure the bunch charged product, the ATLAS Collaboration has determined that the uncertainty on its luminosity measurement is $\delta L = \pm 3.7\%$. For a complete description of the methods used and the systematic errors evaluated see reference [15].

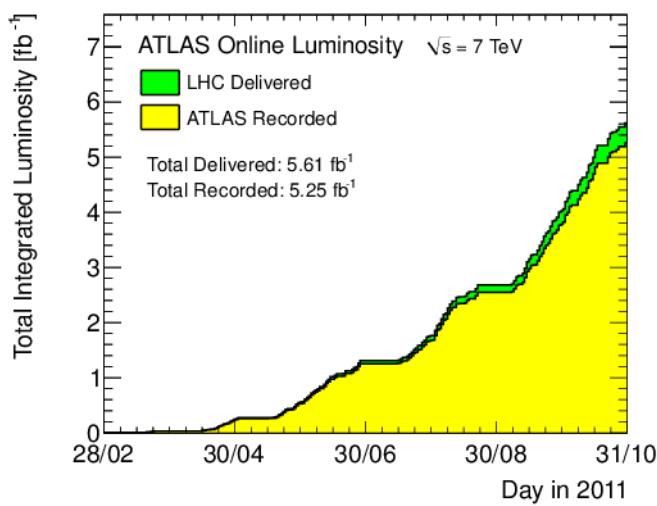


Figure 3.3: Total luminosity delivered by the LHC and recorded by ATLAS during the 2011 $\sqrt{s} = 7$ TeV proton-proton run

As anticipated, due to the cross-section for interaction and the number of protons per bunch, the possibility to observe multiple pp interactions per bunch crossing increases proportionally. This phenomenon, referred to as “pile-up”, can really occur in two distinct forms. The first form is the presence of multiple pp collisions (different from the interaction of interest) in the same bunch crossing, referred to as “in-time” pile-up. The second form of pile-up takes place due to electronic integration times within the detector. Certain detector components are actually sensitive to multiple bunch cross-

ings due to the long electronic signals generated in the response to energy depositions or charge collection. One or more pp collisions in a bunch-crossing different from that which produced the collision of interest can then affect the measurement. This form of pile-up is referred to as “out-of-time” pile-up and will become more and more important as the LHC bunch spacing gets closer to the nominal value, 25 ns.

The fraction of events with pile-up increased significantly since the data taking started. The experimental signature of this fact is obtain via the number of reconstructed primary vertices, or NPV. The effect of the event NPV is an important concern for the measurement of jet properties and will be discussed in the next chapters.

3.2 The ATLAS Detector

The ATLAS detector [17] is one of the two general purpose particle detectors built for probing pp collisions at the LHC. As it was described in the previous section, inside the LHC, bunches of up to 10^{11} protons will collide 40 million times per second to provide 14 TeV proton-proton collisions at a nominal luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. These high interaction rates and energies, as well as the requirements for high precision physics measurements set the standars for the design of the detector. At even 7 TeV center-of-mass energy, the LHC interactions result in high particle multiplicity, requiring fine detector granularity, and particle production at forward rapidity, requiring large detector angular coverage.

To achieve these performance goals, a design consisting of multiple detector sub-systems with cylindrical symmetry around the incoming beams is used as shown in Fig. 3.4. Closest to the interaction point the inner tracking

detector is placed, providing charged particle reconstruction. The magnet configuration comprises a thin superconducting solenoid surrounding the inner detector cavity, and three large superconducting toroids (one barrel and two end-caps) arranged with an eight-fold azimuthal symmetry around the calorimeters. This fundamental choice has driven the design and size (44 m in length and 25 m in height) of the rest of the detector. Outside the solenoid, a calorimeter system performs electron, photon, tau, and jet energy measurements. Finally, the calorimeter is surrounded by the muon spectrometer where an array of muon drift chambers perform muon identification and momentum measurements.

The ATLAS detector coordinate system is used to describe the position of particles as they traverse these subdetectors. It is a right-handed coordinate system, with z pointing along the beam direction, positive x pointing toward the center of the LHC ring, and positive y pointing up. The $x-y$ plane is referred to as the transverse plane, and the z direction as the longitudinal direction. The azimuthal angle ϕ is measured as usual around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as $\eta = \ln \tan(\frac{\theta}{2})$, regions of low η are referred to as “central”, and regions of high η are referred to as “forward”. The transverse momentum p_T is defined in the $x-y$ plane unless stated otherwise. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

To meet the extremely high demands that the LHC luminosity places on the speed with which ATLAS must record data, a dedicated trigger and data acquisition (TDAQ) system is used. The interaction rate at the design luminosity is approximately 1 GHz, while the event data recording, based on technology and resource limitations, is limited to about ~ 200 Hz. This requires a high rejection of minimum-bias processes while maintaining max-

