

# TRAPS:

*Topological Reconstruction Algorithm for Parton Scatters*

**K.V. Ellis**

Thesis submitted for the degree of Doctor of Philosophy

**Supervisor:** Prof. Graham Thompson

**Co-Supervisors:** Dr. D. Traynor and Mr. M. Landon

May 2012



*To my Father, who  
led me to believe that everyone does a PhD, and my Mother, who let him get  
away with it*

## Declaration

No portion of the work documented in this thesis has been submitted in support of an application for another degree or qualification at this or any other institute of learning.

## Author's Contribution

This work is largely my own, but has been carried out as part of the ATLAS collaboration, and hence input has come from several other sources. These are acknowledged here:

### TRAPS

- The TRAPS algorithm was written into code solely by the author, with inspiration, ideas and feedback from Graham Thompson.
- The skeleton code to generate events using Rivet was supplied by Dan Traynor.
- The Athena wrapper for TRAPS was devised and implemented by John Morris.

### L1Calo

- The timing analysis was written by the author with early coding assistance by Murrough Landon.
- The Gauss-Landau fit function was devised by Paul Seidler, with further developments by Felix Mueller, Michael Henke and Valerie Lang.
- The simple fit method was suggested by Stephen Hillier.

# Abstract

There is strong motivation to study standard model physics using the highest-energy data provided by the Large Hadron Collider. This is aided by the process of defining clusters of hadrons to form ‘jets’. Existing jet-finders are dependent on pre-defined parameters which, to some extent, influence their properties. This thesis introduces a novel algorithm which aims to reconstruct partons outgoing from hard interactions, *prior* to any splitting, by concentrating solely on the highest momentum transfer scale. In this way parton properties such as fragmentation and structure functions from hadron colliders may be compared directly with results from DIS and  $e^+e^-$  annihilation.

This original, standalone tool is named ‘TRAPS’ - the Topological Reconstruction Algorithm for Parton Scatters. The algorithm was developed using PYTHIA Monte Carlo QCD events, under a pragmatic approach that assumes the model provides a good approximation to reality at both hadronic and partonic level.

Various tests were made to gauge the performance of the algorithm against standard jet-finders. The infrared safety and algorithm speed were also assessed. The objective of TRAPS is to have low sensitivity to parameters, and to be fast and robust. A high event acceptance is necessary, as maximum statistics are required where cross-sections are at their lowest.

A chapter of this thesis is dedicated to a description of the author’s studies in calibration and monitoring of the timing of the ATLAS Level-1 Calorimeter Trigger system. Pulses from triggered energy are sent via largely  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  granularity ‘trigger towers’. Synchronous triggering with 1-2 ns precision is required for the system to make an accurate energy estimate.

## Acknowledgements

First and foremost I would like to thank my supervisor, Prof. Graham Thompson, whose enthusiasm for our work together was always inspirational. Even when he claimed to be ‘depressed’, it remained an enthusiastic depression which served to spur me on! His detailed explanations continue to be very much appreciated, and I have greatly enjoyed our many wide-ranging chats about physics. I wish him well in his retirement, but happily, Graham will always be a physicist.

Thanks go to my secondary supervisors Dr Dan Traynor and Mr Murrough Landon. Dan supported my use of Rivet and provided useful input on fragmentation functions. I was very appreciative of Murrough’s patient explanations on the workings of L1Calo, and also his help with coding.

I am extremely grateful to Dr John Morris for his interest in my work and programming support. Without John, TRAPS in Athena would not look nearly as professional (or would not exist at all!). Also for his attempts to improve my coding practice. And to Kyle Stevenson who was often helpful with programming.

I thank the entire L1Calo sub-group for their welcoming nature and friendly attitude to collaboration. Particular thanks are to Stephen Hillier, who was always interested in my investigations. Also to Paul Seidler, Felix Mueller, Michael Henke and Valerie Lang from Heidelberg who shared their L1Calo timing results and code.

Thanks to all the other PhD students who have made office life so enjoyable. Top mentions go to James Poll, Alex Hyndman, Matilde Castanheira, Elisa Piccaro, Tom Macey and Jack Goddard. Also to the LTA crew at CERN, who always knew how to provide a good time. And particularly to

Mark Stockton for his endless thesis-writing advice!

Best wishes to the particle physics group at QMUL, which is going from strength to strength. Thanks to everyone for making the group such a friendly place to work. To the computing guys who solve even the most stupid of problems.

Thanks to my parents who are there for me when I need them. To Dad, who is always interested in my work, and to Mum who has to put up with physics-chat over dinner.

Thanks to everyone who read chapters from this thesis and helped bash it into shape, particularly Graham, who will need a new red pen for Christmas, and Murrough who corrected me on lots of L1Calo technical details.

Finally, I acknowledge STFC who financially supported this PhD.

Katy Ellis

23rd of September 2011

# Contents

<b>1</b>	<b>Introduction</b>	<b>19</b>
1.1	Partons and QCD . . . . .	23
1.1.1	Flavour and Colour . . . . .	24
1.1.2	Parton Interactions . . . . .	25
1.2	Thesis Layout . . . . .	29
<b>2</b>	<b>Phenomenology of Parton Scattering</b>	<b>30</b>
2.1	Kinematics . . . . .	30
2.1.1	$x$ . . . . .	30
2.1.2	Mandelstam Variables . . . . .	31
2.1.3	Momentum Transfer . . . . .	33
2.1.4	Rapidity and Pseudorapidity . . . . .	33
2.2	Luminosity and Cross-section . . . . .	34
2.3	The Structure of the Proton . . . . .	36
2.3.1	PDFs: Parton Density Functions . . . . .	36
2.3.2	Fragmentation Function . . . . .	39
2.4	Underlying Event . . . . .	40
2.4.1	Initial and Final State Radiation . . . . .	40
2.4.2	Proton Remnant . . . . .	43
2.4.3	Multiple Parton Interactions and Pile-up . . . . .	43

<i>CONTENTS</i>	8
2.5 Monte Carlo Modelling . . . . .	44
2.5.1 Parton Showering . . . . .	45
2.5.2 Hadronisation . . . . .	46
2.5.3 Pythia . . . . .	47
2.5.4 Rivet . . . . .	49
2.5.5 Athena . . . . .	49
2.6 Final State Particles . . . . .	50
<b>3 LHC and ATLAS</b>	<b>52</b>
3.1 The Large Hadron Collider . . . . .	52
3.1.1 LHC in Numbers . . . . .	54
3.2 Brief Description of ATLAS . . . . .	56
3.2.1 ATLAS in Numbers . . . . .	57
3.2.2 Co-ordinate System . . . . .	57
3.2.3 Detection Concepts . . . . .	58
3.2.4 Detector Components . . . . .	59
3.2.5 Accessing ATLAS Data . . . . .	66
<b>4 L1Calo Timing Studies</b>	<b>67</b>
4.1 The ATLAS Trigger . . . . .	67
4.2 Brief Description of L1Calo . . . . .	69
4.2.1 Physical Layout . . . . .	71
4.2.2 L1Calo Timing . . . . .	72
4.3 Timing Verification Using Cosmic Rays . . . . .	75
4.3.1 Method I . . . . .	76
4.3.2 Results I - Interpartition Timing . . . . .	79
4.4 Timing Monitoring Using Collision Data . . . . .	83
4.4.1 Method II - ‘Simple’ Method . . . . .	84



