Subjet Multiplicities and α_s Determination with the ATLAS Experiment

Joel Beaudry

Master of Science

Department of Physics

McGill University

Montreal, Quebec

2012-12-11

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science

DEDICATION

To my family for their constant love and support.

ACKNOWLEDGEMENTS

First and foremost all those that helped me throughout my studies, in particular, my supervisor, Dr. François Corriveau and all members of the McGill ATLAS group. Special mention needs to be made for Robert Keyes, Phil Lagogiannis, Olivier Shelbaya, and Eliot Hijano for the many interesting discussions and for making my stay at McGill that much more enjoyable. To my office mates and friends, Kuhan Wang, Micheal Stoebe, Bertrand Chapleau, Philippe Giguere, and Grant Salton, for making my office a pleasure to work in. Finally a special thanks to my twin, Justin, whose ability to put up with me at times is nothing beyond amazing.

ABSTRACT

At sufficiently high energies and jet resolution, the fragmentation effects of Quantum Chromodynamics (QCD) become small and the number of jet-like substructures within a jet, known as the subjet multiplicity, can be parametrised by means of perturbative QCD (pQCD). This thesis investigates such a relationship by measuring the mean subjet multiplicity $\langle n_{sbj} \rangle$ as a function of the jet resolution scale y_{cut} with the ATLAS experiment. Jets were reconstructed using the anti-k_T jet clustering algorithm and subjets were resolved by application of the k_T algorithm at a smaller resolution scale y_{cut}. The value of the strong coupling constant α_s, extracted from $\langle n_{sbj} \rangle$ with a resolution scale from $y_{cut} = 5 \cdot 10^{-3}$ to 0.1 with jets with transverse momentum $p_{T,jet}$ greater than 5 GeV and pseudorapidity η from -2.8 to 2.8, is $\alpha_s = 0.148 \pm 0.022$ (stat.)$^{+0.005}_{-0.008}$ (syst.) $\pm$ 0.005 (theory). The subjet multiplicity behaviour and subjet reconstruction were also studied as a function of $p_{T,jet}$.

ABRÉGÉ

Lorsque l'énergie et la résolution des jets sont suffisamment élevées, les effets de fragmentation de la Chronodynamique Quantique (QCD) deviennent négligeables. Dans ce régime, il devient possible de paramétriser la multiplicité des sous-jets, soit le nombre de structures ressemblant à un jet à l'intérieur d'un jet, en utilisant les méthodes perturbatives de QCD (pQCD). Cette thèse étudie ces relations en mesurant la multiplicité moyenne des sous-jets $\langle n_{sbj} \rangle$ en fonction de la résolution des jets y_{cut}, avec ATLAS. Les jets sont reconstruits avec l'algorithme anti-k_T, et la résolution des sous-jets est faite par l'application de l'algorithme k_T à une échelle de résolution y_{cut}. La valeur de la constante de l'interaction forte, α_s, extraite à partir de $\langle n_{sbj} \rangle$ avec une résolution de y_{cut} allant de $5 \cdot 10^{-3}$ à 0.1 et avec des jets dont le quantité de mouvement transversale $p_{T,jet}$ est plus grand que 5 GeV et une pseudorapidité η entre -2.8 à +2.8, est $\alpha_s = 0.148 \pm 0.022$ (stat.)$^{+0.005}_{-0.008}$ (syst.) $\pm$ 0.005 (théorie). La comportement de la multiplicité des sous-jets et leur reconstruction ont aussi été étudiés en fonction de $p_{T,jet}$.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction

The ATLAS (A Toroidal LHC ApparatuS) experiment is an international scientific collaboration at the particle physics laboratory, CERN, located near Geneva, Switzerland. The goal of ATLAS is to investigate proton-proton collisions with the use of the ATLAS detector at the Large Hadron Collider (LHC). The experiment that will be probing at a new energy frontier, currently at a center of mass energy of 8 TeV, intends to discover new physics and validate the existing theories about matter.

One of the most successful theories in physics is the Standard Model (SM), a collection of theoretical models that provide an accurate description of the current state of particle physics. The SM is a gauge quantum field theory that describes 3 of the 4 fundamental forces of physics: the electromagnetic, strong, and weak forces. Gravity is not included [1]. The theory describes matter and its interactions by elementary particles, the fermions and the bosons. The SM bosons, along with the forces they mediate, are listed in Table 1-1 [2].

Interactions in the SM occur via the exchange of integer spin bosons. The photon, both massless and neutrally charged, is responsible for mediating the electromagnetic force. The electromagnetic force is described by quantum electrodynamics (QED), a $U(1)$ theory. The weak force mediated by the massive W and Z bosons is described by the SU(2) weak theory. The electromagnetic and weak force can be

Table 1–1: Standard Model gauge bosons and their current empirically measured properties. In lieu of experimental measurements for the gluon, theoretical values are used for its mass and charge.

Boson	Force Mediated	Mass $[GeV/c^2]$	Charge $[e]$
Photon (γ)	Electromagnetism	$< 10^{-27}$	$< 10^{-35}$
$W^{\pm}$	Weak	80.4	± 1
Z^0	Weak	91.2	0
Gluon (g)	Strong	0	0

unified to give the electroweak interaction which is described by the $SU(2) \times U(1)$ symmetry group. The strong force is mediated by the massless gluons which carry a strong charge analogous to electric charge, called "colour". As described by the $SU(3)$ theory, quantum chromodynamics (QCD), there are three possible charges labelled red, blue, or green.

The forces of the nature can be further characterized by their coupling constant, a dimensionless value that represents the "strength" of interactions between the fermions and bosons. A list of the physical forces and their approximate relative coupling constants is given in Table 1-2.

Table 1–2: Forces and the relative strength of their coupling constants.

Force	Coupling constant
Strong	$\alpha_s \approx 1$
Weak	$\alpha_W \approx 10^{-6}$
Electromagnetic	$\alpha_{EM} \approx 1/137$
Gravity	$\alpha_g \approx 10^{-39}$

The other fundamental particles in the SM are the building blocks of matter, the half-integer spin fermions, which can be further divided into two sub-categories: leptons and quarks.

Leptons are particles that interact through the electromagnetic and weak force, whereas quarks are particles that interact through all the SM forces (electromagnetic, weak, and strong). The leptons and quarks can be organized into three generations, identical to one another aside from their masses that increase in each generation, as shown in Table 1-2 [2].

Table 1–3: Standard Model leptons and quarks and their empirically measured properties.

	Mass $[GeV/c^2]$	Charge $[e]$
Leptons		
electron (e)	0.000511	-1
electron neutrino (ν_e)	$< 2 \times 10^{-9}$	0
muon (μ)	0.0106	-1
muon neutrino (ν_μ)	$< 1.9 \times 10^{-4}$	0
tau (τ)	1.777	-1
tau neutrino (ν_τ)	$< 18.2 \times 10^{-3}$	0
Quarks		
up (u)	1.5 to 3.3 $\times 10^{-3}$	$+2/3$
down (d)	3.5 to 6.0 $\times 10^{-3}$	$-1/3$
charm (c)	1.27	$+2/3$
strange (s)	70 to 130	$-1/3$
top (t)	171.3	$+2/3$
bottom (u)	4.2	$-1/3$

A consequence of colour is that the strong force exhibits two properties different from the other SM forces: asymptotic freedom and confinement. As two coloured particles are brought close together, the force is lessened such that at high energies, the quarks act as free particles. This is known as asymptotic freedom. However, as the quarks are pulled apart, the force between them increases. This is called colour confinement and it prevents coloured free particles. It is by the latter process by which the quark bound states (hadrons) are always colourless [1].

The SM formalism is written for massless particles and requires the Higgs mechanism of spontaneous symmetry breaking for generating the non-zero boson and fermion masses [3]. The symmetry breaking introduces a new field that also leads to a new massive boson, the Higgs boson. The Higgs boson is the only SM elementary particle yet to be verified experimentally. A recent discovery of a boson at a mass of 126 GeV may prove to be the Higgs but has yet to be validated [4]. To conclude, the SM also allows for antiparticles, which are nearly identical to their particle counterparts except for having opposite quantum numbers. However, it is not required that every particle is distinguished from its antiparticle. For example, the photon is its own anti-particle. The complete picture of all the elementary particle of the SM is shown in Figure 1-1 [7].

Although the SM has been very successful, there are many areas of physics that cannot be explained by concepts currently in the SM. Some of the issues not addressed by the SM include the hierarchy problem, strong CP violation, neutrino oscillations, antimatter-matter asymmetry, the nature of dark matter and dark energy, and the incompatibility of the SM to general relativity and gravity. Physics not explained in the Standard Model is sometimes labelled as "new physics" or "beyond the standard model" (BSM). Aside for looking for possibilities of new physics at ATLAS, there is also a need to verify and refine the already established ideas present in the SM [5].

1.1 Motivation

The number of quarks and gluons produced at the LHC provide a detailed testing ground for QCD. At high energies and not too fine a resolution scale, QCD becomes a perturbative theory (pQCD) [6]. This is indicated in the value of the strong coupling

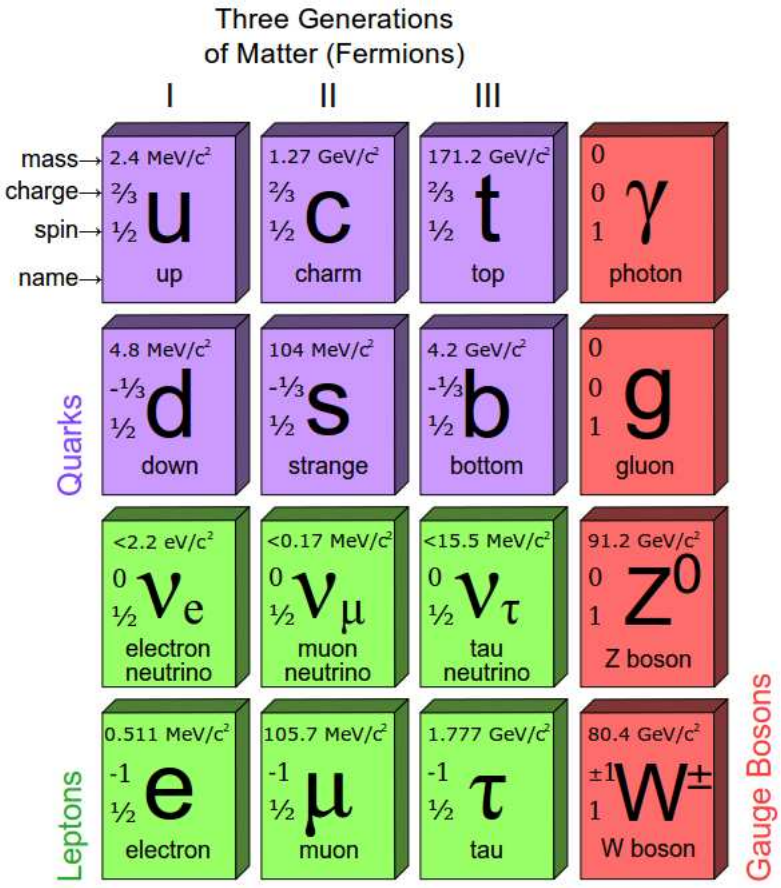

Figure 1–1: The elementary particles of the Standard Model [7].

constant, α_s. The value for this constant is energy dependent: At high energy the value drops below 1, at lower energies (below the GeV scale) the value is of order 1. It is when this value is less than one that the theory becomes perturbative as each higher order term in α_s is smaller than the previous one, allowing for the theory to not diverge. The running (energy dependence) nature of this coupling constant is depicted in Figure 1-2 [8].

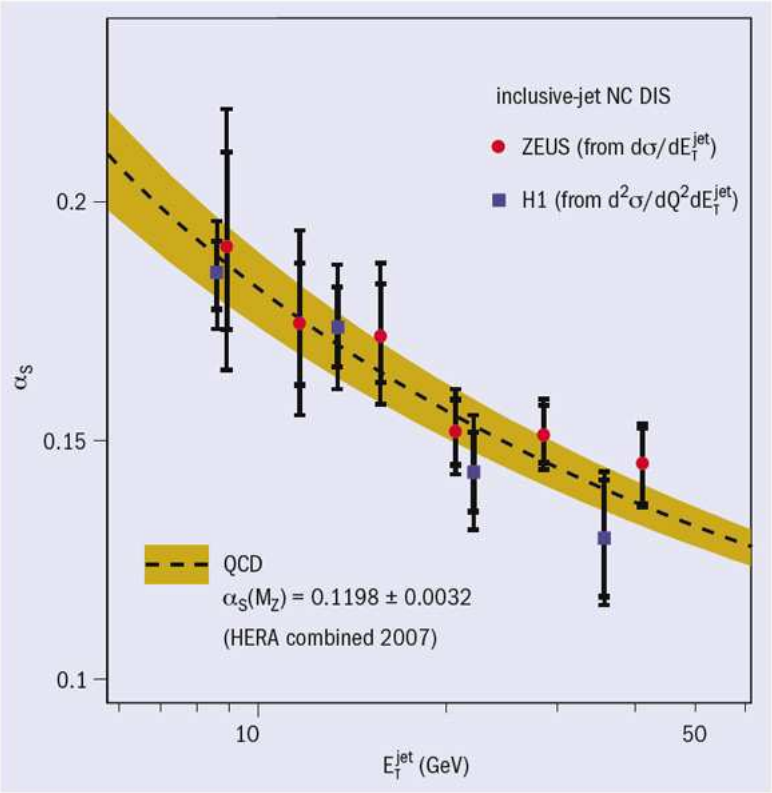

Figure 1–2: The running of the strong coupling constant α_s as measured by H1 and ZEUS [8].

The bare quarks produced at the LHC, that are ideal for studying QCD, undergo showering and cluster together to form a large number of hadrons. This process is called hadronization. The initial quark is now represented by a collimated shower of particles known as a jet. More details on jets and how they are produced and seen at the ATLAS detector will be discussed in subsequent chapters. Jets are structurally complicated and their production is understood through parton shower evolution [9]. One such technique that allows to study jets and how they are formed is by looking inside the jets, i.e. at jet substructure. It is this idea of jet substructure that will be

further explored in later chapters. By looking inside jets, and in particular, at the subjets within, relations to QCD and its coupling constant can be made.

In the zeroth order pQCD, a jet consists of only one parton and the number of subjets, jets found within a jet, is trivially equal to one. The first non-trivial contribution to the subjet multiplicity is given by $\mathcal{O}(\alpha)$ processes, for example, when a quark radiates a gluon at a small angle. The deviation from unity is proportional to the rate of parton emission and thus to α_s. The next-to-leading-order (NLO) QCD corrections or $\mathcal{O}(\alpha^2)$ processes allow for more precise measurements.

More than one subjet reconstructed within a jet indicates that the higher order processes are involved. The relationship between pQCD and subjet production can allow for a determination of the strong coupling constant, α_s, and forms the motivation for this study.

1.2 Description of the Large Hadron Collider

The Large Hadron Collider (LHC) is a 27 km long circular particle collider located underground in a tunnel at CERN [10]. Straddling the Swiss-France border, the LHC is located just outside Geneva, Switzerland. With an intended centre of mass energy of 14 TeV, it is the world's most energetic particle collider.

The collider is made of two counter-rotating rings which contain protons accelerated close to the speed of light. The protons are grouped together and delivered in bunches rather than a continuous beam as a consequence of the radio frequency acceleration scheme [11]. Each beam may contain up to 2808 bunches with each bunch containing about 10^{11} protons. The beam intensity at the LHC at full capabilities will have a luminosity of order $10^{34}\text{cm}^{-2}\text{s}^{-1}$, with the luminosity defined as the rate

of interactions per unit area. The luminosity, L, of a collider is directly proportional to event generation by

$$N_{event} = L\sigma_{event} \ , \tag{1.1}$$

where N_{event} is the number of events recorded per second, and σ_{event} is the cross-section for such an event. To obtain the total number of events, one integrates over the luminosity to get the total integrated luminosity, $\int Ldt$.[1] At the design luminosity, the proton bunches will cross the interaction points every 25 ns, causing on average 23 proton-proton collisions per bunch crossing and nearly 600 million collisions per second [12]. The current total integrated luminosity is 6.63 fb^{-1} as of mid 2012 and is shown in Figure 1-3.

To reach the required center of mass energy, the protons are accelerated in a series of stages. Protons are first obtained by removing the electrons from hydrogen atoms. The protons are then boosted in a linear accelerator up to 50 MeV, followed by injection into three successive synchrotrons: the Booster, the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS), which accelerate them to 1.4, 26, and 450 GeV, respectively. The protons are then injected into the LHC ring where they are accelerated to their nominal energy. Currently, as of 2012, the center of mass energy is 8 TeV (4 TeV beams), with the final goal of the LHC being 14 TeV. A schematic of this process is shown in Figure 1-4.