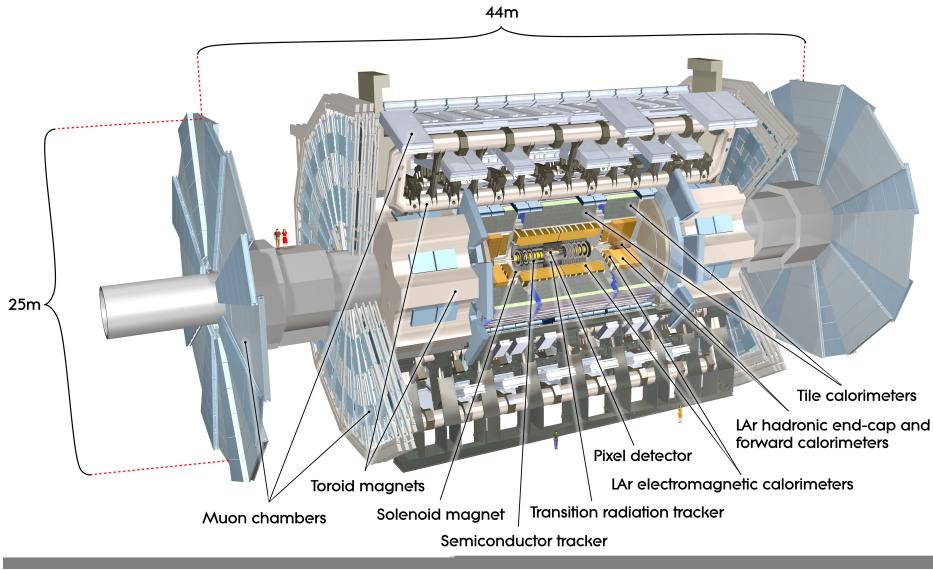


Figure 3.4: El detector de ATLAS

imum efficiency for the new physics. The Level-1 (L1) trigger system uses a subset of the total detector information to make a decision on whether or not to continue processing an event, reducing the data rate to approximately 75 kHz (limited by the bandwidth of the readout system, which is upgradeable to 100 kHz). The subsequent two levels, collectively known as the high-level trigger (HLT), are the Level-2 (L2) trigger and the event filter. They provide the reduction to a final data-taking rate of approximately 200 Hz.

3.2.1 Inner tracking system

The inner tracking system or inner detector (ID) is composed of three sub-detectors: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). The goal of these three is to provide charged

particle trajectory reconstruction and momentum measurements with an overall acceptance in pseudorapidity of $|\eta| < 2.5$ and full ϕ coverage.

The sensors which built this system register signals, referred to as “hits”, in response to the passage of charged particles. The ID is immersed in a 2 T magnetic field, generated by the central solenoid. The positions of the registered hits are combined to form tracks, with the radius of curvature of the tracks (caused by the presence of the magnetic field) providing a measurement of the particles transverse momentum. The track reconstruction efficiency ranges from 78% at $p_T^{track} = 500$ MeV to more than 85% above 10 GeV, averaged across the full η coverage [18]. The transverse momentum resolution of $\sigma_{p_T}/p_T = 0.05$ [19] (upper bound) and a transverse impact parameter resolution of $\sim 20 \mu m$ for high momentum resolution particles in the central η region[20].

The pixel detector, SCT, and TRT sensors are arranged on concentric cylinders around the beam axis, known as barrel layers, and on disks perpendicular to the beam at either end of the barrel, known as end-caps. A more complete description of these systems is given below. The overall layout of the inner detector is shown in Fig. 3.5.

The Pixel detector

The pixel detector consists of three concentric barrel layers. The innermost one, the so called “b-layer” due to its role in identifying b -quarks initiated jets, is located at 5 cm from the interaction region. Three additional disks are located at each end-cap, producing typically three pixel position measurements per charged particle track. Each layer or disk is instrumented with modules that form the basic unit of data acquisition, each with 47232 pixels. All pixel sensors are identical and have a minimum pixel size in r

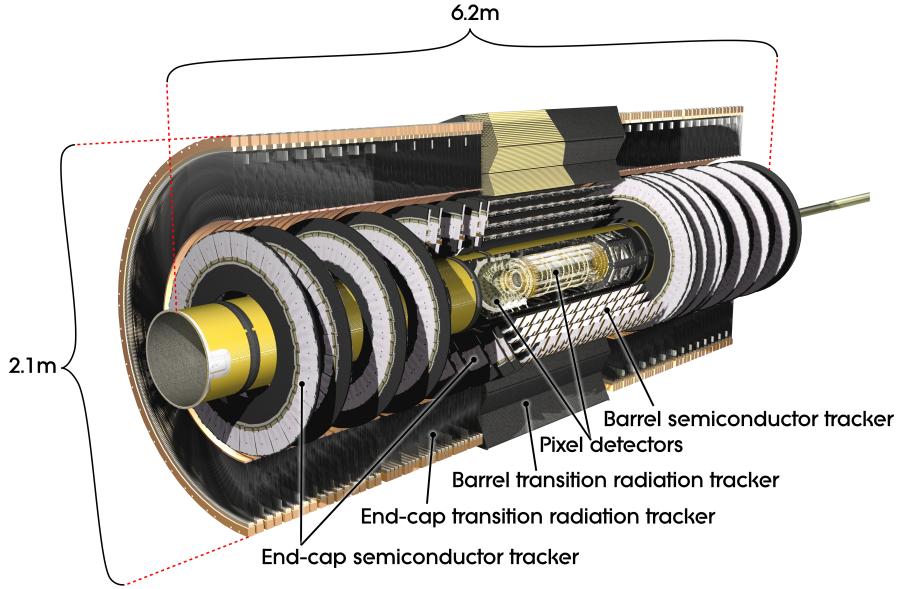


Figure 3.5: Layout of the ATLAS inner detector.

