MEASURING THE QUARK AND GLUON JET ENERGY RESPONSE IN PROTON-PROTON COLLISIONS AT 7 TEV CENTER-OF-MASS ENERGY WITH THE ATLAS DETECTOR

by

David J Shinkaruk
B.Sc.(Hons.), University of British Columbia, 2010

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN THE DEPARTMENT OF PHYSICS FACULTY OF SCIENCE

© David J Shinkaruk 2013

SIMON FRASER UNIVERSITY
Spring 2013

All rights reserved. However, in accordance with the Copyright Act of Canada, this work may be reproduced, without authorization, under the conditions for Fair Dealing. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review, and news reporting is likely to be in accordance with the law, particularly if cited appropriately.
APPROVAL

Name:       David J Shinkaruk
Degree:     Master of Science
Title of thesis:  Measuring the Quark and Gluon Jet Energy Response in Proton-Proton Collisions at 7 TeV Center-Of-Mass Energy with the ATLAS Detector
Examinining Committee:  Dr. J. Steven Dodge, Associate Professor (Chair)

__________________________________________
Dr. Michel Vetterli
Senior Supervisor
Professor

__________________________________________
Dr. Dugan O’Neil
Supervisor
Associate Professor

__________________________________________
Dr. Sarah Johnson
Internal Examiner
Senior Lecturer

Date Approved:       April 25, 2013
Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website (www.lib.sfu.ca) at http://summit.sfu.ca and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, British Columbia, Canada

revised Fall 2011
Abstract

The jet energy scale is one of the largest systematic uncertainties in physics analyses at ATLAS, making it vital to understand and minimize. This thesis provides a comprehensive study of the quark and gluon jet responses, by comparing results of the missing transverse energy projection fraction method in the established gamma+jet events and the newly developed analysis of dijet events. A likelihood discriminator is used to tag jets according to their calorimeter response, and correct the response to one for the dijet study. The mean energy of quark and gluon tagged jets is shown to differ by 4-7%, depending on the energy of the jet.
This thesis is dedicated to my wife, Kelly.
For her endless love, support and encouragement.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approval</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Contents</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2 Calorimetry</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Electromagnetic calorimetry</td>
<td>4</td>
</tr>
<tr>
<td>2.1.1 Photon interactions</td>
<td>5</td>
</tr>
<tr>
<td>2.1.2 Electrons and positrons</td>
<td>8</td>
</tr>
<tr>
<td>2.1.3 Electromagnetic cascade</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Hadronic showers</td>
<td>11</td>
</tr>
<tr>
<td>2.2.1 Electromagnetically decaying particles</td>
<td>12</td>
</tr>
<tr>
<td>2.2.2 Visible non-electromagnetic energy</td>
<td>13</td>
</tr>
<tr>
<td>2.2.3 Invisible energy</td>
<td>14</td>
</tr>
<tr>
<td>2.3 Detecting energy deposition</td>
<td>15</td>
</tr>
<tr>
<td>2.3.1 Calorimeter response</td>
<td>17</td>
</tr>
</tbody>
</table>
3 Experimental setup

3.1 The LHC ......................................................... 19
3.2 Coordinate system ............................................... 20
3.3 Inner detector ..................................................... 21
   3.3.1 Semiconductor tracking ................................... 22
   3.3.2 Straw tube tracking ....................................... 22
   3.3.3 Material distribution of the inner detector .............. 23
3.4 Calorimeter ....................................................... 25
   3.4.1 Electromagnetic barrel calorimeter ....................... 25
   3.4.2 Hadronic calorimeter .................................... 27
3.5 Muon spectrometer system ...................................... 27
3.6 Triggers and readout ............................................ 32
3.7 Pileup ............................................................ 33

4 Jet construction and measurement .......................... 36

4.1 Jet formation ..................................................... 36
   4.1.1 Jet flavor ................................................ 37
   4.1.2 Hadronization and fragmentation ......................... 40
   4.1.3 Jet reconstruction ....................................... 41
4.2 Calorimeter response ........................................... 45
   4.2.1 Flavour response ......................................... 46
   4.2.2 Monte-Carlo response .................................... 48
   4.2.3 MPF response ............................................ 49

5 Defining hadronic tag jet candidates ....................... 54

5.1 Optimizing the datasets ........................................ 54
   5.1.1 Dijet selection .......................................... 61
5.2 Statistical methods to identify a tag jet ................... 64
   5.2.1 Frequentist .............................................. 65
   5.2.2 Binning ................................................ 67
   5.2.3 The \( V_i \) templates .................................... 74
5.3 The \( \alpha \) templates ........................................... 76
   5.3.1 Ensuring the MC describes the data variables .......... 81
6 Calorimeter response 93
  6.1 Quark and gluon fraction calculation ............................... 93
  6.2 Validating the dijet MPF $\alpha$ correction ........................... 94
  6.3 MPF ................................................................. 97
    6.3.1 MET ............................................................ 99
    6.3.2 The dijet and $\gamma$+jet MPF response ........................... 99
    6.3.3 Quark and gluon calculated response curve ...................... 103

7 Conclusion 108

Appendix A Error Analysis 110
  A.1 MPF bins .......................................................... 110
  A.2 Quark and Gluon Fraction ............................................ 114
  A.3 Fitting ............................................................. 114
  A.4 Quark and Gluon $R_{MPF}$ .......................................... 115

Bibliography 117
List of Tables

1.1 Properties of leptons. The given masses are approximate, and the neutrino masses are greater than zero. ............................................. 3
1.2 Properties of quarks, where the given masses are approximate. In addition, each quark has an associated colour (red, green, blue), which allows them to couple through the strong interaction. .............................. 3
1.3 Forces of the Standard Model are described by gauge bosons, where the given masses are approximate. The gluon has eight different colour-anticolour configurations. ........................................... 3

2.1 This table shows radiation and interaction lengths. ............................ 12

3.1 Various radial layers a particle traverses through the barrel (|η| < 1.375) of the detector [6]. ................................................................. 31

4.1 Colour factors for the various strong processes. ............................... 38
4.2 Squared matrix elements for $2 \rightarrow 2$ parton-parton subprocesses in QCD. $q$ and $q'$ denote distinct flavors of quark, and $\bar{q}$ denotes an anti-quark. $\theta^*$ is the c.m. scattering angle, $|M|^2$ is the squared matrix element (averaged over spin and colour), and $g^4 = 4\pi\alpha_s$ [20]. ................................................. 40
4.3 The maximum jet energy scale systematic uncertainties from a MC based study for anti-$k_T$ jets with $R = 0.6$ [9]. ................................. 46
4.4 Fundamental differences between quark and gluon jets. ......................... 48
5.1 After determining the bin in which the leading jet $p_T$ falls, the event is checked to see if the corresponding event trigger fired. If not, the event is not considered for analysis [19]. .............................................................. 56
5.2 The luminosity calculated per period and event level trigger [19]. ........... 57
5.3 Cleaning cuts for anti-$k_T$, $R = 0.6$ jets [9]. .................................................. 59
5.4 Standard jet properties given by ATLAS .......................................................... 66
## List of Figures

2.1 Photon interactions with matter as a function of incoming energy. $\sigma_{p.e.}$ is the photoelectric effect, $\sigma_{Rayleigh}$ is Rayleigh scattering, $\sigma_{Compton}$ is Compton scattering, $\kappa_{nuc}$ is pair production off a nucleus, and $\kappa_{e}$ is pair production off an atomic electron [11]. ............................... 6

2.2 The three dominant processes for photon interaction are both energy and $Z$ dependent. The lines indicate the boundaries between the dominant processes. For example, figure 2.1 describes photon interactions for lead ($Z=82$), where $\sigma_{p.e.} = \sigma_{Compton}$ at $\approx 0.95$ MeV and $\sigma_{Compton} = \kappa_{nuc}$ at $\approx 1.2$ MeV [3]. ............................... 7

2.3 Fractional energy loss per radiation length in lead as a function of electron or positron energy. $X_o(Pb) = 6.37 g/cm^2$ [3]. ............................... 9

2.4 Model of a simple EM shower. ............................... 11

2.5 Schematic representation of the response function of a non-compensating calorimeter to the EM and non-EM components of hadronic showers. The ratio of the mean values of these distributions is the $e/h$ value of this calorimeter (1.8) [3]. ............................... 18

3.1 Cut away view of the ATLAS detector [21]. ............................... 20

3.2 Eta is Lorentz invariant under addition and subtraction. ............................... 21

3.3 A three dimensional view of the ATLAS inner detector [21]. ............................... 23

3.4 A single track traverses the beryllium beampipe, the three cylindrical silicon-pixel layers, the four cylindrical double layers of barrel silicon-microstrip sensors (SCT), and approximately 36 axial straws of 4mm diameter contained in the barrel transition-radiation tracker [21]. ............................... 24
LIST OF FIGURES

3.5 Cut-away view of the ATLAS calorimeter system [21] .......................... 26
3.6 Sketch of a barrel module of the EM calorimeter where the various layers are clearly visible. The granularity in eta and phi of the cells of each of the three layers and of the trigger towers is also shown [21]. ......................... 28
3.7 The radiation length ($X_0$) of each EM calorimeter layer is given, as well for the material in front of the accordion [21]. ............................... 29
3.8 Schematic showing how the mechanical assembly and the optical readout of the tile calorimeter are integrated. The various components of the optical readout, namely the tiles, the fibres and the photomultipliers, are shown [21]. 30
3.9 Event with four reconstructed primary vertices, consistent with an event with four primary proton-proton collisions in the same beam-crossing. The z-coordinates of the vertices are (from left to right in the plot, in mm from the origin) -32.1, -22.9, 4.6, 18.8. The lines indicate tracks reconstructed from the inner detector [24]. .............................. 35
4.1 The fraction of the cross section from gluon-gluon, gluon-quark and quark-quark processes at $\eta_1 = \eta_2 = 0$ and $\sqrt{s} = 2\ TeV$ (solid) and $\sqrt{s} = 14\ TeV$ (dashed) as a function of $x_T = \frac{E_T}{\sqrt{s}}$, where $E_T$ is the outgoing parton’s transverse energy [13] .................................................. 39
4.2 The cluster (left) and string (right) hadronization models. PYTHIA uses string fragmentation, while HERWIG uses cluster fragmentation [9] .................. 41
4.3 The three regimes describing the evolution of a jet: parton level, particle level, and calorimeter level ......................................................... 42
4.4 A sample parton-level event (generated with HERWIG), superimposed on random soft energy simulating the performance of the anti-$k_T$ algorithm. The circular boundaries can be seen, and the shapes are independent of the soft energy distribution [16] ................................. 43
4.5 Collinear problem - replacing any massless parton with an exactly collinear pair of massless partons should not change energy of a jet, as it does in this example. ......................................................... 45
4.6 Infrared problem - adding a soft parton should not change the number of jets as it does in this example. ......................................................... 45
4.7 Quark and gluon jets radiate proportionally to their colour factors. $C_F$ denotes the strength of gluon coupling to quarks and $C_A$ denotes the strength of gluon coupling. 47

4.8 The fraction of the final state dijet partons flavor as a function of $p_T$ [15]. 48

4.9 Fraction of jet flavor from $\gamma+1$jet events as a function of $p_T$ [15]. 48

4.10 The leading order Feynman diagrams for parton-parton (dijet) scattering, where the time arrow can be to the left or right. 50

4.11 leading order Feynman diagrams for $\gamma+jet$ production. 51

5.1 The number of reconstructed primary vertices for dijet MPF events. 55

5.2 The mean number of interactions per bunch crossing for dijet MPF events. 55

5.3 The highest $p_T$ jet within a dijet event, with the MC shown in black and data in red. 58

5.4 The second highest $p_T$ jet within a dijet event, with the MC shown in black and data in red. 58

5.5 This drawing shows a low $p_T$ jet with a dead cell in the center. The true value is 10 GeV, and the cell level correction calculates the estimated value to be 5 GeV. This gives an underestimate of the jet $p_T$ by 25% [17]. 60

5.6 This drawing shows a high $p_T$ jet with a dead right cell. The true value is 25 GeV, and the cell level correction calculates the estimated value to be 50 GeV. This gives an overestimate of the jet $p_T$ by 17% [17]. 60

5.7 A central dijet event with the highest-$p_T$ jet collected by the end of April 2012: two central high-$p_T$ jets have an invariant mass of 3.62 TeV and the highest $p_T$ jet has a $p_T$ of 1.96 TeV. The event was collected on 9 April 2012. The lines indicate tracks reconstructed from the inner detector data. The green section is the EM calorimeter, with the hadronic calorimeter highlighted in red. The outer blue section is the muon spectrometer [24]. 62

5.8 The dijet balance is defined as $1 - (\text{Probe Jet } p_T)/(\text{Tag Jet } p_T)$, where the $p_T$ is measured at particle level. A lower measurement indicates a better momentum balance between the two jets in the transverse plane. 63

5.9 The number jets selected in PYTHIA as signal and background, plotted on a log scale as a function of $p_T$. 65
5.10 A signal jet favours a higher energy deposition in the EMB1 and EMB2 layer, and a lower energy deposition in the pre-sampler layers. 68
5.11 A signal jet favours lower energy deposition in the EMB3, TILE0 and TILE1 layers. 69
5.12 A signal jet favours lower energy deposition in the TILE2 layer. The combination of all calorimeter layers shows that signal jets favour a high electromagnetic fraction (EMFRAC) and a low longitudinal energy deposition profile (\((\text{Tile0+Tile2+EMB3+PRE}) / (\text{EMB1+EMB2})\)) 70
5.13 The EMB2 layer (3) absorbs the most energy over 90% of the time. The percent of energy deposited in the maximum layer and the jet length are excellent discriminating variables between signal and background, suggesting well measured jets are denser than the average jet. 71
5.14 The \(n_{90}\) variable confirms that well measured jets are denser than the average jet. The width and mass of a jet are both correlated to the jet opening angle, with signal jets favouring a smaller radius. 72
5.15 The number of towers is another way to measure the width of the jet, using the physical calorimeter structure and providing the same conclusion as the width variable. Signal jets favour fewer than average tracks, shown by the number of tracks and percentage of jet \(p_T\) carried by tracks. 73
5.16 The mean width variable is weakly correlated to surrounding activity, calculated by ISO08. 75
5.17 As the energy increases, the mean number of tracks associated with a jet becomes loosely correlated to ISO08. 75
5.18 The probability that a jet is a signal jet, \(P(S)\), is plotted for the likelihood inputs variables \(V_i\)'s\) for two different energy bins, and for ISO08 < 5%. 77
5.19 The probability that a jet is a signal jet, \(P(S)\), is plotted for the likelihood inputs variables \(V_i\)'s\) for two different energy bins, and for ISO08 < 5%. 78
5.20 The MC jet response after the frequentist likelihood has been applied. 79
5.21 Number of events corresponding to the frequentist likelihood shown in figure 5.20. 79
5.22 The mean MC response (black) is plotted along with the number of jets (red), with the calculated likelihood on the x-axis. No cuts on the jet $p_T$ are applied to produce this figure. ......................... 80
5.23 The $\alpha$ correction is the MC calorimeter response of a tag jet. The jets are binned according to their calculated likelihood, and binned in $p_T$. Each $p_T$ bin is fit according to a gaussian curve as shown in figure 5.24. The mean values from the gaussian fits are then fit with equation 5.11. ................. 82
5.24 Two $p_T$ bins from figure 5.23 are shown, demonstrating the gaussian nature of the calorimeter response in bins of likelihood. ......................... 83
5.25 The calculated width of a jet. ................................. 85
5.26 The first EM calorimeter layer beyond the Pre-Sampler. ................. 86
5.27 The primary EM calorimeter layer. ............................... 87
5.28 A ratio of calorimeter layers is used as a powerful discriminator. ......... 88
5.29 The percent of the jets $p_T$ that corresponds to tracks. ................... 89
5.30 From the maximum energy deposition calorimeter sampling layer, the percent of energy deposited in that layer is shown. ......................... 90
5.31 The shower length, where 1-7 indicate what layer the jet finished showering 96% of its energy. 1 corresponds to the pre-shower, and 7 corresponds to TILE2. ................................. 91
5.32 This figure is the projection of figure 5.21 on the likelihood axis for data and the PYTHIA MC simulation. The number of tag jets selected for a MPF study are binned in likelihood, where the likelihood is calculated as the negative logarithm of equation 5.8. ......................... 92
5.33 ATLAS’s pile-up estimation tool for pile-up for the tag jet in data and the Monte Carlo simulation. ................................. 92
6.1 The probe jet quark fraction for the $\gamma$+jet sample ....................... 95
6.2 The probe jet quark fraction for the dijet sample ........................... 95
6.3 The weighed average (according to sample size) of the probe-jet quark fraction. ................................. 96
6.4 The PYTHIA MC tag jet response, black is the parton response ($p_T^{\text{calorimeter}}/p_T^{\text{parton}}$), and red is the particle jet response ($p_T^{\text{calorimeter}}/p_T^{\text{particle}}$) ................................. 98
6.5 The HERWIG MC tag jet response, black is the parton response ($P_{T,\text{calorimeter}} / P_{T,\text{parton}}$), and red is the particle jet response ($P_{T,\text{calorimeter}} / P_{T,\text{particle}}$). ............................................. 98

6.6 The weighed average (according to sample size) of the tag jet MC response ($P_{T,\text{calorimeter}} / P_{T,\text{particle}}$), calculated from figures 6.4 and 6.5. ............................................. 99

6.7 Comparing the final MC response ($P_{T,\text{calorimeter}} / P_{T,\text{particle}}$) of the $\alpha$ corrected tag jet (black) to the MPF response for photons (red). These figures are for the PYTHIA MC generator. ............................................. 100

6.8 The angle in radians between the tag jet and the MET for data and Monte Carlo simulation; good agreement is seen. ............................................. 101

6.9 MET in dijet events for data and Monte Carlo simulation; good agreement is seen. ............................................. 101

6.10 The $\gamma$+jet MPF response curves. ............................................. 102

6.11 The dijet MPF response curves. ............................................. 102

6.12 The four inputs to calculate a quark and gluon MPF response curves from data. ............................................. 104

6.13 The four inputs to calculate a quark and gluon MPF response from the PYTHIA simulation. ............................................. 105

6.14 The calculated quark and gluon MPF response from data, alongside the PYTHIA $\gamma$+jet MPF curve. The difference in response between the two jet flavours range from 4-7%. The $\gamma$+jet MPF curve closely follows the calculated MPF quark curve. ............................................. 106

6.15 The calculated quark and gluon MPF response from PYTHIA, alongside the PYTHIA $\gamma$+jet MPF curve. The difference in response between the two jet flavours range from 4-6%. The $\gamma$+jet MPF curve closely follows the calculated MPF quark curve. ............................................. 107

A.1 The normalized MPF distributions for the $\gamma$+jet (red) and dijet (black) data sample, with MPF response on the x-axis. ............................................. 111

A.2 The normalized MPF distributions for the $\gamma$+jet (red) and dijet (black) PYTHIA sample, with MPF response on the x-axis. ............................................. 112

A.3 The normalized MPF distributions for the $\gamma$+jet (red) and dijet (black) HERWIG sample, with MPF response on the x-axis. ............................................. 113
Chapter 1

Introduction

The Standard Model of Particle Physics is a description of the fundamental constituents of matter and their interactions, providing a nearly complete framework$^1$ for the properties and interactions of all known sub-atomic particles. The Standard Model (SM) is built on fermions (leptons and quarks) and gauge bosons (mediating forces between fermions), shown in tables 1.1, 1.2, and 1.3. All SM particles are considered to be point-like, while possessing an internal spin (angular momentum) degree of freedom. Spin-1/2 fermions obey Fermi statistics, having the property that no two fermions can be in the same quantum state. Spin-0 and spin-1 particles obey Bose-Einstein statistics, which allow many particles in the lowest energy or ground state.