[1] The common unit of area used in luminosities and cross sections is the *barn*, 1 barn = 10^{-24}cm^2

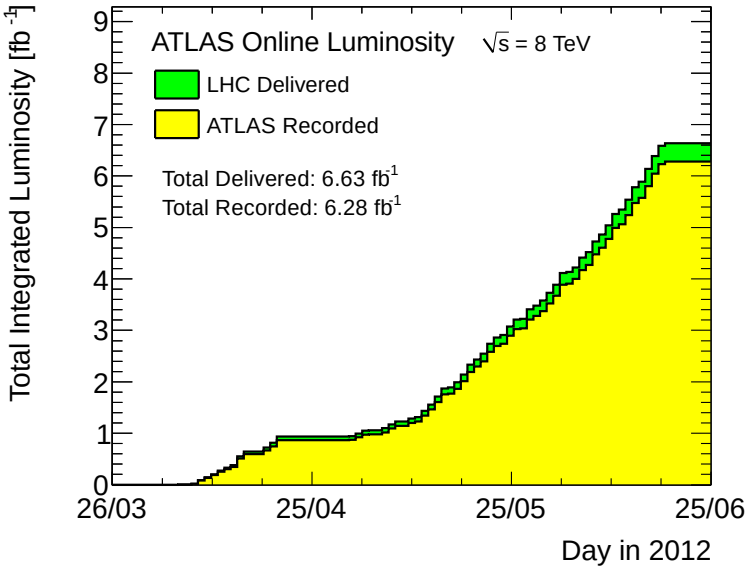

Figure 1–3: Total integrated luminosity versus day delivered by the LHC and recorded by ATLAS for proton-proton (p-p) collisions at 8 TeV centre-of-mass energy during stable beams in 2012. The ATLAS detector's total efficiency currently is 94.0%.

There are 4 main experiments operating at the LHC: two general purpose experiments (ATLAS and CMS) and two more physics process specific experiments (LHCb, with the intent to study B-physics, and ALICE, with the intent to study heavy ion physics) located at four interaction points, as seen in Figure 1-5.

This section gave a brief overview of the LHC and how it operates. For a more in-depth description see the LHC Design Report [10].

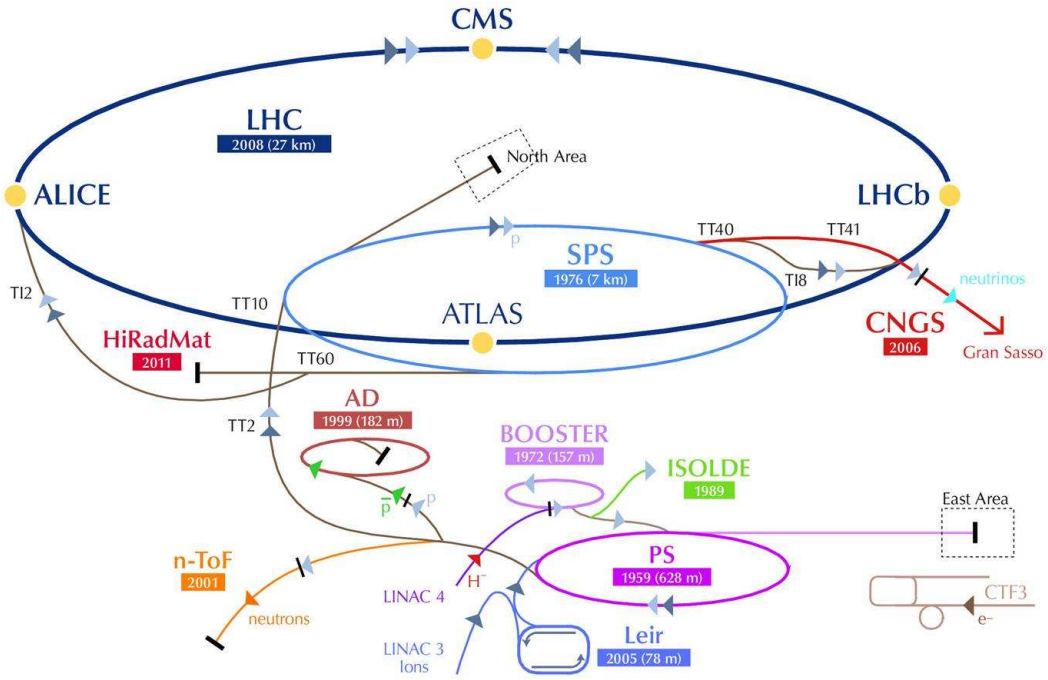

Figure 1–4: Protons are obtained by removing an electron from Hydrogen atoms. They are injected from LINAC4 into the PS Booster, Proton Synchrotron (PS), Super Proton Synchrotron (SPS), and finally into the Large Hadron Collider ring. The protons circulate in the rings for about 20 minutes before reaching their maximum energy of 4 TeV. Apart from the 4 major detectors in the LHC ring itself (ATLAS, CMS, LHCb, and ALICE), several other experiments also take place at the CERN. A non-exhaustive list includes: the Leir accelerator that is used for boosting ions used for the ion-ion collisions, the CLIC Test Facility (CTF3) for a building a compact linear collider, and the neutron-Time of Flight experiment (n-TOF) that studies neutron physics [13].

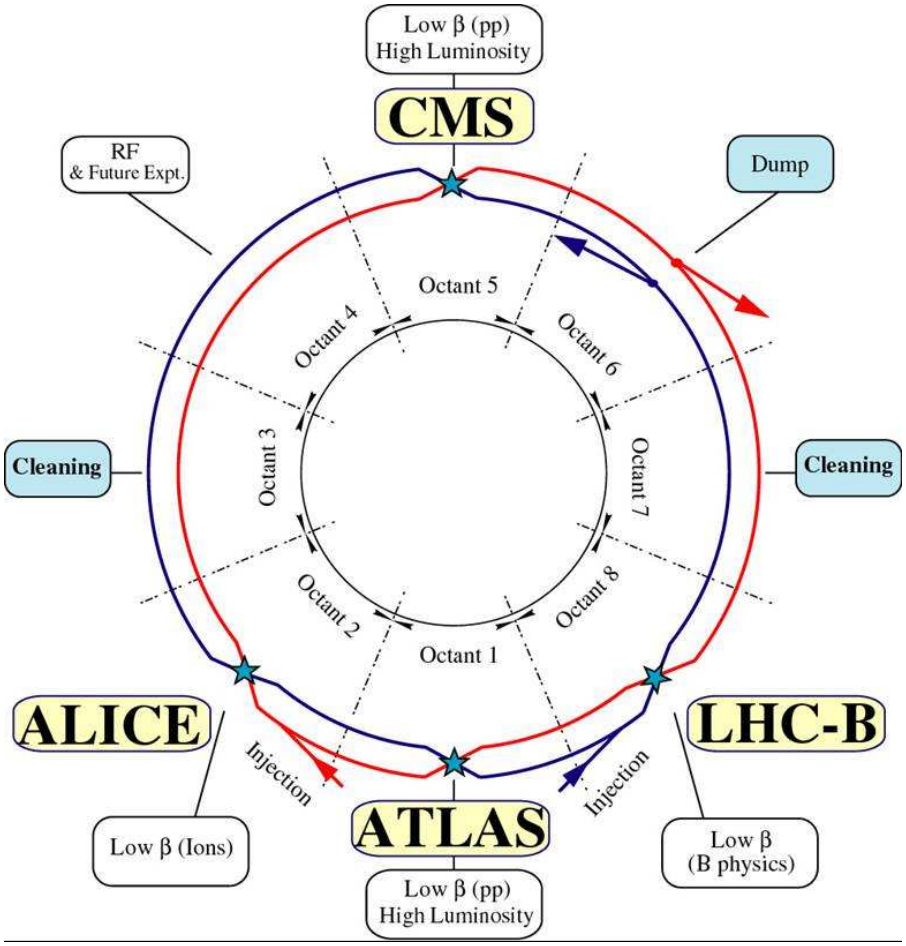

Figure 1–5: A bird's eye view of the LHC and its interaction points. At each interaction point one of the 4 main experiments, ATLAS, CMS, ALICE, and LHCb reside. TOTEM and LHCf, much smaller LHC experiments, are not shown in this figure. CERN Experiment Image: Copyright CERN [14]

CHAPTER 2
The ATLAS Detector

The ATLAS detector consists of several main components: an inner detector and tracking system, calorimeters, and muon spectrometers. A large magnet system is responsible for the bending the trajectories of charged particles. A solenoid with a magnetic field of 2T is used in the inner detector, and large toroids generate the field for the muon spectrometer. The ATLAS detector is shown in Figure 1-5.

The detector weighs over 7000t and is 44m long with a radius of 11m. The ATLAS coordinate system is a right-handed system defined about the nominal interaction point. The z-axis is directed along the beam line, whereas the transverse plane to the beam direction makes up the $x-y$ plane. The positive x direction is defined as pointing to the center of the LHC ring and the positive y direction is upwards. The polar angle θ is measured from the beam axis (along z) and the azimuthal angle ϕ is measured starting from the x axis in the $x-y$ plane. Since the polar angle is not invariant under Lorentz boosts a more convenient unit is desired. One such observable that is invariant under such boosts, is the rapidity $y = 1/2 \ln[(E+p_z)/(E-p_z)]$, which can be further simplified in the high energy limit as the pseudorapidity, defined as $\eta = -\ln(\tan(\theta/2))$. The transverse momentum p_T, the transverse energy E_T, and the missing transverse energy E_T^{miss}, unless otherwise stated are defined in the $x-y$ plane. A common measurement of distance occurs in the $\eta-\phi$ space and is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

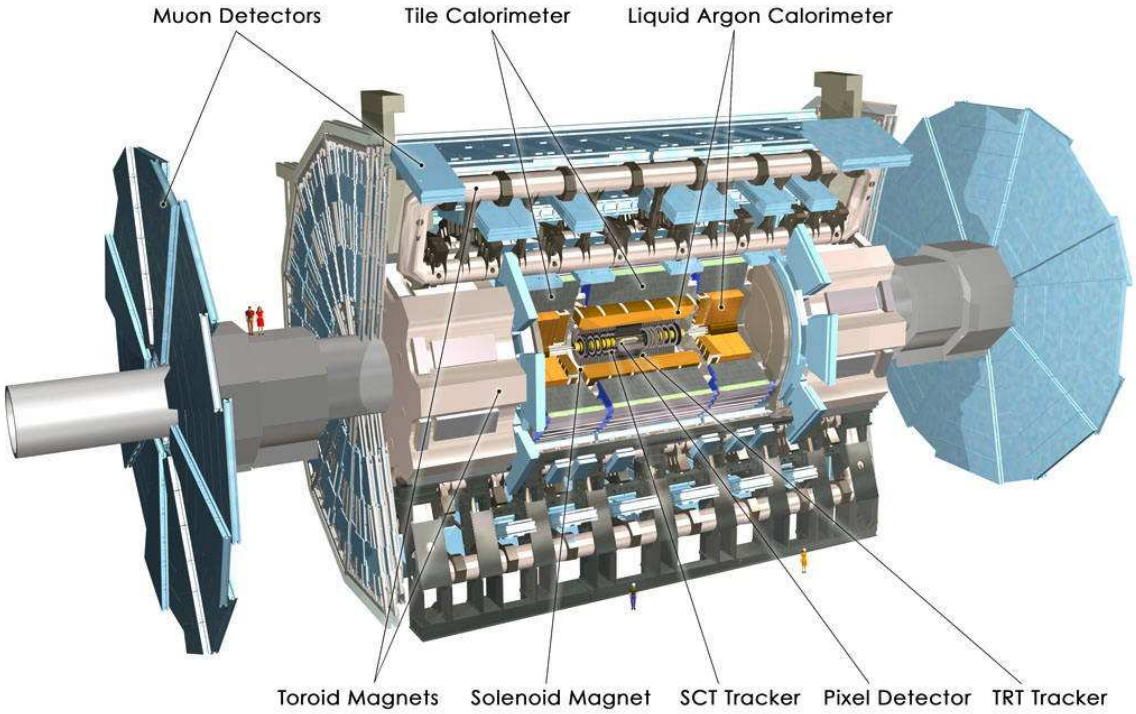

Figure 2–1: A visualisation of the ATLAS detector. Located closest to the beam pipe is the ATLAS tracking system (The TRT, Pixel, and SCT detectors, collectively referred to as the inner detector), followed by the calorimeter system, and finally the muon spectrometer. The locations of the ATLAS magnets are also indicated [12] .

The ATLAS detector was created with several goals in mind [11]. These are:

- Fast, radiation-hard electronics and sensors with high detector granularity to handle particle fluxes and overlapping events.
- Precise electromagnetic calorimetry for photon and electron detection, along with full coverage in the hadronic calorimeter for accurate energy measurements.
- Good muon identification and momentum resolution.

- Efficient tracking for charged particles and reconstruction efficiency, even at high luminosity.
- Almost complete solid angle coverage for detector subsystems.
- Sufficient background rejection and efficient triggering on interesting signals.

2.1 Inner Detector

The ATLAS Inner Detector (ID) was designed to perform high precision measurements at fine granularity of the large track density events produced by the LHC. The ID is located directly around the beam axis with a length and diameter of 6.2m $\times$ 2.1m, respectively. The goal of the detector is to provide robust pattern recognition, high momentum resolution, and primary and secondary vertex measurements for charged tracks above a minimum p_T threshold. In order to achieve the necessary momentum and vertex resolution requirements, high-precision measurements must be made with the ID's three independent but complementary sub-detectors.

The innermost component is the Pixel detector followed by the Semiconductor Tracker (SCT). These silicon based detectors cover the range $|\eta| < 2.5$ and contain 80.4 million and 6.3 million read out channels respectively. Further out radially from the beam axis is the Transition Radiation Tracker (TRT) comprised of many gaseous straw tubes and transition radiation material. The TRT is able to provide continuous tracking with its capability of having many hits per track and serves to enhance and improve both the pattern recognition and momentum resolution of the other sub-detectors. It covers the range $|\eta| < 2.0$ and contains 351,000 readout channels. The design is such that each sub-component has an endcap and barrel region to minimize the amount of material that a particle coming from the interaction point

must pass through. All the components are inside the 2T magnetic field generated by the superconducting solenoid which surrounds the ID. The curvature of the charged particles caused by the magnet allows for the momentum measurement to be made. The layering and design of the ID can be seen in Figure 2-2 and Figure 2-3.

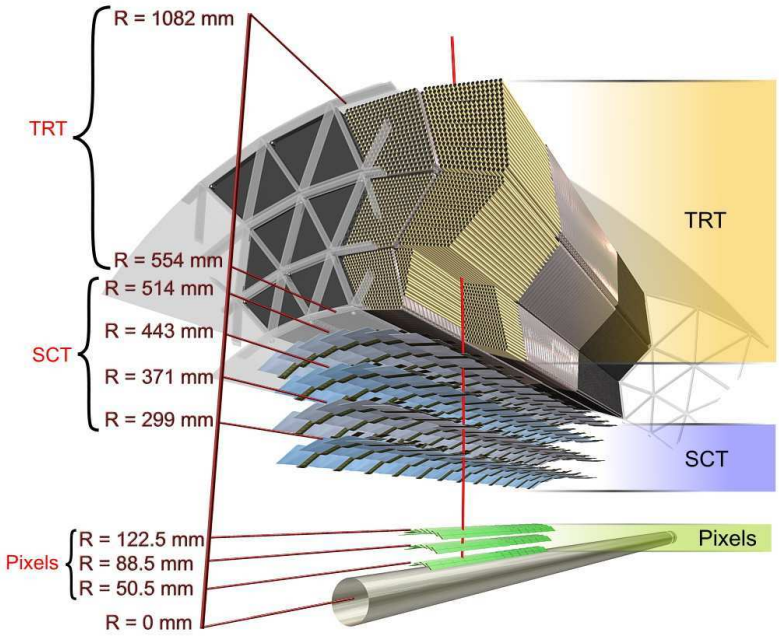

Figure 2–2: Cross-section view of the ATLAS Inner Detector. The Pixels, SCT, and TRT detectors are shown with their distance from the beam axis.

2.2 Calorimeter

The ATLAS calorimeter system [5], shown in Figure 2-4, is comprised of five subsystems divided into barrel and endcap regions. It encloses the ID system, covering the full ϕ-space and $|\eta| < 4.9$, and extends out to 4.25 m radially from the beam pipe.