<i>CONTENTS</i>	9
4.4.2 Method Comparison . . . . .	86
4.4.3 Results II - Timing Monitoring . . . . .	90
4.5 Timing Problems in the Overlap . . . . .	97
4.5.1 Results III - Calibration of Overlap Regions . . . . .	98
4.6 Summary . . . . .	102
<b>5 Jet-finding Algorithms</b>	<b>103</b>
5.1 Infrared and Collinear Safety . . . . .	104
5.2 Sliding Window . . . . .	107
5.3 Cone-type . . . . .	107
5.3.1 SIScone . . . . .	111
5.4 Recombinatorial . . . . .	111
5.4.1 $k_T$ algorithm . . . . .	113
5.4.2 Variations on the $k_T$ algorithm . . . . .	113
5.5 Comparison of Jet-finding Algorithms . . . . .	115
<b>6 Introducing TRAPS</b>	<b>118</b>
6.1 Motivation: Why Develop TRAPS? . . . . .	118
6.2 Development Techniques . . . . .	120
6.2.1 Algorithm Testing . . . . .	121
6.3 Input Objects . . . . .	122
6.4 Technical Description of TRAPS . . . . .	122
6.4.1 Preparation - Setting $\eta_{range}$ . . . . .	122
6.4.2 Step I - Splitting the Event in Two . . . . .	124
6.4.3 Step II - Floating Boundary Mean . . . . .	128
6.4.4 Step III - Calculation of Standard Deviation . . . . .	131
6.4.5 Step IV - Ellipse Method . . . . .	131
6.4.6 Step V - Underlying Event Subtraction . . . . .	135

<i>CONTENTS</i>	10
6.4.7 Step VI - Virtuality Correction . . . . .	137
6.4.8 Visual Summary . . . . .	138
6.5 Running TRAPS . . . . .	138
6.5.1 Event Flags . . . . .	140
6.5.2 Primary TRAPS outputs . . . . .	141
6.6 Development Tests . . . . .	141
6.6.1 Analysing the Delta Plots . . . . .	143
6.7 Alternative TRAPS Sub-methods . . . . .	146
6.7.1 No Floating Boundary . . . . .	147
6.7.2 4-jet Sink . . . . .	148
<b>7 Using TRAPS</b>	<b>149</b>
7.1 Adaption for Athena . . . . .	149
7.1.1 Official Production MC . . . . .	151
7.1.2 Comparison with Rivet . . . . .	152
7.2 Algorithm Resolution . . . . .	153
7.3 The $n$ -parameter . . . . .	157
7.4 UE Sample Region Size . . . . .	159
7.5 Consideration of Selection Cuts . . . . .	161
7.5.1 Acoplanarity Cut . . . . .	162
7.5.2 Other Cut Possibilities . . . . .	162
<b>8 Evaluating TRAPS</b>	<b>167</b>
8.1 Comparisons with Standard Jet-Finders . . . . .	167
8.1.1 Invariant Mass Calculation Results . . . . .	171
8.1.2 Performance as a Function of Hard Scale . . . . .	174
8.2 TRAPS and IRC Safety . . . . .	176
8.3 Speed of TRAPS . . . . .	181

<i>CONTENTS</i>	11
8.4 Sensitivity to Monte Carlo . . . . .	184
8.5 Exercise on $d\sigma(x)$ . . . . .	188
8.5.1 Employment of Migration Matrix . . . . .	189
8.6 Detector Scaling and Resolution . . . . .	194
<b>9 Conclusions</b>	<b>196</b>
9.1 The Future for TRAPS . . . . .	198

# List of Figures

1.1	The six quarks arranged as three weak isospin-spin doublets. . .	24
1.2	Simple diagram of an undisturbed proton. . . . .	25
1.3	Feynman diagrams of $2 \rightarrow 2$ QCD processes. . . . .	26
1.4	Diagrams of virtual loop corrections in QCD. . . . .	27
1.5	Feynman diagrams of $2 \rightarrow 3$ parton QCD processes. . . . .	28
2.1	A proton-proton collision illustrating the composite nature. . .	31
2.2	Schematic illustrating quarks interacting via gluon exchange. . .	32
2.3	Diagram of two-body scatter labelled with particle four-momenta. .	32
2.4	Schematic illustrating partons forming jets of hadrons. . . . .	34
2.5	The CTEQ 6.6. Parton Distribution Function. . . . .	38
2.6	Fragmentation function as a function of hard scale. . . . .	41
2.7	A simplified view of a proton-proton collision, based on events modelled in PYTHIA. . . . .	42
2.8	Fractional energy of jets by particle type as a function of jet transverse energy. . . . .	51
3.1	Cartoon of LHC, indicating the positions of the main detectors. .	54
3.2	Drawing of the ATLAS detector. . . . .	56
3.3	Diagram indicating relative radii of the detector layers, the remaining 3 m of height being composed of structural supports. .	60

3.4	Diagram illustrating subsystems where particles are observed.	60
3.5	Computer generated image of the full calorimeter. . . . .	62
4.1	Data flow within the ATLAS trigger . . . . .	68
4.2	Block diagram of the Level-1 Trigger. . . . .	69
4.3	Trigger tower granularity for $\eta > 0$ and one quadrant in $\phi$ . . .	73
4.4	Elements used for the $e/\gamma$ and $\tau$ /hadron algorithms . . . . .	73
4.5	A digitised LAr pulse with a functional fit. . . . .	74
4.6	Atlantis event display of a cosmic particle . . . . .	77
4.7	Gauss-Landau fit on a digitised pulse. . . . .	78
4.8	Diagram illustrating combinations for adjacent tower pairs. . .	79
4.9	Diagram illustrating adjacent towers across the $\eta = 0$ boundary.	80
4.10	Deviation in timing from the mean for EM towers across $\eta = 0$ .	81
4.11	Deviation in timing from the mean for had towers across $\eta = 0$ .	81
4.12	Deviation in timing from the mean for EM towers across $\eta = 0$ following corrections. . . . .	82
4.13	Diagram indicating the pulse slices used to calculate finetime.	85
4.14	Schematic examples of L1Calo pulses. . . . .	87
4.15	EM $\eta - \phi$ maps comparing timing results of methods I and II.	88
4.16	Hadronic $\eta - \phi$ map comparing timing results from methods. .	89
4.17	Online monitoring plot of $\Delta finetime$ . . . . .	90
4.18	Difference from ideal in finetime in runs 177924/9. . . . .	91
4.19	Hadronic $\eta - \phi$ map for hits, average timing and $\sigma$ . . . . .	93
4.20	EM $\eta - \phi$ map for hits, average timing and $\sigma$ . . . . .	94
4.21	$\Delta finetime$ tracking of single EM tower. . . . .	95
4.22	EM layer $\Delta finetime$ partition tracking. . . . .	96
4.23	Hadronic layer finetime partition tracking. . . . .	97
4.24	Diagram of the ‘overlap’ region in L1Calo (EM). . . . .	98

4.25	$\eta - \phi$ maps of EM trigger towers timing during the overlap test.	99
4.26	$\eta - \phi$ maps of the timing difference (t(barrel on) - t(endcap on)). . . . .	100
4.27	Deviation in timing of the overlap region. . . . .	100
4.28	Results of timing analysis following the overlap corrections. . .	101
5.1	Diagram illustrating the problem of infrared unsafety. . . . .	106
5.2	Diagram illustrating the problem of collinear unsafety. . . . .	106
5.3	Diagram illustrating the sliding window principle. . . . .	107
5.4	Diagram illustrating the principle of the cone jet algorithm. . .	110
5.5	General approach to the SISCone method. . . . .	112
5.6	Flowchart describing the key steps of the $k_T$ algorithm. . . . .	114
5.7	Diagram showing the different jet areas when using a cone and the $k_T$ algorithm. . . . .	115
6.1	Diagram of a simple 2-jet event, (un)rotated. . . . .	125
6.2	Histogram of acoplanarity, various $\hat{p}_T$ . . . . .	127
6.3	Diagram illustrating ‘positive’ and ‘negative’ semicylinders. . .	128
6.4	Schematic of a PYTHIA event on an $\eta$ - $\phi$ map. . . . .	129
6.5	Schematic demonstrating the floating boundary method. . . . .	130
6.6	Scatter plot of the standard deviations in $\eta$ and $\phi$ . . . . .	131
6.7	Diagram of the TRAPS ellipses placed on the example event. .	133
6.8	Diagram to illustrate the movement of the ellipse during the iterative process. . . . .	134
6.9	Ellipse convergence histogram in two $p_T$ ranges. . . . .	134
6.10	UE assessment positions for case of TRAPS-objects having the same $\eta$ sign. . . . .	136