The six leptons and six quarks are divided into three generations, according to their masses. The known leptons are the electron (e), muon (μ), tau (τ), and their associated neutrinos (ν_e, ν_μ, ν_τ). The known quarks are the up (u), down (d), strange (s), charm (c), bottom (b), and top (t). Stable particles are built by combining quarks into baryons and mesons, composed of three quarks and a quark-antiquark pair respectively. The four known force-mediating particles of the SM are the photon, W^± and Z bosons, and the gluon. The photon is responsible for the electromagnetic interactions, while the W^± and Z bosons mediate the weak interactions. The gluon mediates the strong interactions, responsible for baryon and meson formation.

The Standard Model has been able to predict the top quark, tau neutrino, W^± and Z

---

$^1$Gravity is not part of the Standard Model.
One major aspect of the Standard Model that has not been experimentally verified is the Higgs boson, created by an excitation of the Higgs field\(^2\). The Higgs field has two charged and two neutral components, the \(W^\pm\) and \(Z\) bosons, and a Higgs boson. The Higgs boson was predicted almost 50 years ago, and is crucial for the Standard Model to explain where the mass of the \(W^\pm\) and \(Z\) bosons come from. The ATLAS (A Toroidal LHC Apparatus) experiment is one of two multipurpose detectors located at the LHC (Large Hadron Collider) at CERN in Geneva, Switzerland. ATLAS has been designed specifically to search for the Higgs particle. In 2012, ATLAS discovered a Standard Model Higgs boson candidate, with a mass of 126 GeV. The SM predicts the Higgs boson to have no internal spin, electric charge, or colour charge\(^3\). Extensive work must be undertaken in 2013 to confirm if the new particle is the SM Higgs boson.

One of the main components of the ATLAS detector is the calorimeter, the apparatus that measures the energy from the particles produced by the proton-proton collisions. It is the focus of this thesis to explore new techniques to improve the calibration of the calorimeter, and understand the particles as they traverse the detector. Chapter 2 introduces how calorimeters detect particles, with a detailed description of particle behaviour in a medium. The proton accelerator (LHC) and the detector (ATLAS) are described in Chapter 3. Chapter 4 defines a hadronic jet and the Jet Energy Scale (JES), specifically focusing on the difference between quark and gluon jets. Chapter 5 provides detailed information on the setup of the analysis, followed by the results in Chapter 6. It is the goal of this thesis to help reduce the Jet Energy Scale uncertainty at ATLAS by applying an existing technique, the Missing \(E_T\) Projection Fraction (MPF). Two types of events will be analyzed, making it possible to split the measured calorimeters response into quark and gluon samples.

\(^2\)The Standard Model predicts that the Higgs field has a non-zero value in all space.

\(^3\)Colour charge is unique to the strong force, discussed in section 4.1.1.
CHAPTER 1. INTRODUCTION

<table>
<thead>
<tr>
<th>Generation</th>
<th>Particle</th>
<th>Symbol</th>
<th>Spin</th>
<th>Charge (e)</th>
<th>Mass (MeV/c^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>Electron</td>
<td>e</td>
<td>1/2</td>
<td>-1</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Electron Neutrino</td>
<td>ν_e</td>
<td>1/2</td>
<td>0</td>
<td>&lt; 3 × 10^{-6}</td>
</tr>
<tr>
<td>2nd</td>
<td>Muon</td>
<td>μ</td>
<td>1/2</td>
<td>-1</td>
<td>105.7</td>
</tr>
<tr>
<td></td>
<td>Muon Neutrino</td>
<td>ν_μ</td>
<td>1/2</td>
<td>0</td>
<td>&lt; 0.19</td>
</tr>
<tr>
<td>3rd</td>
<td>Tau</td>
<td>τ</td>
<td>1/2</td>
<td>-1</td>
<td>1777.0</td>
</tr>
<tr>
<td></td>
<td>Tau Neutrino</td>
<td>ν_τ</td>
<td>1/2</td>
<td>0</td>
<td>&lt; 18.2</td>
</tr>
</tbody>
</table>

Table 1.1: Properties of leptons. The given masses are approximate, and the neutrino masses are greater than zero.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Particle</th>
<th>Symbol</th>
<th>Spin</th>
<th>Charge (e)</th>
<th>Mass (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>down</td>
<td>d</td>
<td>1/2</td>
<td>-1/3</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>up</td>
<td>u</td>
<td>1/2</td>
<td>2/3</td>
<td>0.003</td>
</tr>
<tr>
<td>2nd</td>
<td>strange</td>
<td>s</td>
<td>1/2</td>
<td>-1/3</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>charm</td>
<td>c</td>
<td>1/2</td>
<td>2/3</td>
<td>120</td>
</tr>
<tr>
<td>3rd</td>
<td>bottom</td>
<td>b</td>
<td>1/2</td>
<td>-1/3</td>
<td>430</td>
</tr>
<tr>
<td></td>
<td>top</td>
<td>t</td>
<td>1/2</td>
<td>2/3</td>
<td>174</td>
</tr>
</tbody>
</table>

Table 1.2: Properties of quarks, where the given masses are approximate. In addition, each quark has an associated colour (red, green, blue), which allows them to couple through the strong interaction.

<table>
<thead>
<tr>
<th>Force</th>
<th>Mediator</th>
<th>symbol</th>
<th>Spin</th>
<th>Charge (e)</th>
<th>Mass (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>gluons</td>
<td>g</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Electromagnetic</td>
<td>photon</td>
<td>γ</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Weak</td>
<td>charged</td>
<td>W^+, W^-</td>
<td>1</td>
<td>±1</td>
<td>80.4</td>
</tr>
<tr>
<td>Weak</td>
<td>neutral</td>
<td>Z^0</td>
<td>1</td>
<td>0</td>
<td>91.2</td>
</tr>
</tbody>
</table>

Table 1.3: Forces of the Standard Model are described by gauge bosons, where the given masses are approximate. The gluon has eight different colour-anticolour configurations.
Chapter 2

Calorimetry

An incoming particle can interact (through the strong, electromagnetic, or weak force) with a detector, losing a fraction of its energy. It is the function of the calorimeter to measure the energy and position of particles. Beyond energy and position measurements, the calorimeter is responsible for triggering and determining the missing energy of the event, discussed in sections 3.6 and 4.2.3 respectively.

A calorimeter shower\footnote{A collimated spray of particles from an incident parton or particle.} initiated by hadrons, such as protons or pions, is distinctly different and more complicated than an electromagnetic (EM) shower initiated by electrons or photons. The simpler EM shower is discussed in section 2.1, followed by the hadronic shower in section 2.2. The reader is referred to Ref. [3] for a more complete description of calorimetry.

2.1 Electromagnetic calorimetry

Electromagnetic calorimeters are designed to measure the energy of incoming photons, electrons, or positrons. An incoming photon, electron, or positron interacts with the Coulomb field of the atoms, initiating a cascade of particles. One multi-GeV electron may radiate thousands of photons on its way through the detector. The majority of these photons are very soft (low energy), and absorbed through Compton scattering or the photoelectric effect.
CHAPTER 2. CALORIMETRY

2.1.1 Photon interactions

The four processes responsible for energy loss of photons in matter are:

- Photoelectric effect
- Rayleigh scattering
- Compton scattering
- Pair Production

These processes dominate in different energy regimes, which can be seen in figure 2.1. The mean free path ($\lambda$) is defined as the average distance a photon travels between interactions. The probability that a photon will interact over a given distance $x$ is given by:

$$P(x) = 1 - e^{-x/\lambda} \quad (2.1)$$

**Photoelectric effect - low-energy phenomenon**

In the photoelectric effect, the photon is absorbed by an atom, and an electron is emitted. The now excited atom returns to the ground state by emission of an x-ray or Auger\textsuperscript{2} electron. This effect is dominant for energies below 0.7 MeV. The cross section\textsuperscript{3} falls as $\sigma \approx E^{-3}$ and has a $Z^5$ dependence on the material, where $Z$ is the atomic number.

**Rayleigh scattering - low-energy phenomenon**

In Raleigh scattering, the photon is deflected by atomic electrons. However, the photon does not lose any energy, only its direction is affected.

\textsuperscript{2}When a core electron is removed, an electron from a higher energy level falls into the vacancy. An energy transfer occurs between the electron that is falling to a lower energy level and a valence electron, ejecting the valence electron from the atom.

\textsuperscript{3}A measure of the probability that an encounter between particles will result in the occurrence of a particular atomic or nuclear reaction.
CHAPTER 2. CALORIMETRY

Figure 2.1: Photon interactions with matter as a function of incoming energy. \( \sigma_{p.e.} \) is the photoelectric effect, \( \sigma_{\text{Rayleigh}} \) is Rayleigh scattering, \( \sigma_{\text{Compton}} \) is Compton scattering, \( \kappa_{\text{nuc}} \) is pair production off a nucleus, and \( \kappa_{e} \) is pair production off an atomic electron [11].
Figure 2.2: The three dominant processes for photon interaction are both energy and $Z$ dependent. The lines indicate the boundaries between the dominant processes. For example, figure 2.1 describes photon interactions for lead ($Z=82$), where $\sigma_{p.e.} = \sigma_{\text{Compton}}$ at $\approx 0.95$ MeV and $\sigma_{\text{Compton}} = \kappa_{\text{nuc}}$ at $\approx 1.2$ MeV [3].

**Compton scattering - mid-energy phenomenon**

In Compton scattering, an incoming photon interacts with an atom, ejecting an electron. The photon is scattered at an angle $\theta$, and the difference in wavelength is given by

$$\lambda' - \lambda = \frac{h}{m_e c} (1 - \cos(\theta))$$

(2.2)

Compton scattering becomes the dominant process in the energy range between 0.1-10 MeV. The cross section falls as $\sigma \approx E^{-1}$.

**Pair production - high-energy phenomena**

In the field of a charged particle, a photon may create an electron-positron pair. The energy of the incident photon must be larger than twice the electron rest mass ($2 \times 0.511$ MeV). The cross section rises with energy and $Z$, reaching its asymptotic value at energies greater than 1 GeV.
2.1.2 Electrons and positrons

The two dominant processes for energy loss of electrons and positrons are ionization and bremsstrahlung, with minor contributions from other processes (Møller scattering \( (e^- + e^- \rightarrow e^- + e^-) \), Bhabha scattering \( (e^- + e^+ \rightarrow e^- + e^+) \)), as seen in Fig. 2.3. Ionization is the complete removal of an electron from an atom following the transfer of energy from a passing charged particle. If the knocked out atomic electron has enough energy to ionize further atoms through subsequent interaction of their own, they are referred to as delta rays. The energy loss (per distance \( dx \)) from ionization for electrons is given by

\[
\frac{dE}{dx} = \frac{2\pi n_e e^4}{mc^2} \left(2 \ln \frac{2mc^2}{I} + 3 \ln \gamma - 2 \right)
\]

where \( n_e \) is the density of electrons in the medium, \( I \) the ionization potential of the medium, \( \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} \) with \( v \) denoting the incoming velocity of the electron, \( e \) and \( m \) denote the electric charge and mass of the electron, and \( c \) is the speed of light.

Bremsstrahlung is the process of photon emission from an accelerated charged particle. The electron interacts with a Coulomb field (from a nearby atom), resulting in radiation due to the acceleration of a charged particle. Fernow [4] derives the average rate of energy loss due to bremsstrahlung as

\[
\frac{dE}{dx} = \frac{n_e}{Z} E \sigma
\]

with \( \sigma \) denoting the cross-section for the process. The ionization loss rate varies logarithmically with the electron energy through the \( \gamma \) term, where an electron will lose energy by bremsstrahlung at a rate nearly proportional to its energy [5]. The critical energy, \( E_C \), is defined as the energy where the loss due to radiation (bremsstrahlung) is equal to the loss from ionization [4],

\[
E_C = \frac{1600}{Z} mc^2
\]

where \( m \) is the particle mass. It can be seen from Eq. 2.5 that an electron (\( m = 0.511 \text{ MeV} \)) has \( E_C \approx \frac{800}{Z} \text{ MeV} \).
Figure 2.3: Fractional energy loss per radiation length in lead as a function of electron or positron energy. $X_o(Pb) = 6.37g/cm^2$ [3].
2.1.3 Electromagnetic cascade

The standard distance used to characterize the interaction of electromagnetic particles is a radiation length \(X_0\)^4. The radiation length describes the energy loss of photons, electrons and positrons in a material independent way. It is defined as the distance over which a high-energy \((\gg 1\text{GeV})\) electron or positron loses, on average, all but \((1 - e^{-1})\) of its energy to bremsstrahlung [3]

\[
X_o = \frac{716.4A}{Z(Z+1) \ln(287/\sqrt{Z})} \propto \frac{A}{Z^2}
\]  

where \(A\) is the atomic mass number. Equation 2.6 states that an electron will lose the same fraction of its energy in 18cm of water as in 2.8 mm of lead; both described by \(0.5X_o\).

A simple model of an electromagnetic shower can be built under the assumption that high-energy particles \((e^\pm\text{ and } \gamma)\) undergo an interaction within the same distance, illustrated in figure 2.4. The high-energy electrons or positrons \((E_o \gg E_c)\) undergo bremsstrahlung after one radiation length, emitting a photon.

\[
e^\pm \rightarrow e^\pm + \gamma
\]  

After another radiation length, the electron or positron undergoes another bremsstrahlung and the photon decays to an electron-positron pair.

\[
\gamma \rightarrow e^+ + e^-
\]  

In this simple model, any two daughter particles has half the energy of their parent. If the initial electron has energy \(E_o\), then after \(N\) radiation lengths a particle in the shower has an energy of

\[
E_N = \frac{E_o}{2^N}
\]  

The cascading process will stop abruptly when \(E_N \leq E_c\), and ionization is more probable than bremsstrahlung. At this point, the length of the shower can be written in terms of the initial and critical energies

\(^4\)The radiation length is usually measured in \(g/cm^2\), where dividing \(X_0\) by the density of the material (in \(g/cm^3\)) gives the length in cm.
CHAPTER 2. CALORIMETRY

Figure 2.4: Model of a simple EM shower.

\[ N_{\text{max}} \approx \ln \frac{E_o}{E_c} \]  \hspace{1cm} (2.10)

giving a total length of \( X_0 N_{\text{max}} \).

The width of an electromagnetic shower can be understood by multiple scattering, mostly in the early stages of the shower development, and Compton scattering beyond the shower maximum. The Moliere radius (\( \rho_M \)) is defined as the radius of a cylinder containing on average 90% of the shower’s energy deposition. It can be defined through the ratio of the radiation length and the critical energy:

\[ \rho_M = mc^2 \sqrt{\frac{4\pi}{\alpha}} \frac{x_0}{E_c} \approx 0.0265x_0(Z + 1.2) \approx \frac{A}{Z} \]  \hspace{1cm} (2.11)

Equations 2.6 and 2.11 show that the longitudinal development (\( X_0 \approx A/Z^2 \)) is more sensitive to the material’s atomic number than the lateral shape (\( \rho_M \approx A/Z \)).

2.2 Hadronic showers

A hadron is a composite particle made of quarks (and/or antiquarks) held together by the strong force. Hadrons are categorized into two families: baryons (made of three quarks) and mesons (made of one quark and one antiquark). The physical processes that cause the propagation of hadronic showers are considerably different from electromagnetic showers.
In general, the energy resolution of hadronic calorimeters is worse than electromagnetic calorimeters, primarily due to the complexity of the interactions through the strong force. The hadronic showering process is dominated by a succession of inelastic\textsuperscript{5} nuclear interactions, and can be subdivided into four components.

Due to the relatively frequent production of \(\pi^0\) and \(\eta\) that in turn decay to photons, there exists an electromagnetic component to hadronic showers. A sizeable amount of the non-EM component of the shower is converted into excitation and breakup of nuclei. A small fraction of this energy will eventually appear as a detectable signal. The remaining fraction of the shower’s energy from the incident particle is lost and does not result in an observable signal.

The longitudinal profile of hadronic showers is characterized by the nuclear interaction length (\(\lambda_{\text{int}}\)), which is the average distance a hadron travels before undergoing a nuclear interaction, analogous to the mean free path of a high-energy photon. Table 2.1 shows that the nuclear interaction length is longer than the radiation length, and that higher Z materials give a larger \(\lambda_{\text{int}}/X_0\) ratio. The larger the ratio, the larger the difference between the shower lengths for hadronic and electromagnetic showers. The shower length will become important in chapter 5.2.

### 2.2.1 Electromagnetically decaying particles

Pions are the most abundant type of particle produced in nuclear interactions: \(\pi^+\), \(\pi^-\) and \(\pi^0\) are created in roughly equal proportions in each interaction. Hadronic showers therefore have an EM component, primarily through the decay to photons of \(\pi^0\) and \(\eta\) particles which are also produced in the nuclear interactions. The fraction of the hadron

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|}
\hline
Material & Atomic No. & Radiation Length & Interaction Length & $\lambda_{\text{int}}/X_0$ \\
 & (Z) & (g/cm\textsuperscript{2}) - (cm) & (g/cm\textsuperscript{2}) - (cm) & \\
\hline
Aluminum & 13 & 24.0 - 8.9 & 106.4 - 39.4 & 4.4 \\
Iron & 26 & 13.8 - 1.8 & 131.9 - 16.8 & 9.5 \\
Lead & 82 & 6.37 - 0.56 & 194.0 - 17.1 & 30.5 \\
\hline
\end{tabular}
\caption{This table shows radiation and interaction lengths.}
\end{table}

\textsuperscript{5}The kinetic energy of an incident particle is not conserved.
shower that is electromagnetic varies significantly from event to event. On average, $\frac{1}{3}$ of the mesons produced in a given stage of the shower are $\pi^0$. With this assumption, the EM content of the shower ($f_{em}$) can be written as

$$f_{em} = 1 - (1 - f_{\pi^0})^n$$

(2.12)

where $n$ is the generation of the reaction, and $f_{\pi^0}$ is on average $\frac{1}{3}$. Wigmans [3] derives the average $f_{em}$ as

$$f_{em} \approx 0.10 \log(E)$$

(2.13)

showing that $f_{em}$ scales logarithmically with energy (GeV). Equations 2.12 and 2.13 show that the electromagnetic component of a hadronic shower increases with the number of stages in the shower development and therefore with the energy of the incoming hadron. The EM sub-showers initiated by $\pi^0$s develop in the same way as the electromagnetic showers described in section 2.1. Thus the $\pi^0$s require a much smaller detector volume to deposit their energy than other shower components carrying the same energy. Therefore, the energy density is considerably larger in the areas near $\pi^0$s.