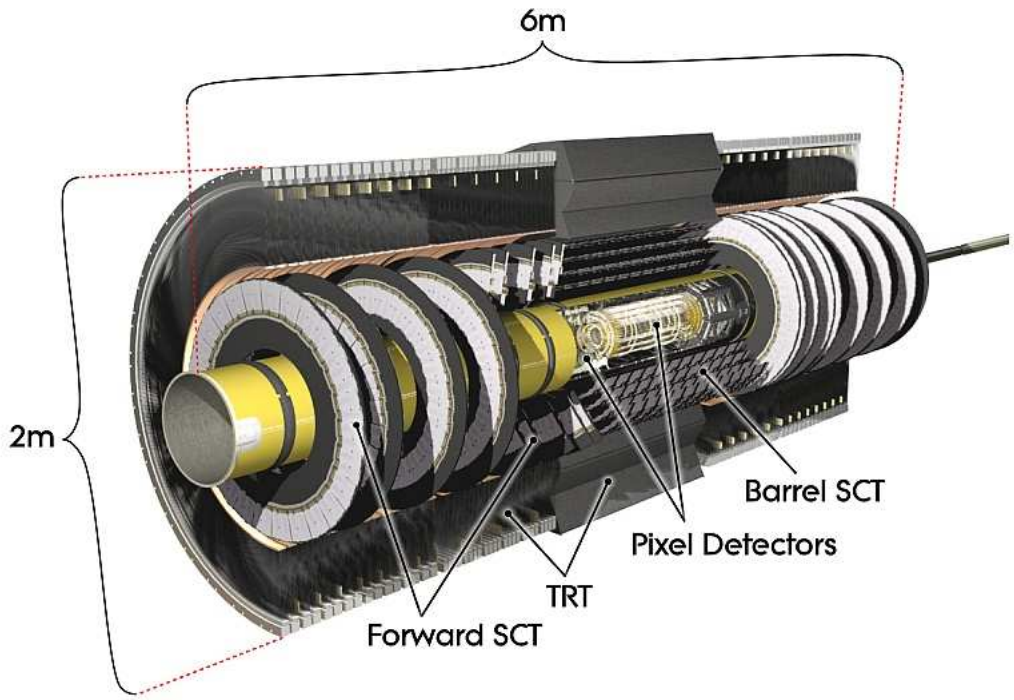

Figure 2–3: The ATLAS Inner Detector. The subdetectors, each with an endcap and barrel region, are shown with their location with respect to the nominal interaction point.

Calorimeters can be categorized as either electromagnetic calorimeters, for measurements of electrons and photons, and hadronic calorimeters, that measure the energy of hadrons. The granularities of the electromagnetic and hadronic calorimeter in the $\eta - \phi$ space are 0.025×0.025 and 0.1×0.1, respectively. The way in which particles interact and their energy response in calorimeters will be discussed in more detail in Chapter 3.

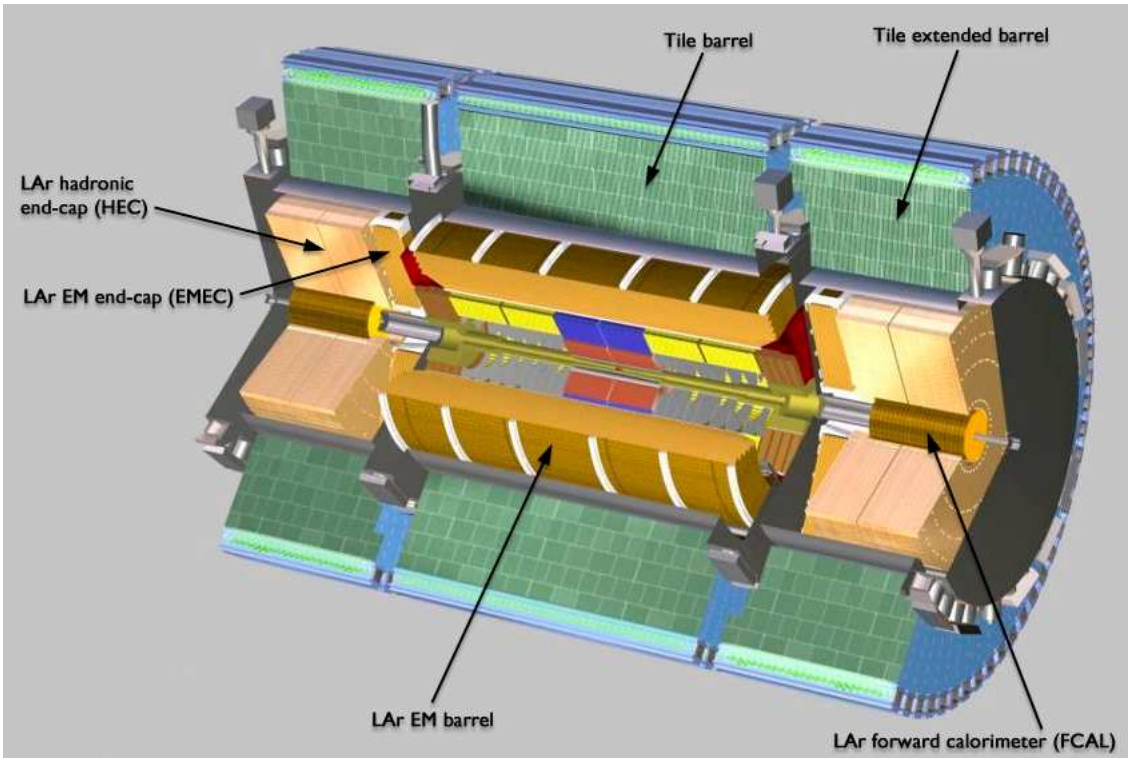

Figure 2–4: ATLAS calorimeter diagram. All of the calorimeter systems (Hadronic Tile, Hadronic endcap, EM barrel, EM endcap, and Forward calorimeters) are depicted.

2.2.1 Electromagnetic Calorimeter

The electromagnetic calorimeter is a Pb-LAr detector (liquid argon sampling calorimeter with lead absorbers) with an accordion-shaped geometry. The geometry provides a complete coverage in ϕ with no gaps between cells. The calorimeter is made of three sections: a barrel component covering the range $|\eta| < 1.475$, and two end-cap calorimeters with range $1.375 < |\eta| < 3.2$. A sketch of a section of the LAr calorimeter can be seen in Figure 2-5.

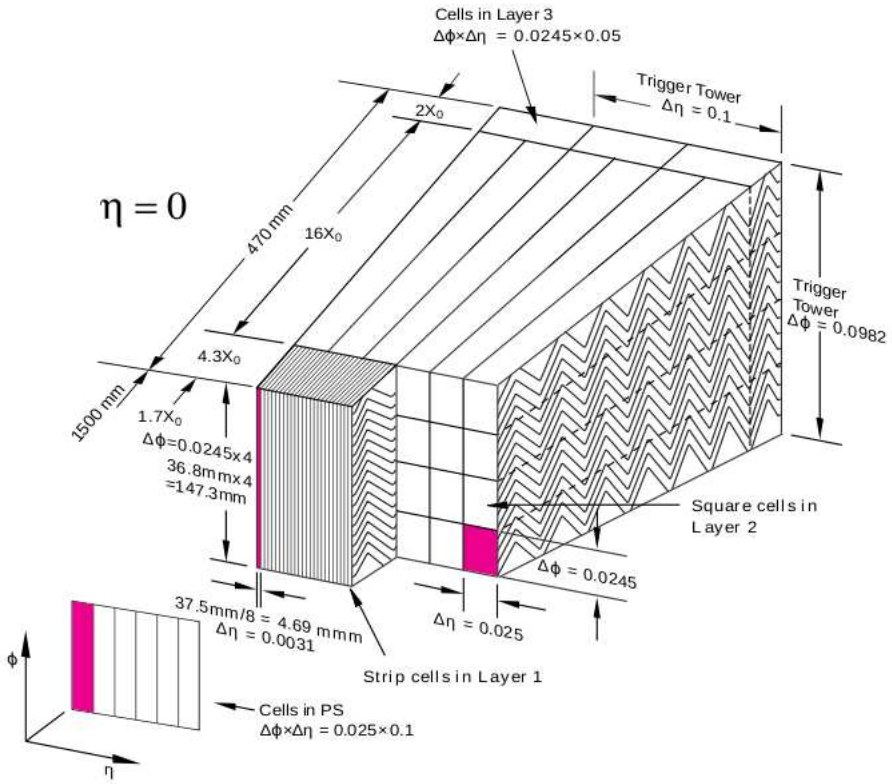

Figure 2–5: ATLAS Liquid Argon calorimeter module diagram. PS indicates the pre-sampler [15].

2.2.2 Hadronic Calorimeters

Tile Calorimeter

The Tile Calorimeter is a sampling calorimeter (absorber material is iron and active scintillating plastic) covering the EM calorimeter. The tiles are layered in an alternating pattern such that there is overlap between adjacent tiles. The tiles are connected to photomultiplier tubes at both ends. The Tile calorimeter barrel covers the range $|\eta| < 1.0$ and its extensions $0.8 < |\eta| < 1.7$. The function of the Tile

calorimeter is to provide good measurement of the energy of jets, which are able to pass through the electromagnetic calorimeter and to also determine any missing transverse energy, E_T^{miss}.

Hadronic Endcap Calorimeters

The Hadronic endcaps (HEC) are sampling calorimeters made out of copper plates and LAr gaps. They are located beside the electromagnetic calorimeter and cover the region $1.5 < |\eta| < 3.2$

Forward Calorimeter

The Forward Calorimeters (FCal) are located at the far ends of the detector close to the beam pipe, covering the region $3.1 < |\eta| < 4.9$. The FCal is made up of three components per endcap: The first section made out of copper is for electromagnetic measurements, while the other two, made out of tungsten, measure the energy of hadronic interactions. In all three, LAr is the active medium.

2.3 Muon Spectrometer

The muon spectrometer [5] is designed to identity muons and measure their track curvature to perform precise momentum measurements. The muon trajectories are caused by a large magnetic field generated by the air-core toroid system. The location of the muon spectrometer is past the calorimeter system making it, radially, the farthest out subdetector. The open structure of the muon spectrometer minimizes multiple scattering effects and an excellent momentum resolution is achieved. The spectrometer has three high-precision tracking chambers and is made up of monitored drift chambers, cathode strip chambers, resistive plate chambers, and thin gap chambers. The presence of a muon is indicated by having the muon

ionize a gas inside the chambers, liberating electrons which drift generating a current that can be detected.

2.4 Trigger

The ATLAS trigger system, known as the Trigger and Data Acquisition (TDAQ), is designed to filter interesting signals on an event by event basis. Due to the very high rate of collisions in ATLAS, it is not possible to record every event, nor is it desirable to. The TDAQ, therefore, has the task of reducing the rate to something more manageable but not to reject any events that are physically interesting.

The trigger is made up of hardware and software components with three different levels: Level 1 (L1), Level 2 (L2), and the Event Filter (EF) [11]. Each level filters the decisions made at the previous step and is able to apply additional selection criteria, called trigger signatures. These criteria are typically a combination of a physics object (e.g. a muon) and a threshold (e.g. $p_T \geq 20$ GeV/c). Numerous trigger signatures can be applied in combination depending on the expected sensitivity of a given physics process wanted. In order to optimize event storage rate even further, some trigger signatures are prescaled, which suppresses a given trigger by some multiplicative factor.[1]

The first level, L1, makes decisions using only partial detector information, e.g. scanning for high p_T physics objects or regions with interesting physical features. These regions of interest (RoI) in η-ϕ space are then passed down to the L2 tier for

[1] A prescale of 10 will imply that 1 in every 10 events for that trigger will be recorded.

further processing. The L1 trigger monitors all events during run-time and is only inactive when recording an event (deadtime). The L1 trigger lowers the event rate to approximately 75 kHz from the nominal bunch crossing rate of 40 MHz. The L2 system applies further selection criteria to the data passed on from L1. All data belonging to an event are collected and gathered under a single identifier. Therefore all cells, tracks, vertices, and trigger information belonging to one event are collected and tagged as being collectively from one event. The passing events, as deemed by L2 filtering criteria, are then sent to the EF level for final processing. Overall, the L2 trigger will bring down the event rate to around 3.5kHz. The final trigger level, the EF, reduces the rate to 200Hz and is also responsible for categorizing events. By combining all of the subdetectors' information on the event, the EF is able to process what particles and processes took place and subsequently place the events into different streams for data recording. The time process for each level on average from L1 to L2 and finally to EF takes about 2.5 μs, 40 ms, and 4s, respectively [11]. The trigger system and the processing stream are shown in Figure 2-6.

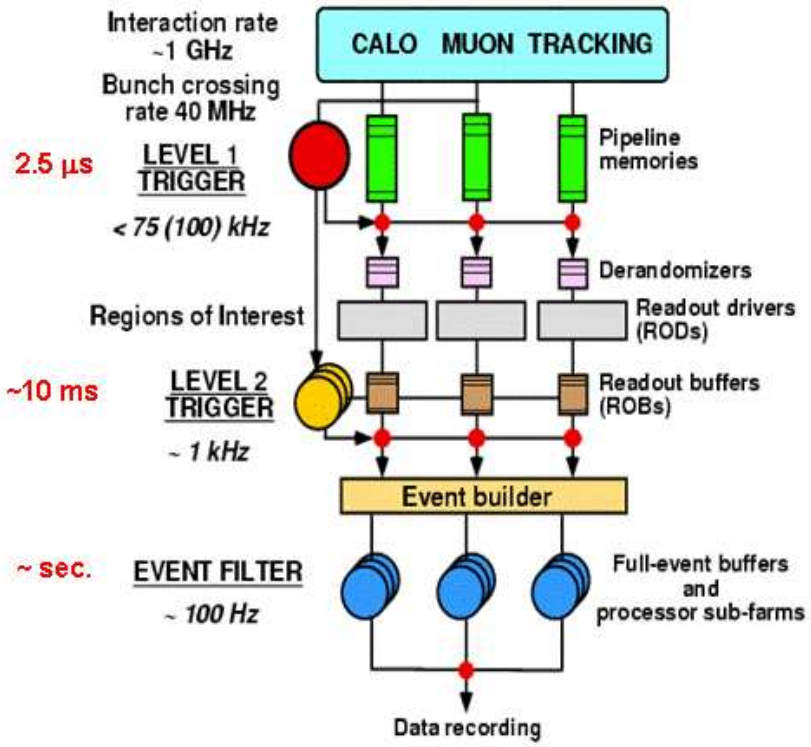

Figure 2–6: The processing stream of the ATLAS trigger system.

CHAPTER 3
Calorimetry

Calorimetry, or energy measurement, is one of the fundamental techniques used in modern particle collider experiments. Designed to estimate or measure the energy of an incident particle, a calorimeter works by completely stopping a particle, absorbing its energy, and converting that into an electrical signal, from which the energy of the particle can be determined [16]. In addition, calorimetry can be used to provide particle identification and use momentum conservation imbalances to determine the presence of undetected particles, such as neutrinos.

There are two groups of calorimeters based upon the types of particle interactions with matter: electromagnetic, optimized to measure the energy of incident photons, electrons and positrons[1] , and hadronic, optimized to measure the energy of incident hadrons. As a particle's energy is measured through its interaction with the calorimeter, the particle no longer has the same energy or momentum as it did before. Therefore, calorimeters are located further out radially as to not interfere with other subdetectors located closer to the beam pipe. Muons are minimum ionizing particles and therefore interact weakly with calorimeters. This allows for the muon spectrometer to be placed outside of the calorimeter system.

[1] From this point onward, electrons will be taken to mean both electrons and positrons.

The two types of calorimeters and the mechanics behind such processes are explained in sections 3.1 and 3.2. In section 3.3 calorimeter linearity and compensation will be discussed. Finally, with the knowledge of calorimetry and their signals covered, the concept of a "jet" will be introduced in detail in Chapter 4.

3.1 Electromagnetic Calorimeter

3.1.1 Particle Interactions

Photons and electrons both interact with matter electromagnetically in a variety of ways dependent upon initial particle energy. These interactions are well understood by the theory of Quantum Electrodynamics (QED). The main processes by which protons interact are the Photoelectric effect, Compton scattering, and pair production (which are covered extensively in [17]). For highly energetic photons, pair production is the dominant interaction. Pair production is the process by which a photon interacts with a nucleus and is converted into an e^+/e^- pair. When a charged particle, such as an electron, travels through matter, the main interactions are ionization, scattering, and bremsstrahlung (the radiation of photons from charged particles as they are accelerated (deflected) during interaction with atomic nuclei). At energies above the critical energy of $E_C \approx 550$ MeV$/Z$, where Z is the atomic number of the absorber, bremsstrahlung is the dominant form of energy loss for electrons. At lower energies, there is also a contribution from electron-electron (Møller scattering) and electron-positron (Bhabha scattering) scattering processes. The fractional energy loss per radiation length (in lead) as a function of incident energy is given in Figure 3-1.