- $\phi \times z$ of $50 \times 400 \mu\text{m}^2$. The intrinsic accuracies in the barrel are $10 \mu\text{m}$ in $r - \phi$ and $115 \mu\text{m}$ along z , or along r in the end-caps. The pixel detector has approximately 80.4 million readout channels, an order of magnitude more readout channels than the rest of ATLAS combined, and it extends to a total length of $z \sim \pm 650 \text{ mm}$ and radius of $r \sim 150 \text{ mm}$, providing good reconstruction efficiency for tracks up to $|\eta| < 2.5$.

The SCT

The SCT consists of four barrel layers and nine end-cap layers surrounding the pixel detector, resulting in at least four hits along every charged particle track. The SCT barrel reaches to $z \sim \pm 750 \text{ mm}$ and $r \sim 515 \text{ mm}$, while the end-cap covers out to $z \sim \pm 2720 \text{ mm}$ and $r = 560 \text{ mm}$. There are 15,912 SCT module sensors, each 12.8 cm long and approximately $285 \mu\text{m}$ thick.

In the barrel region, these modules use small-angle (40 mrad) stereo strips to measure both coordinates, with one set of strips in each layer parallel to the beam direction, measuring the ϕ coordinate directly . In the end-cap region, the detectors have a set of strips running radially and a set of stereo strips at an angle of 40 mrad. The mean pitch of the strips is 80 μm . The intrinsic accuracies per module in the barrel are 17 μm in $r - \phi$ and 580 μm in z (or r in the end-caps). The total number of readout channels in the SCT is approximately 6.3 million. A hit is registered only if the pulse height in a channel exceeds a preset threshold (~ 1 fC). The charged measured in the strip is then recorded into a memory buffer that is only read out and used for tracking if a trigger is received signaling that the event should be considered in more detail.

The TRT

The TRT surrounds the silicon detectors and is comprised of up to 76 layers of longitudinal straw tubes in the barrel, extending to $z \sim \pm 710$ mm and $r \sim 1060$ mm, and 160 radial straw planes in each end-cap cylinders, reaching $z \sim \pm 2710$ mm and $r \sim 1000$ mm.

The TRT sensors are thin drift tubes consisting of cathode metal straws filled with an ionizing gas mixture of xenon, oxygen, and CO₂, with an anode wire running down the center of the straw. The passage of a charged particle through the gas produces positive ions and free electrons, which travel to the cathode and anode, respectively, under the influence of an applied voltage of 1600 V. Comparing the time that the signals are received at the cathode and the anode gives a drift time measurement that can be used to calculate the impact parameter of the particle. This method gives no information on the position along the length of the straw.

To give the best resolution of particle trajectories as they bend in the solenoidal field, the straws lie along the beam direction in the barrel and radially in the end-caps. The straw diameter of 4 mm causes a maximum drift time of approximately 48 ns and an intrinsic accuracy of 130 m along the radius of the straw.

In addition to directly detecting charged particles produced by the collision, the TRT also measures the transition radiation induced by the passage of these particles through polypropylene sheets placed between the drift tube straws. Transition radiation refers to the photons emitted by charged particles as they pass from one material into another with a different dielectric constant. These photons yield a much larger signal amplitude than the charged particles, so separate thresholds in the electronics can be used to distinguish the two.

One of the most important tasks of the inner detector is to provide accurate collision vertex identification, exploiting the excellent position resolution and tracking efficiency. Vertices are reconstructed by matching inner detector tracks with $p_T > 150$ MeV back to a common origin.

3.2.2 The Calorimeter System

The purpose of the ATLAS calorimeter system is to measure the energy of electrons, photons, taus and jets, within the pseudorapidity region of $|\eta| < 4.9$ and with full ϕ symmetry and coverage around the beam axis. It also provides fast position and energy measurements to serve as trigger signals for these objects as well as the missing transverse energy.

The calorimeter detector consist of electromagnetic (EM) calorimeter and hadronic calorimeter components. The EM calorimter provides fine granularity measurements of electrons and photons. Each calorimeter is segmented

both transverse to the particle direction, to give position information, and along the particle direction, to chart the development of the particle shower. This permits detailed mapping of EM and hadronic showers in the calorimeter, allowing for studies of the internal structure of hadronic jets and partially giving rise to the high resolution measurements of their energy.

The EM and hadronic calorimeters are sampling calorimeters meaning that they utilize alternating layers of absorber material, composed of heavy atoms that interact with energetic particles and cause them to lose energy, and an active material, that produce a signal in response to the deposited energy.

The calorimeters closest to the beam-line are housed in three cryostats, one barrel and two end-caps. The barrel cryostat contains the electromagnetic barrel calorimeter, and the two end-cap cryostats each contain an electromagnetic end-cap calorimeter (EMEC), a hadronic end-cap calorimeter (HEC), located behind the EMEC, and a forward calorimeter (FCal) to cover the region closest to the beam. These calorimeters use liquid argon as the active detector medium and need to be maintained at a constant temperature of $\sim 88\text{K}$. Liquid argon (LAr) has been chosen for its intrinsic linear behaviour (production of ionization charge as a function of incident charge), its stability of response over time and its intrinsic radiation-hardness.