### 2.2.2 Visible non-electromagnetic energy

The non-electromagnetic fraction of a hadronic shower can be approximated by $f_{\text{non-em}} \approx 1 - 0.10 \log(E)$ from Eq. 2.13, with its energy deposit approximately divided as:

1. Ionization by pions $\approx 20\%$
2. Ionization by protons $\approx 25\%$
3. Kinetic energy of evaporation neutrons\textsuperscript{6} $\approx 15\%$
4. Total invisible energy $\approx 40\%$

The majority of hadronically showering particles in ATLAS, $\approx 90\%$, are pions. The charged hadrons first ionize the medium, creating measurable energy. For lead ($\lambda \approx 194 \text{ cm}^2$), the average amount of energy that a charged hadron loses through ionization before

\textsuperscript{6}The soft or slow component of a hadronic shower.
interacting with a nucleus is \( \approx 180 \text{ MeV} \). A semi-classical calculation was done by Fer-now [4], showing that the average rate of ionization energy loss for a heavy, spin-0 particle is given by:

\[
\frac{dE}{dx} = c_o T \left( \ln \frac{2p^2}{mI} - \beta \right)
\]

(2.14)

where \( T, p, \) and \( m \) are the incoming hadron’s kinetic energy, momentum, and mass respectively. \( I \) is the ionization potential of the medium, \( \beta = v/c \) and \( c_o \) is a constant that is dependent on the material and the hadron’s properties.

### 2.2.3 Invisible energy

The term *invisible-energy* refers to energy that is not measurable by a calorimeter. Invisible energy can originate from delayed emitted photons in nuclear reactions, soft neutrons, binding energy, or neutrinos. The invisible-energy contribution of hadronic showers causes the hadronic calorimeter response and energy resolution to be worse than for an electromagnetic calorimeter.

### Nuclear spallation reactions

When a high-energy hadron interacts with a nucleus, the most probable reaction is *spallation*. Spallation is a two step process, a fast intranuclear cascade, followed by a slower evaporation stage.

1. (fast) The high energy hadron collides with nucleons inside the struck nucleus. Internal collisions occur (creating fast moving nucleons within the nucleus); pions and other unstable hadrons may be created at this step. Some of the particles taking part in this cascade reach the nuclear boundary and escape. Others get absorbed and distribute their kinetic energy among the remaining nucleons in the nucleus, resulting in the production of an excited intermediate nucleus.

2. (slow) The second step of the spallation reaction consists of the de-excitation of the intermediate nucleus. This is achieved primarily by evaporating nucleons, until the excitation energy is less than the binding energy of a single nucleon. The remaining energy, typically a few MeV, is released in the form of photons.
The energy spectrum of the evaporation nucleons from step two is considerably softer than the escaping nucleons from step 1. Most of the nucleons from step two are neutrons because the Coulomb barrier prevents soft protons from being released [3]. The evaporation neutrons carry an average kinetic energy of 2 - 3 MeV [3]. The absorption of soft neutrons depends on the strong (and sometimes weak) interaction. The energy needed to eject nucleons in spallation reactions (nuclear binding energy), is unmeasurable for the calorimeter.

**Neutrinos (ν)**

Neutrinos also contribute to invisible energy, almost always escaping the detector unmeasured. They do not carry electric charge, which means they do not leave an ionization trail. Neutrinos interact only through the weak force, with a typical neutrino capable of passing unaffected through several million kilometres of lead.

### 2.3 Detecting energy deposition

Calorimeters are subdivided into two broad classes, homogeneous and sampling. A homogeneous calorimeter is such that the absorber and the active (signal producing) medium are the same, while in a sampling calorimeter the absorbing and active media are interleaved. ATLAS uses sampling calorimeters, allowing only some fraction of the shower’s energy to be read by the active material.

The passive medium is a high-density material (iron, copper, lead, or tungsten) used to absorb energy and produce secondary particles, while the active medium detects the secondaries by generating the light or charge that forms the basis for the signals for the calorimeter. A sampling calorimeter is smaller and cheaper than a homogeneous calorimeter, at the cost of poorer energy resolution. In the barrel (central) section of the calorimeter, ATLAS uses a plastic scintillator\(^7\) for the active material in the hadronic calorimeter, and liquid argon to measure ionization in the electromagnetic calorimeter. The detector specific details are discussed in section 3.4.

---

\(^7\)A flash of light is produced in a transparent material by the passage of a particle.
CHAPTER 2. CALORIMETRY

Scintillation

Scintillation is an inherent property of many organic \(^8\) molecules and arises from their electronic structures. Scintillation also occurs in many inorganic materials, including salts, gases, and liquids. As the particles from the calorimeter shower enter the active medium, molecules and atoms are excited, and then emit light as they return to the ground state. The scintillation light is picked up by wavelength shifting plastic fibers \(^9\) and carried to photomultiplier tubes \(^10\) (PMTs). The PMTs then produce an electronic signal, which is digitized and recorded.

Ionization

As the energy of the shower particles drops below the critical energy, the ionization process dominates. Ionization can happen any time the incoming electron’s kinetic energy exceeds the ionization potential of the medium. ATLAS measures the ionization in the EM calorimeter from a lead / liquid argon sampling calorimeter, through the reaction

\[
Ar + e^- \rightarrow Ar^+ + 2e^-
\]  

(2.15)

where charged particles ionize the liquid. The ionization energy of argon is:

- Ar I Ground state (\(\ldots 3p^6 1S_0\)): 15.76 eV
- Ar II Ground state (\(\ldots 3p^5 2P_2\)): 27.63 eV

An applied field causes ions and electrons to separate and move to charged plates. The calorimeter then measures the energy deposited within the calorimeter from the induced current on the plates. If the ionizing electron has significantly more energy than the ionization potential of argon, a fragmentation reaction can occur

\[
Ar + e^- \rightarrow Ar^+ + 2e^- \rightarrow Ar^{39+} + n + 2e^-
\]  

(2.16)

---

\(^8\) Any gaseous, liquid, or solid chemical compounds whose molecules contain carbon.

\(^9\) A photofluorescent material that absorbs higher frequency photons and emits lower frequency photons, making it possible for the light to be transported to the PMT.

\(^10\) A vacuum tube with electrodes converts light into electrons that are collected to form an electronic signal.
where $Ar^{39+}$ denotes the argon isotope (18 protons, 21 neutrons, and 17 electrons), and $n$ is a neutron. The fragmentation described in equation 2.16 can continue and a molecule can fragment multiple times until all the excess energy is distributed.

### 2.3.1 Calorimeter response

The calorimeter response is usually defined as the average calorimeter signal per unit of deposited energy\(^\text{11}\). In a linear calorimeter, the response is constant, independent of the deposited energy; suggesting that a 100 GeV photon will have the same response as a 1000 GeV photon. Non-linearity can be an indication of instrumental problems such as signal saturation or shower leakage. More commonly, signal non-linearity is a result of hadronic shower detection. The invisible energy discussed in section 2.2.3 is a major contributing factor to non-linearity, which may easily lead to a response difference of 10% over one order of magnitude in energy [3]. The measured response, $r$, of a hadronic shower is given by [10] as:

\[
    r = h(1 - f_{em}) + e f_{em} \\
    = (e - h)f_{em} + h
\]

(2.17)

where $h$ denotes the hadronic energy response, $e$ is the electromagnetic energy response, and $f_{em}$ is given by Eq. 2.13. If the ratio of the EM and non-EM responses is not equal to one, the non-linear $f_{em}$ term survives.

Compensating calorimeters can be designed to make the $e/h$ ratio equal to 1. This can be done by using uranium as a medium, introducing fission reactions. Compensating calorimeters have the benefits of improving the linearity and resolution of calorimeters. ATLAS uses a non-compensating calorimeter, with an $e/h$ ratio of 1.36. Figure 2.5 illustrates the response function of a non-compensating calorimeter, where the distribution of the normalized signals around the mean value are plotted separately for the EM and non-EM components. For more details on calorimeter response, refer to section 4 or to Ref. [3].

\(^{11}\text{e.g. Number of photoelectrons/GeV, picoCoulombs/MeV}\)
Figure 2.5: Schematic representation of the response function of a non-compensating calorimeter to the EM and non-EM components of hadronic showers. The ratio of the mean values of these distributions is the e/h value of this calorimeter (1.8) [3].
Chapter 3

Experimental setup

The ATLAS detector is a multi-purpose detector, looking at proton-proton collisions from the Large Hadron Collider (LHC). It is the largest of the four detectors at the LHC, measuring 44 m in length, 25 meters in diameter, and weighing in at nearly 7000 tons. Figure 3.1 is a cut-away diagram of the detector, emphasizing the key components. ATLAS has three primary detection systems: the inner tracker, the calorimeter, and the muon system. This thesis focuses on the central (barrel) region of the detector, so that the forward detector systems are not covered in detail. For full details of the detector, see Ref. [6]

3.1 The LHC

The Large Hadron Collider (LHC) is a proton-proton synchrotron\(^1\) located at CERN in Geneva, Switzerland. The 27 km circumference ring collides protons at an unprecedented center of mass energy of \(\sqrt{s} = 8\) TeV in 2013. The LHC does not collide a steady stream of protons, but discrete bunches, with each bunch containing approximately \(10^{11}\) protons. The bunches are set to collide every 25 or 50 ns, requiring fast detector readout times.

\(^1\)A particle accelerator that accelerates charged particles to very high energies in the presence of an alternating electric field, while confined to a circular orbit.
CHAPTER 3. EXPERIMENTAL SETUP

3.2 Coordinate system

ATLAS uses a right-handed Cartesian coordinate system to define the position of particles within the detector, with the z-axis along the beampipe, the x-axis pointing to the centre of the LHC ring, and the y-axis pointing up. The azimuthal angle $\phi$ is the angle in the xy-plane with respect to the positive x-axis. The polar angle $\theta$ is measured from the positive z-axis. The origin of the coordinate system is defined as the center of the ATLAS detector, which coincides with the nominal interaction point. Transverse momentum ($p_T$) and energy ($E_T$) are defined in the (x,y) plane. In the relativistic limit, $\beta \to 1$, the pseudorapidity ($\eta$) is defined as,

$$\eta = \frac{1}{2} \left( \ln \frac{1 + \cos(\theta)}{1 - \cos(\theta)} \right) = -\ln \left( \tan \frac{\theta}{2} \right)$$

(3.1)

$\eta$ has the advantage of being Lorentz invariant under addition and subtraction. This becomes important because the reaction products are relativistically boosted and the lab frame does not coincide with the CM frame for the fundamental process being studied. Figure 3.2
Figure 3.2: Eta is Lorentz invariant under addition and subtraction.

shows how η in the lab frame is related to η* in the center of mass frame through the relationship

\[ η^* = \frac{1}{2}(η_1 - η_2). \]

3.3 **Inner detector**

Approximately 1000 particles emerge from the collision point every 50 ns within \( |η| ≤ 2.5 \), creating a very large track density in the detector [6]. It is the job of the inner detector to

- Identify tracks of individual particles,
- Match tracks to the interaction vertex (collision point),
- Measure charge and momentum,
- Trace the decay of very short-lived particles such as b quarks,
- Disturb the particles as little as possible, not to influence the subsequent measurement of the energy.

The inner detector is located in a 2T solenoid magnet and is composed of three sub-detectors: the Pixel Detector, the Silicon Tracker (SCT), and the Transition Radiation Tracker (TRT). The inner detector is further divided into barrel and endcap regions as seen in figures 3.3 and 3.4, providing coverage for \( |η| < 2.5 \).
CHAPTER 3. EXPERIMENTAL SETUP

3.3.1 Semiconductor tracking

The first two layers of the inner detector are based on semiconducting sensors. Absorbed energy from a traversing particle creates electron-hole pairs within the medium, the number of which is proportional to the material’s bandgap\(^2\). When electron-hole pairs form in silicon, the holes drift to measurement strips and the electrons drift to the backplane. The holes induce a negative charge in the strips, which gives rise to a current.

The Pixel Detector can be thought of as a very specialized 82 megapixel camera, capable of taking a picture every 25 nanoseconds. With 1.8 m\(^2\) of silicon and 80 million readout channels distributed over three barrel layers and 3 endcap disk layers, the excellent resolution of the Pixel Detector next to the beam pipe allows for pattern recognition, precise estimation of the impact parameter\(^3\) of tracks and lifetime tagging of particles. The thickness of each layer is 1.7 percent of a radiation length at normal incidence, with intrinsic accuracy for the position measurement in the barrel of 10 \(\mu m\) in the \(R\phi\) direction and 115 \(\mu m\) in \(z\).

The Semi-Conductor Tracker (SCT) is similar to the Pixel Detector but with long, narrow strips rather than small pixels. The SCT detector has 61 m\(^2\) of silicon, and 6.2 million readout channels distributed over four barrel layers and 9 disk layers. The intrinsic accuracies in the barrel are 16 \(\mu m\) in \(R\phi\) and 580 \(\mu m\) in \(z\).

3.3.2 Straw tube tracking

The final layer of the inner detector is based on drift tube technology. A cylindrical straw is filled with a gas that is ionized when a charged particle traverses it. The ions drift towards the walls (cathode) and the electrons diffuse towards the anode wire that is stretched along the axis of the cylinder. The Transition Radiation Tracker (TRT) radiates energy as a particle moves across the interface of two media with different dielectric constants; for ATLAS it is a xenon based gas mixture within 4mm diameter straws. A charged particle will momentarily polarizes the material nearby, and if the particle crosses a boundary where the

---

\(^2\)The energy difference (in electron volts) between the top of the valence band and the bottom of the conduction band.

\(^3\)The impact parameter of a track is defined as the transverse distance of closest approach of the track to the primary vertex point.
index of refraction changes, the change in polarization gives rise to the emission of electromagnetic transition radiation\textsuperscript{4} [22]. The TRT provides only $R\phi$ information with 420,000 electronic channels giving an intrinsic precision of 130 $\mu m$, up to $|\eta| = 2$. In the barrel region, the 1.44 meter long straws are parallel to the beam axis, with their wires divided into two halves at $\eta=0$. The detector geometry guarantees that particles cross 35-40 straws in a pseudorapidity interval from 0 to 2, providing continuous tracking at larger radii of the inner detector while enhancing its pattern recognition ability to better than 50 $\mu m$ at the LHC design luminosity.

\textbf{3.3.3 Material distribution of the inner detector}

It can be seen from table 3.1 that a particle travels though approximately 0.5 $X_0$ in the inner detector (within the barrel). This is followed by another 1$X_0$ from the liquid-argon

\textsuperscript{4}About one photon is emitted for every 100 boundaries crossed, for transitions between air and matter of ordinary density.
Figure 3.4: A single track traverses the beryllium beampipe, the three cylindrical silicon-pixel layers, the four cylindrical double layers of barrel silicon-microstrip sensors (SCT), and approximately 36 axial straws of 4mm diameter contained in the barrel transition-radiation tracker [21].
cryostats. The consequences of introducing material in front of the calorimeter are [6]:

1. Many electrons lose a significant fraction of their energy through bremsstrahlung before reaching the electromagnetic calorimeter.

2. Approximately 40 percent of photons convert into an electron-positron pair before reaching the LAr cryostat and the electromagnetic calorimeter.

3. A significant fraction of low-energy charged pions undergo an inelastic hadronic interaction inside the inner detector.

### 3.4 Calorimeter

The ATLAS calorimeter is composed of five subsystems, located in the barrel ($|\eta| < 1.475$), endcap ($1.375 < |\eta| < 3.2$), and forward ($1.7 < |\eta| < 5.0$) regions as seen in Fig. 3.5. The barrel region consists of a lead/liquid-argon sampling electromagnetic calorimeter, and a steel/scintillating plastic tile sampling hadronic calorimeter. The endcap calorimeters are all liquid-argon technology.

#### 3.4.1 Electromagnetic barrel calorimeter

The ATLAS electromagnetic calorimeter is a lead/liquid-argon (LAr) sampling detector with accordion geometry, giving $\phi$ uniformity. The absorbers are made of lead plates, with a thickness of 1.53 mm for $|\eta| < 0.8$. Liquid argon was chosen as the active calorimeter medium due to its well-known linear response to ionization, and its inherent resistance to radiation damage. As discussed in section 2.3, the lead plates generate particle showers and the LAr measures the ionization of the particles.

The solenoid that supplies the 2 T magnetic field for the inner detector is integrated into the barrel cryostat and is located in front of the EM calorimeter. The EM calorimeter is finely segmented both laterally ($\eta \times \phi$ space) and radially, and split into four radial segments (Table. 3.1 and Fig 3.6). At high energies, most of the EM shower energy is collected in the second layer. The first layer consists of finer segmented strips in the $\eta$-direction (but coarser
Figure 3.5: Cut-away view of the ATLAS calorimeter system [21].
in $\phi$), providing excellent $\gamma - \pi^0$ discrimination\(^5\) [6]. These two layers are preceded by a very finely segmented pre-sampler layer. The pre-sampler is installed immediately behind the cryostat cold wall, and is designed to correct for particles that have pre-showered within the inner detector. It is also an excellent discriminator between photons ($\gamma$) and $\pi^0$. The last layer of the EM calorimeter is designed to catch the tails of very energetic EM showers, and does not require fine segmentation. The average jet energy resolution is designed to be $\Delta E/E = 11.5\%/\sqrt{E} \oplus 0.5\%$ (E in GeV). Figure 3.7 shows how the radiation lengths of the ID and EM calorimeter are fairly uniform within the $\eta$ region studied in this thesis ($|\eta| < 0.8$).

### 3.4.2 Hadronic calorimeter

The hadronic barrel calorimeter ($|\eta| < 1.7$) surrounds the EM calorimeter, and uses organic plastic scintillator plates (tiles) embedded in an iron absorber. The calorimeter is 2 meters deep, corresponding to 7.4 radiation lengths. There are three radial sampling layers, approximately $1.5\lambda$, $4.1\lambda$, and $1.8\lambda$ thick at $\eta = 0$ respectively. The scintillators consist of 3 mm thick tiles, sandwiched between 4 and 5 mm thick iron plates. The scintillator tiles and absorber plates are stacked parallel to the incident particles; figure 3.8 shows the geometry. The barrel tile calorimeter has 10,000 readout channels, with a fast 50 ns readout time thanks to the plastic scintillators and photomultiplier tubes. The hermiticity of the ATLAS calorimeter system also allows for an excellent measurement of the missing transverse energy, a crucial measurement for the techniques used in this thesis. The average jet energy resolution is designed to be $\Delta E/E = 50\%/\sqrt{E} \oplus 3\%$ (E in GeV).