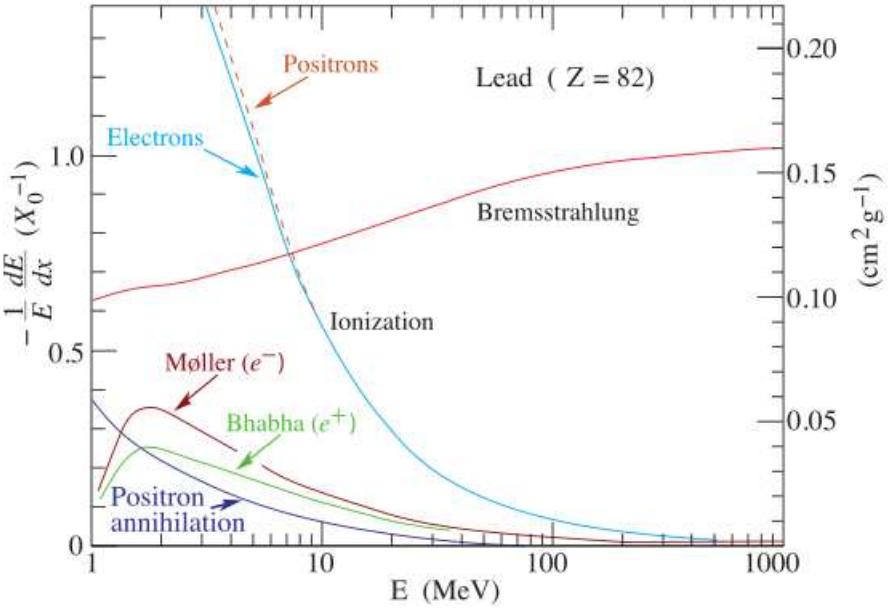

Figure 3–1: Fractional energy loss per radiation length in lead as a function of e^+/e^- energy. The main losses of energy for electrons are ionization at low energy and bremsstrahlung at high energy [18].

The radiation length, X_0, is defined in terms of energy loss, ΔE, for a particle with initial energy E over a length scale Δx such that

$$-\frac{\Delta E}{E} = \frac{\Delta x}{X_0}. \tag{3.1}$$

The radiation length to first order is material independent, i.e. a particle in any medium is likely to interact within one X_0. A useful approximation for the radiation length of a single elemental material is $X_0 \approx 180A/Z^2$, where A is the mass number and Z the atomic number. The rate of energy loss [19] for an electromagnetically interacting particle with initial energy E_0 as a function of depth x, within the calorimeter is given by

$$-X_0 \frac{dE}{dx} = E_0 b \frac{(bx/X_0)^{a-1} e^{-bx/X_0}}{\Gamma(a)}, \tag{3.2}$$

where $\Gamma(a)$ is the Gamma function[2] and a and b are obtained from empirical fits to the data.

3.1.2 Electromagnetic Showers

When an energetic electron or photon interacts with matter the two dominant processes are bremsstrahlung (with an end result of an electron and photon) and pair production (with an end result of an electron and positron), respectively. These particles are then free to interact with the calorimeter medium and if of sufficient energy undergo themselves bremsstrahlung and pair production. This chain reaction causes a cascade of secondary particles which collectively are known as an electromagnetic shower. This simple model (Rossi-Heitler) which begins with one electron or photon shows that at each interaction step the number of particles doubles [17]. This process will continue until the energy of the charged particles drops below E_C, at which point bremsstrahlung/pair production no longer occurs, multiplication halts, and the shower reaches a maximum. The remaining low-energy particles interact by ionization (electrons) or Compton scattering or photoelectric effect (photons) until they are absorbed by the calorimeter medium. The total energy in the final state particles, which is the total energy of the shower, is measured, and from this the energy of the incident particle can be inferred.

[2] The gamma function is defined to be $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$.

The initial (and the subsequent) interactions occur on average after traveling one radiation length, X_0, through the material. With equation (3.2) the length of the shower can be estimated and is usually expressed in radiation lengths. The lateral spreading of the shower is controlled primarily by the radiated photons from bremsstrahlung and multiple scattering of the electrons. A parametrisation using material properties defines the Moliere radius as

$$\rho_M = mc^2 \sqrt{\frac{4\pi}{\alpha}} \frac{X_0}{E_C}, \tag{3.3}$$

where α is the fine structure constant $\frac{1}{137}$. Over 90% of the shower energy is contained within the Moliere radius. Figure 3-2 shows the development of a ideal electromagnetic shower.

3.2 Hadronic Calorimeter

3.2.1 Particle Interactions

The process behind hadronic calorimeters is more complicated than electromagnetic calorimetry due to the different types and ways the particles interact with matter. While charged hadrons are able to interact electromagnetically, in general, all hadrons are also able to interact via the strong force. This allows for complex nuclear interactions where there is large fluctuation on the types of particles produced on a event-by-event basis. The large number of possible interactions adds to the complexity.

Similar to the radiation length, X_0, a nuclear interaction is likely to take place after a certain distance. This occurs on average after a distance called the nuclear interaction length, λ_I, and is much larger than the radiation length for a given

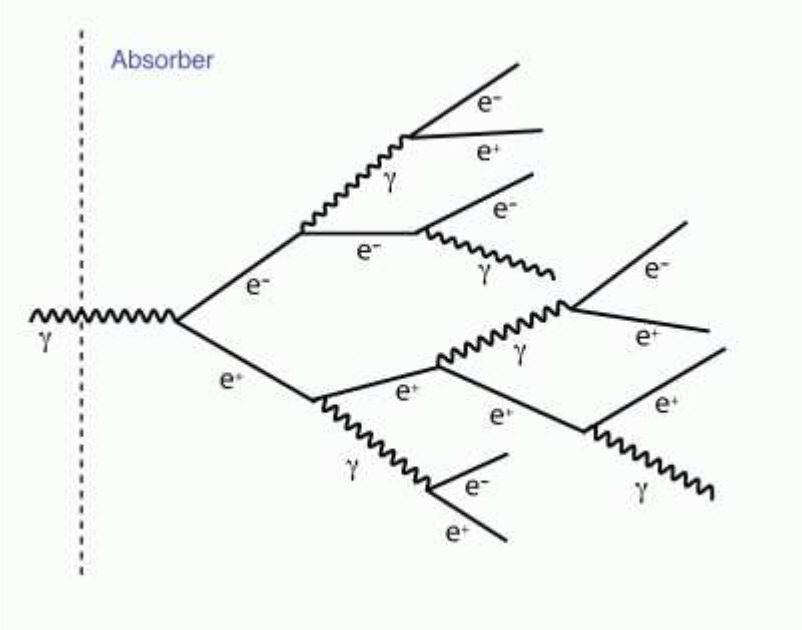

Figure 3–2: The development of an electromagnetic shower as explained by the Rossi-Heitler model. At each interaction step the number of particles doubles. The shower will continue to develop until the particles drop below a critical energy in which they no longer interact by the highly energetic electromagnetic processes [20].

material ($\lambda_I > X_0$). Therefore a hadron must travel a greater depth in order to deposit all of its energy compared to an electromagnetic interacting particle. Due to this, hadronic calorimeters are often made of denser material (to have a reasonable λ_I value) and placed on the outside of the electromagnetic calorimeters. A table of nuclear interaction lengths, radiation lengths, and other properties can be seen in Table 3-1.

3.2.2 Hadronic Showers

When a high energy hadron interacts with matter, its interaction produces a number of nuclear fragments and secondary hadrons which can then interact with

Table 3–1: A short list of particle shower quantities such as the electromagnetic radiation length, X_0, nuclear interaction length, λ_I, atomic mass, A, and atomic number, Z, in some common materials. For a more complete list, including common calorimeter materials, see [2].

Material	Z	A	$\langle Z/A \rangle$	λ_I (g cm^{-2})	X_0 (g cm^{-2})
H_2	1	1.00794(7)	0.99212	52.0	63.04
D_2	1	2.0141017(1)	0.49650	71.8	125.97
He	2	4.002602(2)	0.49967	71.0	94.32
Li	3	6.941(2)	0.43221	71.3	82.78
Be	4	9.012182(3)	0.44384	77.8	65.19
C	6	12.0107(8)	0.49955	85.8	42.70
N_2	7	14.0067(2)	0.59976	89.7	37.99
O_2	8	15.9994(3)	0.50002	90.2	34.24
F_2	9	18.9984032(5)	0.47372	97.4	32.93
HNe	10	20.1797(6)	0.49555	99.0	28.93
Al	13	26.9815386(8)	0.48181	107.2	24.01
Si	14	28.0855(3)	0.49848	108.4	21.82

other nuclei, similarly as was shown for electromagnetic showers. Likewise, a hadronic shower will develop until the hadrons do not have enough energy to produce more secondaries, at which point the hadrons will mainly lose energy through ionization or will be absorbed in a nuclear process. Conceptually then, a hadronic shower is similar to an electromagnetic one, as far as the cascade of particles is concerned. However, a hadronic shower is in general much more complicated structurally. Since both electromagnetic and hadronic interactions take place, a hadronic shower will also have two components: an electromagnetic component (distinguished by X_0) and a hadronic component (located at a further depth and characterized by λ_I). The lateral extension of a hadronic shower will be larger than an electromagnetic shower as its particle constituents are able to travel further before interacting. An example of a hadronic shower is shown in Figure 3-3. The electromagnetic component

is caused by hadrons that decay into two photons (such as π^0, η), which, inside of the hadronic shower, then develop electromagnetic showers.

The parametrisation of hadronic showers is non-trivial and the energy loss equation as a function of depth, x can be given by [19],

$$-\frac{1}{E_0}\frac{dE}{dx} = \alpha\frac{b^{a+1}}{\Gamma(a+1)}x^a e^{-bx} + (1-\alpha)ce^{-cx}, \tag{3.4}$$

where the constants α, a, b, and c are empirically determined. Comparing equation (3.5) with equation (3.2) and redefining some variables the first term of the hadronic energy loss equation can be written in terms of the electromagnetic equation for energy loss, thus showing the evident two component nature of hadronic showers.

A fraction of the energy in a hadronic shower will be absorbed by nuclear breakups and excitations and cannot be made into a measurable signal. In addition there may be muons or neutrinos created that escape the calorimeter without interacting and take away some of the shower's energy. This undetectable energy is known as invisible energy.

3.2.3 Compensation

The total hadronic energy is made up of invisible energy, electromagnetic energy, and hadronic energy. The fraction of energy deposited by the latter two is energy dependent and varies for each event. The ratio between the efficiency in measuring these energies is given by e/h, where e is the electromagnetic response and h is the

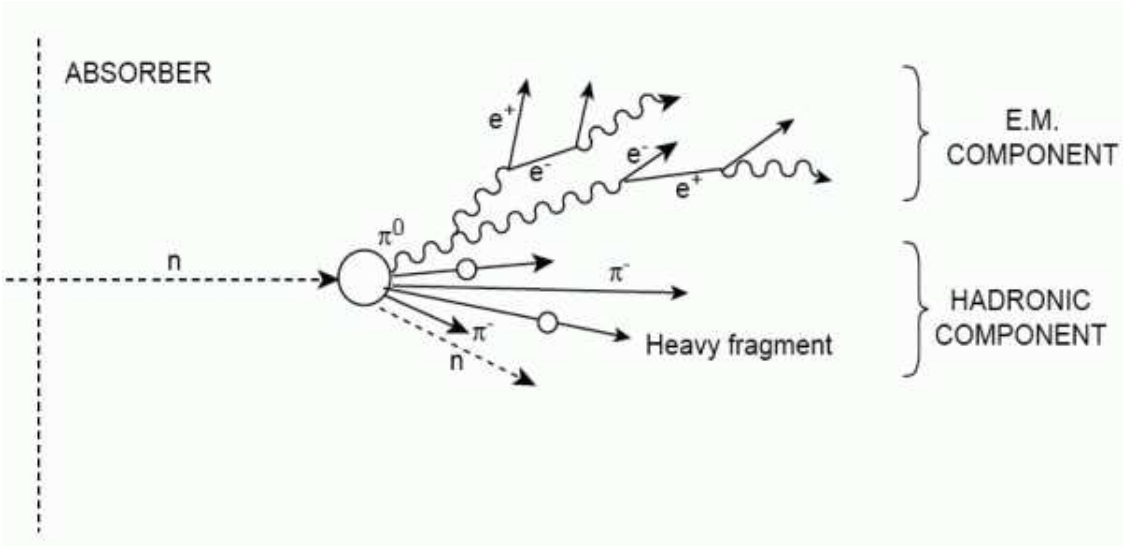

Figure 3–3: A simplified example of a hadronic shower showcasing the different components that contribute to energy loss. Invisible energy also makes up a portion of the total energy that is not depicted in the figure, such as the energy carried away from created muons or neutrinos, or energy undetected as nuclear bindings/excitations [20].

hadronic response of the calorimeter[3] . Since in most calorimeters the hadronic energy is not fully detected, $e/h > 1$. One way to have this ratio approach 1 is to build a compensating calorimeter, which increases h and lowers e. An increase in hadronic measurements is usually accomplished by calorimetry modifications, e.g using Uranium for the detection of slow moving neutrons that are not usually accounted for. The energy response ratio also affects energy resolution of a calorimeter, $\sigma(E)$ [21]:

$$\frac{\sigma(E)}{E} \propto \frac{k_1}{\sqrt{E}} + k_2 \, |e/h - 1| \tag{3.5}$$

[3] A calorimeter response is defined as $E_{measured}/E_{true}$

with the constants $k_i > 0$. For a compensated calorimeter the ratio $e/h = 1$ and so the second term vanishes thus improving the resolution. Compensation also improves the linearity (see next section) of a calorimeter.

3.2.4 Linearity

The linearity of a calorimeter refers to its response being proportional to energy, E_{true}. The average fraction of EM energy f_{em} in a hadronic shower grows with energy [22]:

$$f_{em}(E) = \alpha_0 \ln \frac{E}{E_{scale}}, \tag{3.6}$$

where E_{scale} is of the order of 1 GeV and α_0 is a constant. As the shower depth (which is dependent on the incident energy E) increases, the hadrons that interact electromagnetically contribute less and less to the hadronic signal. On the other hand, hadrons that are produced in hadronic showers can continue to interact electromagnetically. Therefore the fraction of hadronic energy, f_h is also dependent on the incident energy, albeit in a different relation than f_{em}. Although these fractions are subject to change with the incident energy the sum of the fractions $f_{em} + f_h = 1$ is fixed. If the response of a calorimeter to hadronic showers is written as

$$\begin{aligned} R_{had} &= f_{em}e + f_h h & (3.7)\\ &= f_{em}e + (1 - f_{em})h & (3.8)\\ &= (e - h)f_{em} + h & (3.9) \end{aligned}$$

then for a compensated calorimeter the non-linear term f_{em} cancels and the response depends only on h, improving the linearity of the calorimeter significantly [23] [21].

CHAPTER 4
Jets

The quarks and gluons that are produced from the hard scattering of protons undergo parton showering and hadronisation each resulting in a collimated flow of hadrons. This collection of hadrons is defined as a single object called a jet. The abundance of quarks and gluons produced makes jets the most common reaction product at ATLAS and therefore an important object to study. This chapter will discuss how jets are defined as well as go into some detail of the different algorithms used to reconstruct them at particle detectors.

4.1 Jet Theory and Production

The strong force is responsible for the interaction of particles that carry the colour charge, as mentioned in Chapter 1. For colour to be conserved, it is required that quarks carry one of three possible colours (red, blue, or green), whereas gluons carry two: one colour and one anti-colour. Colour confinement forbids that quarks and gluons appear in isolation, being coloured, and therefore they must exist as hadrons, composite particles that are colour-neutral. The observed hadrons can be further classified as baryons that are made up of three cancelling coloured or anti-coloured quarks or mesons that have one coloured and anti-coloured quark, both resulting in a net colour charge of neutral. During the collisions of protons at the LHC, highly energetic partons are able to overcome the strong force and separate from one another. As the distance between these partons increases so will the strong

force, up to a point in which the energy required to move the quarks any further apart is more likely to be used to create a quark-antiquark pair from vacuum. This process will produce a shower of quarks and gluons. Due to colour confinement, these particles will recombine into hadrons in a process known as hadronisation. The resulting shower of hadrons, known as a jet, will retain the same direction as the initial parton.