6.11	UE assessment positions for case of TRAPS-objects having opposite $\eta$ sign. . . . .	136
6.12	Flowchart of TRAPS steps. . . . .	139
6.13	Delta plots for 560-1120 GeV PYTHIA MC. . . . .	145
6.14	Spreadsheet showing delta plot statistics when using or excluding the floating boundary method. . . . .	147
7.1	Block diagram of sequence of packages for TRAPS in Athena. . . . .	151
7.2	Delta plots comparing four PYTHIA JX samples. . . . .	154
7.3	Log scale delta plots comparing four PYTHIA JX samples. . . . .	155
7.4	$\Delta\eta$ as a function of $\eta_{parton}$ . . . . .	156
7.5	Graphs showing dependence of the delta plots on $n$ . . . . .	160
7.6	Graphs showing mean and RMS $p_T$ of the UE regions. . . . .	161
7.7	$\Delta p_T/p_T$ for all events, and with acoplanarity cuts. . . . .	162
7.8	Histograms of possible selection variables. . . . .	165
8.1	Invariant mass of decay products at parton level from excited quarks, 2 TeV/ $c^2$ sample. . . . .	169
8.2	Invariant mass of decay products at parton level from excited quarks, 1 TeV/ $c^2$ sample. . . . .	170
8.3	Histogram of the invariant mass of 2 TeV/ $c^2$ $q^*$ as calculated by the anti- $k_T$ algorithm. . . . .	171
8.4	Histogram of the invariant mass of 1 TeV/ $c^2$ $q^*$ as calculated by various algorithms. . . . .	172
8.5	Histogram of the invariant mass of 2 TeV/ $c^2$ $q^*$ as calculated by various algorithms. . . . .	173
8.6	Graphs of statistics relating to invariant mass calculations. . . . .	175
8.7	Infrared safety test using 1 GeV additions. . . . .	178

8.8	Infrared safety test using 10 MeV additions. . . . .	179
8.9	Infrared safety test using 0.1 MeV additions. . . . .	180
8.10	Scatter plot comparing the speed of TRAPS with contemporary jet-finders. . . . .	183
8.11	Delta plots comparing HERWIG with PYTHIA, J6. . . . .	185
8.12	Delta plots comparing HERWIG with PYTHIA, J3. . . . .	186
8.13	Graphs comparing PYTHIA and HERWIG delta plot statistics.	187
8.14	Inclusive differential cross-section with respect to $x$ , log-log scale. . . . .	192
8.15	Inclusive differential cross-section with respect to $x$ , linear-log scale. . . . .	193
8.16	Inclusive differential cross-section vs. $\hat{p}_T$ . . . . .	194
9.1	Diagram illustrating the un-implemented ‘tilted ellipse’ method.	199



# List of Tables

2.1	Examples of LHC subprocess cross-sections at design luminosity.	36
2.2	Cross-sections of QCD processes. . . . .	37
2.3	Examples of parameters used within PYTHIA. . . . .	48
3.1	LHC Tunnel Characteristics. . . . .	55
3.2	LHC proton beam characteristics. . . . .	55
3.3	ATLAS characteristics. . . . .	57
4.1	Electromagnetic and hadronic calorimeter partitions. . . . .	72
5.1	Relative duration of commonly used jet-finding algorithms. . .	117
6.1	Mean UE $p_T$ summed, per unit area. . . . .	137
6.2	TRAPS default inputs. . . . .	140
6.3	TRAPS event flags. . . . .	140
6.4	Member functions of the TRAPS class. . . . .	142
6.5	Resolution statistics relating to example delta plots. . . . .	146
7.1	Number of events in official production MC samples. . . . .	152
7.2	Resolution statistics relating to PYTHIA J3 delta plots. . . . .	158
7.3	Resolution statistics relating to PYTHIA J4 delta plots. . . . .	158
7.4	Resolution statistics relating to PYTHIA J5 delta plots. . . . .	158

7.5	Resolution statistics relating to PYTHIA J6 delta plots. . . . .	158
7.6	J3 $\Delta p_T/p_T$ statistics for all events and those with acoplanarity < 0.1. . . . .	163
7.7	Cut quantities for example selections. . . . .	166
8.1	Invariant mass statistics for the 1 TeV/ $c^2$ $q^*$ sample. . . . .	172
8.2	Invariant mass statistics for the 2 TeV/ $c^2$ $q^*$ sample. . . . .	173
8.3	Statistical results from the infrared safety test with the addi- tion of a 1 GeV particle. . . . .	178
8.4	Statistical results from the infrared safety test with the addi- tion of a 10 MeV particle. . . . .	179
8.5	Statistical results from the infrared safety test with the addi- tion of a 0.1 MeV particle. . . . .	180
8.6	Resolution statistics relating to HERWIG J6 delta plots. . . . .	188
8.7	Resolution statistics relating to HERWIG J3 delta plots. . . . .	188

# Chapter 1

## Introduction

The Large Hadron Collider (LHC) [1] has extended the accessible energy for which to study fundamental processes. Since 2009 the LHC has been providing proton-proton collisions at centre of mass energy 7 TeV and is scheduled to continue doing so until the end of 2012. A complete shutdown is planned for 2013, after which the collider should begin operating at its design energy of 14 TeV. The multi-TeV scale is required to search for completely new phenomena. However, it is equally necessary to confirm the predicted, extrapolated behaviour of Standard Model observables at these increased energies.

Evidence for quark universality has been shown [2][3][4][5] between experiments involving electron-positron ( $ee$ ) and electron-proton ( $ep$ ) collisions. But there is no definitive proof that these are the same partons as those involved in proton-proton ( $pp$ ) interactions, although they are strongly expected to be so. This is not entirely a given, since the hard process in each case involves different forces - electromagnetic force interactions for  $ee$  and  $ep$  and strong force interactions for  $pp$ . In any case, the ‘reality’ of the partonic nature of the proton can become so much clearer at the higher momentum

attainable by the LHC.

It is generally agreed within the community that the relatively clean environment of  $ep$  scattering is preferable for the study of parton fragmentation functions and PDFs via structure functions (described in chapter 2). However the considerable increase in centre of mass energy at the LHC could give new information and bring to light anomalies providing that suitable techniques are developed to cope with the increased background associated with strong processes. A *jet* fragmentation function has already been measured within the ATLAS collaboration [6].

It is highly beneficial to make fragmentation function and PDF measurements at ATLAS, and not only to investigate QCD at high momentum transfer,  $Q$ . The sizable cross-section of totally hadronic channels allows study of a large range of  $x$ , the proportion of beam momentum taken by the colliding particle. In principle, access is available to low values of  $x$ , thanks to the large rapidity coverage provided by the calorimetry, although often in practice only a limited range is currently analysed.

Due to the emphasis on rare processes and the noisy background normally associated with hadron-hadron scattering, coupled with the energy record during the 1990's and early 2000's being held by the Tevatron machine ( $p\bar{p}$ ) at Fermilab, these kind of standard model measurements have been somewhat neglected. The most recent results for parton fragmentation functions and structure functions are from the H1 and Zeus experiments at the HERA laboratory<sup>1</sup> [7][8] at a CMS energy of 318 GeV, so there is a considerable energy range at the LHC, currently up to 7 TeV, but soon 14 TeV, open for discovery.

---

<sup>1</sup>H1 and Zeus are experiments that operated at the HERA particle accelerator at DESY in Hamburg between 1992 and 2007. HERA collided protons with electrons or positrons.

In terms of the structure function, we may even be in a position to observe the beginnings of a deviation from the quark's point-like structural behaviour, should any exist within the available energy spectrum. Hence we are interested only in jets with the highest transverse momentum that statistics will allow. However our results should overlap with the upper limit of existing measurements which reliably reach a hard scale of approximately 100 GeV.