### 3.5 Muon spectrometer system

The muon spectrometer measures the momentum of muons based on track curvature induced by a azimuthal magnetic field. The magnetic field for the Muon Spectrometer is divided into barrel ($|\eta| \leq 1$) and endcap ($1.4 \leq |\eta| \leq 2.7$) regions, with a transition region spanning the gap ($1.0 \leq |\eta| \leq 1.4$). The spectrometer uses four detector technologies: mon-

\(^5\)The position resolution of the detector makes it possible to distinguish the two photons from the $\pi^0$ decay from a single photons.
Figure 3.6: Sketch of a barrel module of the EM calorimeter where the various layers are clearly visible. The granularity in eta and phi of the cells of each of the three layers and of the trigger towers is also shown [21].
Figure 3.7: The radiation length ($X_0$) of each EM calorimeter layer is given, as well for the material in front of the accordion [21].
Figure 3.8: Schematic showing how the mechanical assembly and the optical readout of the tile calorimeter are integrated. The various components of the optical readout, namely the tiles, the fibres and the photomultipliers, are shown [21].
### Table 3.1: Various radial layers a particle traverses through the barrel (|\(\eta\)| < 1.375) of the detector [6].

<table>
<thead>
<tr>
<th>longitudinal Segmentation</th>
<th>Granularity</th>
<th>Radiation Lengths ((\eta = 0))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inner Detector</strong></td>
<td>(R(\phi) x Z ((\mu)m))</td>
<td>(X_0)</td>
</tr>
<tr>
<td>beam-pipe</td>
<td></td>
<td>0.005</td>
</tr>
<tr>
<td>pixel layers</td>
<td>10 x 115</td>
<td>0.1</td>
</tr>
<tr>
<td>SCT</td>
<td>16 x 580</td>
<td>0.1</td>
</tr>
<tr>
<td>TRT</td>
<td>50 x N/A</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>Electromagnetic Calorimeter</strong></td>
<td>((\Delta\eta) x (\Delta\phi))</td>
<td>(X_0)</td>
</tr>
<tr>
<td>PreSampler</td>
<td>0.025 x 0.1</td>
<td>2.5</td>
</tr>
<tr>
<td>EMB1</td>
<td>0.003 x 0.1</td>
<td>4.3</td>
</tr>
<tr>
<td>EMB2</td>
<td>0.025 x 0.025</td>
<td>16</td>
</tr>
<tr>
<td>EMB3</td>
<td>0.05 x 0.025</td>
<td>2</td>
</tr>
<tr>
<td><strong>Hadronic Calorimeter</strong></td>
<td>((\Delta\eta) x (\Delta\phi))</td>
<td>(\lambda)</td>
</tr>
<tr>
<td>TILE1</td>
<td>0.1 x 0.1</td>
<td>1.5</td>
</tr>
<tr>
<td>TILE2</td>
<td>0.1 x 0.1</td>
<td>4</td>
</tr>
<tr>
<td>TILE3</td>
<td>0.2 x 0.1</td>
<td>1.8</td>
</tr>
</tbody>
</table>
CHAPTER 3. EXPERIMENTAL SETUP

Monitored drift tubes (MDT), cathode strip chambers (CSC), resistive plate chambers (RPC) and thin gap chambers (TGC). The MDT ($|\eta| < 2.7$) and CSC ($2.0 < |\eta| < 2.7$) subsystems are the primary measurement detectors for the muon tracks. Due to the higher event rate in the forward region, the CSC uses multiwire proportional chambers, while the MDT uses drift tube chambers. Both systems measure the ionization trail left by a muon. The RPC ($|\eta| < 1.05$) and TGC ($1.05 < |\eta| < 2.4$) muon subsystems are designed primarily for triggering, but do assist in momentum measurements. The designed momentum resolution ranges from 2 percent for a 100 GeV muon, to 8 percent for a 1 TeV muon within the central barrel.

3.6 Triggers and readout

Due to the high bunch crossing rate in the LHC, 40 MHz, the amount of data measured in ATLAS is too large to record every event. A three level trigger system is used to select only potentially interesting events: Level-1 (L1), Level-2 (L2), and event filter (EF).

The L1 trigger searches for signatures from high-$p_T$ particles, as well as high missing transverse energy events with large total transverse energy. Custom designed hardware implements a sliding window algorithm on groups of $0.2 \times 0.2$ ($\eta \times $ $\phi$) calorimeter towers, defining regions of interest (RoI) for the high level triggers (L2 and EF). The maximum L1 acceptance rate is 75 kHz. The L2 trigger is seeded by RoI from the L1 trigger, and in addition is given access to information from the entire detector. An algorithm groups the calorimeter cells into cones with a radius of 0.4 in $\eta \times $ $\phi$ space, allowing the L2 trigger to reduce the acceptance rate to below 3.5 kHz. Finally, the EF uses offline analysis procedures on fully-built events to further select events down to a rate of 200-400 Hz. The EF algorithm uses the full granularity and precision of the inner detector and the calorimeter to refine the trigger selection. A flexible trigger menu is used in ATLAS to allow for various types of events to be selected, depending on the analysis.
3.7 Pileup

Luminosity ($\mathcal{L}$) is a measure of the beam current and size, determining the event rate:

$$N_i = \mathcal{L} \sigma_i$$

It is measured in units of inverse-cross section per unit time ($cm^{-2}s^{-1}$), where $\sigma_i$ and $N_i$ are the cross section and the number of events for a given process $i$. The LHC delivers protons in a bunch-train scheme, where approximately $10^{11}$ protons are bunched together, separated by 50 ns in 2011. The LHC achieved a peak luminosity of $3.65 \times 10^{33} cm^{-2}s^{-1}$ and ATLAS recorded a total luminosity of $5.25 fb^{-1}$ for 2011, corresponding to approximately $3.57 \times 10^{14}$ collisions.

A proton-proton collision can be broken down into two distinct processes: the hard collision and the underlying event. Events studied in this thesis define the hard collisions as a $2 \rightarrow 2$ reaction, where two incoming partons collide within the two parent protons, producing two outgoing partons. Table 4.2 and figures 4.10 and 4.11 show the $2 \rightarrow 2$ processes considered for this thesis. The underlying event consists of the remaining partons, often continuing down the beam pipe.

When two tightly packed bunches of protons cross, it is highly probable that multiple hard collisions occur within the same bunch. When hard scattering or underlying event particles from different interaction points overlap within the calorimeter, the convolution of energy is known as pileup. Pileup can take two forms, in-time and out-of-time. In-time pileup occurs when the overlapping energy is from the same bunch, while out-of-time pileup is from the leading or trailing bunches which can also overlap because of the response time of the detectors.

One metric to describe the amount of in-time pile-up that occurs in any given event is the number of reconstructed primary vertices, $N_{PV}$. If $N_{PV}$ is large, then it can be assumed the detector is measuring a lot of background energy in addition to the hard scattering event of interest. Another way to gauge the amount of in-time pile-up is to measure the number of tracks ($N_{trkout}$) not associated with the primary vertex. $N_{trkout}$ is less sensitive to vertex reconstruction efficiencies, and may be a more precise measure of pile-up activity. Figure 3.9 shows an event with four primary proton-proton collisions in the same beam-crossing, where the lines indicate the particle tracks reconstructed by the inner detector.
Out-of-time pile-up is due to residual signal within the detector from previous events. Currently, there is no way in ATLAS to directly measure the number of interactions in preceding and following bunch crossings. Instead, the mean number of inelastic interactions per bunch crossing ($\mu$) is calculated directly from the instantaneous luminosity, inelastic cross-section, and beam parameters that contains the event of interest [9].

Signal shaping in the calorimeter readout is used to compensate for both types of pile-up. The pulses are designed such that the negative tail of the signal from out-of-time pileup cancels the positive signal from in-time pile-up, on average. However, this only works if the bunch intensity is constant over the integration time of the calorimeter. For example, at the beginning of a bunch train there is insufficient out-of-time pile-up to cancel the in-time pile-up, so the calorimeter response is systematically higher for these events. One problem in 2011 data is that the electronics are designed to provide the pile-up cancellation for 25 ns bunch spacing, while 50 ns was used. To correct this problem, a pile-up correction tool was created to correct jets, taking into consideration $N_{PV}$, $\mu$, and the location of the jet within the detector. This correction is implemented in the following analysis, and can be seen from figure 5.33 to have up to 10% impact on the measured jet $p_T$. 
Figure 3.9: Event with four reconstructed primary vertices, consistent with an event with four primary proton-proton collisions in the same beam-crossing. The z-coordinates of the vertices are (from left to right in the plot, in mm from the origin) -32.1, -22.9, 4.6, 18.8. The lines indicate tracks reconstructed from the inner detector [24].
Chapter 4

Jet construction and measurement

Quantum chromodynamics (QCD) is a theory in the Standard Model that describes the strong interaction between colour charged objects (quarks and gluons), providing the framework needed to understand jet formation. An essential feature of QCD is asymptotic freedom, stating that the coupling strength decreases at short distances. A complementary concept is quark/colour confinement which states that the coupling increases at longer distances. As a result, quarks behave as almost free particles inside hadrons but are never seen in isolation.

4.1 Jet formation

When two asymptotically free partons collide in a proton-proton collider, the final state partons will attempt to escape from their parent hadrons. However, quarks and gluons are not observed in nature as free particles due to colour confinement, which states that only colourless QCD objects (baryons and mesons) can be observed in nature. The parton will create a collimated spray of particles, referred to as a jet. Jets can be assigned a flavour in Monte Carlo simulations according to the parton that created the jet.
4.1.1 Jet flavor

Quarks and gluons are assigned an extra quantum number, colour, that is unique to the strong force, required to satisfy the Pauli Exclusion Principle\(^1\). Quarks will have an associated colour, and anti-quarks will have an anti-colour. The strong force gauge boson, the gluon, comes in eight different colour-anticolour configurations to accommodate any quark colour combination.

The colour factor becomes important in understanding the difference between quark and gluon jets. Quarks come in three colours (red, green, and blue), assigned to a triplet of the SU(3) colour group. The gluons, which mediate the QCD force between colour charges, come in eight different colour combinations: \(R\bar{G}, R\bar{B}, G\bar{R}, G\bar{B}, B\bar{R}, B\bar{G}, \sqrt{\frac{1}{2}}(R\bar{R} - G\bar{G}), \sqrt{\frac{1}{2}}(R\bar{R} + G\bar{G} - 2B\bar{B})\) belonging to an SU(3) colour octet. The remaining combination, the SU(3) colour singlet, \(\sqrt{\frac{1}{2}}(R\bar{R} + G\bar{G} + B\bar{B})\), does not carry colour and cannot mediate between colour charges.

In QED, the strength of the electromagnetic coupling between two quarks is given by \(e_1 e_2 \alpha\), where \(e_i\) is the electric charge of quark \(i\), listed in table 1.2 and \(\alpha\) is the fine structure constant (1/137). In QCD, the strength of the (strong) coupling for single-gluon exchange between two colour charges is \(\frac{1}{2}c_1 c_2 \alpha_s\), where \(c_i\) are the colour coefficients associated with the vertices and \(\alpha_s\) is the strong coupling coefficient. The colour factor

\[
C_F = \frac{1}{2} |c_1 c_2| \tag{4.1}
\]

has become a conventional notation to denote the strength of the process. The three possible strong-interaction couplings are: a quark radiating a gluon, a gluon radiating a gluon, or a gluon splitting into a quark anti-quark pair. The associated colour-averaged coupling strengths are listed in table 4.1, where a higher value indicates a stronger probability of occurring.

\(^1\)For example, the \(\Delta^{++}\) baryon requires all three \(u\) quarks to have identical spin quantum numbers, which would violate the Pauli Exclusion Principle
CHAPTER 4. JET CONSTRUCTION AND MEASUREMENT

<table>
<thead>
<tr>
<th>Process</th>
<th>Symbol</th>
<th>Coupling Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gluon coupling to a quark</td>
<td>$C_F$</td>
<td>4/3</td>
</tr>
<tr>
<td>Gluon self coupling</td>
<td>$C_A$</td>
<td>3</td>
</tr>
<tr>
<td>Gluon splitting into a quark anti-quark pair</td>
<td>$T_F$</td>
<td>1/2</td>
</tr>
</tbody>
</table>

Table 4.1: Colour factors for the various strong processes.

Dijet flavor

The lowest order parton-parton QCD scattering processes that contribute to two jet production are listed in Table 4.2. The squared matrix element $|M(\eta^*)|^2$ is averaged over spin and colour, and written in terms of the Mandelstam variables$^2$, $\hat{s}, \hat{t},$ and $\hat{u}$. The rightmost column in Table 4.2 shows the squared matrix elements calculated at the center-of-mass scattering angle of $\theta^* = \pi/2$. The largest parton-parton cross sections are elastic scattering ($ab \rightarrow ab$), dominated by the gluon interactions.

The parton-parton cross section for hard $2 \rightarrow 2$ processes, as a function of the $p_T^2$ scale, is given by

$$\frac{d\hat{\sigma}}{dp_T^2}(ab \rightarrow cd) = \frac{|M|^2}{16\pi\hat{s}^2\cos(\theta)}$$

(4.2)

and the lowest order QCD cross section for two-jet production from a proton-proton collision is given by:

$$\frac{d\hat{\sigma}}{dp_T^2}(AB \rightarrow 2Jets) = \sum_{abca} \int_0^1 \int_0^1 dx_A d\hat{x}_B f_{a/A}(x_A) f_{b/B}(x_B) \frac{d\hat{\sigma}}{dp_T^2}(ab \rightarrow cd)$$

(4.3)

where $f_{i/I}$ represents the parton distribution function$^3$ for parton $i$ within proton $I$, and $x_i$ is the longitudinal momentum fraction. If $0< x_i < 1$, then it is an inelastic collision, and $x_i = 1$ refers to an elastic collision. The calculated fractional cross sections for gluon-gluon, gluon-quark and quark-quark processes are shown in figure 4.1, where it can be seen that the quark-gluon and gluon-gluon initial states are dominant at lower energies, and the quark-quark initial state process dominates at higher energy [13].

$^2$for the parton-parton process $ab \rightarrow cd$, $\hat{s} = (a + b)^2, \hat{t} = (a - c)^2, \hat{u} = (a - d)^2$, where the letters denote four-momenta and the usual sign convention is used for the four vector scalar product.

$^3$The momentum distribution functions of the partons within the proton.
Figure 4.1: The fraction of the cross section from gluon-gluon, gluon-quark and quark-quark processes at $\eta_1 = \eta_2 = 0$ and $\sqrt{s} = 2\, TeV$ (solid) and $\sqrt{s} = 14\, TeV$ (dashed) as a function of $x_T = \frac{2E_T}{\sqrt{s}}$, where $E_T$ is the outgoing parton’s transverse energy [13].
4.1.2 Hadronization and fragmentation

After a hard interaction, a parton shower will be produced from the final-state parton. As quarks and gluons are separated, the energy in the field between them grows to the point that a quark-antiquark pair can be produced out of the "strong" or QCD vacuum. Hadronization is the mechanism by which the 'bare' quarks or gluons from a parton shower transform into a hadron. The overall process of parton showering and hadronization is referred to as fragmentation.

Because QCD is a non-perturbative theory at low energies, hadronization cannot be calculated analytically. As an alternative, Monte-Carlo (MC) statistical models have been developed to describe these effects. They contain theory and models for a number of physics properties, such as hard and soft interactions, parton distributions, initial and final state parton showers, multiple interactions, fragmentation and decay. The two hadronization models explored in this analysis are cluster and string, corresponding to the HERWIG and PYTHIA MC generators respectively. Figure 4.2 pictorially displays the fundamental differences between the two models.

The cluster model makes the splitting $g \rightarrow q\bar{q}$, for all gluons after the parton shower.
All colour-singlet\(^4\) \(q\bar{q}\) combinations are then grouped together. These colour-singlet combinations are assumed to form clusters, which undergo decay into hadrons.

The string model is based on the dynamics of a relativistic string, representing the colour flux stretched between the initial \(q\bar{q}\). At the end of the parton shower, all gluons are grouped within the quark-antiquark pairs. Neighbouring pairs of \(q\bar{q}\) form colour neutral clusters which (usually) decay into hadrons. The hadron type is determined by the available phase space, and baryon production is modelled through quark-diquark and antiquark-antidiquark colourless combinations for both models.

### 4.1.3 Jet reconstruction

Jet reconstruction is a key element in the description of physics at high-energy particle colliders; as QCD jets are observed within the particle detectors, and used as a tool to infer the properties of the quarks and gluons produced in the collisions. Figure 4.3 illustrates the different levels at which a jet can be defined: parton, particle, or calorimeter level. A parton level jet is defined using the final state partons\(^5\) from the proton-proton collision. A particle jet is defined after the parton has hadronized into baryons and mesons. Finally,

---

\(^4\)The colour singlet state can be viewed as \((r\bar{r} + b\bar{b} + g\bar{g})/\sqrt{3}\), saying it is equally probable that the \(q\bar{q}\) state is \(r\bar{r}, b\bar{b}, g\bar{g}\)

\(^5\)The parton may radiate multiple gluons.
the calorimeter level jet is constructed from the energy deposits in the calorimeter, after the particles have interacted with the detector.

Parton and particle jets are only defined in MC simulations, while calorimeter jets are available in both MC generators and ATLAS datasets. Because theorists can calculate to particle level, an ideal algorithm reconstructs the measured calorimeter jets to particle level, taking account of calorimeter response, with 100\% efficiency. The first step in going from the measured calorimeter level to the particle level is to cluster the measured signal from the calorimeter cells. The cells are first selected using a topological clustering algorithm to filter electronic noise.

**Topological clustering algorithm**

Calorimeter jet reconstruction begins when a cell (the smallest calorimeter readout object) measures energy such that \( E_{\text{cell}} > 4\sigma_{\text{noise}} \), where \( \sigma_{\text{noise}} \) is a pre-defined threshold given by the RMS of the energy distribution for random noise. All neighbouring cells to the measured cell are checked to see if \( E_{\text{cell}} > 2\sigma_{\text{noise}} \). If the cell satisfies this criterion then it is added to the cluster; this algorithm operates in three dimensions. Once all neighbour cells...
with \( E_{\text{cell}} > 2\sigma_{\text{noise}} \) have been iteratively added, all nearest neighbour cells surrounding the cluster are also included. This process, which occurs in three dimensions, is known as the topological clustering algorithm. Overlapping clusters can be split or merged depending on local maxima or minima within the clusters. The cell energies are summed to give the cluster energy, and the clusters are treated as massless.