The overall evolution of a jet can be divided into three sections as shown in Figure 4-1. The first stage, named the parton level, consists of the bare parton. Partons at this level may radiate off gluons known as final state radiation. The second stage, known as the particle level jet, allows for the partons to undergo hadronisation and will contain hadrons, i.e. colourless baryons and mesons. Finally, the last stage is when the final state particles, the hadrons, interact with the detector's calorimeter and produce the hadronic showers described in section 3.2. For parton level studies, the calorimeter signals must be reconstructed and calibrated back down to the parton level. The first step in reconstructing a jet is to cluster together the calorimeter signals.

4.2 Jet Reconstruction/Algorithms

Jets can be reconstructed by several different algorithms that attempt to cluster together nearby final state particles [5]. The purpose of jet reconstruction is to build an object, the jet, with the kinematics and other characteristics that reflect that of the initial parton which initiated the creation of the jet. There are two main types of jet algorithms: cone algorithms, that rely on the spatial separation of the particles to produce the jet, and sequential recombination algorithms, that focus on

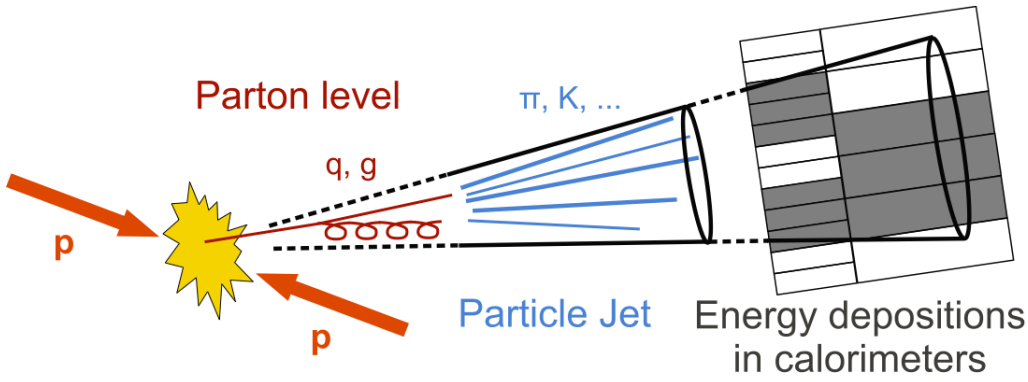

Figure 4–1: The three stages of jet evolution: parton level, particle (hadron) level, and calorimeter (detector) level [24].

the relative transverse momentum or closeness in momentum space. For the purpose of this study only the latter will be discussed.

The inputs to the ATLAS jet algorithms are the topological calorimeter clusters (TopoClusters). TopoClusters are three-dimensional clusters of calorimeter cells that are grouped together using a nearest-neighbour algorithm, subject to an energy threshold for noise suppression[25]. These inputs are also called preclusters. The generic sequential jet algorithm is as follows [26]:

- for each precluster i, define

$$d_i = p_{T,i}^{2p}, \tag{4.1}$$

  and for each pair (i, j) of preclusters,

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2}, \tag{4.2}$$

  where $p_{T,i}$ is the transverse momentum for precluster i, $R_{ij} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, and R is a parameter of the algorithm that approximately controls the radius

of the jet. The variable p is the jet algorithm parameter that determines which jet sequential mode is used: $p = 1$ specifies the k_t algorithm, $p = 0$ specifies the Cambridge/Aachen algorithm, and $p = -1$ specifies the anti-k_t algorithm.

- if $d_{ij} < d_i$, combine preclusters i and j into a single object
- else if $d_i < d_{ij}$, the precluster is considered to be isolated and can be removed from the list of preclusters and labelled as a jet.
- for inclusive jet reconstruction this process is repeated with all remaining objects until no preclusters remain.

In other words, the sequential jet algorithm first calculates a distance between each pair of calorimeter topoclusters, d_{ij}, and a parameter defined by the objects transverse momentum, d_i. The minimum between these two values are compared and if d_{ij} is smaller the preclusters i and j are combined. Else, if d_i is found to be smaller the precluster i is defined as a jet, and removed from the list of preclusters. These steps are repeated until all preclusters have either been merged or identified as a jet.

The different p values represent the emphasis the algorithm places upon the momentum of the preclusters. The anti-k_t will favor to cluster hard particles (particles with high p_T), the k_t will favor to cluster soft particles (low p_T) and the Cambridge/Aachen algorithm favors momentum independent clusterings. The inverse relation to p_T for anti-k_t jets results in soft particles being combined with hard particles, compared to k_t jets that would tend to cluster the soft particles together. A beneficial property that comes out of this relation is that the shape of anti-k_t jets in $\eta \times \phi$ space will be sensitive to hard radiation but insensitive to soft radiation. This

results in the anti-k_t algorithm reconstructing circular hard jets, as opposed to the irregularly shaped jets obtained from the others, which is an appealing property for experimental jet reconstruction.

The sequential combination algorithms are the predominant use of reconstruction algorithms because of their convenient advantages: theoretically there are both infrared and collinear safe, and experimentally, they are computationally fast. For a jet algorithm to be safe it should yield the same jet regardless whether a soft gluon is emitted or an outgoing parton splits into two collinear partons. One drawback of these algorithms is that they are conceptually difficult to visualize in comparison to the older and less theoretically safe cone algorithms. Also, in events with large contributions from underlying events and pileup, the k_t algorithm may have the undesirable effect of combining preclusters which are not physically from the same jet. Another advantage in anti-k_t jets is that due to the nature of the inverse momentum dependence, the jets resemble a cone-like shape in (η, ϕ) space. For such reason, the anti-k_t jet finding algorithm is currently the default algorithm used at the LHC. A diagram showing the different jet algorithms with the same input particles is shown in Figure 4-2 [26].

All jets that were analyzed in this study were reconstructed with the anti-k_T clustering algorithm with $R = 0.6$. The ATLAS experiment has chosen values of $R = 0.4$ and 0.6 for jet analysis studies. The value of 0.6 was chosen as it allows for larger jets, and therefore more jet substructure, to be examined. Larger values of R are avoided because of the difficulty to distinguish nearby jets. In the event of

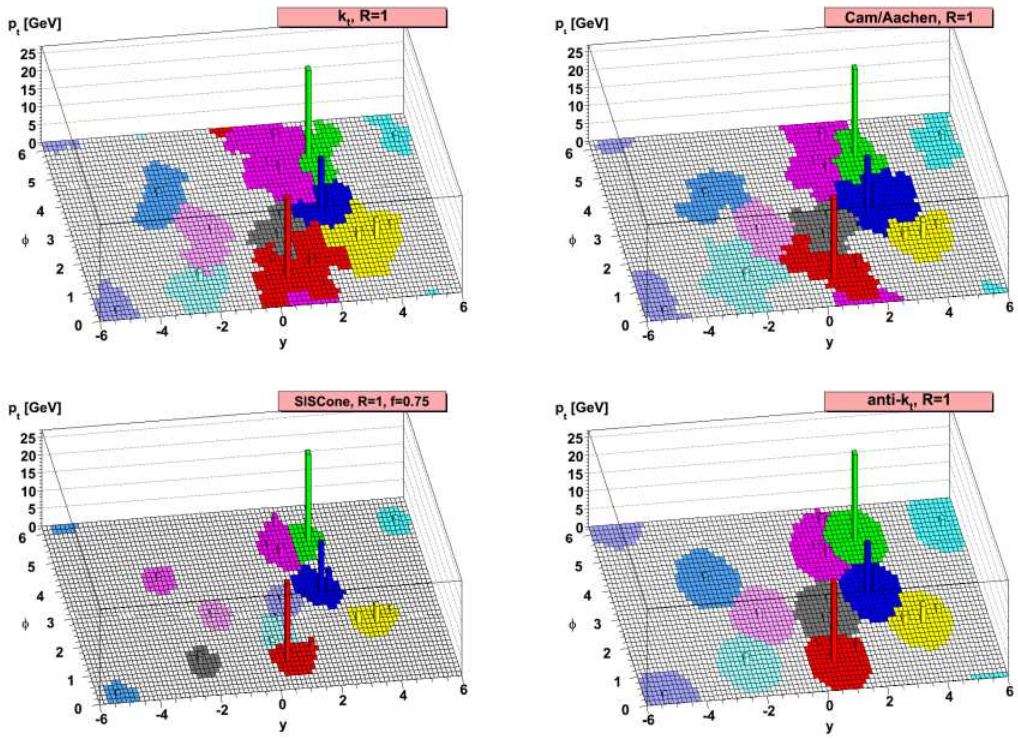

Figure 4–2: Comparison of different jet algorithms with $R = 1$ at the parton level (figure taken from [26]).

jet overlap, jets may be merged in reconstruction, which is undesired. To avoid such behaviour strict exclusion cuts must be designed and applied [27].

CHAPTER 5
Subjets

5.1 Definition of Subjet Multiplicity

Subjet reconstruction has a similar approach as jet reconstruction but rather than looking at all preclusters in an event, the subjet analysis is limited to objects only within a jet. The subjets in this analysis were resolved within a jet by use of the k_T clustering jet algorithm on the jet constituents. The algorithm was repeatedly applied for every pair of objects i and j, until the quantity d_{ij} was found to be greater than $d_{cut} = y_{cut} \cdot (p_{T,jet})^2$ [28]. The remaining objects were defined as subjets. This type of algorithm is an exclusive type due to the cut on the separation variable d_{ij}, in comparison to the inclusive type used in jet reconstruction that continues until all objects are resorted. The k_T algorithm is the only jet algorithm that correctly identifies the resulting substructure as physical objects and therefore is the algorithm used for jet substructure analysis [29].

Due to the systematic effects largely cancelling in the ratio $d_{ij}/(p_{T,jet})^2$ as seen in Eq. (4.2), the subjet reconstruction was performed with uncorrected calorimeter cells and jet energies [30]. The size, shape, and multiplicity of the subjets depend upon the chosen value for the resolution parameter y_{cut}. The mean subjet multiplicity, $\langle n_{sbj} \rangle$, is defined as the average number of subjets contained in a jet at a given y_{cut} value,

$$\langle n_{sbj}(y_{cut})\rangle = \frac{1}{N_{jets}} \sum_{i=1}^{N_{jets}} n^i_{sbj}(y_{cut}) \tag{5.1}$$

where N_{jets} is the total number of jets in the sample and $n^i_{sbj}(y_{cut})$ is the number of subjets in jet i. As the value of y_{cut} is increased the subjets increase in size and lower in number. The limiting case occurs when a subjet fully encompasses the jet that contains it. Therefore, by definition, $\langle n_{sbj}(y_{cut})\rangle \geq 1$. A depiction of this behaviour is shown in Figure 5-1. In this study, values of y_{cut} were taken to be from $5 \cdot 10^{-4}$ up to 0.1.

$$d_{ij} > d_{\text{cut}} = y_{\text{cut}} \left(E_T^{\text{jet}}\right)^2$$

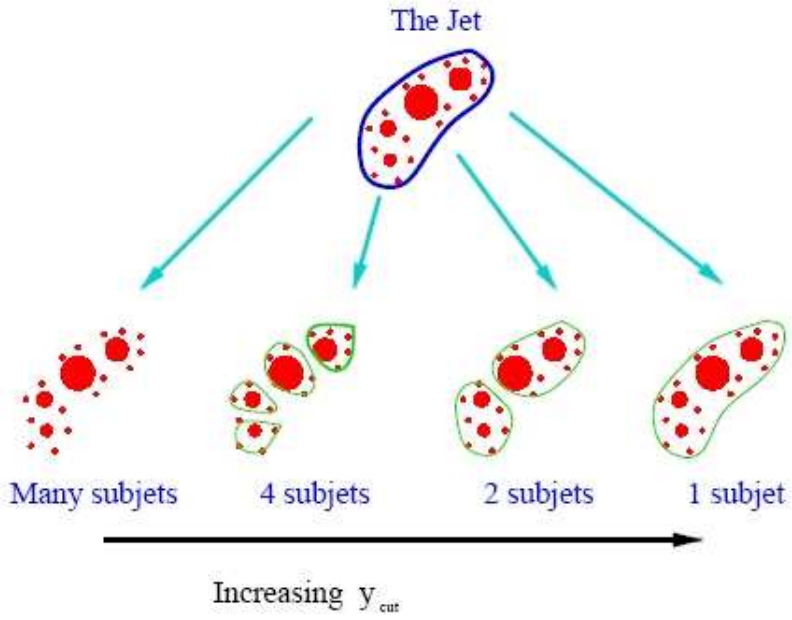

Figure 5–1: Subjet behaviour with increasing y_{cut}

5.2 QCD calculations

The perturbative QCD value for $\langle n_{sbj} \rangle$ can be calculated as the ratio of the cross section for subjet production to that of inclusive jet production (σ_{jet}) [30]:

$$\langle n_{sbj}(y_{cut}) \rangle = 1 + \frac{1}{\sigma_{jet}} \sum_{j=2}^{\infty} (j-1) \cdot \sigma_{sbj,j}(y_{cut},) \tag{5.2}$$

where $\sigma_{sbj,j}(y_{cut})$ is the cross section for the producing jets with j subjets with a resolution scale of y_{cut}. The QCD predictions for the mean subjet multiplicity were derived from Eq. (5.2) by computing the subjet cross section to $\mathcal{O}(\alpha_s^2)$ and the inclusive jet cross section to $\mathcal{O}(\alpha_s)$. The mean subjet multiplicity has an α_s dependence up to $\mathcal{O}(\alpha_s^2)$ and is therefore given by

$$\langle n_{sbj} \rangle = 1 + C_1 \alpha_s + C_2 \alpha_s^2. \tag{5.3}$$

C_1 and C_2 are coefficients whose values depend on y_{cut} and the jet and kinematic variables. This parametrisation shows the explicit α_s dependence up to NLO. A LO calculation performed up to $\mathcal{O}(\alpha_s)$ would severely restrict the accuracy of the predictions.

The NLO QCD calculations were performed using the CT10 parametrisations [31] of the parton distribution functions (PDFs). A parton distribution function is defined as the probability density for finding a certain parton inside a hadron with a certain longitudinal momentum fraction x at momentum transfer Q^2. The currently known PDFs are obtained from experimental data. The PDFs used in this study were used to describe the partons found within a proton. An example of the CT10 PDF can be seen in Fig. 5-2.

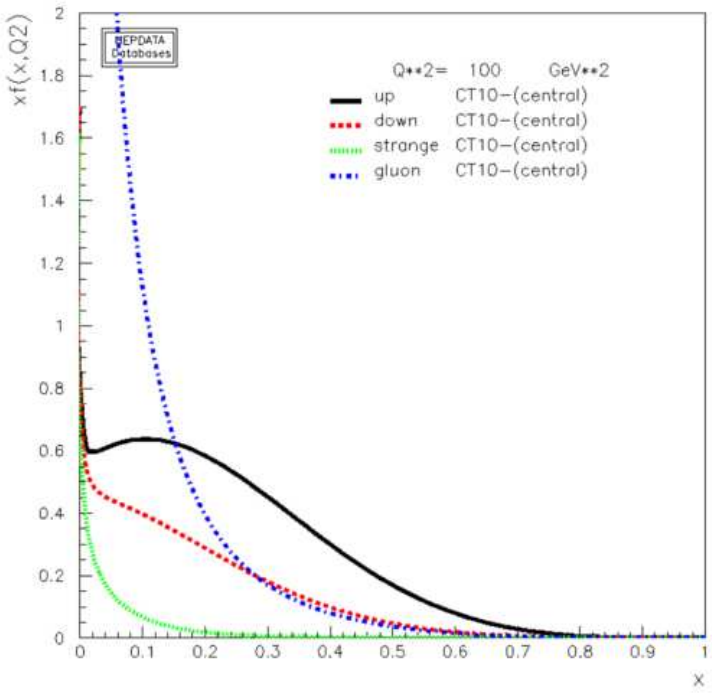

Figure 5–2: CT10 parton distribution function for several parton flavors as a function of the momentum fraction x with $Q^2 = 100$ GeV.