At the highest  $Q^2$  accessible, events tend to be relatively clean and planar, even at the LHC, as a large fraction of the available momentum has entered into the hard interaction, leaving less for underlying event effects. This implies high  $x$ , since the scattered parton takes a large fraction of the overall proton momentum. Nevertheless, underlying event remains a significant challenge for all analyses. These terms are defined in chapter 2.

The key ingredient for calculations of both structure and fragmentation functions, is the vector momentum of the partons exiting the hard interaction. This information is particularly difficult to obtain from hadron-hadron collisions due to the wide range of possible interaction energies and the comparatively messy final state observed in the detector. Existing jet-finders recombine deposits of calorimeter energy and/or tracks measured by a detector, within parameters that are pre-defined by the user. In combination with jet triggers this gives a good self-consistent definition of a *jet*. This is then assumed to represent a parton, although not necessarily the exiting parton from a  $2 \rightarrow 2$  scatter. Jet finders work well for identifying multi-jets or the signatures of rare events such as the decay of a top quark for example - although the method is flawed if soft radiation alters the final result, as discussed in chapter 5.

Many measurements may become biased by their selection cuts; to find

the parton momentum, the least troublesome method would be to choose a clear sample of back-to-back di-jet events with balanced transverse momentum. This would largely remove next to leading order effects and events with large amounts of underlying event, but it would also greatly reduce the amount of data available to the study. The result could not be said to be correct for *all* QCD events - only those selected di-jets with no significant bremsstrahlung or background, although naturally this could be accounted for. However, it is noted that for the highest momentum jets, where statistics are low, events are typically planar in nature.

For these reasons a novel tool specific to the above requirements has been developed for use with data from the LHC. This option became particularly attractive after the delay to the experimental data from the LHC. Specifically it has been adapted for use on the ATLAS experiment [9].

The principal subject of this thesis is the Topological Reconstruction Algorithm for Parton Scatters (TRAPS). TRAPS is a ‘top down’-type jet-finder for  $2 \rightarrow 2$  scatters, which make up the majority of hard interaction events at the LHC. A brief review of current jet-finding methods can be found in chapter 5. The distribution of final state particle energy deposits within an event is used in an attempt to measure the parton direction and momentum, as provided in the first instance by the Monte Carlo simulation, and then tested using a more robust method of reconstructing the invariant mass of a fictitious ‘excited quark’ simulated at a TeV-scale single mass. This is necessary as partons are considered model-dependent entities, and hence their precise definition may vary between MC models.

From the beginning, it was decided to apply an ethos which would guide development of the algorithm. The ethos stated that the algorithm should be fast, simple and as inclusive of as many events as possible. Even more

importantly, it should also be seedless, and have no fixed parameters. If any were deemed essential, they should be proven arbitrary and insensitive. Thus a ‘top-down’ topological approach was chosen, with internal parameters dictated by the event geometry.

## 1.1 Partons and QCD

A parton is a general term for a constituent of presumed composite hadrons, as recognised by J.D Björken in his 1969 paper [10] and subsequently developed into the parton model by R.P. Feynman [11]. Partons comprise of fermionic quarks and antiquarks, and bosonic gluons, and are among the most fundamental particles known. The upper limit on parton size dictated by current maximum energy attainable is less than  $6 \times 10^{-5} \text{fm}$  [12]. Quarks/antiquarks have electric charge of  $\pm 2/3$  or  $\mp 1/3$  that of an electron. They are held together within hadrons by the strong (colour) force, which is described by an exchange of gluons and easily dominates the electromagnetic force experienced by the charged quarks. Balance between the strong force and the Pauli Exclusion Principle results in the size of a hadron to be of the order of one femtometre.

At close-range, the interactions between partons are weak and the quarks are ‘free’ to move apart but as they do so, their attraction increases linearly, confining them within the colourless hadron. This concept is called infrared slavery [13]. As parton momentum transfer increases the strong coupling constant is reduced, albeit on a logarithmic scale, to give the reverse concept of asymptotic freedom.

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \quad (1.1)$$

Figure 1.1: The six quarks arranged as three weak isospin-spin doublets.

### 1.1.1 Flavour and Colour

There are six quark ‘flavours’ comprising in order of increasing mass, up (u), down (d), strange (s), charm (c), bottom(b) and top (t). They are arranged within the standard model in three weak-isospin doublets as illustrated in figure 1.1 [14].

Quarks are created from the vacuum in quark/anti-quark pairs, e.g.  $u\bar{u}$ ,  $s\bar{s}$ , as are leptons, e.g.  $e^-e^+$  pairs. Quarks and gluons are ‘coloured’ objects, and, except at the very high temperatures known as the quark-gluon plasma, are not thought to exist in a stable form as individual entities<sup>2</sup>. The theory of Quantum Chromodynamics (QCD) states that quarks must combine to create stable ‘white’ hadrons [15], as illustrated in figure 1.2. Each quark has a ‘colour charge’ of red, green or blue; anti-particles have anti-red, anti-green and anti-blue. When a quark/anti-quark pair is created it is considered to have equal probability of being red/anti-red, green/anti-green or blue/anti-blue. Gluons are thought to carry one colour and one anti-colour. There are eight antisymmetric combinations observed, with net colour charge commonly given as  $\frac{2}{3}$  that of the quark. As colour is conserved, the gluons must have the appropriate colour combination to link up the quarks and form hadrons.

It is the conservation of colour charges as a result of a perfect SU(3)

---

<sup>2</sup>On a measurable scale this is true, however a parton ‘liberated’ during a multi-TeV scatter travels a considerable distance compared to its own (measured maximum) diameter, before hadronisation.



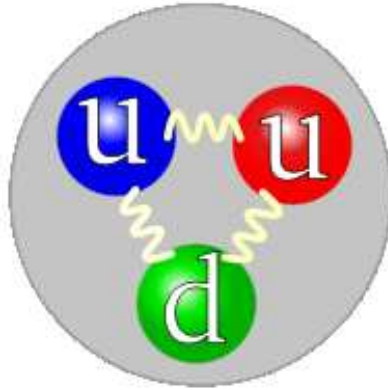


Figure 1.2: Simple diagram of an undisturbed proton, comprising three differently coloured quarks exchanging coloured gluons. [16]

symmetry which manifests as the strong force. Being a perfect symmetry, colour charge cannot be measured directly. The colour charge is only inferred through the statistical nature of quarks and gluons combining to form hadrons and in cross-section calculations such as  $e^-e^+ \rightarrow$  hadrons. According to various models the unknown process of hadronisation could take place by combination with nearby coloured objects or the splitting of gluonic ‘tubes’ of force between them.

### 1.1.2 Parton Interactions

The least complex interaction for a given final state is represented by a Feynman diagram drawn at ‘tree’ level. Tree diagrams for QCD  $2 \rightarrow 2$  parton scatters are shown in figure 1.3. Calculation at this level is known as a first order approximation. Each of the six groupings have different initial and final states. Bracketed combinations illustrate both  $t$ - and  $s$ -channel processes