**Anti-\( k_T \) algorithm**

The anti-\( k_T \) algorithm constructs a calorimeter jet from the topological clusters, using rapidity\(^6\) \((y)\) as a spatial variable. The predefined parameters of the algorithm are

- minimum separation distance in \((y \times \phi)\) space, \(D\)
- possible \(d_{\text{cut}}\) to stop the algorithm when \(d_{\text{min}} \leq d_{\text{cut}}\)

For each cluster \(i\) from the topological clustering algorithm, the beam distance is defined as

\[ y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right), \text{ where } E \text{ is the energy and } p_z \text{ is the momentum along the z-axis} \]
\[ d_{iB} = p_{T,i}^2 \]  

Then, for each pair of clusters, the distance (in $y \times \phi$ space) between the two clusters is defined as

\[ d_{ij} = \min(p_{T,i}^2, p_{T,j}^2) \frac{\delta R_{ij}^2}{D^2} \]

where $p_{T,i/j}$ is the transverse momentum, and $\delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ with $y_i$ and $\phi_i$ denoting the rapidity and azimuth of particle $i$ respectively. Then a recursive algorithm best pairs the clusters:

- Find the minimum, $d_{\text{min}}$, of all $d_{iB}$ and $d_{ij}$
- if $d_{\text{min}}$ is a $d_{ij}$, merge cluster $i$ and $j$ into a single cluster, adding their four-momenta.
- if $d_{\text{min}}$ is a $d_{iB}$, cluster $i$ is a jet, and it is added to the final list of jets
- repeat until no clusters remain

This algorithm tends to group highest $p_T$ objects first. As seen from figure 4.4, soft particles tend to cluster with hard ones long before they cluster among themselves. If a hard particle has no hard neighbours within a distance $2D$, then it will simply accumulate all the soft particles within a circle of radius $D$ in $y$-phi space, resulting in a perfectly conical jet [7].

The more common situation is two or more hard (high $p_T$) particles are present within $2D$, with $p_{T1} \gg$ all other jets. In these situations, the highest $p_T$ reconstructed jet within $2D$ will be constructed as a cone, and the other jets will be partly conical, missing the piece clipped out by the higher $p_T$ jet. Finally for the situation where two jets have approximately the same momentum, they will share clusters at the boundary. An important aspect of this algorithm is that the soft particles do not modify the shape of the jet, while hard particles do.

**Infrared and collinear safety**

Infrared and collinear (IRC) safety is the property that if an event is modified by a collinear splitting or a soft emission gluon, the set of hard jets should remain unchanged. These two
properties are fundamental requirements for jet algorithms, making it possible to compare theory calculations to experimental measurements. Collinear safety requires the same jets to be reconstructed given one particle or two collinear particles with same total energy. Infrared safety requires the same jets to be reconstructed in the presence of soft radiation in the event. Figures 4.5 and 4.6 give examples of each property, where the parton splitting or the addition of soft radiation has affected the jet reconstruction. The anti-$k_T$ algorithm does satisfy infrared and collinear safety.

Figure 4.5: Collinear problem - replacing any massless parton with an exactly collinear pair of massless partons should not change energy of a jet, as it does in this example.

Figure 4.6: Infrared problem - adding a soft parton should not change the number of jets as it does in this example.

### 4.2 Calorimeter response

The ATLAS experiment measures jets at the calorimeter level, and applies the jet energy scale (JES) to calibrate the measured calorimeter jet energy to the particle level. It is the design goal of the ATLAS experiment to achieve a 1% uncertainty on the JES, where currently the JES for 2011 data ranges between 2-7%, shown in table 4.3 [9]. In order to reach the goal of 1% precision, it will be necessary to treat gluon and quark jets with independent JES corrections because the response varies by > 2% depending what parton initiated the shower, as explained below.

The electromagnetic calorimeter has been calibrated to give a linear response to photons...
Table 4.3: The maximum jet energy scale systematic uncertainties from a MC based study for anti-$k_T$ jets with $R = 0.6$ [9].

| $|\eta|$ region | $p_T^{\text{jet}} = 20$ GeV | $p_T^{\text{jet}} = 200$ GeV | $p_T^{\text{jet}} = 1.5$ TeV |
|-----------------|-----------------|-----------------|-----------------|
| [0.0, 0.3]      | 4.6%            | 2.3%            | 3.1%            |
| [0.3, 0.8]      | 4.5%            | 2.2%            | 3.3%            |
| [0.8, 1.2]      | 4.5%            | 2.4%            | 3.4%            |
| [1.2, 2.1]      | 5.5%            | 2.5%            | 3.5%            |
| [2.1, 2.8]      | 7.1%            | 2.5%            |                  |

and electrons. This calibration was initially done through test beams, and is monitored in-situ using the decay of the $Z$ boson\(^7\). The ATLAS calorimeter is non compensating (e/h $> 1$), thus the hadronic response must be calibrated. A jet’s response is a measure of the calorimeter response to the hadrons within a jet\(^8\), derived as a function of incoming particle energy [10] as

\[
j(E) = b_0 + b_1 \ln(E) + b_2 \ln^2(E) \tag{4.6}\]

The coefficients $b_0$, $b_1$, and $b_2$ depend on the fraction of hadronic and electromagnetic particles in the jet, and the calorimeter response to these particles. The power of equation 4.6 is that any measured jet can be calibrated back to particle level after the coefficients have been calculated.

### 4.2.1 Flavour response

Jets are manifestations of hard quarks or gluons produced at very short distances, which shower and fragment into collections of collinear particles [15]. The character of these showers is determined by the fragmentation properties of the original parton, quark or gluon. The colour factors given in Table 4.1 suggest that hard scattering quarks and gluons have different probability of radiating a soft gluon. Because of these different colour factors, as seen in figure 4.7, the showering profiles become dependent on the original parton.

\(^7\)Z $\rightarrow ee$, with a reconstructed mass of 91.876$\pm$0.002 GeV

\(^8\)Other showering contributions such as leakage, dead calorimeter cells, and pileup also play a roll
CHAPTER 4. JET CONSTRUCTION AND MEASUREMENT

Figure 4.7: Quark and gluon jets radiate proportionally to their colour factors. $C_F$ denotes the strength of gluon coupling to quarks and $C_A$ denotes the strength of gluon coupling.

To leading order, the ratio

$$ r = \frac{C_A}{C_F} = \frac{9}{4} $$

(4.7)

suggests that gluon induced jets radiate more soft particles than quark jets. The magnetic field within the inner detector amplifies the broadness of gluon-jets, because the soft radiation is more susceptible to particle bending than the relatively higher-$p_T$ particles from quark jets. The ratio in Eq. 4.7 also means that gluon jets have a higher multiplicity of particles than quark induced jets, typically between 1 and 3 additional tracks for jets with transverse momentum between 20 GeV and 100 GeV [18]. The harder particles from the quark jets provide a denser energy profile in the calorimeter and tend to penetrate further than the gluon jet particles [18].

The major factor in the response difference between quark and gluon jets is the larger number of soft particles in gluon jets. It is more probable for low $p_T$ particles to bend outside the jet cone due to the magnetic field, resulting in a mis-measurement of the calorimeter-level jet. Furthermore, a calorimeter’s response is energy dependent, where lower $p_T$ particles have a worse response, on average [23]. The above differences lead to the 3-8% response difference between quark and gluon jets, with the latter having the lower response [14]. Table 4.4 summarizes all of the important differentiating properties between quark and gluon jets.

---

9A gluon jet of 100 GeV has, on average, the same width as a quark jet of 60 GeV.
Differences in Jet

<table>
<thead>
<tr>
<th>Quark</th>
<th>Gluon</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_F = \frac{4}{3}$</td>
<td>$C_A = 3$</td>
</tr>
<tr>
<td>Harder particles</td>
<td>Softer particles</td>
</tr>
<tr>
<td>Lower multiplicity jets</td>
<td>Higher multiplicity jets</td>
</tr>
<tr>
<td>Higher response (on average)</td>
<td>Lower response (on average)</td>
</tr>
</tbody>
</table>

Table 4.4: Fundamental differences between quark and gluon jets.

Figure 4.8: The fraction of the final state dijet partons flavor as a function of $p_T$ [15].

Figure 4.9: Fraction of jet flavor from $\gamma + 1$jet events as a function of $p_T$ [15].

4.2.2 Monte-Carlo response

The MC datasets used in this thesis are produced with the programs HERWIG\textsuperscript{10} and PYTHIA. The objective of event generators is to provide a description of what happens in a particle collision, giving access to parton, particle, and calorimeter level jets. It is up to the user to match the parton and particle level jets to the calorimeter jet.

Truth level matching

Because the MC samples give access to the $2 \rightarrow 2$ parton information (the true incoming and outgoing partons), parton level matching is done by selecting the highest energy parton

\textsuperscript{10}Hadron Emission Reactions With Interfering Gluons
that points to the calorimeter jet within a $\Delta \phi < 0.6^{11}$. This choice is made because all events considered for this analysis have been selected to favour hard scattering events, with two partons emerging back to back in phi space. Once the $\phi$ space matching has been done between the calorimeter jet and the partons, the parton MC response is defined as

$$R_{MC}^{parton} = \frac{P_T(\text{Calorimeter Level})}{P_T(\text{Parton Level})} \quad (4.8)$$

Monte Carlo simulation particle jets are built from stable particles defined to have proper lifetimes longer than 10 ps excluding muons and neutrinos. The particle jet level matching is done by selecting the largest $p_T$ particle jet that lies within a $\Delta \phi$ of 0.3 from the calorimeter jet. Once the matching is done, the particle MC response can be defined as:

$$R_{MC}^{particle} = \frac{P_T(\text{Calorimeter Level})}{P_T(\text{Particle Level})} \quad (4.9)$$

### 4.2.3 MPF response

The Missing $E_T$ Projection Fraction (MPF) technique was introduced at the DØ experiment$^{12}$ with great success. The objective of the technique is to calibrate the hadronic response of a jet against the well measured photon using transverse momentum balance. MPF studies using $\gamma$+jet and $Z$+jet have been used at ATLAS, both using a well measured tag object ($\gamma, Z$) to probe the jet response. This thesis explores the possibly of using jets with a good response (high $\pi^0$ content) as tag objects, and implementing a jet+jet (dijet) MPF study from the tree level diagrams in figure 4.10.

### Missing energy

Applying conservation of momentum to the transverse plane, the final state parton momentum must satisfy:

$$\sum_{\text{parton}} \vec{P}_T^{\text{parton}} = 0 \quad (4.10)$$

---

$^{11}\Delta \phi = \phi_{\text{jet}} - \phi_{\text{parton}}$

$^{12}$The experiment was located at the high-energy accelerator, the Tevatron Collider, at the Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois, USA.
where it is assumed that the incoming beam carries all of its momentum along the $z$-axis. Equation 4.10 holds for parton level and must be modified for the calorimeter level

$$ \vec{E}_T = \sum_{\text{cells}} -\vec{E}_T $$

where the missing transverse energy ($E_{T}^{\text{Miss}}$ or $\vec{E}_T$) can be defined at particle or calorimeter level. At particle level, the $E_{T}^{\text{Miss}}$ gives information about non-interacting particles such as neutrinos. At calorimeter level, the $E_{T}^{\text{Miss}}$ also includes effects from the hadronic response, and any calorimeter defects (dead cells, ...). ATLAS has defined the $E_{T}^{\text{Miss}}$ using a variety of inputs, ranging from cell information to the reconstructed and calibrated jets themselves. This thesis focuses on the $E_{T}^{\text{Miss}}$ constructed from topological clusters.

**$\gamma$+jet events**

The following definition assumes that $l$ denotes the photon, and 2 denotes the jet being measured. At the parton level, momentum conservation between the photon and jet gives
Neglecting fragmentation and hadronization effects, the momentum balance equation still holds at particle level,

\[ \vec{P}_{\text{parton}}^{T_1} + \vec{P}_{\text{parton}}^{T_2} \approx 0 \quad (4.13) \]

In order to move to the calorimeter level, the calorimeter response must be accounted for

\[ e \vec{P}_{\text{particle}}^{T_1} + j(E_T) \vec{P}_{\text{particle}}^{T_2} + \vec{E}_T \approx 0 \quad (4.14) \]

where \( e \) and \( j(E_T) \) are the photon and jets calorimeter responses respectively. Knowing the electromagnetic response \( e \) to photons is approximately one, \( e \vec{P}_{\text{particle}}^{T_1} \) can be replaced with \( \vec{P}_{\text{Calorimeter}}^{T_1} \). Equation 4.14 can be projected into the direction of the photon,

\[ \vec{P}_{\text{Calorimeter}}^{T_1} + j(E_T) \vec{P}_{\text{particle}}^{T_2} \cdot \hat{n}_\gamma = -\vec{E}_T \cdot \hat{n}_\gamma \quad (4.15) \]

and from Eq. 4.13, Eq. 4.15 can be simplified to

\[ R_{\text{MPF}}^{\gamma} = j(E_T) = 1 + \frac{\vec{E}_T \cdot \hat{n}_1}{\vec{p}_{\text{Calorimeter}}^{T_1}} \quad (4.16) \]

Thus the MPF gives an experimental way to calculate the hadronic response parametrized by equation 4.6. The above derivation uses \( E_T \) and \( p_T \) interchangeably, as photons are massless (and the jet is assumed to have no mass), thus \( E_T = p_T \).
Dijet events

Photons provide an excellent reference for jet calibration, because EM showers are very well-measured and they are well-calibrated in test beams. However, the photon plus jet sample is dominated by quark jets. One way to assess the bias is to create a gluon dominated sample, and measure the jet response. As seen in section 4.1.1, dijet events produce mostly gluon jets. The cross section for dijet events within the barrel (|η| < 2.5) and with $p_T > 25$ GeV dominate over direct photons, with

- dijets $\sigma = 0.5$ mb
- Direct photons $\sigma = 0.2 \mu$b

which implies that $\sigma_{dijets}/\sigma_{DirectPhoton} \approx 2500$ [12]. This large cross section for dijet events makes it possible to consider replacing the photon by a "tag" jet that has fluctuated to have a larger fraction of $\pi^0$ than average. If the jet does fluctuate to a high $\pi^0$ fraction, and fully showers within the EM calorimeter, the measured response should be close to one. One motivation for replacing the photon by a QCD jet is that QCD probe jets are dominated by gluons at low $p_T$, making it possible to study the calorimeter response for quark and gluon jets separately.

This derivation follows closely the standard derivation for $\gamma$+jet events, however this time 1 refers to the tag jet while 2 still refers the probe jet. The criteria for a tag jet will be explained in section 5.2. The derivation can be picked up from Eq. 4.14, replacing the photon with a tag jet:

$$R_1(P_T)\frac{\vec{p}_{\text{Particle}}}{p_{\text{Calorimeter}}^1} + j(P_T)\frac{\vec{p}_{\text{Particle}}}{p_{\text{Calorimeter}}^2} + \vec{E}_T \approx 0$$ (4.17)

where $R_1(P_T)$ is the calorimeter response (at particle level) to the tag jet, $j(P_T)$ is the hadronic response (at particle level) to the probe jet. Projecting equation 4.17 in the direction of the tag jet gives

$$R_1\frac{\vec{p}_{\text{Particle}}}{p_{\text{Calorimeter}}^1} + j(P_T)\frac{\vec{p}_{\text{Particle}}}{p_{\text{Calorimeter}}^2} \cdot \hat{n}_1 = -\vec{E}_T \cdot \hat{n}_1$$ (4.18)

and from Eq. 4.13, Eq. 4.18 can be simplified to
\[ R_{MPF}^{jet} = j(E_T) = R_1(1 + \frac{\hat{E}_T \cdot \hat{n}_1}{P_{Calorimeter}}) \]  

(4.19)

Where a non-zero \( R_1 \) calorimeter correction term has been introduced into the MPF definition. This correction is derived in chapter 5 and denoted \( \alpha \), which accounts for the non-perfect response of the tag jet.

**Quark and gluon MPF response**

The purpose of separating responses into jet flavour is to determine a separate JES correction for quark and gluon jets, allowing physics groups to calculate a modified JES correction, depending on the mix of jet types in the process being studied. The hypothesis of this thesis is that the MPF responses given by equations 4.16 and 4.19 can be decomposed in the following form:

\[ R_{MPF}^{\gamma} = f_{\gamma}^{quark} R_{MPF}^{quark} + f_{\gamma}^{gluon} R_{MPF}^{gluon} \]  

(4.20)

\[ R_{MPF}^{\gamma} = f_{\gamma}^{jet} R_{MPF}^{quark} + f_{\gamma}^{gluon} R_{MPF}^{gluon} \]  

(4.21)

where \( f_{\gamma}^{quark}, f_{\gamma}^{gluon}, R_{MPF}^{quark}, R_{MPF}^{gluon} \) are the fractions of and responses to quark and gluon jets respectively\(^1\). Both light (u,d,s) and heavy (c,b,t) quarks are identified as a quark jet in this study. Equations 4.20 and 4.21 can be solved for \( R_{MPF}^{quark} \) and \( R_{MPF}^{gluon} \)

\[ R_{MPF}^{quark} = \frac{R_{MPF}^{\gamma} - f_{\gamma}^{gluon} R_{MPF}^{\gamma} - f_{\gamma}^{quark} R_{MPF}^{\gamma}}{f_{\gamma}^{jet} - f_{\gamma}^{quark}} \]  

(4.22)

\[ R_{MPF}^{gluon} = \frac{R_{MPF}^{\gamma} - f_{\gamma}^{quark} R_{MPF}^{\gamma} - f_{\gamma}^{gluon} R_{MPF}^{\gamma}}{f_{\gamma}^{jet} - f_{\gamma}^{quark}} \]  

(4.23)

using the identity \( f_{\gamma}^{gluon} = 1 - f_{\gamma}^{quark} \).

---

\(^1\)The fractions \( f \) are calculated from Monte Carlo simulations.
Chapter 5

Defining hadronic tag jet candidates

This chapter describes the hadronic tag jet candidates for a dijet MPF response study, using proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV, measured with the ATLAS detector. Section 5.3 builds the $\alpha$ correction required to calculate a dijet MPF, implementing the likelihood separator presented in section 5.2.1.