An illustration of all these components can be found in Fig. 3.6. Further specifications are given in the next sections.

Liquid argon EM calorimeter

The EM calorimeter uses lead as the absorber and liquid Argon as the active material. A photon traversing the absorber will interact with the heavy nucleus via Compton scattering or the photo-electric effect, producing low-

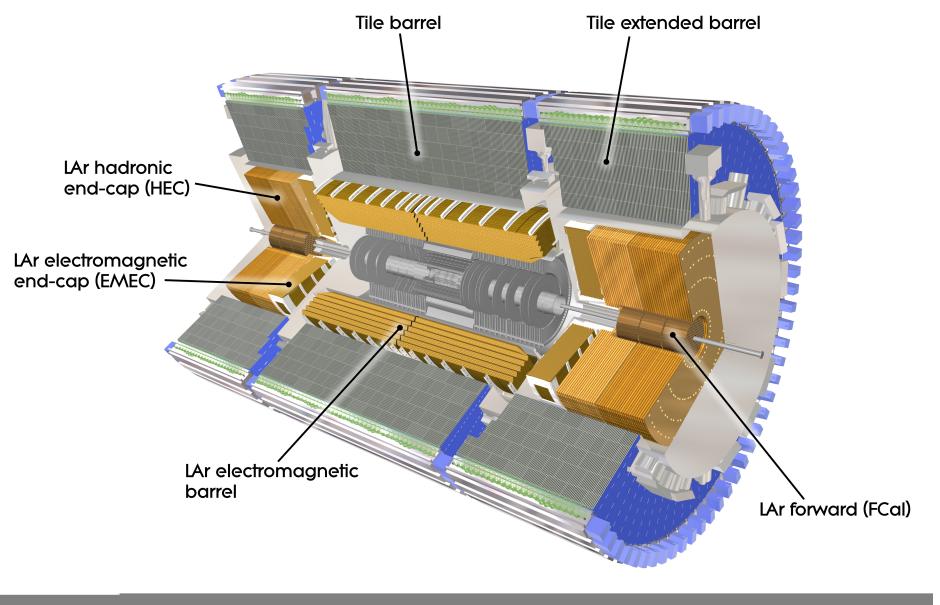


Figure 3.6: Layout of the ATLAS electromagnetic and hadronic calorimeter systems. The total length is ~ 12 m, extending to a maximum radius of 4.25 m.

energy electrons, or pair production, producing electron/positron pairs. An electron or positron, in turn, can produce bremsstrahlung photons as it is deflected by the nuclei or produce more charged particles via ionization. Thus each incident photon, electron, or positron produces a shower of photons, electrons, and positrons that lose their energy through successive interactions in the absorber. The produced particles ionize the liquid argon, and the charge is collected by electrodes located in the liquid argon gap. These electrodes consist of three layers of copper sheets, the outer two kept at high-voltage potential and the inner one used to readout the signal.

To provide full coverage in ϕ without any cracks, an accordion-shaped absorber and electrode geometry is used, shown in Fig. 3.7. This design was chosen to ensure high azimuthal uniformity, a regular liquid argon ionization gap, and a constant sampling fraction within a given detector region. The figure highlights how this geometry is divided among rectangular cells in $\eta \times \phi$ space, the individual readout elements of varying size, finely segmented both laterally and longitudinally. Such fine segmentation $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ in the second layer of the EM barrel, for example permits a detailed mapping of the electromagnetic and hadronic showers.

The position resolution of the EM is driven by the readout geometry (rectangular cells). There are three layers of cells, segmented along the particle’s direction of motion. The ϕ segmentation comes from grouping the accordion-shaped electrodes together into a common read out channel.

In the region $0 < |\eta| < 1.8$ the electromagnetic calorimeters are complemented by a “presampler” detector, an instrumented argon layer, which provides a measurement of the energy lost in the solenoid and the outer wall of the barrel cryostat.

The EMEC uses the same accordion geometry as the EMB, whereas the

granularity is typically slightly larger than in the barrel.

The signal readout chain for the LAr calorimeter (indeed for all calorimeter systems) is divided into a fast analog readout for the trigger system and a slower digital readout used for more redefined trigger decisions and the offline reconstruction. However, regardless of the readout path, the signal is initiated within the active LAr medium. To minimize noise and increase speed the first level of readout is located on the detector (both for LAr and Tile calorimeter, see 3.2.2). The front-end electronics amplify and shape the signal. Shaping electronics induce a bipolar pulse shape in the ionization signal. This shape is characterized by having both a positive and a negative component, which renders the integral of the signal exactly equal to zero.

The performance of the shaping electronics is critical for a correct energy calibration of the detector since the energy is primarily determined from the peak height of the pulse. In each calorimeter region, the overall pulse shape and duration are optimized to approximately cancel a constant injection of energy into the detector. The motivation for this approach is to effectively redefine the baseline of the energy measurement. In the high luminosity environment of the LHC, this reduces the sensitivity to the background from multiple pp interactions on average.