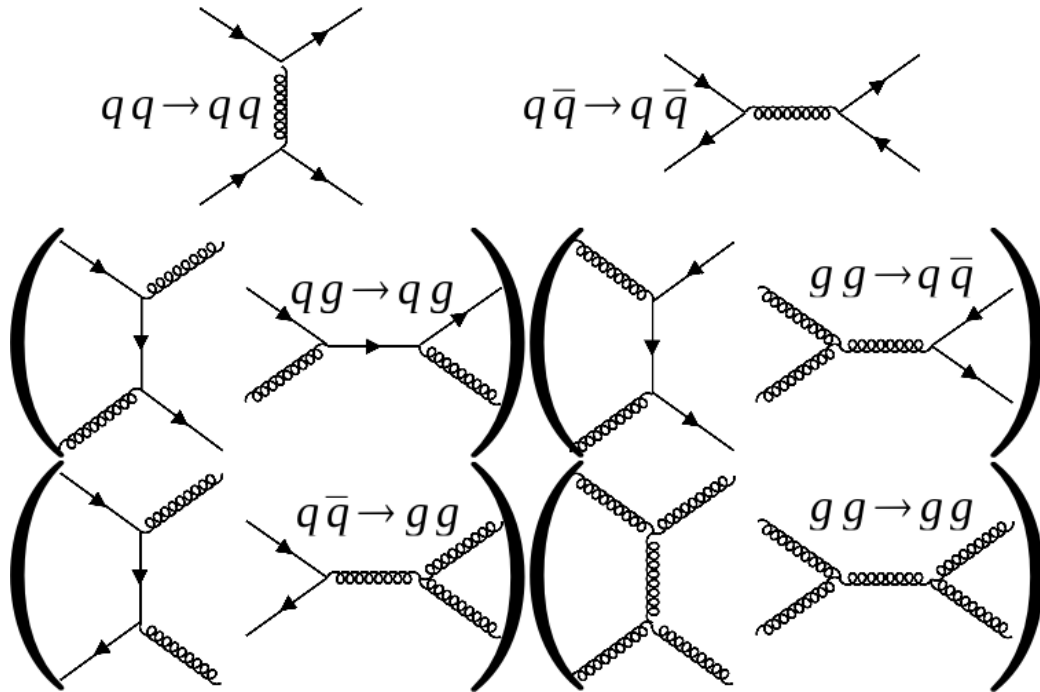


Figure 1.3: Feynman diagrams of  $2 \rightarrow 2$  parton QCD processes. Each quark/anti-quark may be exchanged to make new diagrams. Bracketed combinations have the same initial and final state.

with the same initial and final states which interfere with each other, forming  $u$ -channel processes. However in each case the  $t$ -channel process strongly dominates, as an exchange is more likely than an annihilation.

Parton interactions are dominated by the strong force, the strength of which is determined by the strong coupling constant,  $\alpha_s$ . This varies, or ‘runs’ with the transferred momentum scale of the interaction if higher order processes known as loop corrections, such as those shown in figure 1.4 are suitably absorbed.

The observed size of the *electric charge* is dependent on  $Q^2$ , through a process known as ‘screening’. As a result, the electromagnetic coupling,  $\alpha$ , also runs with  $Q^2$ , although this variation is extremely small. The EM

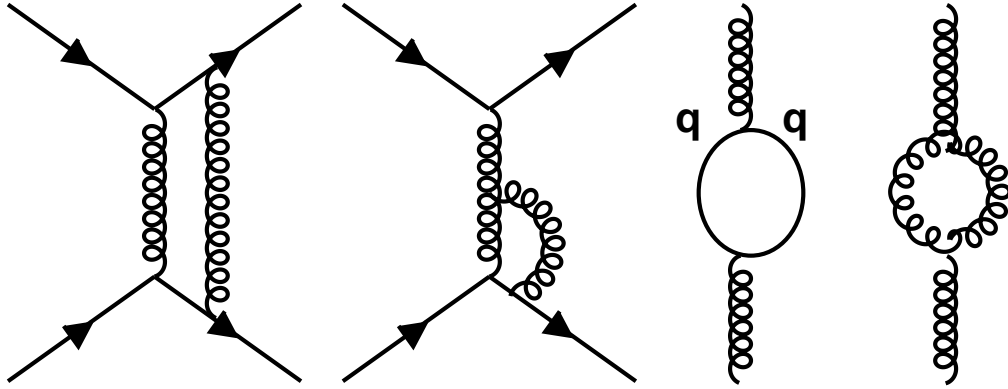


Figure 1.4: Diagrams of example virtual loop corrections in QCD. From left to right the diagrams are known as ‘box’ diagram, a vertex correction, and propagator loops formed of quarks and gluons.

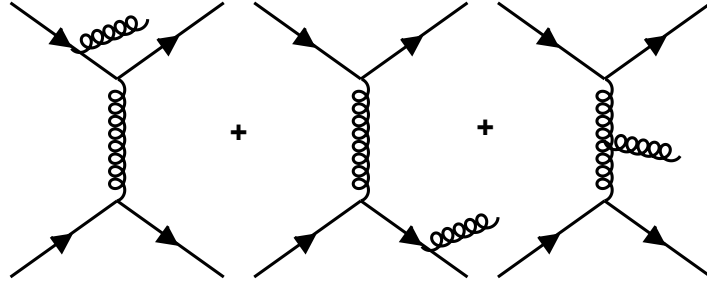
coupling is calculated via equation 1.2, in which  $\alpha_0$  is the size of the coupling at the fixed scale  $M$  - a ‘cut-off’ scale.

$$\alpha(Q^2) = \frac{\alpha_0}{1 - \frac{\alpha_0}{3\pi} \log(Q^2/M^2)} \quad (1.2)$$

The form of the equation for the strong coupling constant resembles that of the EM, but with a different coefficient in the denominator. This alters the behaviour of  $\alpha_s$  to allow for the non-abelian nature of the strong force. The expression for calculation of  $\alpha_s$  is given by equation 1.3 [13], where  $\mu$  is a renormalisation or reference scale to be chosen and  $n_f$  is the number of accessible quark flavours.

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi} (33 - 2n_f) \log(Q^2/\mu^2)} \quad (1.3)$$

It is customary to re-write the expression as in equation 1.4, where  $\Lambda$  is the (low)  $Q^2$  scale at which the coupling would become large, and thus is an absolute limit for perturbative QCD as a viable theory.  $\Lambda$  is determined by

Figure 1.5: Feynman diagrams of  $2 \rightarrow 3$  parton QCD processes.

experiment and given by equation 1.5.

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)\log(Q^2/\Lambda^2)} \quad (1.4)$$

$$\Lambda^2 = \mu^2 \exp\left(\frac{-12\pi}{(33 - 2n_f)\alpha_s(\mu^2)}\right) \quad (1.5)$$

Screening effects are provided by quark/anti-quark pairs, however anti-screening is also contributed by virtual gluon loops. The fact that gluons carry twice the colour charge of quarks increases the influence of the anti-screening. Multiple virtual loop terms are not shown here, but may be found with more detail in [17].

Processes that are  $2 \rightarrow 3$  or more partons are also expected. The simplest examples of  $2 \rightarrow 3$  tree level diagrams are shown in figure 1.5, where a gluon is radiated from any of the other partons (initial, final or propagator). Higher orders are calculated in a perturbative expansion, with smaller corrections summed with increasing order of  $\alpha_s$ . The coupling increases with decreasing energy so that when  $Q$  is low, the coupling is large, and perturbation theory can not be used.

## 1.2 Thesis Layout

This chapter gave a brief introduction to partons and their fundamental interactions. The following chapter lays out some of the background phenomenology of parton scattering and specifies variables relevant to the work in this thesis. Chapter 3 briefly summarises the experimental apparatus, comprising the Large Hadron Collider and the ATLAS detector. A description of the ATLAS trigger subsystem, and particularly the Level-1 Calorimeter Trigger, can be found in chapter 4, as well as details on the author's work on the timing of the subsystem.

In chapter 5, a review of existing jet-finding techniques is given, with some remarks on their relative advantages. The final version of the TRAPS algorithm itself is described in detail in chapter 6, alongside discussion of alternate sub-methods that were investigated and rejected. Chapter 7 addresses the use of TRAPS within the Athena framework, and gives final results from the testing that was done during development. It also addresses dependence on the main parameter involved and investigates what measurement improvements may be made through selection cuts. Chapter 8 is an evaluation of the TRAPS method, including comparison with commonly used jet-finders, assessment of its infrared safeness and a speed test. In addition, the sensitivity of TRAPS to alternative Monte Carlo is observed, an example usage of TRAPS in the measurement of  $x$  within MC is given, as well as a first-look at the effect of detector resolution. Conclusions and thoughts on further work can be found in chapter 9.

# Chapter 2

## Phenomenology of Parton Scattering

This chapter summarises the background theory and phenomenological tools of parton scattering. Some kinematic variables and concepts which will be used throughout this thesis are defined, and the structure of the proton is discussed. Expected sources of coherent and incoherent background noise are detailed. Finally, modelling techniques are outlined.