5.1 Optimizing the datasets

ATLAS provides a variety of datasets, skimmed\(^1\) by different physics groups to optimize the efficiency of their analyses. For example, some may be interested in leptons, while others may be focussed on $E_T^{\text{Miss}}$. The dataset for this study is taken from the JetTauEtmiss stream from 2011, focusing on jets and $E_T^{\text{Miss}}$. The events are recorded by run number, a unique number that identifies the time and conditions under which the data were recorded. Run numbers can be grouped into larger classifications according to pileup conditions. For 2011, the larger classifications (periods) are denoted $D$ to $M$. The data periods can be segmented into four main groups. Period $D$ corresponds to a relatively low pileup environment, with the average number of events per bunch crossing, $\mu \approx 5.5$. Periods $E - H$ have the same pileup conditions as period $D$, however a section of the calorimeter was not working correctly, so a veto is applied to all jets that fall within this region. In periods $I - K$, the calorimeter has been fixed, and the pileup has been increased to $\mu \approx 6.5$. Finally

\(^1\)Skimming is the process of selecting only events that match a set of given criteria.
for periods $L - M$ the pileup is $\mu \approx 10$. Figures 5.1 and 5.2 show the pileup conditions at ATLAS in 2011.

**Triggers**

The trigger used to select events for this analysis are single jet event filter triggers. The jet trigger scheme used in this analysis follows the 2011 ATLAS jet cross section recommendations [19], where the highest $p_T$ jet is taken as the trigger jet. The trigger strategy is as follows:

1. Determine the appropriate trigger based on the leading jet $p_T$ from table 5.1

2. Ask if the event passes the trigger
   
   (a) If it doesn’t pass, fail the event
   
   (b) If the appropriate trigger is fired, give the event a weight of $1/\text{luminosity}$, taken from table 5.2

This procedure produces a leading jet $p_T$ spectrum cross section in $pb^{-1}$. Figures 5.3 and 5.4 show the two highest $p_T$ jet distributions, showing some disagreement between the PYTHIA MC generator and data. Because the MPF analysis is binned in $p_T$, the cross section discrepancy that begins at approximately 250 GeV has no effect on the results.
Table 5.1: After determining the bin in which the leading jet $p_T$ falls, the event is checked to see if the corresponding event trigger fired. If not, the event is not considered for analysis [19].

The efficiency of the jet trigger has been evaluated for offline jets reconstructed with the anti-$k_T$ algorithm with radius parameter $R=0.4$. The efficiency is defined as the fraction of offline reconstructed jets that match a trigger jet with $\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2} < 0.4$ and which pass the corresponding trigger threshold. The recommendation of Ref. [19] to accept only dijet events where the leading jet $p_T$ is greater than 70 GeV due to trigger and jet reconstruction efficiencies.

**Jet cleaning**

A jet must pass a number of quality cuts to be considered for analysis. Jets are classified in three categories: good, bad, and ugly. Bad jets arise from various sources, including hardware problems, LHC beam conditions, and cosmic-ray showers. They do not correspond to energy depositions in the calorimeters from jet showers and must be removed from the analysis. The ugly jets are physical jet showers; however, they occur in problematic calorimeter regions where the impact of inoperative tiles in the calorimeter is still not well understood. The good jets are jets that are neither bad nor ugly. Jet cleaning uses the following discriminating variables to minimize the number of bad and ugly jets used in the
<table>
<thead>
<tr>
<th>Period</th>
<th>j10</th>
<th>j15</th>
<th>j30</th>
<th>j40</th>
<th>j55</th>
<th>j75</th>
<th>j100</th>
<th>j135</th>
<th>j180</th>
<th>j240</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>4.54e3</td>
<td>2.67e4</td>
<td>1.25e5</td>
<td>8.43e4</td>
<td>2.81e5</td>
<td>1.05e6</td>
<td>3.55e6</td>
<td>8.60e7</td>
<td>1.66e8</td>
<td>1.66e8</td>
</tr>
<tr>
<td>E</td>
<td>2.78e2</td>
<td>1.37e3</td>
<td>6.01e3</td>
<td>8.92e3</td>
<td>3.14e4</td>
<td>1.12e5</td>
<td>3.89e5</td>
<td>1.98e7</td>
<td>4.87e7</td>
<td>4.87e7</td>
</tr>
<tr>
<td>F</td>
<td>6.74e2</td>
<td>3.27e3</td>
<td>1.43e3</td>
<td>2.13e4</td>
<td>7.59e5</td>
<td>2.67e5</td>
<td>9.33e5</td>
<td>3.58e6</td>
<td>1.32e8</td>
<td>1.32e8</td>
</tr>
<tr>
<td>G</td>
<td>2.48e3</td>
<td>1.05e4</td>
<td>4.66e4</td>
<td>6.73e4</td>
<td>2.34e5</td>
<td>8.35e5</td>
<td>2.84e6</td>
<td>1.08e7</td>
<td>4.77e8</td>
<td>4.97e8</td>
</tr>
<tr>
<td>H</td>
<td>1.25e</td>
<td>5.05e3</td>
<td>2.52e4</td>
<td>3.65e4</td>
<td>1.27e5</td>
<td>4.54e5</td>
<td>1.53e6</td>
<td>5.89e6</td>
<td>2.53e8</td>
<td>2.59e8</td>
</tr>
<tr>
<td>I</td>
<td>1.66e3</td>
<td>7.00e3</td>
<td>2.82e4</td>
<td>4.11e4</td>
<td>1.43e5</td>
<td>5.07e5</td>
<td>1.72e6</td>
<td>6.49e6</td>
<td>2.65e7</td>
<td>3.86e8</td>
</tr>
<tr>
<td>J</td>
<td>9.36e2</td>
<td>3.96e3</td>
<td>1.25e4</td>
<td>1.83e4</td>
<td>6.39e4</td>
<td>2.27e5</td>
<td>7.71e5</td>
<td>2.94e6</td>
<td>1.21e7</td>
<td>2.26e8</td>
</tr>
<tr>
<td>K</td>
<td>2.50e3</td>
<td>1.06e3</td>
<td>3.24e4</td>
<td>4.73e4</td>
<td>1.65e5</td>
<td>5.86e5</td>
<td>1.99e6</td>
<td>7.61e6</td>
<td>3.15e7</td>
<td>6.00e8</td>
</tr>
<tr>
<td>L</td>
<td>3.15e3</td>
<td>9.53e3</td>
<td>3.11e4</td>
<td>5.24e4</td>
<td>2.13e5</td>
<td>8.15e5</td>
<td>2.89e6</td>
<td>1.10e7</td>
<td>4.73e7</td>
<td>1.29e9</td>
</tr>
<tr>
<td>M</td>
<td>8.16e2</td>
<td>3.59e3</td>
<td>2.52e4</td>
<td>3.91e4</td>
<td>1.40e5</td>
<td>5.21e5</td>
<td>1.90e6</td>
<td>7.38e6</td>
<td>3.16e7</td>
<td>8.93e8</td>
</tr>
<tr>
<td>Totals</td>
<td>1.83e4</td>
<td>8.15e4</td>
<td>3.47e5</td>
<td>4.16e5</td>
<td>1.47e6</td>
<td>5.37e6</td>
<td>1.85e7</td>
<td>1.62e8</td>
<td>1.23e9</td>
<td>4.50e9</td>
</tr>
</tbody>
</table>

Table 5.2: The luminosity calculated per period and event level trigger [19].
CHAPTER 5. DEFINING HADRONIC TAG JET CANDIDATES

Figure 5.3: The highest $p_T$ jet within a dijet event, with the MC shown in black and data in red.

Figure 5.4: The second highest $p_T$ jet within a dijet event, with the MC shown in black and data in red.

analysis.

- $f_{\text{HEC}}$: energy fraction of the jet in the Hadronic Endcap Calorimeter (HEC)
- $Q_{\text{HEC}}$: fraction of LAr cells (within the HEC) with a cell Q-factor$^2$ greater than 4000
- $Q_{\text{LAr}}$: fraction of LAr cells with a cell Q-factor greater than 4000
- $f_{\text{EM}}$: fraction of the jet energy in the EM calorimeters
- $f_{\text{charged}}$: the ratio of the sum of the $p_T$ of tracks associated with the jet, divided by the jet $p_T$ determined from the calorimeter
- $f_{\text{Max}}$: maximum energy fraction in any one calorimeter layer
- $t$: jet time computed as the mean time of the cells weighted by the square of the energy

The cuts in table 5.3 aim at removing jets with the following flaws: electronic noise spikes in the hadronic end-cap calorimeter (HEC), coherent noise in the electromagnetic calorimeter, and non-collision backgrounds. If either of the two jets in the dijet event fails a cleaning cut, the event is rejected.

$^2$The cell Q-factor is the difference between the measured pulse shape $q_{i,\text{meas}}$ and the predicted pulse shape
CHAPTER 5. DEFINING HADRONIC TAG JET CANDIDATES

<table>
<thead>
<tr>
<th>Source</th>
<th>Identifier</th>
<th>Flagged as Bad Jet if:</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEC energy spike</td>
<td>Loose</td>
<td>$f_{HEC} &gt; 0.5$ AND $</td>
</tr>
<tr>
<td>HEC energy spike</td>
<td>Loose</td>
<td>$</td>
</tr>
<tr>
<td>HEC energy spike</td>
<td>Medium</td>
<td>$f_{HEC} &gt; (1 -</td>
</tr>
<tr>
<td>Noise</td>
<td>Loose</td>
<td>$f_{EM} &gt; 0.95$ AND $</td>
</tr>
<tr>
<td>Noise</td>
<td>Medium</td>
<td>$f_{EM} &gt; 0.9$ AND $</td>
</tr>
<tr>
<td>Cosmic muons</td>
<td>Loose</td>
<td>$</td>
</tr>
<tr>
<td>Cosmic muons</td>
<td>Loose</td>
<td>$f_{EM} &lt; 0.05$ AND $f_{charged} &lt; 0.05$ AND $</td>
</tr>
<tr>
<td>Cosmic muons</td>
<td>Loose</td>
<td>$f_{EM} &lt; 0.05$ AND $</td>
</tr>
<tr>
<td>Cosmic muons</td>
<td>Loose</td>
<td>$f_{Max} &gt; 0.99$ AND $</td>
</tr>
<tr>
<td>Cosmic muons</td>
<td>Medium</td>
<td>$f_{EM} &lt; 0.05$ AND $f_{charged} &lt; 0.1$ AND $</td>
</tr>
<tr>
<td>Cosmic muons</td>
<td>Medium</td>
<td>$f_{EM} &gt; 0.95$ AND $f_{charged} &lt; 0.05$ AND $</td>
</tr>
</tbody>
</table>

Table 5.3: Cleaning cuts for anti-$k_T$, $R = 0.6$ jets [9].

Estimate energy in dead cells

Bad or misread cells are a dominant source of fake $E_T^{Miss}$, and poor jet energy resolution. Cell-neighbour-averaging is the current method to correct dead calorimeter cells. A dead cell is given the average energy density of surrounding cells, as shown in figures 5.5 and 5.6. Studies have shown this correction can under-correct by up to 6% for central jets in dead tile regions, and over-correct by up to 4% for central jets just next to dead tile regions [17]. Two tools exist to estimate the energy in dead cells based on the cell averaging correction, cell based and jet based.

The cell-level correction ($BCH\_CORR\_CELL$) is the default correction in ATLAS, applied before jet and $E_T^{Miss}$ reconstruction, calculated as

$$BCH\_CORR\_CELL = \frac{\sum E_{cell}^{corr}}{E_{calo}}$$

(5.1)

where $E_{cell}^{corr}$ is the corrected cell energy of dead cells that are fully contained within the jets core; and $E_{calo}$ is the total energy in the cone, including $\sum E_{cell}^{corr}$. Any boundary cells or multiple neighbouring dead cells are not included in $E_{cell}^{corr}$.
CHAPTER 5. DEFINING HADRONIC TAG JET CANDIDATES

Figure 5.5: This drawing shows a low $p_T$ jet with a dead cell in the center. The true value is 10 GeV, and the cell level correction calculates the estimated value to be 5 GeV. This gives an underestimate of the jet $p_T$ by 25% [17].

Figure 5.6: This drawing shows a high $p_T$ jet with a dead right cell. The true value is 25 GeV, and the cell level correction calculates the estimated value to be 50 GeV. This gives an overestimate of the jet $p_T$ by 17% [17].

Jet level correction ($BCH\_CORR\_JET$) uses the jet shape from MC, and can only be computed after jet reconstruction. This correction builds a cone of radius 0.6 around the jet axis, and all the estimated energy within the bad cells inside the cone is summed up, calculated as

$$BCH\_CORR\_JET = \sum E_{\text{corr}}^{\text{jet}} / E_{\text{calo}}$$

(5.2)

where $E_{\text{corr}}^{\text{jet}}$ estimates bad cells in cone, including boundary cells and multiple neighbouring dead cells. Typically the jet-level correction is larger than the cell-level correction. This difference arises because the cell-level correction does not calculate a dead area that is too large to be filled from immediate neighbours. While both corrections are well modelled by the MC generators, studies in Ref. [17] have shown that the jet-level correction is a better estimator. Thus the following corrections are made to the jet $p_T$ and $E_T^\text{Miss}$:

$$P_T^{\text{corr}} = P_T^{\text{meas}} \times (1 - \frac{BCH\_CORR\_CELL}{BCH\_CORR\_JET})$$

$$\vec{E}_T^{\text{corr}} = \vec{E}_T - (\vec{P}_T^{\text{corr}} - \vec{P}_T)$$

(5.3)

(5.4)

with the requirement that the tag jet correction is less than 5 percent of the jet $p_T$, as
CHAPTER 5. DEFINING HADRONIC TAG JET CANDIDATES

Jet-Vertex fraction

The Jet-Vertex Fraction (JVF) measures the probability that a jet originated from a particular vertex. For a single jet, $i$, the JVF with respect to the vertex ($vtx$) $j$ is given by:

$$JVF(i, j) = \frac{\sum_k P_T(trk_{jet}^k, vtx_j)}{\sum_m \sum_n P_T(trk_{jet}^n, vtx_m)}$$

where $trk$ denotes a track measured by the inner detector. The JVF allows the discrimination of jets from the hard scattering and jets from pileup interactions. The values that the JVF can take on are:

- $JVF = 1$ : Jets where no pile-up contribution is observed (hard scatter jet)
- $JVF = -1$ : Jets without matched tracks
- $JVF = 0$ : Jets with all charged tracks originating from a pile-up interaction
- $JVF \in (0,1)$ : Degree of pile-up contributing to the jet, with a large number denoting low pileup

The standard JVF value for 2011 data is to accept jets with a JVF greater than 0.75 as hard scattering jets.

5.1.1 Dijet selection

Figure 5.7 shows a dijet event measured by the ATLAS detector in April of 2012. This event has two hard scattering partons emerging from one of the six interaction points. The two partons quickly hadronize, forming well defined calorimeter jets in the electromagnetic and hadronic barrel calorimeters.

ISR and FSR

Two important processes that can affect a clean dijet event are initial state radiation (ISR) and final state radiation (FSR). When a parton radiates from the incoming (ISR) or outgoing
Figure 5.7: A central dijet event with the highest-$p_T$ jet collected by the end of April 2012: two central high-$p_T$ jets have an invariant mass of 3.62 TeV and the highest $p_T$ jet has a $p_T$ of 1.96 TeV. The event was collected on 9 April 2012. The lines indicate tracks reconstructed from the inner detector data. The green section is the EM calorimeter, with the hadronic calorimeter highlighted in red. The outer blue section is the muon spectrometer [24].
Measuring DiJet Balance

Figure 5.8: The dijet balance is defined as 1 - (Probe Jet $p_T$)/(Tag Jet $p_T$), where the $p_T$ is measured at particle level. A lower measurement indicates a better momentum balance between the two jets in the transverse plane.

(FSR) parton participating in a hard scattering event, an additional jet may be produced and introduce a bias to the dijet event. An extensive study of these two phenomena has been performed in $\gamma$+jet events. It was found that requiring the jets to be back-to-back in phi space and restricting the energy of all other jets in the event, minimizes the effect of ISR and FSR [10]. Figure 5.8, produced using the PYTHIA generator for dijet events, shows one minus the ratio of the transverse momenta of the two highest $p_T$ particle level jets.

Through the recommendations of [10], and the confirmation from figure 5.8, the following cuts are applied to all events:

1. $|\phi_{Jet_1} - \phi_{Jet_2}| > 2.8$

2. $P_T^{Jet_3} < \max(12 \text{ GeV}, 0.2 \times P_T^{Jet_1})$

Where $P_T^{Jet_3}$ is the third highest $p_T$ jet in the event with JVF > 0.6.
CHAPTER 5. DEFINING HADRONIC TAG JET CANDIDATES

5.2 Statistical methods to identify a tag jet

For the tag and probe technique used in this analysis, the two crucial requirements for the tag jet are:

- Good energy response: \( R = \frac{E_{\text{Meas}}}{E_{\text{Truth}}} \) is as high as possible
- Good energy resolution: \( \sigma_E/E \) is as small as possible

A well-known tag jet response is required as a reference to correct the average response of the probe jet to one. Furthermore, a tag with a good response contributes minimally to the \( E_{\text{Miss}}^{T} \), allowing the MPF technique to accurately measure the hadronic activity in the probe jet. The best energy resolution possible is required to keep the tag jet uncertainty to a minimum, in turn minimizing the MPF uncertainty. To achieve these goals, two templates have been built using the PYTHIA dataset. The first template is discussed in section 5.2.2, and used to calculate a likelihood according to the jet MC calorimeter response. Section 5.3 builds a second template from the first template, calculating the energy correction required to calculate the dijet MPF from equation 4.19. The goal of this section is to identify and implement a method to separate tag jet candidates according to their calorimeter response.

Defining 'signal' and 'background'

Ideally, a tag jet would be fully contained within the EM calorimeter, and consist solely of \( \pi^0 \)'s, \( \gamma \)'s, and \( e^\pm \)'s, because it would have a response of one. In reality, jets of this nature are extremely rare, thus the criteria must be loosened. Signal jets are defined as the MC jets with the best calorimeter response, consisting of a minimum of 10% of the MC dataset, while the rest are background. The critical value \( C_{\text{value}} \) that defines the boundary between signal and backgrounds jets is

\[
C_{\text{value}} = p0 + p1 \times \log(p_T) + p2 \times \log(p_T)^2
\]  

(5.6)

where \( p0 \approx 0.41, p1 \approx 0.13, \) and \( p2 \approx -0.0074 \). For example, a jet of 100 GeV \( (p_T) \) must have a MC calorimeter response higher than \( C_{\text{value}} = 0.84 \) to be a signal jet, while a 600 GeV \( (p_T) \) jet corresponds to \( C_{\text{value}} = 0.92 \). Equation 5.6 was made through trial and error, balancing the need for signal jet statistics with high responding jets. Figure 5.9 shows that
Figure 5.9: The number jets selected in PYTHIA as signal and background, plotted on a log scale as a function of $p_T$.

The signal jets account for approximately 10% of the statistics at low energy, and 30% at higher energies.