To translate these analog signals to digital signals that can be transmitted long distances to the next stage of the readout system, the pulse shape is measured over several 25 ns (nominal) time intervals, known as samples. The challenge of calorimeter calibration is to map these measured signals to the energy deposited in the active detector medium, known as the visible energy. This calibration is established using test-beam measurements of electrons in the EMB (REFERENCES) and EMEC calorimeters (REFERENCES).

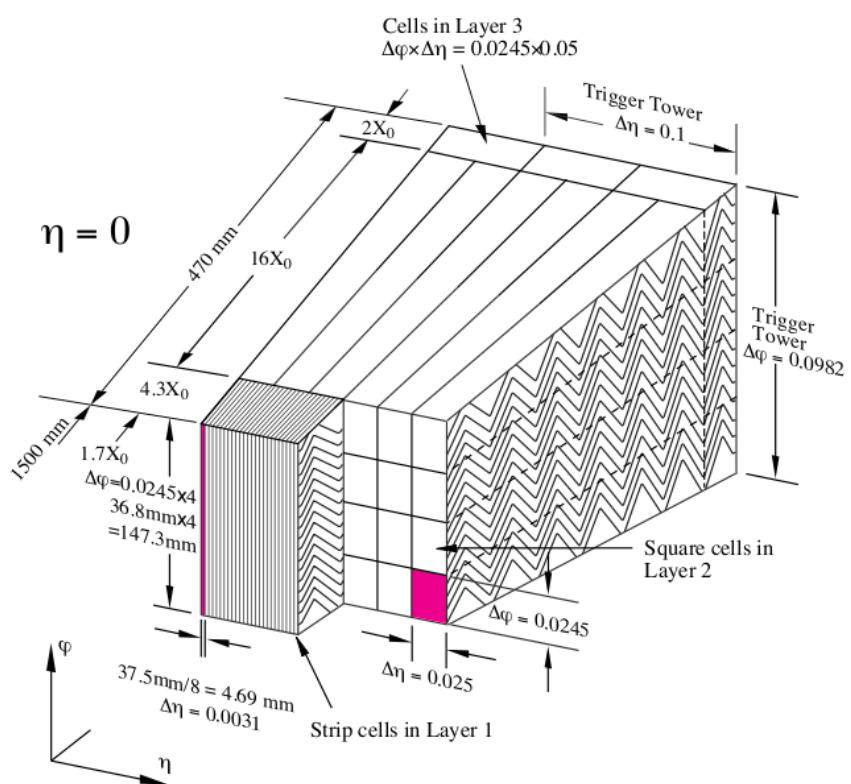


Figure 3.7: Cross section of the LAr barrel calorimeter where the different layers are visible. The granularity in η and ϕ of the cells of each of the three layers is also shown.

The hadronic calorimeter

Outside the EM calorimeter lies the system of hadronic calorimeters. The barrel portion, known as the Tile calorimeter, uses iron absorber slabs interspersed with scintillating tiles. The Tile calorimeter is most notable for its depth of 7.4 radiation lengths (λ^1). The hadronic end-cap and the forward calorimeter, which need to absorb the more energetic particles that are produced at large $|\eta|$, are made of copper and tungsten absorbers, respectively, with liquid argon as the active material.

The tile calorimeter is composed of 3 mm thick scintillating tiles, arranged to lie parallel to the incoming particle direction, interleaved with 14 mm thick iron plates. It is divided into the barrel calorimeter, covering $|\eta| < 1.0$, and two extended barrel calorimeters, covering $0.8 < |\eta| < 1.7$. Each tile is read out by two wavelength-shifting fibers, which convert the scintillator signal to visible light. The readout fibers of several tiles are grouped to a single photomultiplier tube forming cells in $\eta \times \phi$ space. As in the EM calorimeter, these cells are segmented into three layers, the first two of size $\Delta\eta = 0.1$ and $\Delta\phi = 0.1$ and the last of size $\Delta\eta = 0.2$ and $\Delta\phi = 0.1$. Towers to provide information to the trigger systems are formed from 0.1×0.1 grouping of all three layers.

The HEC uses the LAr active readout design due to the higher radiation tolerance required for the forward regions. Although housed in the same cryostat as the accordion geometry EMEC, the HEC implements a flat-plate design.

The forward calorimeter extends to cover the region $3.1 < |\eta| < 4.9$. Since

¹To quantify the amount of material needed to capture a particle's energy, the unit of an interaction length, which is the distance over which a high energy charged particle loses $1 - \frac{1}{e} \sim 63\%$ of its energy, is commonly used.

it is the only calorimeter that covers this very forward region, it must provide both electromagnetic and hadronic measurements. In addition, the high particle fluxes in this region necessitate a finely granulated design. The FCal is approximately 10 interaction lengths deep, and consists of three modules in each end-cap: the first, made of copper, is optimised for electromagnetic measurements, while the other two, made of tungsten, measure predominantly the energy of hadronic interactions.

The hadronic calorimeters are calibrated using muons in test-beam experiments (REFERENCIAS) and those muons produced by cosmic-rays in situ (REFERENCIA). The invariant mass of the Z boson in $Z \rightarrow ee$ events measured in-situ in the 2010 pp collisions is then used to adjust the calibration derived from test-beams and cosmic-muons.