### 2.1 Kinematics

#### 2.1.1 $x$

The interaction of two protons yields a far more complex picture than electron-positron or even electron-proton collisions, the parameters of which can be calculated by Deep Inelastic Scattering (DIS) techniques. In the case where both objects are composite, as indicated schematically in figure 2.1, it is not intrinsically obvious which fundamental particles have interacted, and with

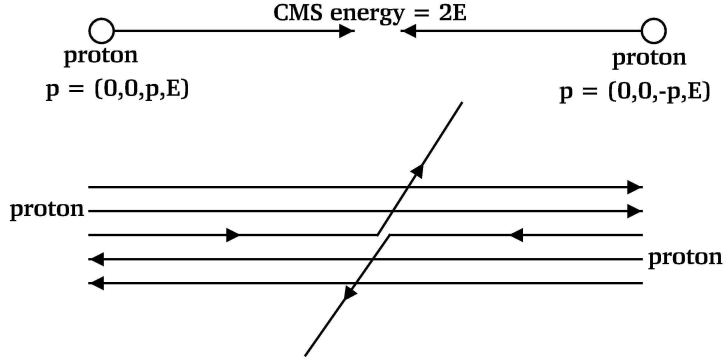


Figure 2.1: A proton-proton collision illustrating the composite nature. An individual quark or gluon from each proton collide elastically, leaving the (coloured) remainder to continue in approximately its original direction.

how much momentum. There is no constraint for momentum to be shared equally between all partons within a proton. The proportion of the proton’s 4-momentum taken by a particular scattered parton is given, at the LHC, by the Lorentz invariant  $x$  as indicated by the parton interaction in figure 2.2.

### 2.1.2 Mandelstam Variables

The three Mandelstam variables,  $s$ ,  $t$  and  $u$  fully describe the kinematics of a two-body scatter in a Lorentz invariant manner. They may be calculated from the four-momenta of the particles before and after the interaction, as given by equations 2.1 - 2.3 where the particle momenta are defined in figure 2.3. The sum of these three variables gives the summed rest masses squared of the four particles involved in the scatter, as given by equation 2.4.

$$s = (p_1 + p_2)^2 \tag{2.1}$$

$$t = (p_1 - p_3)^2 \tag{2.2}$$

$$u = (p_1 - p_4)^2 \tag{2.3}$$

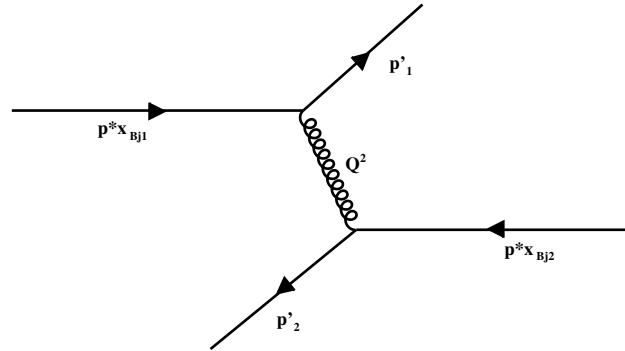


Figure 2.2: Schematic illustrating quarks interacting via gluon exchange.

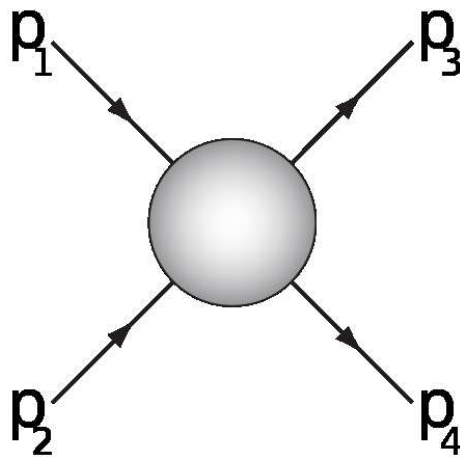


Figure 2.3: Diagram of two-body scatter labelled with particle four-momenta.



$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2 \quad (2.4)$$

### 2.1.3 Momentum Transfer

Two partons interact with one another via gauge boson exchange, which is characterised by a momentum transfer squared  $Q^2$  as shown in figure 2.2. The momentum transfer,  $Q$ , is equal to  $\sqrt{-t}$ . In high energy QCD-type collisions, the propagator is gluon or quark. Interactions with higher  $Q^2$  have greater resolving power; these interactions probe the hadrons to smaller dimensions. A photon with  $Q^2$  of 1 GeV<sup>2</sup> has a wavelength of the order of one femtometre and at this energy a proton is seen by the photon as a point particle. Events with greater  $Q^2$  probe ‘deeper’ into the proton, revealing inner structure.

At higher  $Q^2$  the probe may scatter from one of the partons within the proton, and in the process knock the parton out. At the scale of the proton, this is an inelastic collision. In the course of this interaction, other particles may come into existence, producing jets of hadrons that can be detected as illustrated in figure 2.4. However, from the parton point of view the interaction is essentially an elastic scatter and the effective mass of each individual jet is usually very much smaller than  $Q$ . In hadron-hadron collisions the transverse component of the momentum transfer is sometimes referred to as  $\hat{p}_T$ .

### 2.1.4 Rapidity and Pseudorapidity

Due to the differing  $x$  of the interacting particles the resulting object and its products will be boosted along the axis, often significantly. Hence the  $\theta$  angle between the z-axis and the object is not appropriate to describe the

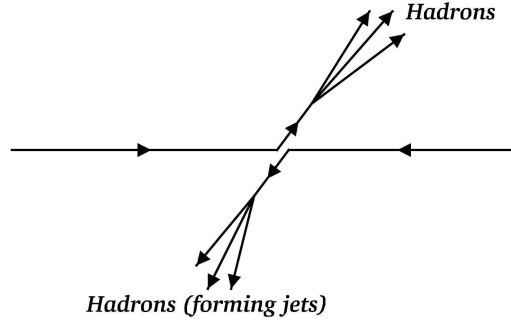


Figure 2.4: Schematic illustrating struck partons ejected from the proton forming jets of hadrons.

kinematics, and rapidity,  $Y$  is used instead. An object's rapidity is defined by its energy and longitudinal momentum, as given in equation 2.5. Rapidity differences between two particles ( $\Delta Y$ ) are an invariant quantity with respect to boosts along the  $z$ -axis.

$$Y = \frac{1}{2} \ln \left[ \frac{E + p_z}{E - p_z} \right] \quad (2.5)$$

Pseudorapidity is equivalent to rapidity in the relativistic limit (i.e. for massless particles), and is directly calculable from the angle between the track and the beam axis,  $\theta$ , via equation 2.6. At the energies being studied and large  $\theta$  it is almost equivalent to rapidity.

$$\eta = -\ln \left[ \tan\left(\frac{\theta}{2}\right) \right] \quad (2.6)$$

## 2.2 Luminosity and Cross-section

Luminosity,  $\mathcal{L}$ , for a single beam, is the number of beam particles passing through one unit area per second. Machine luminosity for a Gaussian beam distribution in a collider is calculated from beam parameters as shown in equation 2.7 [1], where  $N_b$  is the number of particles per bunch,  $n_b$  the number

of bunches per beam,  $f_{rev}$  the revolution frequency,  $\gamma_r$  the relativistic gamma factor,  $\epsilon_n$  the normalised transverse beam emittance,  $\beta^*$  the beta function at the collision point and  $F$  the geometric luminosity reduction factor due to the crossing angle at the interaction point.

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r F}{4\pi \epsilon_n \beta^*} \quad (2.7)$$

Cross-section ( $\sigma$ ) is an imaginary area associated with target particles, that have to intersect for a particular interaction to occur [18]. Its ratio to a unit area of beam represents the probability for a specific interaction to occur and it is related to luminosity by equation 2.8 where  $R$  is the rate of the interaction.

$$\mathcal{L}\sigma = R \quad (2.8)$$

Cross-sections in particle physics are typically quoted in barns, where one barn is equivalent to  $1 \times 10^{-24}$  cm<sup>2</sup>, although even the most common processes have cross-sections of the order of millibarns (mb). Certain key processes are extremely rare and have comparatively low cross-sections. For example at the LHC, pair production of  $t$ -quarks is 100 million times less probable than an inelastic proton-proton scatter (see table 2.1). Example cross-sections of QCD processes which are most relevant to this thesis can be found in table 2.2. These values are based on 100,000 events generated at CMS energy ( $\sqrt{s}$ ) of 7 TeV and with a  $Q$  of 100 and 1000 GeV.