5.2.1 Frequentist

The Frequentist approach is concerned with counting the number of times one observes a signal jet given N jets, where N is large. To be able to identify a signal jet, the jet variables ($V_i$) listed in table 5.4 are used. Figures 5.10, 5.11, 5.12, 5.13, 5.14, and 5.15 plot the normalized number of signal jets ($N_S$) and number of background jets ($N_B$) for the variables listed in table 5.4 in two $p_T$ bins. These figures are indicators of what variables provide good discrimination power between signal and background jets.

A signal-like jet deposits maximal energy in layers EMB1 and EMB2, and minimal
energy in EMB3, TILE0, TILE1, TILE2, and the pre-sampler. One interesting combination can be taken using the calorimeter layers, \((\text{PRE}+\text{EMB3}+\text{TILE0}+\text{TILE2})/(\text{EMB1}+\text{EMB2})\); where the numerator is small, and denominator large for signal jets. The width, mass, and number of towers are highly correlated, all favouring small values. These variables indicate that signal jets prefer minimal hadronic activity, as spallation reactions are a primary cause for wide jet development. Jet track variables that make use of the inner detector information are a powerful separator. Both the number of tracks in a jet and the percentage of \(p_T\) carried by tracks is preferred to be small for signal jets. Finally, the jet length, defined as the number of layers a jet must traverse before it has deposited 97% of its energy (\(\text{PRE} = 1, \ldots, \text{TILE2} = 7\)) can also be used as a discriminator. Signal jets have a higher probability to be contained within the first four layers. The probability \(P(S)\) of the jet being a signal jet can be approximated by

\[
P(S(V_i,x_j)) = \frac{N_S(V_i,x_j)}{N_{\text{data}}(V_i,x_j)} \quad \text{(5.7)}
\]

where \(N_{\text{data}}\) is the total number of jets and \(N_{\text{signal}}\) is the number of signal jets, for variable \(i\) and bin \(j\). In the limit that \(N_{\text{data}} \to \infty\), \(P(S)\) will converge to the true probability. Assuming

<table>
<thead>
<tr>
<th>Moment name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIDTH</td>
<td>(\sum \Delta R(\text{jet axis, constituents}) \times p_T(\text{constituents})) where sums run over constituents</td>
</tr>
<tr>
<td>PRE, EMB(X), TILE(X)</td>
<td>Fraction of jet energy in calorimeter sampling layer (X)</td>
</tr>
<tr>
<td>n90</td>
<td>Number of cells that contain 90% of the jet energy</td>
</tr>
<tr>
<td>Number of Tracks</td>
<td>Number of tracks associated to the jet</td>
</tr>
<tr>
<td>(\sum p_T^{\text{Trk}} / p_T^{\text{calorimeter}})</td>
<td>Percentage of jet (p_T) associated with tracks</td>
</tr>
<tr>
<td>maximum (p_T) layer</td>
<td>Maximum calorimeter sampling layer energy deposition</td>
</tr>
<tr>
<td>Shower length</td>
<td>Pre-sampler = 1, \ldots, \text{TILE2} = 8</td>
</tr>
<tr>
<td>Number of Towers</td>
<td>Number of layer traversed until 97% of the jets energy is deposited</td>
</tr>
<tr>
<td></td>
<td>Pre-sampler = 0, \ldots, \text{TILE2} = 7</td>
</tr>
<tr>
<td>Mass</td>
<td>The number of towers spanned by the jet</td>
</tr>
<tr>
<td>EMFRAC</td>
<td>Percentage of the energy deposited in the Electromagnetic Calorimeter</td>
</tr>
</tbody>
</table>

Table 5.4: Standard jet properties given by ATLAS
independence between the jet variables, combined probability can be written as

$$L_{\text{Freq}} = \prod_{i \in \{V\}} P(S(V_i, x_j))$$  \hspace{1cm} (5.8)$$

However, the $V_i$'s used for the analysis (section 5.2.3) can be shown to have correlations. The correlations between the variables weaken the separation power of equation 5.8.

### 5.2.2 Binning

The variation of jet properties with energy, the activity in the calorimeter surrounding the jet, and the jet development itself must be taken into account when calculating a likelihood. To do this, jet variables have been binned in $p_T$ and the quantity ISO08, which is a measure of jet isolation (see Eq. 5.10). These templates are created for signal and background jets, and used to determine the tag jets and correction for the MPF analysis.

#### $p_T$ binning

A sliding window approach is implemented in order to look at jet properties as a function of $p_T$. The idea behind the matching is to find the true average per bin. The first step is to find a function that accurately describes the number of jets as a function of $p_T$, as seen in figures 5.3 and 5.4.

$$\text{Number of Jets} = 1.25 \times 10^{22} \times P_T^{-7.1} - 1.5 \times 10^{24} \times P_T^{-8.4} + 6.7 \times 10^{17} \times P_T^{-4.8}$$  \hspace{1cm} (5.9)$$

where equation 5.9 has been fit for the two highest $p_T$ jets in a dijet event. The sliding window will range between 60 - 800 GeV, in steps of 10 GeV, with a window size increasing linearly from 115 GeV to 170 GeV. An algorithm is then applied to select the bin centre, calculated by weighting the $p_T$ (x-axis) by the function in Eq. 5.9; therefore taking account of the non-symmetric distribution in the bins. For example, the first window has a range of 60 - 175 GeV with a mean $p_T$ of 70 GeV, while a higher window of 450 - 600 GeV has a mean $p_T$ of 500 GeV.
Figure 5.10: A signal jet favours a higher energy deposition in the EMB1 and EMB2 layer, and a lower energy deposition in the pre-sampler layers.
Figure 5.11: A signal jet favours lower energy deposition in the EMB3, TILE0 and TILE1 layers.
Figure 5.12: A signal jet favours lower energy deposition in the TILE2 layer. The combination of all calorimeter layers shows that signal jets favour a high electromagnetic fraction (EMFRAC) and a low longitudinal energy deposition profile (\((\text{Tile0+Tile2+EMB3+PRE}) / (\text{EMB1+EMB2})\))
Figure 5.13: The EMB2 layer (3) absorbs the most energy over 90% of the time. The percent of energy deposited in the maximum layer and the jet length are excellent discriminating variables between signal and background, suggesting well measured jets are denser than the average jet.
Figure 5.14: The n90 variable confirms that well measured jets are denser than the average jet. The width and mass of a jet are both correlated to the jet opening angle, with signal jets favouring a smaller radius.
Figure 5.15: The number of towers is another way to measure the width of the jet, using the physical calorimeter structure and providing the same conclusion as the width variable. Signal jets favour fewer than average tracks, shown by the number of tracks and percentage of jet $p_T$ carried by tracks.
ISO08 binning

A simple and effective method to look at the external activity surrounding a jet is to sum the energy of all reconstructed jets within a cone in eta-phi space, such that the jet boundaries overlap. ISO08 is defined as:

$$ISO08 = \left( \sum_{j} P_{T}^{j} \right) / P_{T}^{i}, i \neq j$$ (5.10)

for jet $i$, where the $j$ jets must lie within a cone radius 0.8 from the jet axis. Thus, ISO08 is a measure of the surrounding jet activity, scaled with respect to the reference jet $p_{T}$. One major reason to look at the surrounding activity of the tag jet is to control the width variable that is used for the likelihood discussed in section 5.2.1. It is possible to have a jet with a larger width, and no surrounding activity. An example of such a jet would be a jet where a large number of spallation interactions have occurred. However, events can occur where a jet is highly collimated, but surrounding activity gives a misleadingly large width. Thus the ISO08 variable controls the external noise, giving a truer width value. Figures 5.16 and 5.17 show the isolation variable ISO08 is weakly correlated to the width and number of tracks associated to a jet.

The binning separates jets into two different classes, corresponding to low and medium levels of surrounding energy. The isolated jets are defined to have ISO08 < 5%, and non-isolated jets have 5% < ISO08 < 25%. Any jet measuring an ISO08 > 25% is not considered as a tag jet candidate, ensuring that only soft energy surrounds the tag jet. Figure 5.33 shows the tag jet pile-up calculated using an ATLAS tool after the MPF code has run, with ISO08 of the tag jet being less than 25%. The ATLAS pile-up tool including both in-time ($N_{PV}$) and out-of-time ($\mu$) pileup, shows that the majority of the jets have 3-4% of their energy from external noise.

5.2.3 The $V_{i}$ templates

All jets considered in creating the templates must pass all cleaning criteria, to ensure that the likelihood template jets are of the same quality as the MPF jets. The following algorithm is applied, using the two highest $p_{T}$ jets in the event:
Figure 5.16: The mean width variable is weakly correlated to surrounding activity, calculated by ISO08.

Figure 5.17: As the energy increases, the mean number of tracks associated with a jet becomes loosely correlated to ISO08.

1. Determine the correct bin for the input jet $p_T$ and ISO08. The jet is ignored if either $p_T$ or ISO08 is outside the allowed range.

2. Use the following jet variables as inputs ($V_i$'s):

   WIDTH: $\frac{\sum \Delta R(\text{Jetaxis, constituents}) \times P_{\text{constituents}}}{\sum P_{\text{constituents}}}$ where the sums run over all the constituents

   Energy fraction in EMB1: $\frac{\text{EMB1}}{\text{Energy of Jet}}$

   Energy fraction in EMB2: $\frac{\text{EMB2}}{\text{Energy of Jet}}$

   Percent of jet $p_T$ from tracks: $\frac{\sum P_{\text{Trk}}^T}{P_T^{\text{calorimeter}}}$

   Longitudinal energy deposition profile: $(\text{Tile0+Tile2+EMB3+PRE}) / (\text{EMB1+EMB2})$

   Layer with maximum energy deposition: $\max(EMB_X, TILE_X)$

   Length of Jet: Number of layers until 98% of the jets energy is deposited

3. Save the signal and background templates separately.

4. Calculate $P(S(V_i,x_j))$ from equation 5.7 for each $V_i$ from the saved templates.

These functions are built using the truth information given from the PYTHIA MC, relying on the assumption that the data $V_i$'s are well described by the MC, and validated in sec-
CHAPTER 5. DEFINING HADRONIC TAG JET CANDIDATES

5.3 The $\alpha$ templates

Armed with the likelihood function to classify jets with respect to their calorimeter response, it is possible to define the $\alpha$ correction required to calculate the MPF given by equation 4.19. Figure 5.20 displays the MC response of a jet, as a function of energy and likelihood. It can be seen that the likelihood separator does a good job separating the jets according to their MC calorimeter response. The number of jets is shown in figure 5.21, emphasizing that a small number of jets are given a low likelihood value (high MC response). Figure 5.22 combines figure 5.20 and 5.21 for all energy ranges, with the mean MC response profile plotted together with the number of jets. While the best tag jet candidates have a likelihood less than five, the minimal statistics would prevent an accurate MPF calculation.

Likelihood template algorithm

In an effort to increase the sample size and have the best resolution for the MC response, the $\alpha$ correction is calculated in bins of likelihood. Thus a jet with a likelihood of 3 is treated much differently than a jet with a likelihood of 10. This deconvolution of the total MC response allows for improved resolution of the tag jets, reducing the overall uncertainty of the MPF. The following algorithm, which is applied to the two highest $p_T$ jets in the event, will be used to identify a tag jet candidate

- Determine the correct ISO08 bin for the jet.
  - If the jet has an ISO08 value outside the binned range, the jet is rejected.
- For each $V_i$ discussed in section 5.2.3:
  - determine the correct $p_T$ bin
- Calculate $L_{Freq}$ from equation 5.8 using the $P(S(V_i,x_j))$ as input
Figure 5.18: The probability that a jet is a signal jet, P(S), is plotted for the likelihood inputs variables ($V_i$'s) for two different energy bins, and for ISO08 $< 5\%$. 
CHAPTER 5. DEFINING HADRONIC TAG JET CANDIDATES

Figure 5.19: The probability that a jet is a signal jet, \( P(S) \), is plotted for the likelihood inputs variables (\( V_i \)'s) for two different energy bins, and for ISO08 < 5%.
Figure 5.20: The MC jet response after the frequentist likelihood has been applied.

Figure 5.21: Number of events corresponding to the frequentist likelihood shown in figure 5.20.
Figure 5.22: The mean MC response (black) is plotted along with the number of jets (red), with the calculated likelihood on the x-axis. No cuts on the jet $p_T$ are applied to produce this figure.
• Determine the correct bin for $L_{Freq}$ ([0, 7], [7, 10], [10, 12], [12, 14], or [14, 16])

• Save the MC response, jet $p_T$, and $L_{Freq}$

Once this algorithm has run, an average MC response ($\alpha$) is determined for each $p_T$ bin. For each $L_{Freq}$ bin, a gaussian fit of the MC response is performed for each $p_T$ bin. The value of the $\alpha$ correction is taken from the curve fit using the equation

$$\alpha_i = A_i + B_i \times \log(p_T) + C_i \times \log(p_T)^2 + D_i \times \log(p_T)^3$$  \hspace{1cm} (5.11)

where $A_i$, $B_i$, $C_i$, and $D_i$ are the fitting parameters, and $i$ corresponds to the likelihood bin.

The power of the likelihood binning can be seen in figure 5.23 and 5.24. Figure 5.23 has the final $\alpha$ correction values, where the mean values are calculated from the gaussian fits that are shown in figure 5.24, and finally fit to equation 5.11. Figure 5.24 clearly shows that using the likelihood to classify jets greatly improves the resolution of the tag jet response. The $\alpha$ correction is determined as follows when calculating the MPF:

• Determine the correct ISO08 bin for the jet

• Determine the correct $p_T$ bin for the jet, and calculate the likelihood

• Determine the correct likelihood bin for the jet

• Fetch the $\alpha$ correction value from the templates shown in figures 5.24

The performance of the likelihood algorithm is discussed in section 6.2, validated against the HERWIG and PYTHIA MC datasets.

### 5.3.1 Ensuring the MC describes the data variables

While the $\alpha$ correction can be directly calculated in MC, the ATLAS dataset cannot be used directly to determine if $\alpha$ corrects the average tag jet response to one. The underlying assumption is that if the MC models the input variables to the likelihood well, then the $\alpha$ correction derived in the MC simulation can be applied to the data. Figures 5.25, 5.26, 5.27, 5.28, 5.29, 5.30 and 5.31 plot the likelihood variable distributions before and after the likelihood is applied. The figures are split into four different $p_T$ bins and plotted for
The $\alpha$ correction

Figure 5.23: The $\alpha$ correction is the MC calorimeter response of a tag jet. The jets are binned according to their calculated likelihood, and binned in $p_T$. Each $p_T$ bin is fit according to a gaussian curve as shown in figure 5.24. The mean values from the gaussian fits are then fit with equation 5.11.
Figure 5.24: Two $p_T$ bins from figure 5.23 are shown, demonstrating the gaussian nature of the calorimeter response in bins of likelihood.
PYTHIA and data. All variables are within uncertainties before and after the likelihood selection is applied, showing that the MC describes the data for the seven input variables. As expected from section 5.2.3, tag jets favour higher energy deposition EMB1 and EMB2, a smaller width and shorter shower length, and less $p_T$ in the tracks.

Figure 5.32 is a projection of figure 5.21 onto the likelihood axis, showing that the MC simulation and data calculate the same number of tag jets for each likelihood value. Finally, figure 5.33 is produced using the ATLAS pile-up calculation tool, with the number of primary vertices ($N_{PV}$) and the average number of interactions per bunch crossing ($\mu$) as inputs. The fraction of the measured tag-jet energy coming from pileup is shown to be 3-4% on average. Since the Monte Carlo simulation reproduces the data, it is a reasonable to assume that the $\alpha$ correction derived in Monte Carlo is valid for data.
Figure 5.25: The calculated width of a jet.
Figure 5.26: The first EM calorimeter layer beyond the Pre-Sampler.
Figure 5.27: The primary EM calorimeter layer.
Figure 5.28: A ratio of calorimeter layers is used as a powerful discriminator.
Figure 5.29: The percent of the jets $p_T$ that corresponds to tracks.
Figure 5.30: From the maximum energy deposition calorimeter sampling layer, the percent of energy deposited in that layer is shown.
Figure 5.31: The shower length, where 1-7 indicate what layer the jet finished showering 96% of its energy. 1 corresponds to the pre-shower, and 7 corresponds to TILE2.
CHAPTER 5. DEFINING HADRONIC TAG JET CANDIDATES

Figure 5.32: This figure is the projection of figure 5.21 on the likelihood axis for data and the PYTHIA MC simulation. The number of tag jets selected for a MPF study are binned in likelihood, where the likelihood is calculated as the negative logarithm of equation 5.8.

Figure 5.33: ATLAS’s pile-up estimation tool for pile-up for the tag jet in data and the Monte Carlo simulation.
Chapter 6

Calorimeter response

The Missing $E_T$ Projection Fraction technique determines the calorimeter response, that when combined with a calorimeter showering correction\(^1\), provides the absolute energy scale for jets [23]. This chapter presents a study of the MPF calorimeter response using dijet events in proton-proton collisions at 7 TeV center-of-mass energy with the ATLAS detector. The dijet MPF curve is then combined with the $\gamma$+jet MPF curve, to calculate separate quark and gluon MPF responses. The first step required in calculating a flavour (quark or gluon) response is to predict the flavour content of the probe jet from the MPF calculations.

6.1 Quark and gluon fraction calculation

To calculate the quark or gluon MPF response derived from the dijet and $\gamma$+jet MPF curves, the assumption must be made that the MC correctly predicts the fraction of quark and gluon jets produced. Since it is not known which MC generator gives the more accurate flavour content, the probe jet flavour fraction is calculated as the weighted average between the two models.

For each MPF sample (dijet and $\gamma$+jet) and each MC generator (PYTHIA and HERWIG), the probe jet flavour is identified by the parton level matching discussed in sec-

\(^1\)The showering correction can be described as accounting for the net energy flow across the jet boundary [23].
Figures 6.1 and 6.2 show the quark fraction for the different samples. The two MC samples disagree somewhat, with differences up to 10% for the dijet flavour at highest $p_T$ (figure 6.2). The origin of the flavour difference in the dijet sample at higher energies is not known. To account for the flavour difference, a weighted average is calculated between the two MC samples as

$$f_{\text{quark}} = \frac{N_{\text{PYTHIA}}}{N_{\text{PYTHIA}} + N_{\text{HERWIG}}} f_{\text{PYTHIA}} + \frac{N_{\text{HERWIG}}}{N_{\text{PYTHIA}} + N_{\text{HERWIG}}} f_{\text{HERWIG}}$$  \hspace{1cm} (6.1)$$

where $N_{\text{PYTHIA}}$ and $N_{\text{HERWIG}}$ are the number of events in the Monte Carlo samples used to determine $(f_{\text{PYTHIA}} \text{ quark}, f_{\text{HERWIG}} \text{ quark})$ in a given energy bin. The final weighted average probe jet quark fraction is shown in figure 6.3, with the uncertainty calculated in appendix A.2. To perform an unbiased analysis of MC and data, this combined result for the quark fraction will be used for all MC samples and data.