3.2.3 The Muon System

The muon system gives the ATLAS detector its overall shape and imposing nature, as depicted in Fig. 3.8. Muons have much smaller cross section to interact in material than electrons and hadrons, for this, they do not deposit all their energy in the calorimeters. The muon spectrometer is designed to detect muons within $|\eta| < 2.7$. Because many new physics signatures involve high-momentum muons, the system is also required to provide trigger signals based on the particle p_T for $|\eta| < 2.4$.

To provide a momentum measurement, the muons trajectories are bent in a toroidal magnetic field. This field is provided by one large barrel toroid and two large end-cap toroids, each toroid consisting of eight coils arranged symmetrically around the beam axis. The toroid system produces a magnetic field that is typically oriented in the ϕ direction and that is measured with over 1800 Hall sensors placed through the magnets. Under the influence of

this field, muons are deflected in the $r-z$ plane and the transverse momentum of the muons is given then by the radius of curvature of the tracks. Since the highly-energetic muons bend very little even in this high magnetic field, the muon system is the largest of all the ATLAS sub-detectors, covering a radius from ~ 4.5 m to ~ 12.5 m.

Four primary subsystems comprise the integrated muon spectrometer: monitored drift tubes (MDT), cathode strip chambers (CSC, which are multiwired proportional chambers with cathodes segmented into strips), resistive plate chambers (RPC) and thin gap chambers (TGC). The MDT and CSC subsystems are primarily designed for precision measurements of muon tracks, with the MDT system providing coverage for the more central region ($|\eta| \leq 2.7$, with full coverage only in $|\eta| < 2.0$), whereas the CSC is located in the more forward region ($2.0 < |\eta| < 2.7$) due to its ability to cope with higher background rates. The RPC and TGC muon subsystems are designed to provide fast, robust readout for use in the trigger and data acquisition system.

A description of the subsystems can be found elsewhere (REFERENCES!!!).

3.2.4 Forward detectors

Three smaller detector systems cover the ATLAS forward region. The main function of the first two systems is to determine the luminosity delivered to ATLAS. At ± 17 m from the interaction point lies LUCID (LUminosity measurement using Cerenkov Integrating Detector). The principle of LUCID is to detect inelastic $p - p$ scattering in the forward region, exploiting the fact that the number of particles detected is proportional to the total, both primary and pileup, interactions in a bunch-crossing. LUCID thus provides a

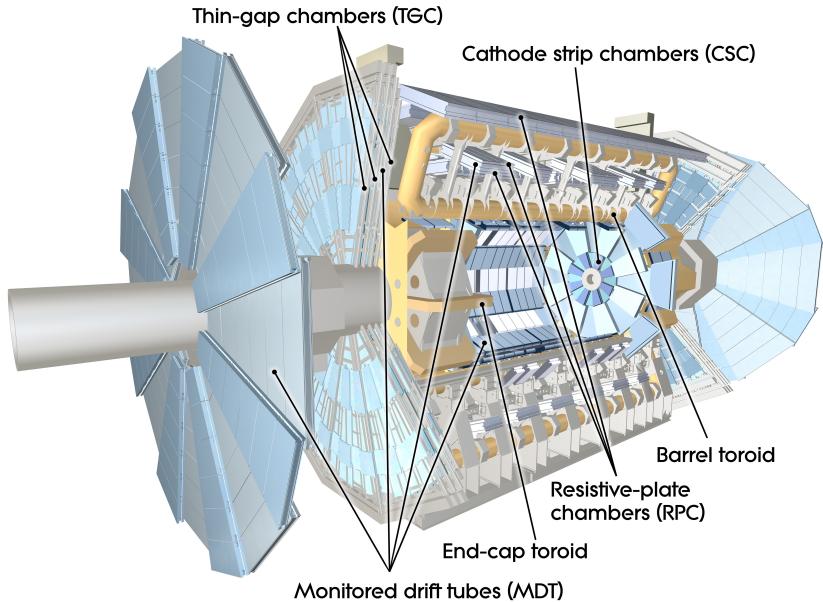


Figure 3.8: Muon Chamber

relative luminosity measurement, in which the detected number of particles must be translated to the total number of proton-proton interactions via calibration runs. The second detector is ALFA (Absolute Luminosity For ATLAS). Located at ± 240 m, it consists of scintillating fibre trackers located inside Roman pots which are designed to approach as close as 1 mm to the beam. The third system is Zero-Degree Calorimeter (ZDC), which place a role in determining the centrality of heavy-collisions.

3.2.5 Trigger and Data Adquisition

At design luminoisty, the LHC will deliver approximately 40 million collision events every second. With an average ATLAS event size of ~ 1.5 MB, this is far more information than can be saved into the finite data storage resources available. The goal of the trigger system is to move interesting physics events

to permanent storage, while rejecting the vast majority of other events.