Experimentally, we measure cross-sections using equation 2.9, in which  $(N_{obs} - N_{bg})$  is the number of events remaining after background subtraction,  $\epsilon$  is the efficiency at which the process is recorded,  $A$  is the detector acceptance and  $BR$  is the branching ratio for the signature being measured.

Process	cross-section
Inelastic pp	80 mb
$t\bar{t}$ production	833 pb
Higgs production	$\sim$ fb

Table 2.1: Examples of LHC subprocess cross-sections at design luminosity. Higgs predictions vary with expected mass and production method [19] [20].

$$\sigma = \frac{N_{obs} - N_{bg}}{\mathcal{L} \cdot \epsilon \cdot A \cdot BR} \quad (2.9)$$

## 2.3 The Structure of the Proton

Undisturbed, a proton is said to be composed of three quarks consisting of two up and one down quarks, bound together by a number of gluons, as in figure 1.2. These are ‘valence’ quarks and contribute additively to the quantum numbers associated with the proton. In addition, pairs of virtual ‘sea’ quarks/antiquarks and copious gluons pop in and out of existence. When the proton is probed at high energies, these virtual partons are observed. Hence, it is not a trivial matter to state the components of the protons in collision.

### 2.3.1 PDFs: Parton Density Functions

A Parton Density Function (PDF) gives the number density of a particular quark type or the gluon within a fast moving proton. This is denoted by  $f_a(x)$ , where the index  $a$  is the parton type. As PDFs are non-perturbative, and difficult to calculate in lattice QCD, their values must be extracted from data [17].

Process	Cross-section, pb	
	$Q = 100 \text{ GeV}$	$Q > 1000 \text{ GeV}$
$qq' \rightarrow qq'$	$3.545 \times 10^4$	$2.313 \times 10^{-1}$
$q\bar{q} \rightarrow q'\bar{q}'$	$6.587 \times 10^2$	$2.851 \times 10^{-3}$
$q\bar{q} \rightarrow gg$	$5.682 \times 10^2$	$1.664 \times 10^{-3}$
$gg \rightarrow qq$	$1.653 \times 10^5$	$1.293 \times 10^{-1}$
$gg \rightarrow q\bar{q}$	$4.810 \times 10^3$	$5.107 \times 10^{-4}$
$gg \rightarrow gg$	$1.184 \times 10^5$	$1.107 \times 10^{-2}$

Table 2.2: Cross-sections of QCD processes taken from 100,000 events modelled with PYTHIA version 6.418 at minimum momentum transfer of 100 and 1000 GeV.

Measurement of the parton density is an important matter in itself. The precision of PDFs dictate the accuracy achievable in many experimental measurements as well as strongly affect the predictive powers of theoretical calculations. At low energies the PDF is a non-calculable and non-perturbative feature which then varies in a QCD calculable manner with the hard scale of the probe.

At lower probe energies only gluons, and up and down sea quark pairs are likely to be created in the virtual sea; at greater energies the possibilities cover increasingly the strange, charmed, bottom and top. Gluons have a larger colour charge so dominate the cross-section, after which up, down and then strange, commonly referred to as the ‘light’ quarks, are most frequently observed. However, the net quantum numbers of the object do not change and hence it remains a proton.

A plot of fitted PDF data for accessible partons at  $Q^2 = 10000 \text{ GeV}^2$  as

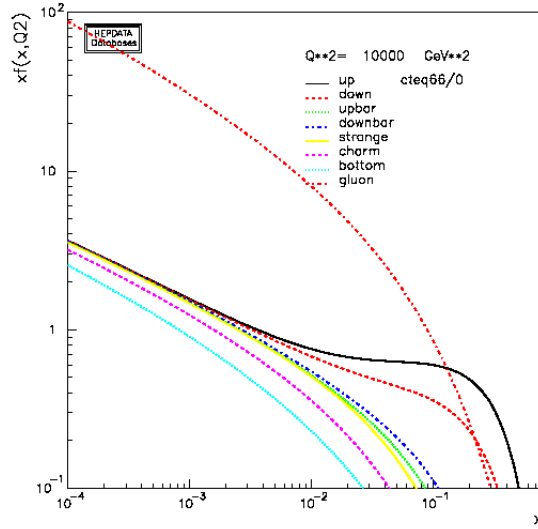


Figure 2.5: The CTEQ 6.6. Parton Distribution Function at  $Q^2 = 10000$   $\text{GeV}^2$  [21].

a function of  $x$  is shown in figure 2.5 [21]. It is clear that at this energy the distribution is dominated by gluons, at all but the highest  $x$ . At  $x = 10^{-4}$  the experimental data ends, and only extrapolations to the fitted data exist beyond this value.

The variation of parton density with probing energy is described by the structure functions,  $F_2(x, Q^2)$  and  $F_L(x, Q^2)$ . These are dimensionless parameters which vary as a function of  $x$  and  $Q^2$ , although both become approximately independent of scaling for  $Q^2$  greater than a few  $\text{GeV}^2$ , but well below the mass of the  $Z$ .

To zeroth order in the strong coupling constant  $\alpha_s$ ,  $F_2(x, Q^2)$  and  $F_L(x, Q^2)$  are related to the PDFs via equations 2.10 and 2.11, in which  $f_a(x)dx$  is the probability to find a quark or anti-quark  $a$  of charge  $e_a$  and momentum fraction between  $x$  and  $x + dx$ . PDFs given independent of  $Q$  correspond to the simplified quark-parton model. A more complex treatment of  $F_2$ , incorpo-

rating higher orders of calculation, can be found in [17].

$$F_2(x, Q^2) = \sum_a e_a^2 x f_a(x) \quad (2.10)$$

$$F_L(x, Q^2) = 0 \quad (2.11)$$

### 2.3.2 Fragmentation Function

The concept of measuring the fragmentation of a high-energy quark to a jet was first documented by R.D. Field and R. Feynman in their 1977 paper [22] following their earlier investigations [23]. A model was described to show how formations of quark pairs and subsequently mesons may occur, although it is emphasised that this model is not considered sound physical theory. The probability  $f(\eta)$  of the of the remaining jet has a fraction  $\eta$  of the original jet is chosen, such that the final charged hadron distribution agrees with experimental data.

A 7 TeV collision picked at random creates a large number of final state hadrons, each with a small proportion of the overall momentum. But sometimes fewer hadrons are found in the final state, each carrying a large fraction of their parent parton momentum.

The scaled momentum,  $x_p$ , defined by equation 2.12, is the fraction of the total available partonic momentum taken by a given, charged, final state hadron. Here,  $p^h$  is the momentum of a given hadron and  $p^p$  is that of the parent parton.

$$x_p = \frac{p^h}{p^p} \quad (2.12)$$

The fragmentation function,  $D(x_p)$  is then the inclusive probability of finding a charged hadron with fraction  $x_p$  of the parent parton, and is given

in equation 2.13. The integral of  $D$  over all  $dx_p$  is the average number of charged final state particles created by a particular interaction at momentum transfer  $Q$ .

$$D(x_p, Q) = \frac{1}{N} \frac{dn^\pm}{dx_p} \quad (2.13)$$

where  $N$  is the number of events and  $n^\pm$  is the charged multiplicity. Figure 2.6 shows the fragmentation function as measured by a number of  $e^+e^-$  and  $ep$  experiments plotted against the appropriate hard scale, and split into nine  $x_p$  bins. The original version of this plot can be found in [24], and demonstrates ‘scaling violations’. Given the logarithmic scale of  $Q$ , these curves change relatively little on a linear scale, but what change they exhibit is to increase the proportion of hadrons with a low share of the parton’s momentum, at the expense of a few of those with a high  $x_p$ .