The NLO QCD calculations were performed on events which were simulated by the Monte Carlo (MC) event generator PYTHIA [32]. PYTHIA is an event generator that simulates the hard scatter between the two colliding particles. From this point the resulting partons undergo a parton showering process that is modelled by the Lund string model [9]. PYTHIA then accounts for hadronisation of the partons and the final state particles are passed through a detector simulator (GEANT4) that simulates the calorimetry interactions [33]. The name given to each step correspond to the parton (*part*), hadron (*had*), and calorimeter (*calo*) level, respectively. For more information on PYTHIA see [32]. The other MC event generator used in

this study, for comparison purposes and to calculate systematics uncertainties, was POWHEG. POWHEG also simulates hard scattering of colliding particles but the subsequent partons, and their showering, are modelled using a cluster model [34].

In order to represent these calculations as predictions of the mean subjet multiplicity, the jet algorithm described in section 4.2 was also applied to the simulated particles from each event. The level of correction applied by hadronisation is shown be comparing the parton level predictions against the hadron level in PYTHIA below. This check was necessary as the measurements of data involve jets of hadrons whereas the NLO QCD calculations were performed at the parton level. The multiplicative correction factor, C_{had} was defined as the ratio of $\langle n_{sbj} \rangle$ for jets of hadrons over that for jets of partons,

$$C_{had} = \frac{\left\langle n_{sbj} \right\rangle_{had}^{MC}}{\left\langle n_{sbj} \right\rangle_{part}^{MC}}, \tag{5.4}$$

and used for our comparison.

The value of C_{had} increases as y_{cut} decreases due to the increasing importance of non-perturbative effects that become present at small scales. The limits of the hadron-level predictions are as follows:

$$\left\langle n_{sbj} \right\rangle_{had}^{MC} \to \left\langle n_{jet}^{had} \right\rangle \qquad \text{as } y_{cut} \to 1, \text{ and} \tag{5.5}$$

$$\left\langle n_{sbj} \right\rangle_{had}^{MC} \to \left\langle n_{calo} \right\rangle \qquad \text{as } y_{cut} \to 0 \tag{5.6}$$

where $\left\langle n_{calo} \right\rangle$ is the average number of topoclusters (at the calorimeter level). Therefore, as the value of y_{cut} approaches 1 the subjets eventually encompass all of the hadrons that make up the jet and the average multiplicity of the subjets approaches the multiplicity of jets of hadrons. On the other hand, when y_{cut} approaches very

small numbers the size of the subjets becomes small, eventually having each subjet containing only one hadron (identified as a topocluster) that was used to reconstruct the jet. The latter scenario is undesirable as at such a small scale there is no physical interpretation behind the subjets and the jet substructure. The ability to model such behaviour with the MC is also lost as the fragmentation effects of QCD take over. The plot of C_{had} was generated in the same manner as was for data, $p_{T,jet} > 5$ GeV and $|\eta| < 2.8$, and is shown in Figure 5-3. The shape that forms as y_{cut} becomes small occurs since in the PYTHIA samples the parton level number of subjets begin to plateau whereas the hadron level values continue to rise. This exemplifies the inability for PYTHIA to model the fragmentation effects at low y_{cut}.

In order to ensure that the NLO QCD calculations can be used in comparisons to data the value of C_{had} must not be too large. The value of $\langle n_{sbj} \rangle$ for NLO QCD performed with strictly parton level generators, such as NLOJET++, is restricted to being less than or equal to 3 [35][36]. This restriction on $\langle n_{sbj} \rangle$ does not hold for more general MC generator (that can simulate showering), such as PYTHIA. Therefore any y_{cut} value that reproduces the parton level to greater than 3 subjets is deemed too heavily influenced by non-perturbative effects which will lead to an improper correction at the hadron level and will not be used in the analysis. The values for both the hadronic and post parton shower $\langle n_{sbj} \rangle$ are given in Table 5-1.

Table 5–1: The mean subjet multiplicity for the PYTHIA hadronic and parton levels.

y_{cut}	0.1	0.05	0.01	0.005	0.001	0.0005
$\langle n_{sbj}^{part} \rangle$	1.0006	1.0928	1.8262	2.2308	3.4427	4.072
$\langle n_{sbj}^{had} \rangle$	1.0011	1.1253	2.0176	2.4903	3.8849	4.616

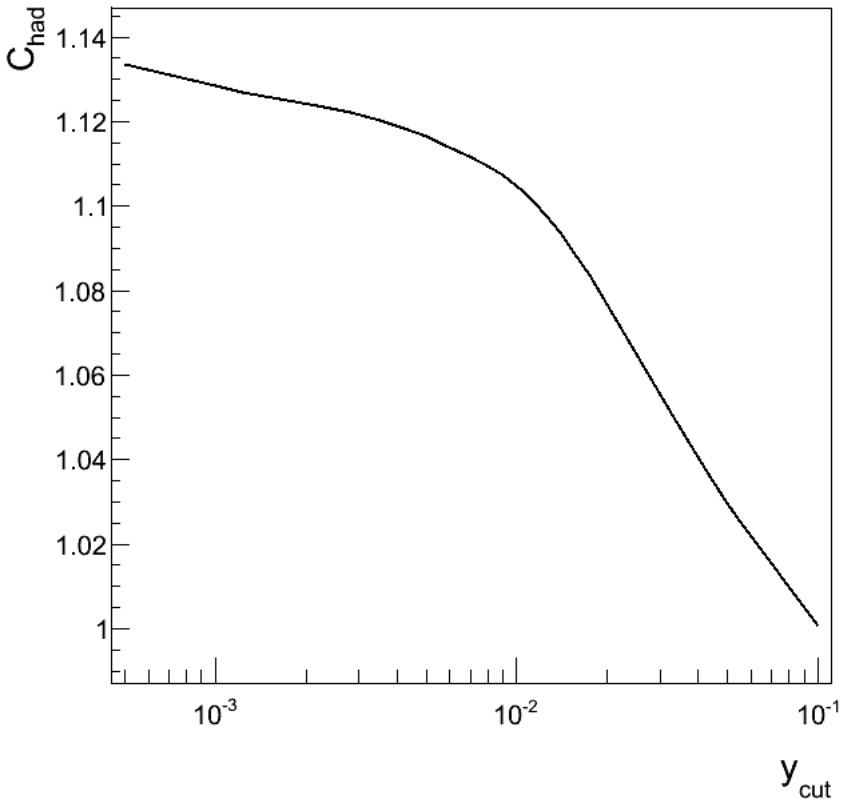

Figure 5–3: The hadronisation correction factor, C_{had}, for comparisons between the MC hadron and parton levels. The mean number of subjets $< n_{sbj} >$ was calculated in the same manner as was for data, $p_{T,jet} > 5$ GeV and $|\eta| < 2.8$. At high y_{cut} values C_{had} is small and grows as y_{cut} is lowered. The plateauing effect is caused by PYTHIA's inability to model the parton level [see text].

This problem was avoided by selecting a sufficiently high $p_{T,jet}$ and a relatively high y_{cut}, i.e. values in which we have an appropriate C_{had} value. For this analysis we restrict it to values of $y_{cut} \geq 5 \cdot 10^{-3}$. In this region the hadronisation correction is still relatively small, without a pleateau shape, and so a reliable comparison of data to NLO QCD can be made and α_s can be extracted.

The following theoretical uncertainties were applied following [30]:

- the uncertainty in the NLO QCD calculations due to terms beyond NLO, estimated by varying MC inputs (μ_R between $Q/2$ and $2Q$), was $\approx 3\%$;
- the uncertainty in the NLO QCD calculations due to that in the hadronisation correction was estimated as half of the difference between the C_{had} obtained from PYTHIA and with another MC generator POWHEG. This values was smaller than 2% at $y_{cut} = 0.01$.
- the uncertainty in the NLO QCD calculations due to uncertainties in the proton PDFs can be estimated by repeating the calculations using additional PDF sets with slight variations from the central value. The differences were negligible.

5.3 Event Selection

The data used in this study was a small subset of the 2011 ATLAS dijet samples. A more complete analysis would include the use of the full dataset but requires extra care and attention for dealing with correction and scale factors. The data used in this analysis was subject to a list of event selection cuts. The order in which these were applied is listed as:

1. Require at least one primary vertex (PV) with at least 5 tracks,
2. Require trigger to pass on the lead jet of the event (ensures the event contains ≥ 1 jet),
3. Require all jets to pass quality cuts or else the event is rejected,
4. Require the LAr error to be avoided [37],
5. Require event to pass the ATLAS Good Run List (GRL)

The first requirements ensures that the event is well reconstructed and that the origin of the collision is well defined. A trigger was then applied to the jet with the

highest p_T, i.e. the lead p_T jet, following the recommendation of the ATLAS jet grooming team as seen in Table 5-2. The trigger cut also ensured that the event contained at least one jet. The name convention (EF_j*_a4tc_EFFS) for the trigger implies that it was applied on the calorimeter topoclusters (tc), generated by the anti-k_T jet algorithm (a4), and used at the EF scale (EFFS). The numeric value (ex: j10) in the tag indicates the p_T threshold trigger value on the lead jet of the event. If the event satisfied the trigger it was kept.

Table 5–2: The trigger applied on an event by event basis according to p_T range of the lead jet of the event.

$p_{T,lead\,jet}$ range (GeV)	Trigger Name
22 - 30	EF_j10_a4tc_EFFS
30 - 40	EF_j15_a4tc_EFFS
40 - 55	EF_j20_a4tc_EFFS
55 - 70	EF_j30_a4tc_EFFS
70 - 95	EF_j40_a4tc_EFFS
95 - 125	EF_j55_a4tc_EFFS
125 - 162	EF_j75_a4tc_EFFS
162 - 215	EF_j100_a4tc_EFFS
215 - 280	EF_j135_a4tc_EFFS
280 - 350	EF_j180_a4tc_EFFS
350 - 1000	EF_j240_a4tc_EFFS

All of the jets reconstructed in the event were subject to jet quality cuts. These standard cuts are based upon calorimeter/reconstruction performance and any event with a jet deemed poor or bad was excluded from the analysis. A small number of events were removed if they were recorded while errors were present in the liquid Argon detector. For each event a LAr flag is used to indicate if any problems were occurring during the recording of the event, in particular issues with noise bursts and data integrity errors in the LAr calorimeter. Finally, the runs used were made to

coincide with the ATLAS GRL which describes the quality of a run and if it should be used in analysis.

All jets used in this study were subject to further constraints. A minimum p_T on the jet reconstruction of 5 GeV was applied. This restriction makes the reconstructed jets both infrared and collinear safe, as per issues described in chapter 4. To avoid low calorimetry resolution, jet analysis was constrained to only the hadronic central barrel and end caps, i.e $|\eta| < 2.8$.

To see how each cut affects the systematics of the study three tables are shown below. Table 5-3, 5-4, 5-5 list the number of events, number of jets, and the value of $\langle n_{sjb} \rangle$ at y_{cut} of 0.01 depending on the cuts applied, respectively. The tables show which event selection most greatly affects the results and number of jets and events. With the entire event selection applied, 18.6% of the events and 17.5% of the jets are kept. As can be seen, the largest criteria for keeping an event was the trigger cuts. This can be caused by either the trigger not passing or simply the event having no jet. For the purpose of statistical errors to be dealt with later the uncertainties on the following values are not considered.

From Table 5-5 the maximum deviation of the mean subjet multiplicity for $y_{cut} = 0.01$ is 2.3% and varies for each y_{cut} value. This maximum deviation for each y_{cut} value will be taken as one of the systematic errors for $\langle n_{sbj} \rangle$.

5.4 Data Corrections

The uncorrected or raw distribution of n_{sbj} in the data is compared to the prediction of PYTHIA simulation for several values of y_{cut} in Fig. 5-4. A satisfactory description of the data is provided by the simulation in the selected y_{cut} regions,

Table 5–3: Percentage of events kept per event selection cut turned off. When a value corresponds to one or more labels, that indicates those cuts are turned off in the analysis.

PV	Trigger	Jet Quality/LAr	% kept
X			19.2
	X		93.4
		X	19.0
X	X		97.3
X		X	19.7
	X	X	95.8
			18.6
X	X	X	100.0

Table 5–4: Percentage of jets kept per event selection cut turned off. When a value corresponds to one or more labels, that indicates those cuts are turned off in the analysis.

PV	Trigger	Jet Quality/LAr	% kept
X			19.0
	X		90.5
		X	17.9
X	X		90.5
X		X	19.5
	X	X	93.2
			17.5
X	X	X	100.0

Table 5–5: Value of $\langle n_{sbj} \rangle$ at $y_{cut} = 0.01$ per event selection cut. All values have yet to have the correction factor applied.

PV	Trigger	Jet Quality/LAr	% kept
X			2.098
	X		2.081
		X	2.101
X	X		2.079
X		X	2.097
	X	X	2.081
			2.101
X	X	X	2.078

thus validating the use of these MC samples to correct the measured mean subjet multiplicity, $\langle n_{sbj} \rangle$, to the hadron level. The number of subjets increases as y_{cut} is reduced, as previously discussed and as is also seen in the figure.

For a comparison to be made between MC and data, the data taken at the detector level (calo) must be unfolded to the hadron level (had). The mean subjet multiplicity corrected for detector effects was determined bin-by-bin as

$$\left\langle n_{sbj} \right\rangle = K \left\langle n_{sbj} \right\rangle_{calo} , \tag{5.7}$$

where the correction factor was defined from MC simulations as

$$K = \frac{\left\langle n_{sbj} \right\rangle_{had}^{MC}}{\left\langle n_{sbj} \right\rangle_{calo}^{MC}}. \tag{5.8}$$

This correction factor was evaluated for each value of y_{cut}.

The subscript *calo* (*had*) indicates that the mean subjet multiplicity was determined using the calorimeter cells (hadrons) at the detector level. The calorimeter cells are produced by running the MC hadrons through the detector simulation. The

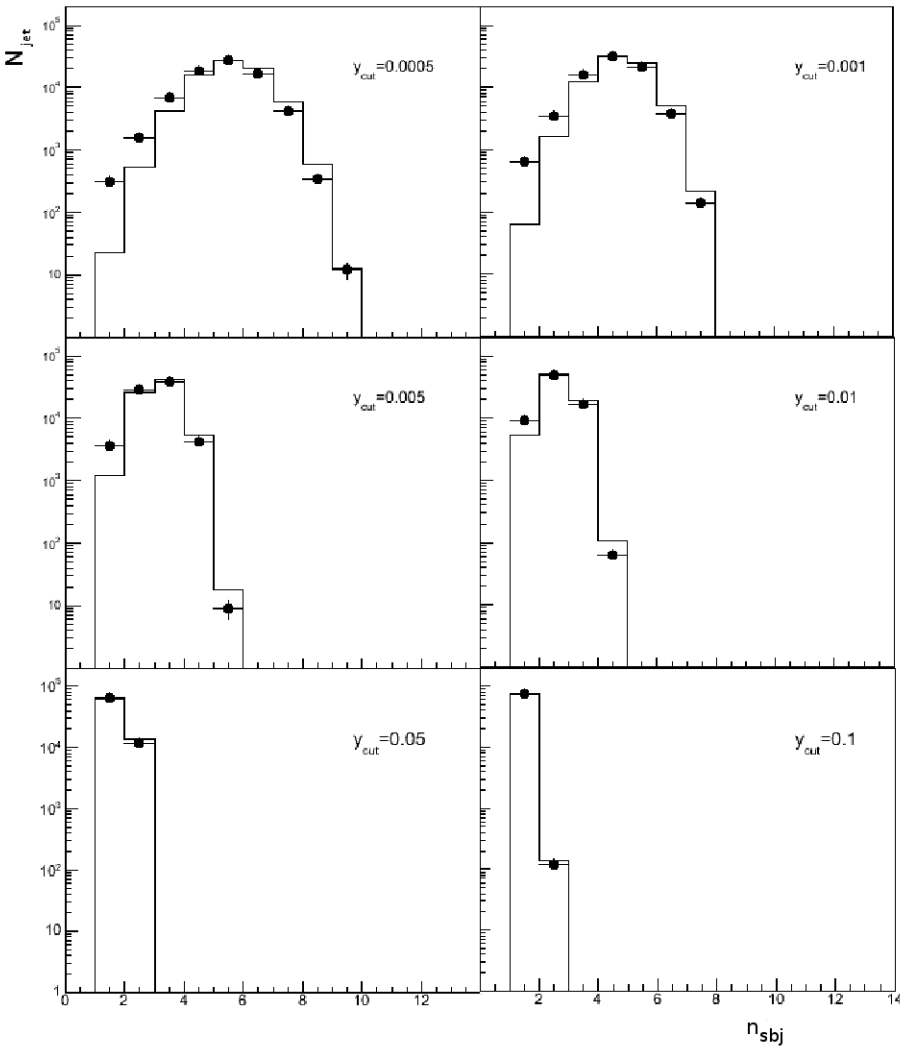

Figure 5–4: Distribution of the number of subjets within a jet at different y_{cut} values. Jets were produced inclusively with $p_{T,jet} > 5GeV$ and $|\eta| < 2.8$ from proton-proton collisions. The bins in n_{sbj} are identified by the values of their left edges. The error bars show the statistical uncertainty, which in most cases are too small to be seen. For comparison, the predictions from the PYTHIA simulation are also shown with their area normalised to the data.

interaction of the hadrons with the calorimeter is well modeled as is seen in the comparison between the calorimeter level MC and data. The deviation of the correction

factor K from unity expresses how well the detector is able to account for real particles, as would occur in our detector. This value was less than 8% for $y_{cut} \geq 5 \cdot 10^{-3}$ and decreased as y_{cut} was increased as seen in Table 5-6.