The online selection is done in three stages: the Level 1 (L1), Level 2 (L2) and Event Filter (EF) stages. Each trigger level refines the decisions made at the previous level and, where necessary, applies additional selection criteria. The data acquisition (DAQ) system receives and buffers the event data from the detector-specific readout electronics, at the L1 trigger accept rate, over 1600 point-to-point readout links. The L1 trigger uses a limited amount of the total detector information (only from the calorimeter and the muon systems) using only simple hardware based algorithms to make a decision in less than $2.5 \mu\text{s}$, reducing the rate to about 75 kHz. The L2 and EF, collectively referred to as the High Level Trigger (HLT), are based on fast software algorithms running on large farms of commercial processors. The L2 is the first stage of the ATLAS DAQ system that has access to data from the ID and is capable of doing partial reconstruction of events up to the L1 accept rate. L2 trigger is designed to reduce the rate to approximately 3.5 kHz, with an event processing time of about 40 ms, averaged over all events. The EF reduces the rate to roughly 200 Hz. Its selections are implemented using offline analysis procedures within an average event processing time of the order of 4 s.

The L1 trigger is designed to accept high- p_T muons, electrons, photons, jets, and taus, as well as events with large missing transverse energy or sum energy. It uses signals from the TGCs and RPCs from muon triggers and reduced granularity calorimeter information for electron, photon, jet, tau and total energy triggers. The calorimeter trigger system, which maintains a fast readout independent from the remainder of the calorimeter is known as the Level-1 Calorimeter. At this level coarse calorimeter information is available in the form of jet elements with $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $|\eta| < 3.2$. Jets are

reconstructed using a square sliding window algorithm. In addition to coarse jets, the total transverse energy is also measured at the L1. The region of the detector corresponding to the location where the L1 thresholds were passed – so called “region of interest” (RoI) – are then delivered to the L2 software algorithms.

The L2 trigger applies additional energy thresholds and multiplicity requirements using the RoI around triggered L1 objects. For example, the L2 jet trigger retrieves the data from cells surrounding the L1 RoI and constructs jets using a simplified cone jet algorithm.

The next step and last stage in the trigger chain is the EF, which receives events that have been selected by the L2 triggers and processes the entire event with the full detector granularity instead of only a restricted region.

The monitoring infrastructure of the HLT supports the real-time accumulation of histograms, and their aggregation across the farm, so that parameters can be extracted from cumulative distributions that contain events from all processor nodes. Beam parameters determined from those live histograms are transmitted online to the LHC and are also available to feed back into the HLT itself for use by its own trigger algorithms that depend on the precise knowledge of the luminous region (such as b -jet tagging).

3.2.6 ATLAS Performance and Data quality

The ATLAS detector has been operational for a number of years collecting large amounts of data. Before the start-up of the LHC, measuring muons from cosmic rays, which were used to test, understand and align the detector. In 2010 and 2011 ATLAS recorded over 5.2fb^{-1} of collision data. Fig. 3.3 presented the luminosity delivered by the LHC in 2011 as well as the recorded luminosity by the detector, showing a good performance of the

ATLAS Experiment.

The fraction of time that each subdetector system was operational during data-taking is shown in Table 3.2.

The Data Quality (DQ) selection within ATLAS is based on the inspection of a standard set of distributions that lead to a data quality assessment which is encoded in so-called DQ flags. DQ flags are issued for each detector, usually segmented in subdetectors like barrel, end-caps and forward. DQ flags are also issued for trigger slices and for each physics object reconstruction. In this way, the state of the ATLAS detector from hardware to physics object reconstruction is expressed through DQ flags, which are saved per luminosity block. A luminosity block is a time interval of typically two minutes.

The DQ information is used in analyses through dedicated lists of good runs/luminosity blocks. Good run lists are formed by DQ selection criteria in addition to other criteria, such as run range, magnetic field configuration and beam energy. A complete list of valid physics runs and luminosity blocks is used in each analysis.

Detector component	operational
Inner Detector	
Pixel	≈96.4%
SCT	≈99.2%
TRT	≈97.5%
Calorimeter	
EM	≈99.8%
Tile	≈96.2%
Hadronic, end-cap	≈99.6%
Forward calorimeter	≈99.8%
Muon Spectrometer	
MDT	≈99.7%
CSC	≈97.7%
RPC	≈97.0%
TGC	≈97.9%

Table 3.2: The approximate fraction of time that each individual subdetector system was operational during data-taking.

Chapter 4

Jet reconstruction and *b*-Tagging

Experimental data and simulated events were both reconstructed using the latest version available of the ATLAS software. In this chapter we briefly describe the reconstruction of the two key objects used in this analysis, namely jets and tracks. We describe how *b*-Tagging is done in ATLAS.

4.1 Jets reconstruction and calibration

Jets are reconstructed using the anti- k_t jet algorithm [21] with a distance parameter $R = 0.4$, using calorimeter topological clusters [22] as input. Several quality criteria are applied to eliminate “fake” jets that are caused by noise bursts in the calorimeters and energy depositions belonging to a previous bunch crossing [23].

The jet energies are corrected for inhomogeneities and for the non-compensating nature of the calorimeter by using p_T – and η -dependent calibration factors determined from Monte Carlo simulation [24]. This calibration is referred to

as the EM+JES scale. Using test beam results, in-situ track and calorimeter measurements, estimations of pile-up energy depositions, and detailed Monte Carlo comparisons, an uncertainty on the absolute jet energy scale was established. This uncertainty is smaller than $\pm 10\%$ for $\eta < 2.8$ and $p_T > 20$ GeV. More sophisticated techniques undergoing commissioning, such as local cluster weighting, are expected to considerably improve the jet energy uncertainty and resolution [25].