### 6.2 Validating the dijet MPF $\alpha$ correction

To test the performance of the $\alpha$ correction on the tag jet, the MPF code is run on samples from both HERWIG and PYTHIA MC generators. Due to the different showering models in these two generators, described in section 4.1.2, one would expect slightly different results. Figures 6.4 and 6.5 show the MC response for the tag jet after the $\alpha$ correction has been applied.

Figure 6.4 shows that the PYTHIA MC generator achieves an $\alpha$ closure to within 0.3% at particle level, while the parton response closure ranges from 1.4% at lowest energies up to 0.8% at highest energies. The HERWIG MC generator produces a larger difference between parton and particle level $\alpha$ closures, indicating that HERWIG calculates a smaller (in $p_T$) particle level jet than PYTHIA (see figure 6.5). HERWIG measures the particle level $\alpha$ closure 1.5% high in the lowest values of $p_T$, decreasing to 0.5% at the highest $p_T$; while the parton level $\alpha$ closure is slightly worse than PYTHIA, at 1.5-2.0%. It is not surprising that the closure is better at particle level for both generators because the alpha correction is calculated at particle level. The difference between particle and parton level is some measure of uncertainties introduced by fragmentation effects.
CHAPTER 6. CALORIMETER RESPONSE

Figure 6.1: The probe jet quark fraction for the $\gamma$ + jet sample

Quark Fraction from the $\gamma$ + Jet Probe Jet

Figure 6.2: The probe jet quark fraction for the dijet sample

Quark Fraction from the DiJet Probe Jet

Figure 6.2: The probe jet quark fraction for the dijet sample
Figure 6.3: The weighed average (according to sample size) of the probe-jet quark fraction.
CHAPTER 6. CALORIMETER RESPONSE

It is not known which model best describes the true physics better, thus figure 6.6 has been produced showing the weighted average (weighted by sample size) between PYTHIA and HERWIG. The mean value of the weighted average $\alpha$ closure is 1 to within 0.3%, with an uncertainty of 1% at the lowest energies, decreasing to 0.2% at the highest energies.

As seen in section 5.3, the likelihood selection applied to tag jets produces jets with a better resolution. Figure 6.7 compares the MPF response for $\alpha$ corrected tag jets from the dijet MPF selection to the photons used in the $\gamma$+jet analysis. The distributions are fit using a Gaussian function.

$$G(\mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where $\sigma^2$ is the variance, and $\mu$ is the mean. The results of the fits using equation 6.2 show the following relationship between the standard deviation ($\sigma$) of the tag jets and the photons

$$\sigma_{dijet} \approx 3\sigma_{\gamma+jet} \quad (6.3)$$

at lowest energy, decreasing to

$$\sigma_{dijet} \approx 2\sigma_{\gamma+jet} \quad (6.4)$$

at highest energy. While binning the likelihood did improve the resolution (figure 5.23), the dijet variance could not be resolved to the precision of the photons. However, the centroid of the distribution can be well determined in both cases.

6.3 MPF

To measure the calorimeter response of jets, the MPF takes advantage of the transverse momentum balance between a well measured tag object and a jet. The final validation required before calculating a MPF response curve is the missing energy in the event, as the MET variable is related to the entire hadronic recoil against the tag jet.
Figure 6.4: The PYTHIA MC tag jet response, black is the parton response \( \left( \frac{P_T^{\text{calorimeter}}}{P_T^{\text{parton}}} \right) \), and red is the particle jet response \( \left( \frac{P_T^{\text{calorimeter}}}{P_T^{\text{particle}}} \right) \).

Figure 6.5: The HERWIG MC tag jet response, black is the parton response \( \left( \frac{P_T^{\text{calorimeter}}}{P_T^{\text{parton}}} \right) \), and red is the particle jet response \( \left( \frac{P_T^{\text{calorimeter}}}{P_T^{\text{particle}}} \right) \).
CHAPTER 6. CALORIMETER RESPONSE

6.3.1 MET

The MET has a direct correlation to the tag jet selection. In pure $\gamma$+jet events, i.e. no ISR or FSR, the MET points in the direction of the jet, as the photon does not contribute to the MET. The tag jets have been selected using the likelihood criterion to have a good jet response, ensuring that the MET will also be in the direction of the probe jet. Figure 6.8 shows that this is indeed the case. Furthermore, figure 6.9 shows that the MC models the MET within uncertainty.

6.3.2 The dijet and $\gamma$+jet MPF response

To ensure a fair comparison between the $\gamma$+jet and dijet MPF, the following selection criteria are applied.

- The tag object is confined to $|\eta| \leq 0.8$
- The probe jet is confined to $|\eta| \leq 1.2$
Figure 6.7: Comparing the final MC response ($p_T^{\text{calorimeter}} / p_T^{\text{particle}}$) of the $\alpha$ corrected tag jet (black) to the MPF response for photons (red). These figures are for the PYTHIA MC generator.
CHAPTER 6. CALORIMETER RESPONSE

Figure 6.8: The angle in radians between the tag jet and the MET for data and Monte Carlo simulation; good agreement is seen.

- $|\phi_{Tag\ Object} - \phi_{Probe\ Jet}| > 2.8$
- $P_{3}^{Jet} < \max(12 \text{ GeV}, 0.2 \times P_{Tag\ Object}^{Jet})$

The tag object must be confined to $|\eta| \leq 0.8$ to use the likelihood variables discussed in section 5.2.3. If the $\eta$ is any larger, then the endcap calorimeter variables would have to be used, increasing the complexity of the likelihood algorithm. Setting a probe jet range of $|\eta| \leq 1.2$ is standard practice from current $\gamma + \text{jet}$ MPF studies. These two criteria are vital to an accurate comparison, as the calorimeter response in ATLAS is a function of $\eta$ [8].

While the $\gamma + \text{jet}$ MPF response curve is capable of accurately measuring a MPF response for $p_T$ as low as 30 GeV, the range of validity for the dijet MPF is above 70 GeV due to jet reconstruction and triggering, as discussed in section 5.1. Figure 6.10 shows the MPF results for PYTHIA, HERWIG, and data for $\gamma + \text{jet}$ events. The two MC generators are in excellent agreement, while the data measures 2% lower. This difference is seen in the in-situ $\gamma + \text{jet}$ calibration of the jet-energy scale and is used to correct the JES derived from Monte Carlo simulation [10].

The dijet MPF response curve is plotted in figure 6.11 for HERWIG, PYTHIA, and data. The two MC generators predict nearly the same MPF response, while the data show the same 1-2% difference seen in the $\gamma + \text{jet}$ MPF response curve. The Gaussian fits used
CHAPTER 6. CALORIMETER RESPONSE

Figure 6.10: The $\gamma$+jet MPF response curves.

Figure 6.11: The dijet MPF response curves.
in each $p_T$ bin for figure 6.11 can be seen in appendix A.1, while the error bars are discussed in appendix A. The dijet MPF technique provides results up to 800 GeV, while the $\gamma$+jet data only goes up to $\approx$ 650 GeV because of the reduced sample size at high energies (appendix A.1).

### 6.3.3 Quark and gluon calculated response curve

Previous sections built the inputs required to calculate quark and gluon calorimeter response curves, i.e. $R_{\text{jet}}^{\gamma}$, $R_{\text{MPF}}^{\gamma}$, $f_{\text{jet}}^{\text{quark}}$, and $f_{\gamma}^{\text{quark}}$. Figures 6.12 and 6.13 show all four input variables for PYTHIA and data, where the $\gamma$+jet response curve is higher than the dijet response curve due to the larger quark fraction in $\gamma$+jet events and the higher response of quark jets. Both graphs show that as the energy increases, the quark fractions and the MPF curves converge. These graphs clearly indicate the correlation between jet flavour and response.

To remove the bin-to-bin fluctuations for the calculated quark and gluon MPF responses, the dijet and $\gamma$+jet MPF curves are fitted using equation 4.6, with the corresponding uncertainties discussed in appendix A.3. Equations 4.22 and 4.23 are used to calculate the quark and gluon MPF response curves in figures 6.14 and 6.15. The difference in the calculated quark and gluon MPF responses ranges from 4-6% for data and 4-7% for PYTHIA. These differences are within the expected range of 3-8% given by Schwartzman [14] in section 4.2.1. The $\gamma$+jet MPF response curve closely follows the calculated quark curve, as the $\gamma$+jet events have 80-90% quark probe jets. These plots show that gluon jets are being over-corrected while quark jets are being under-corrected by using the current $\gamma$+jet MPF as a JES correction.
Figure 6.12: The four inputs to calculate a quark and gluon MPF response curves from data.
Figure 6.13: The four inputs to calculate a quark and gluon MPF response from the PYTHIA simulation.
Figure 6.14: The calculated quark and gluon MPF response from data, alongside the PYTHIA $\gamma$+jet MPF curve. The difference in response between the two jet flavours range from 4–7%. The $\gamma$+jet MPF curve closely follows the calculated MPF quark curve.
Figure 6.15: The calculated quark and gluon MPF response from PYTHIA, alongside the PYTHIA $\gamma$+jet MPF curve. The difference in response between the two jet flavours range from 4-6%. The $\gamma$+jet MPF curve closely follows the calculated MPF quark curve.
Chapter 7

Conclusion

The jet energy scale usually contributes the largest systematic uncertainty in any physics analysis, making it vital to understand and minimize at ATLAS. This thesis provides a comprehensive study of the quark and gluon jet responses, by comparing results of the missing transverse energy projection fraction method in the established gamma+jet events and the newly developed dijet events. This analysis was made possible through a likelihood function that corrected the tag jet calorimeter response to within 0.5%.

An extensive analysis of how the jet response depends on jet properties has been described. The properties of the jets with respect to calorimeter response were explored, exposing powerful discriminating variables in the calorimeter and inner detector. A likelihood function was built, capable of distinguishing jets according to their calorimeter response using Monte Carlo simulation. The likelihood discriminator makes it possible to tag jets that have fluctuated to have a good calorimeter response, and correct the response to one for a dijet MPF study. Using the likelihood to classify jets has been shown to greatly improve the resolution of the tag jet response, and in turn the MPF.

The resulting MPF response for dijets can be combined with the response from gamma+jet events to extract separate calorimeter responses to quark and gluon jets. The responses of quark and gluon jets is shown to differ by 4-7%, depending on jet $p_T$. Applying separate jet responses for quark and gluon jets has the potential of decreasing the contribution to the systematic uncertainty of the jet energy scale. One method to implement a quark/gluon JES correction would be to calculate the final state quark fraction of the process. From the Monte Carlo generated quark fraction, one could calculate the JES from the quark and
gluon MPF response curves. For example, if a process has three jets in the final state, and
the MC generator says two are quark and one is a gluon (at a given energy range), then the
user could calculate a JES correction as $\frac{2}{3} R^{\text{quark}}_{\text{MPF}} + \frac{1}{3} R^{\text{gluon}}_{\text{MPF}}$. Unfortunately,
the large jet by jet fluctuations prevent individual jet flavour tagging, thus one cannot apply
a jet by jet flavoured JES. Nonetheless, the analyses presented here demonstrates that a
quark-gluon JES is possible, giving great promise to achieving a 1% JES at ATLAS in the
coming years.
Appendix A

Error Analysis

A.1 MPF bins

Each MPF bin is fit using a Gaussian, from the form:

\[ G(\mu, \sigma) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}} \]  \hspace{1cm} (A.1)

with \( \mu \) representing the location of the Gaussian mean, and \( \sigma \) the width. The MPF error is not taken from the Gaussian fit \( \sigma \), but from the uncertainty in the fit. This is done by first fitting the data to a Gaussian at the calculated mean location. A second Gaussian fit is then performed where the range is restricted to \( \mu \pm \sigma \). The uncertainty for the MPF calculation is taken as the fitting error of the second Gaussian fit. This technique has been developed in Ref. [10].

Figures A.1, A.2, and A.3 are the distributions for the \( \gamma \)+jet (red) and dijet (black) events, where the energy bins are arranged from lowest energy at the top left, scrolling right to the highest energy at the bottom right. Figure A.1 shows that the \( \gamma \)+jet statistics become poor in the highest energy bin, displaying how the \( \gamma \)+jet events are responsible for the upper limit of \( \approx 650 \) GeV for the quark and gluon MPF calculation.
Figure A.1: The normalized MPF distributions for the $\gamma$+jet (red) and dijet (black) data sample, with MPF response on the x-axis.
Figure A.2: The normalized MPF distributions for the $\gamma$+jet (red) and dijet (black) PYTHIA sample, with MPF response on the x-axis.
Figure A.3: The normalized MPF distributions for the $\gamma$+jet (red) and dijet (black) HER-WIG sample, with MPF response on the x-axis.
A.2 Quark and Gluon Fraction

To calculate the quark and gluon fraction, a weighted mean is implemented for the $\gamma$+jet and dijet samples, using PYTHIA and HERWIG as the two inputs:

$$f^j_i = \frac{N^j_i \text{herwig} + N^j_i \text{pythia}}{N^j_i \text{herwig} + N^j_i \text{pythia}}$$  \hspace{1cm} (A.2)

where $i$ denotes either the $\gamma$+jet or dijet MPF sample, and $j$ denotes quark or gluon. The calculated values $f^j_i$ have an error associated with them from the discrepancy between the two MC generators. For uncorrelated observations with standard deviations $\sigma^j_i$ and weight $w_k$, the weighted sample mean has standard deviation of:

$$\sigma_k = \sqrt{\sum_{i=1}^{n} (w_k^j)^2 (\sigma^j_k)^2}$$  \hspace{1cm} (A.3)

The standard deviations $\sigma_k$ is calculated as:

$$\sigma_k = f^k_{\text{quark}} - f^\text{Weighted Average}_{\text{quark}}$$  \hspace{1cm} (A.4)

and the weighted average is

$$w_k = \frac{(N_{\text{PYTHIA}})}{N_{\text{PYTHIA}} + N_{\text{HERWIG}}}$$  \hspace{1cm} (A.5)

for each energy bin $k$.

A.3 Fitting

When calculating the MPF response, fits are implemented to reduce statistical fluctuations from bin to bin, at the price of increasing the error bars. The uncertainties are increased by the values:

$$\sigma_{fit} = \sqrt{\frac{\sum_{i=1}^{n} (x_{fit} - x_{raw})^2}{n(n-1)}}$$  \hspace{1cm} (A.6)

where $n$ represents the number of bins in the MPF curves. The $\sigma_{fit}$ uncertainties are then convoluted with the original uncertainties, yielding
\[ \sigma_{Total} = \sqrt{\sigma_{fit}^2 + \sigma_{raw}^2} \] (A.7)

where \( \sigma_{raw} \) is the errors from either the MPF or the \( f_i^j \).

### A.4 Quark and Gluon \( R_{MPF} \)

Equations 4.22 and 4.23 have the same form; thus, only a sample calculation for equation 4.23 will be shown here. Equations 4.23 is a function of four variables, \( R_{MPF}^{\gamma}, f_{\gamma}^{q\gamma}, R_{MPF}^{f_{\gamma}^{q\gamma}}, \) and \( f_{\gamma}^{f_{\gamma}^{q\gamma}} \), all with their own errors (\( \Delta R_{MPF}^{\gamma}, \Delta f_{\gamma}^{q\gamma}, \Delta R_{MPF}^{f_{\gamma}^{q\gamma}}, \) and \( \Delta f_{\gamma}^{f_{\gamma}^{q\gamma}} \)). Equations ?? without any uncertainties reads:

\[ R_{MPF}^{\text{gluon}} = \frac{R_{MPF}^{\gamma}(f_{\gamma}^{q\gamma}) - R_{MPF}^{f_{\gamma}^{q\gamma}}(f_{\gamma}^{q\gamma})}{f_{\gamma}^{\gamma} - f_{\gamma}^{q\gamma}} \] (A.8)

assuming all four variables have independent errors, the uncertainty in equation A.8 can be calculated as:

\[ R_{MPF}^{\text{gluon}} \pm \Delta R_{MPF}^{\text{gluon}} = (R_{MPF}^{\gamma} + \Delta R_{MPF}^{\gamma})(f_{\gamma}^{q\gamma} + \Delta f_{\gamma}^{q\gamma}) - (R_{MPF}^{f_{\gamma}^{q\gamma}} + \Delta R_{MPF}^{f_{\gamma}^{q\gamma}})(f_{\gamma}^{f_{\gamma}^{q\gamma}} + \Delta f_{\gamma}^{f_{\gamma}^{q\gamma}}) 
\]

\[ (f_{\gamma}^{q\gamma} + \Delta f_{\gamma}^{q\gamma}) - (f_{\gamma}^{f_{\gamma}^{q\gamma}} + \Delta f_{\gamma}^{f_{\gamma}^{q\gamma}}) \]

(A.9)

Let

\[ N_0 = (R_{MPF}^{\gamma} f_{\gamma}^{q\gamma} - R_{MPF}^{f_{\gamma}^{q\gamma}} f_{\gamma}^{f_{\gamma}^{q\gamma}}) \]

\[ N_1^A = R_{MPF}^{\gamma} f_{\gamma}^{q\gamma} \sqrt{\left( \frac{\Delta R_{MPF}^{\gamma}}{R_{MPF}^{\gamma}} \right)^2 + \left( \frac{\Delta f_{\gamma}^{q\gamma}}{f_{\gamma}^{q\gamma}} \right)^2} \]

\[ N_1^B = R_{MPF}^{f_{\gamma}^{q\gamma}} f_{\gamma}^{f_{\gamma}^{q\gamma}} \sqrt{\left( \frac{\Delta R_{MPF}^{f_{\gamma}^{q\gamma}}}{R_{MPF}^{f_{\gamma}^{q\gamma}}} \right)^2 + \left( \frac{\Delta f_{\gamma}^{f_{\gamma}^{q\gamma}}}{f_{\gamma}^{f_{\gamma}^{q\gamma}}} \right)^2} \]

\[ D_0 = (f_{\gamma}^{q\gamma} - f_{\gamma}^{f_{\gamma}^{q\gamma}}) \]
\[ D_1 = \sqrt{(\Delta f_{\text{jet}}^{\text{quark}})^2 + (\Delta f_1^{\text{quark}})^2} \]

then equation A.9 can be simplified to

\[
R_{\text{gluon MPF}} \pm R_{\text{MPF}}^{\text{gluon}} = \frac{N_0}{D_0} \pm \frac{N_0}{D_0} \sqrt{\left(\frac{\sqrt{(N_1^A)^2 + (N_1^B)^2}}{N_0}\right)^2 + \left(\frac{D_1}{D_0}\right)^2} \quad (A.10)
\]
Bibliography


[17] Yousuke Kataoka, Ginga Akimoto, Naoko Kanaya, Shoji Asai. SUSY WG background meeting, University of Tokyo, Unpublished ATLAS Internal Note.