Table 5–6: The detector level to hadron level correction factor, K, as determined from MC calorimeter and hadronic samples. The last row contains the uncertainty in the correction factor.

y_{cut}	0.0005	0.001	0.005	0.01	0.05	0.1
$\langle n_{sbj} \rangle^{MC}_{calo}$	5.1091	4.2604	2.7039	2.1940	1.1811	1.0019
$\langle n_{sbj} \rangle^{MC}_{had}$	4.6167	3.8846	2.4903	2.0176	1.1253	1.0011
K	0.904	0.912	0.921	0.920	0.953	0.999
$\pm\delta K$	0.002	0.002	0.002	0.002	0.003	0.001

The following sources of systematic uncertainty on the measurement of $\langle n_{sbj} \rangle$ were considered:

- The differences in the results obtained by using PYTHIA or POWHEG to correct for detector effects. POWHEG failed to reproduce the data at low y_{cut} but at higher values resulted in an uncertainty that average around 3%.
- The 4% uncertainty in the jet energy scale (JES) [38] as measured with in-situ techniques was estimated by comparisons in literature resulted in an uncertainty less than 2%.
- The uncertainty in the simulation of the trigger and in data/event selection as investigated in the Section 5.3. This uncertainty was determined to be 2.2% and taken as a single value as there was little deviation on the basis of y_{cut}.

The first contribution is a function y_{cut}, whereas the two other contributions are fixed values. The final systematic uncertainty is obtained by adding the above uncertainties in quadrature.

5.5 Measurement of Mean Subjet Multiplicity

The mean subjet multiplicity was measured for jets in the events that passed the data selection requirements, as described in Section 5.3. All jets were corrected for detector effects as detailed in Section 5.4. For a jet to be included in the study it was required that to have a $p_{T,jet} > 5$ GeV and in the central or end cap hadron calorimeter, $|\eta_{jet}| < 2.8$. These regions offer the highest resolution and avoid gaps found in the calorimeter at transitions from one calorimetry component to another. The lower requirement on p_T was necessary for the study to have small fragmentation effects such that the subjet multiplicity is calculable in pQCD [30].

The mean subjet multiplicity, $\langle n_{sjb} \rangle$, is shown as a function of y_{cut} in Figure 5-5, and increases as y_{cut} decreases. This behaviour is expected as y_{cut} is related to the size of the subjets. Reversedly, an increase in the size of subjets will lead to less subjets being reconstructed. This information is also presented in Table 5-7. In addition, the mean subjet multiplicity is also plotted as a function of $p_{T,jet}$ at $y_{cut} = 0.01$ in Figure 5-6. The measured mean subjet multiplicity decreases as $p_{T,jet}$ increases. This behavior is in agreement with literature on jet substructure [39]. It was observed that jets became more narrow as their momentum increased. Thus low p_T jets will tend to be larger and develop more internal structure allowing for more subjets to be reconstructed in contrast to the high p_T jets that will contain less subjets on average being highly collimated.

The measurements over y_{cut} are compared with MC simulations provided by PYTHIA (with its default PDF tune called MRSTMcal) and also shown in Figure 5-5. The MRSTMcal PDF tune is a specific set of inputs that are used in PYTHIA

with an input α_s values of 0.1149 [40]. The PYTHIA predictions overestimate the observed mean subjet multiplicity at low y_{cut} and converge to the data at high y_{cut}.

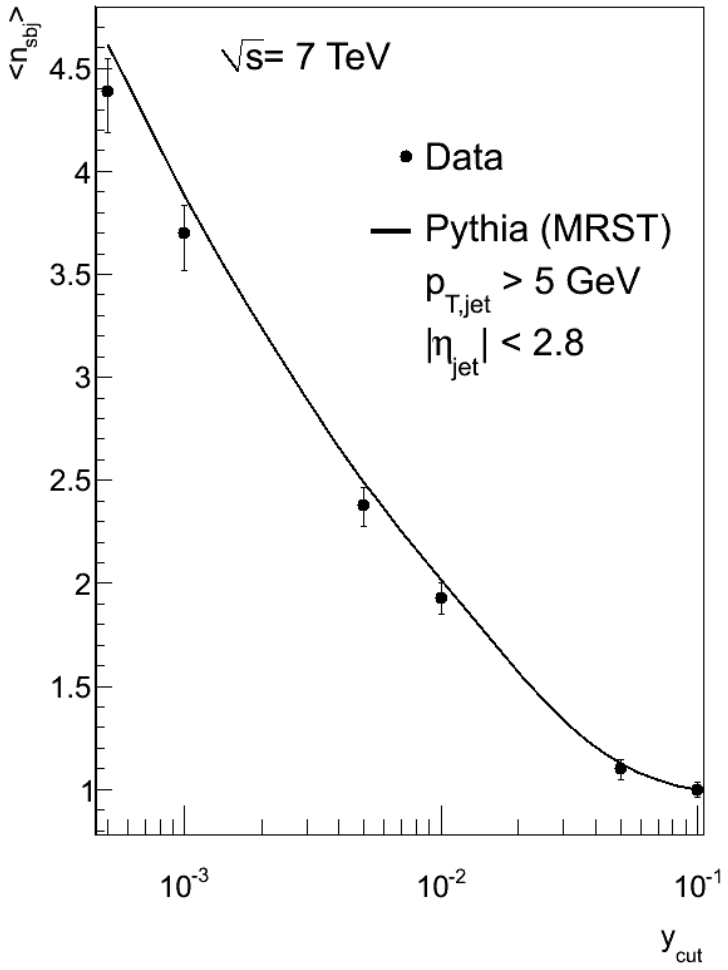

Figure 5–5: The mean subjet multiplicity corrected to the hadron level, $\langle n_{sbj} \rangle$, as a function of y_{cut} for inclusive jet reconstruction with $p_{T,jet} > 5GeV$ and $|\eta| < 2.8$ (represented as dots). Only the statistical uncertainties are shown. For comparison, MC hadron level predictions provided by PYTHIA with the MRSTMcal PDF tune ($\alpha_s = 0.1149$) [31] are shown. (represented as a solid line).

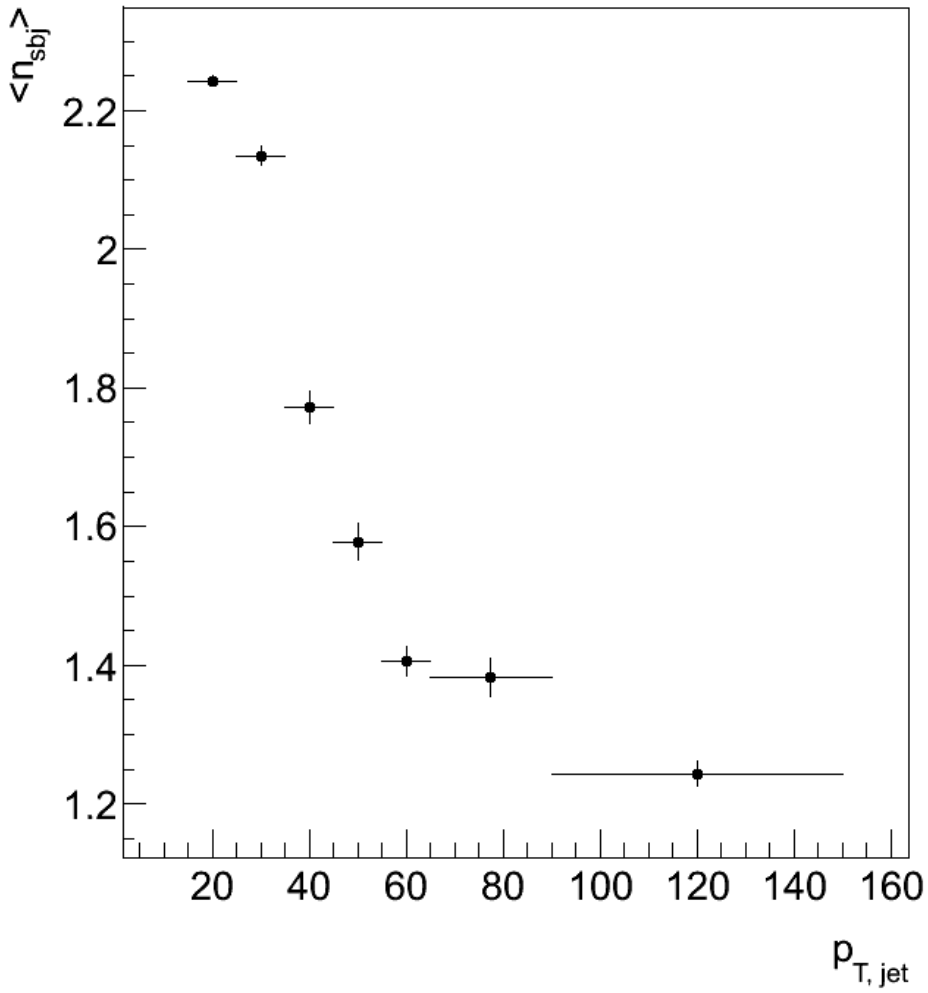

Figure 5–6: The mean subjet multiplicity, $\langle n_{sbj} \rangle$, as a function of $p_{T,jet}$ at $y_{cut} = 10^{-2}$ for inclusive jet reconstruction with $p_{T,jet} > 5GeV$ and $|\eta| < 2.8$ (represented as dots). The mean subjet multiplicity decreases as the p_T of the jet increases.

The PYTHIA predictions are repeated, corrected for hadronisation effects, and reweighted using the CT10 PDFs sets. These PDFs range in their input of α_s from 0.113 up to 0.123. The PDF sets and more information about them are available from the Les Houches accord PDFs (LHAPDF) [31]. To change from the initial PDF

Table 5–7: Mean subjet multiplicity measurement as a function of y_{cut}. Statistical and systematic uncertainties are also shown.

y_{cut} (GeV)	$\langle n_{sbj} \rangle$	Δ_{stat}	Δ_{syst}
0.0005	4.3863	± 0.0099	$^{+0.1582}_{-0.2002}$
0.001	3.7016	± 0.0084	$^{+0.1335}_{-0.1819}$
0.005	2.3803	± 0.0060	$^{+0.0858}_{-0.1033}$
0.01	1.9323	± 0.0052	$^{+0.0697}_{-0.0818}$
0.05	1.1042	± 0.0032	$^{+0.0398}_{-0.0543}$
0.1	1.0008	± 0.0003	$^{+0.0361}_{-0.0361}$

(MRSTMcal) to the final PDFs (CT10) reweighting must be done because each PDF will have different probabilistic outcome depending on the flavors of quarks involved in the MC collision. The weight, calculated on an event-by-event basis, from the two interacting partons is given by

$$w^i = \frac{f_{PDF_f}(f_1, x_1, Q^2)}{f_{PDF_i}(f_1, x_1, Q^2)} \cdot \frac{f_{PDF_f}(f_2, x_2, Q^2)}{f_{PDF_i}(f_2, x_2, Q^2)}, \tag{5.9}$$

where f, x, and Q represent the parton's flavour, the parton's momentum fraction, and the event's energy scale, respectively. The subscripts i and f indicate the initial PDF (MRSTMcal), and final PDF, whereas the numbered subscripts, 1 and 2, indicate the two partons in the scattering process. Finally, f_{PDF} is the function that solves for the PDF's value with given inputs of f, x, and Q. The CT10 PDF tunes with α_s variations of 0.113, 0.118, and 0.123 were used. The PDF reweighting was done as these three PDFs cover the full CT10 α_s range [30]. With the proper

weights applied these predictions were compared to the data and are shown in Figure 5-7.

The PDF reweighting scheme allows for the parametrisation of the pQCD predictions. The tunes do not cover a large range in $< n_{sbj} >$ which implies it is not very sensitive to the PDF tune change. There is however enough of a change to extract the coefficients from Eq. 5.3. The overall description of the data by the QCD calculations in the region $5 \cdot 10^{-3} < y_{cut} < 10^{-1}$ is good and therefore it is expected that the measurements can be used to determine the value of α_s.

As was previously discussed, the hadronisation correction and the n_{sbj} plots were used to validate the analysis cut of $y_{cut} \geq 5 \cdot 10^{-3}$. Theoretical reasons for disagreement in the region of small y_{cut} is that the QCD calculations require a resummation of terms that are enhanced by $\ln y_{cut}$ in order to perform a more precise comparison to data [28]. This disagreement is included in the analysis as part of the theoretical uncertainty in the NLO QCD calculations due to terms beyond NLO.

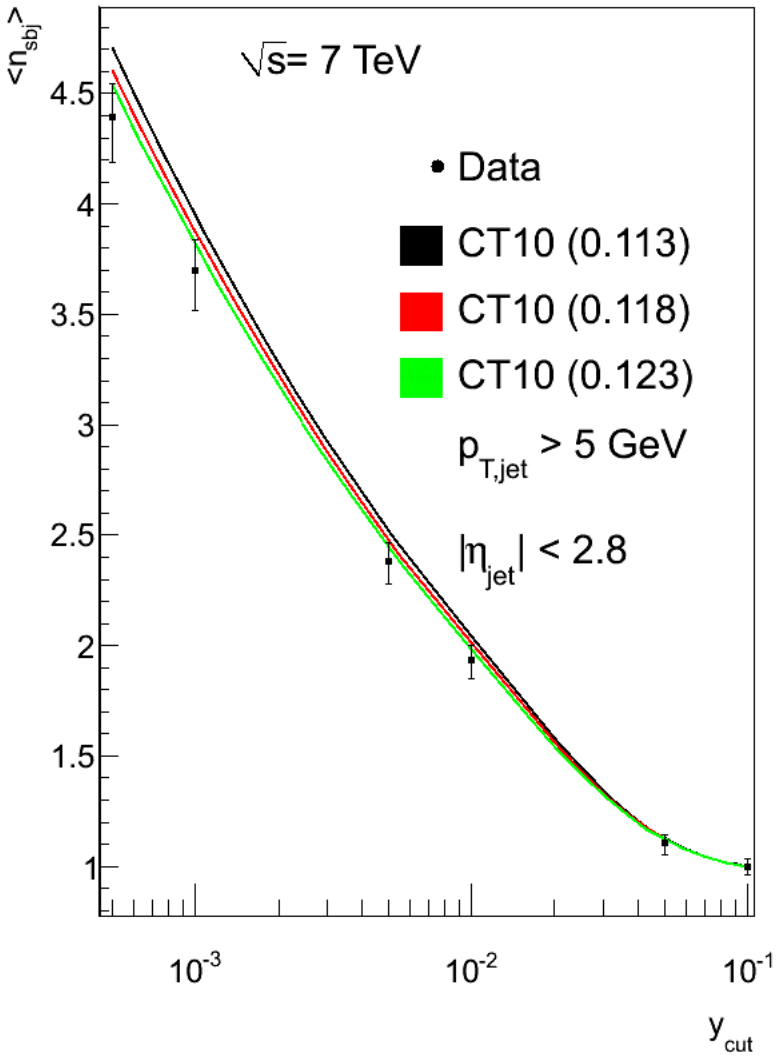

Figure 5–7: The mean subjet multiplicity, $\langle n_{sbj} \rangle$, as a function of y_{cut} for both data corrected for detector effects and the MC hadron level samples reweighted to the CT10 PDFs. The three variations $\alpha_s = 0.113, 0.118, 0.123$ were used. The data and the MC samples have similar shape. The α_s variations will be used to determine the NLO QCD parametrisation coefficients.
i

CHAPTER 6
Determination of α_s

The measurements of $\langle n_{sbj} \rangle$ for $5 \cdot 10^{-3} < y_{cut} < 10^{-1}$ for $p_{T,jet} > 5$ GeV were used to extract α_s. The reasons for the limits on y_{cut} value were justified primarily for hadronisation correction limitations and poor MC agreement to data at lower y_{cut} values, as explained in Section 5.3.