4.2 *b*-jet Tagging

Jets are classified as *b*-quark candidates by the ATLAS MV1 *b*-tagging algorithm, based on a neural network that combines the information from three high-performance taggers: IP3D, SV1 and JetFitter [3]. These three tagging algorithms use a likelihood ratio technique in which input variables are compared to smoothed normalized distributions for both the *b*- and background (light- or in some cases *c*-jet) hypotheses, obtained from Monte Carlo simulation. The IP3D tagger takes advantage of the signed transverse and longitudinal impact parameter significances. The SV1 tagger reconstructs an inclusive vertex formed by the decay products of the *b*-hadron and relies on the invariant mass of all tracks associated to the vertex, the ratio of the sum of the energies of the tracks in the vertex to the sum of the energies of all tracks in the jet and the number of two-track vertices. The JetFitter tagger exploits the topology of the primary, *b*- and *c*-vertices and combines vertex variables with the flight length significance. The *b*-tagging performance is determined using a simulated $t\bar{t}$ sample and is calibrated using experimental data with jets containing muons and with a sample of $t\bar{t}$ events [26].

Bibliography

- [1] ATLAS Collaboration. Performance of Impact Parameter-Based b -tagging Algorithms with the ATLAS Detector using Proton-Proton Collisions at $\sqrt{s} = 7$ TeV. *ATLAS-CONF-2010-091*, 2010.
- [2] ATLAS Collaboration. Performance of the ATLAS Secondary Vertex b -tagging Algorithm in 7 TeV Collision Data. *ATLAS-CONF-2010-042*, 2010.
- [3] ATLAS Collaboration. Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data. *ATLAS-CONF-2011-102*, 2011.
- [4] CDF Collaboration. Measurements of Bottom Anti-Bottom Azimuthal Production Correlations in Proton-Antiproton Collisions at $s^{**}(1/2)=1.8$ TeV. *Phys.Rev.D*, 71:38, 2005.
- [5] S. Frixione and M.L. Mangano. Heavy quark jets in hadronic collisions. *Nucl.Phys.*, B483:321–338, 1997.
- [6] Andrea Banfi, Gavin Salam, and Giulia Zanderighi. Accurate qcd predictions for heavy-quark jets at the tevatron and lhc. *JHEP*, 0707:026, 2007.

- [7] S. Catani, Y.L. Dokshitzer, H. Seymour, and B.R. Webber. Longitudinally invariant $K(t)$ clustering algorithms for hadron hadron collisions. *Nucl. Phys.*, B406:187, 1993.
- [8] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 0101:010, 2001.
- [9] John M. Campbell, R.Keith Ellis, F. Maltoni, and S. Willenbrock. Production of a W boson and two jets with one b^- quark tag. *Phys.Rev.*, D75:054015, 2007.
- [10] ATLAS Collaboration. Search for supersymmetry in pp collisions at $\sqrt{s} = 7\text{TeV}$ in final states with missing transverse momentum, b -jets and no leptons with the ATLAS detector. *ATLAS-CONF-2011-098*, 2011.
- [11] A. Abdesselam et al. Boosted objects: a probe of beyond the standard model physics. *The European Physical Journal C - Particles and Fields*, 71:1–19, 2011.
- [12] A. Altheimer et al. Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks. 2012.
- [13] Jonathan M. Butterworth, Adam R. Davison, Mathieu Rubin, and Gavin P. Salam. Jet substructure as a new Higgs search channel at the LHC. *Phys.Rev.Lett.*, 100:242001, 2008.
- [14] Amos Breskin and Rdiger Voss. *The CERN Large Hadron Collider: Accelerator and Experiments*. CERN, Geneva, 2009.

- [15] ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011. *ATLAS-CONF-2011-116*, Aug 2011.
- [16] S van der Meer. Calibration of the effective beam height in the ISR. *CERN-ISR-PO-68-31. ISR-PO-68-31*, 1968.
- [17] Aad, G. and others. The ATLAS Experiment at the CERN Large Hadron Collider. CERN, Geneva, 2008.
- [18] G. Aad et al. Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC. *New J.Phys.*, 13:053033, 2011.
- [19] ATLAS Collaboration. Estimating Track Momentum Resolution in Minimum Bias Events using Simulation and K_s in $\sqrt{s} = 900$ GeV collision data. *ATLAS-CONF-2010-009*, 2010.
- [20] ATLAS Collaboration. Tracking Studies for b-tagging with 7 TeV Collision Data with the ATLAS Detector. *ATLAS-CONF-2010-070*, 2010.
- [21] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008.
- [22] W Lampl et al. Calorimeter Clustering Algorithms: Description and Performance. (ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003), Apr 2008.
- [23] ATLAS Collaboration. Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data. *ATLAS-CONF-2012-020*, 2012.

- [24] ATLAS Collaboration. Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV and measured with the ATLAS detector. *ATLAS-CONF-2010-056*, 2010.
- [25] G. Aad et al. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. 2009.
- [26] ATLAS Collaboration. Calibrating the b-Tag Efficiency and Mistag Rate in 35 pb^{-1} of Data with the ATLAS Detector. *ATLAS-CONF-2011-089*, 2011.