6.1 Procedure

The procedure for determining α_s is as follows:

- NLO QCD calculations of $\langle n_{sbj} \rangle$ were performed using PYTHIA reweighted with 3 sets of the CT10 PDFs. The input value of α_s for each parton level interaction was that associated with the individual sets of PDFs used. The values used in this study were $\alpha_s = 0.113$, 0.118, and 0.123.
- For each value, i, in y_{cut}, the QCD calculations corrected to the hadron level were used to parametrise the α_s dependence of $\langle n_{sbj} \rangle$ according to

$$\left[\langle n_{sbj} \rangle\right]^i = 1 + C_1^i \alpha_s + C_2^i \alpha_s^2, \tag{6.1}$$

  as was discussed in section 5.2.
- The coefficients C_1^i and C_2^i were determined by fitting Eq. (6.1) to the QCD predictions along with their uncertainties. Figure 6-1 shows the parametrisation coefficients and the corrected data $\langle n_{sbj} \rangle$ for each y_{cut} value.

- The value of α_s was then determined by fitting Eq. (6.1) with the now known coefficients to the measurements of $\langle n_{sbj} \rangle$. The fit was performed over the space $(\langle n_{sbj} \rangle, C_1, C_2)$ with α_s as a free parameter and for $5 \cdot 10^{-3} < y_{cut} < 10^{-1}$. The resulting fit overfit the data with a $\chi^2 = 0.89$ for 3 degrees of freedom, as seen in Figure 6-2.

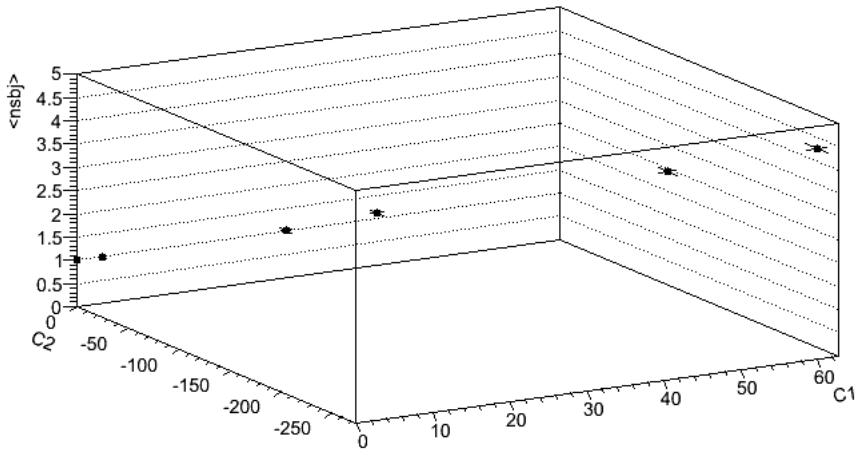

Figure 6–1: The NLO QCD parametrisation plot for $5 \cdot 10^{-4} < y_{cut} < 10^{-1}$. Each point ($C_1$, C_2, $\langle n_{sbj} \rangle$) represents a y_{cut} value that will be used for the determination fit for α_s. The y_{cut} values are shown in decreasing order, i.e the farthest left point is $y_{cut} = 0.1$. The two points farthest right are outside of our valid y_{cut} region and are only included for completeness.

6.2 Extraction of α_s

This method correctly handles the implicit α_s that arrives from use of the PDFs. The systematic uncertainties arise primarily from the measurements on $\langle n_{sbj} \rangle$, as was calculated in section 5-4. The theoretical uncertainties on α_s come primarily from the uncertainties in the hadronisation correction and terms beyond NLO as discussed in section 5.2. The total uncertainties were accounted for by combining individual

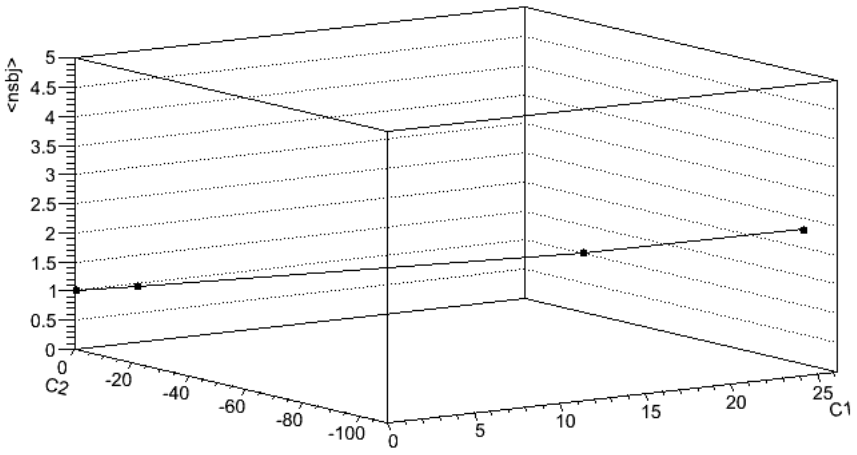

Figure 6–2: The α_s determination fit for $5 \cdot 10^{-3} < y_{cut} < 10^{-1}$. The free parameter, α_s, in the fit, $\langle n_{sbj} \rangle = 1 + C_1\alpha_s + C_2\alpha_s^2$, is given as 0.148 ± 0.022. Each point represents a y_{cut} value which are shown in decreasing order, i.e the farthest left point is $y_{cut} = 0.1$.

components in quadrature. Both the partial results, for single y_{cut} values, and the final result along with uncertainties are listed in Table 6-1.

Reasons for such a large uncertainty primarily arise from a lack of statistics and also the result of PDF reweighting rather than doing a complete re-generation with the appropriate PDF. The overfitting occurs as the fit used predominantly describes the error rather than the underlying relation and is a remnant of the large errors placed on the C_1, C_2 coefficients. The fit used to solve for these coefficients was generated with comparison to the NLO QCD predictions. These values could be improved by reducing the statistical error on such calculations by increasing the number of events.

The final value for α_s for $p_{T,jet} > 5$ GeV over the y_{cut} range $5 \cdot 10^{-3} - 10^{-1}$ is

Table 6–1: The α_s values extracted from the QCD fit to the measured mean subjet multiplicity over $5 \cdot 10^{-3} < y_{cut} < 10^{-1}$ for $p_{T,jet} > 5$ GeV. The last row shows the result from combining the four higher y_{cut} regions. All uncertainties are shown separately.

y_{cut}	α_s	Δ_{stat}	Δ_{syst}	Δ_{theory}
$5 \cdot 10^{-4}$	0.146	0.021	$^{+0.004}_{-0.007}$	± 0.005
10^{-3}	0.146	0.022	$^{+0.004}_{-0.007}$	± 0.005
$5 \cdot 10^{-3}$	0.147	0.029	$^{+0.005}_{-0.006}$	± 0.005
10^{-3}	0.150	0.036	$^{+0.005}_{-0.006}$	± 0.005
$5 \cdot 10^{-2}$	0.165	0.134	$^{+0.005}_{-0.008}$	± 0.006
10^{-1}	0.166	0.120	$^{+0.005}_{-0.006}$	± 0.006
$(5 \cdot 10^{-3}) - 10^{-1}$	0.148	0.022	$^{+0.005}_{-0.008}$	± 0.005

$$\alpha_s = 0.148 \pm 0.022\ (\text{stat.})^{+0.005}_{-0.008}\ (\text{syst.}) \pm 0.005\ (\text{th.}) \tag{6.2}$$

This result is consistent with the Particle Data Group (PDG) value of $\alpha_s = 0.1184 \pm 0.0007$ [41], primarily due to the large uncertainties in this study. There are several explanations for the large uncertainties. Firstly, the extraction of α_s depends largely on all of the PDFs used in our parametrisations of $\langle n_{sbj} \rangle$ that were calculated from both our original and reweighted MC samples, the MRST and CT10 tunes, respectively. As such, this means we are placing all of the PDF's on equal footing and while different PDFs with similar α_s values exhibit similar behavior, they do not produce identical results [31]. This leads to the question of whether a reweighting procedure is sufficient in these regards or whether the MC samples must be re-created with such PDFs as initial inputs through such programs as NLOjet++. Another

reason is that the smaller y_{cut} values bare more impact. Figure 5-7 shows the $\langle n_{sbj} \rangle$ converges to 1 as y_{cut} increases, therefore the lower y_{cut} values, offering more contrast between PDF sets, lead to being the more dominant factor in the parametrisation and hence the α_s extraction. It is interesting to note that comparative studies that look at $\langle n_{sbj} \rangle$ as a function of jet p_T rather than y_{cut} do not have this feature, i.e. the lines do not converge at any point and are approximately uniformly distant from each other. Finally, the NLO QCD theoretical fit, Eq. (5.3), used on the MC samples is more useful as an interpolation tool than as an extrapolating one when predicting the multiplicity behaviour. The extrapolation takes place, to extract the α_s value, because of the insensitivity of the PDF reweighting procedure. It should be noted that issues associated with non-uniformity, reweighting, and extrapolation are not present in studies based on mean subjet multiplicity as a function of jet transverse momentum p_T and thus may be a remnant of the decision to perform the analysis against the resolution parameter y_{cut}. Several methods for improvements and possible solutions to these issues are addressed below.

Improvements that could be done in future work include use of a devoted parton level generator such as NLOjet++ to get the parton level predictions rather than the the PDF reweighting scheme that was shown to be both insensitive, and included additional errors from the corrections needed. For higher precision, more work on the theoretical higher-order contributions would also need to be looked into. One such way to avoid the complexities of dealing with y_{cut} and p_T dependence in the extractions would be to choose a valid y_{cut} value and perform the study as a function of transverse momentum instead, as done in [30]. This approach also benefits in

that the α_s extraction depends on each p_T bin more evenly than was the case using y_{cut}. It can also be hypothesised that the above issues would not appear and a more accurate fit of the MC would be possible. Also having a larger α_s range in MC would alleviate the issue of needing to extrapolate at all, as one could interpolate instead. And finally, being a statistical study, the analysis would need to be performed over many more events. Such improvements would be the next step in such an analysis and may be worked on in future work.

CHAPTER 7
Conclusion

The mean subjet multiplicity for jets produced at proton-proton collisions at a center of mass energy of 7 TeV was measured. The jets were identified with the invariant anti-k_T algorithm in the laboratory frame and were required to have $p_{T,jet} > 5$ GeV and $|\eta| < 2.8$. As the transverse momentum of the jet increases or the resolution parameter of the jet increased, the average number of subjets with a jet was found to decrease.

MC simulations provided the NLO QCD calculations that agreed well with the data. This agreement suggested that the internal structure of the jets were well described by the MC models of the perturbative QCD. The extraction of the strong coupling constant was performed by correcting the parton level predictions to the hadron level and comparing them to the data, that had been corrected for detector effects. The final state parton radiation process has very little dependence upon the proton PDF and therefore, by reweighting the predictions by a range of PDFs, α_s was extracted. Several issues found in the study were identified and possible solutions listed.

A fit of the mean subjet multiplicity for $p_{T,jet} > 5$ GeV for $5 \cdot 10^{-3} \leq y_{cut} \leq 10^{-1}$ gives

$$\alpha_s = 0.148 \pm 0.022 \text{ (stat.)}^{+0.005}_{-0.008} \text{ (syst.)} \pm 0.005 \text{ (th.)} \tag{7.1}$$

References

[1] David Griffiths. *Introduction to Elementary Particles*. Wiley-VCH.

[2] J. Beringer et al. (Particle Data Group). Phys. Rev. D86, 010001 (2012).

[3] Michael E. Peskin Schroeder and Daniel V. *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[4] Aad, G.; Abajyan, T.; Abbott, B.; Abdallah, J.; Abdel Khalek, S. et al. Physics Letters B vol. 716 (1) p.1-29 (2012).

[5] The ATLAS Collaboration. *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*. December 2008.

[6] Yuri Dokshitzer. *Basics of Perturbative QCD*. Atlantica Séguier Frontières, 1991.

[7] Standard Model of Particle Physics. http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[8] C Glasman. *Precision measurements of αs at HERA*, volume 110. Journal of Physics: Conference Series, May 2008.

[9] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. *Parton fragmentation and string dynamics*, July 1983.

[10] LHC Design Report Volume I, 2006. http://ab-div.web.cern.ch/ab-div/Publications/LHC-DesignReport.html.

[11] ATLAS Detector and Physics Performance Technical Design Report, 1999. http://atlasinfo.cern.ch/Atlas/GROUPS/PHYSICS/TDR/access.html.

[12] G Aad, E Abat, and J Abdallah. *The ATLAS experiment at the CERN large hadron collider*. Journal of Instrumentation, 2008.

[13] http://public.web.cern.ch/public/en/research/AccelComplex-en.html *CERN Accelerator Complex Diagram*.

[14] http://lhc machine-outreach.web.cern.ch/lhc-machine ou. *A schematic of the LHC*

[15] R.M. Brown and D.J.A. Cockerill. *Electromagnetic calorimetry.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 666:47–79, February 2012.

[16] C. Fabjan. Calorimetry in high-energy physics. *Techniques and Concepts of High Energy Physics-III*, 1985.

[17] Richard Fernow. *Introduction to Experimental Particle Physics.* Cambridge University Press, New York, 1986.

[18] J. Beringer et al. (Particle Data Group). *Passage of particles through matter.* Physical Review D, D86(010001):19, 2012.

[19] Priscilla B. Cushman. *Instrumentation in High Energy Physics.* World Scientific, 1993.

[20] Thomas Ferbel. *Experimental Techniques in High Energy Physics.* Addison-Wesley Publishing Company, Inc, 1987.

[21] F. Corriveau. "Calorimetry." Slides from PHYS 620 lecture. McGill University (2011).

[22] Richard Wigmans. *Calorimetry - Energy Measurement in Particle Physics.* Clarendon Press, 2000.

[23] D. Schouten. *Jet energy calibration in atlas.* Master's thesis, Simon Fraser University, 2007.

[24] http://cms.web.cern.ch/news/jets-cms-and-determination-their-energy-scale. *Stages of Jet Evolution.* CMS Collaboration

[25] *Jet Reconstruction, Atlas Detector, and Using Proton-proton Collisions.* Atlas note. 2010.

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

[27] ATLAS collaboration. *Performance of large-R jets and jet substructure reconstruction with the ATLAS detector*, ATLAS-CONF-2012-065, 2012

[28] Michael H. Seymour. *The average number of sub-jets in a hadron collider jet.* Nuclear Physics B, 421(3):545–564, 1994.

[29] M. H. Seymour. *Jets in Hadron Collisions.* arXiv:0712.2447. page 13, July 2000.

[30] S. Chekanov, D. Krakauer, J.H. Loizides, and et Al. *Measurement of subjet multiplicities in neutral current deep inelastic scattering at HERA and determination of αs.* Physics Letters B, 558(1-2):41–58, April 2003.

[31] M.R. Whalley, D. Bourilkov, and R.C. Group. *The Les Houches accord PDFs (LHAPDF) and LHAGLUE.* arXiv:hep-ph/0508110. 2005.

[32] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. *A Brief Introduction to PYTHIA 8.1.* arXiv:0710.3820. October 2007.

[33] S. Agostinelli, J. Allison, K. Amako, and Et al. *Geant4a simulation toolkit.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506(3):250–303, July 2003.

[34] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX.* Journal of High Energy Physics, 2010(6):043, June 2010.

[35] S. Catani and M.H. Seymour. *A general algorithm for calculating jet cross sections in NLO QCD.* Nuclear Physics B, 485(1-2):291–419, February 1997.

[36] Zoltán Nagy. *Next-to-leading order calculation of three-jet observables in hadron-hadron collisions.* Physical Review D, 68(9):12, November 2003.

[37] Levêque Jessica Benjamin Trocmé. Talk on "LAr Data Quality: What CP groups need to know about the data", 2011.

[38] D. Schouten, A. Tanasijczuk, M. Vetterli, and for the ATLAS Collaboration. *ATLAS Jet Energy Scale.* arXiv:1201.2429 January 2012.

[39] J. Breitweg et al. *Measurement of jet shapes in high-Q deep inelastic scattering at HERA.* The European Physical Journal C, 8(3):367–380, May 1999.

[40] A. Sherstnev and R. S. Thorne. *Different PDF approximations useful for LO Monte Carlo generators.* arXiv:0807.2132 July 2008.

[41] J. Beringer et al. (Particle Data Group). *Quantum chromodynamics*. Physical Review D, D86(010001):29, 2012.