## 2.4 Underlying Event

‘Underlying Event’ (UE) refers to physics processes not directly related to the main event which is being studied. There are many potential causes of UE associated with the  $pp$  hard interaction signal. Figure 2.7 is a diagram showing the hard scatter labelled M.E. (Matrix Element) as well as UE which is directly related to the main event.

In this thesis UE will occasionally be referred to as ‘background’, in this sense meaning a background to the main, hard scale event.

### 2.4.1 Initial and Final State Radiation

Partons which are off mass shell will radiate gluons, as illustrated in figure 1.5. Initial State Radiation (ISR), also known as Initial State Bremsstrahlung



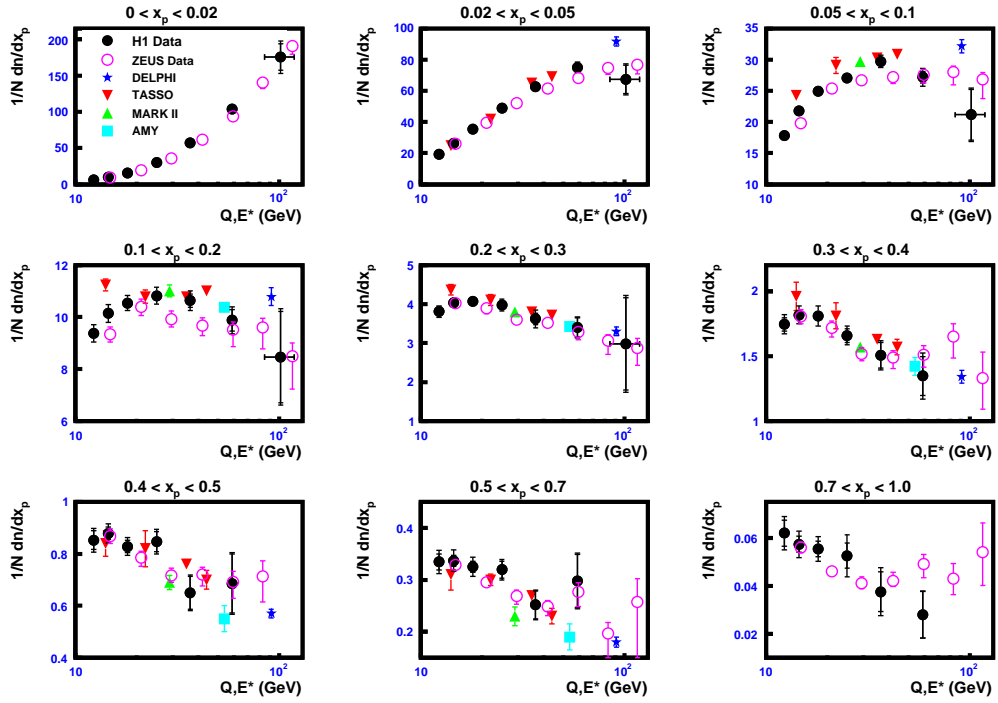


Figure 2.6: Quark fragmentation function,  $D$  for charged particles as a function of the hard scale and split into nine  $x_p$  bins, measured by  $e^+e^-$  and  $ep$  experiments.

(ISB) is, within a pragmatic view, the emission of a gluon *prior* to the hard scatter, i.e. from an incoming parton on the Feynman diagram. In figure 2.7 these are partons coming from particles labelled 3 and 4, but which do not enter the hard interaction. In the scheme of this thesis, ISR *is* a background effect, a factorisable and largely incoherent process, not directly related to the hard scatter - even if their momentum is determined by the scale of that scatter. They are not radiated by the outgoing partons 7 and 8, which are being studied.

Again taking a pragmatic standpoint, the emission of a gluon *after* the main interaction is termed Final State Radiation (FSR). These are repre-

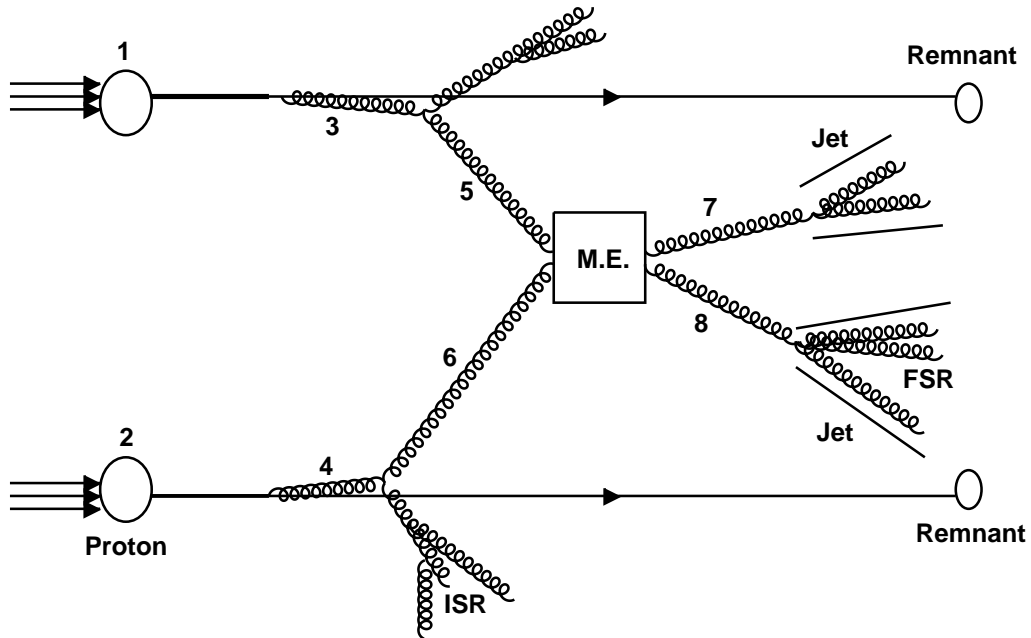


Figure 2.7: A simplified view of a proton-proton collision, based on events modelled in PYTHIA.

sented by the softer partons radiated from particles 7 and 8 in figure 2.7. This is *not* considered a background effect for the purposes of this thesis, as the aim is to know the 4-momentum of the parton exiting the hard interaction i.e. *before* any FSR (particles 7 and 8).

The two types of radiation are indistinguishable given a full quantum mechanical treatment and it is therefore impossible to know whether individual energy deposits came from the hard interaction, ISR, FSR or any other background. However, an FSR parton direction will be correlated with the parton from which it came, and is more likely to be close to it in  $\eta - \phi$  space, whereas ISR would not be strongly correlated with that direction.

### 2.4.2 Proton Remnant

The ‘spectator’ fragments of the two protons continue in much their original direction, and are labelled ‘Remnant’ in figure 2.7. Typically they still carry a large proportion of their original momenta with only a very small deviation from the beam axis with transverse momentum of the order of 1 GeV.

Some models describe a gluon ‘string’ at lower local energies between the ‘struck’, interacting parton and the proton remnant. It is a model of the low energy colour field which tends to be one dimensional because of the mutual attraction of coloured gluons. The string ‘breaks’ into a number of particles, appearing as a ‘fan’ of hadrons between all coloured centres.

### 2.4.3 Multiple Parton Interactions and Pile-up

When two dense beams of particles are incident upon each other it is common that more than one parton interaction will occur within some time interval measurable by the detector. If the secondary partons come from within the same proton as that producing the hard scatter, it is termed a multiple parton interaction (MPI). If the partons are from different protons, although within the same ‘bunch’, the effect is known as pile-up. Both are an unavoidable problem of high-energy and high-luminosity hadron-hadron experiments. Specifically, MPI becomes significant at high- $\sqrt{s}$  but is independent of luminosity, whereas the opposite is true for pile-up.

In any given high-momentum transfer event, there will be a number of low energy interactions between partons not directly associated with the high- $p_T$  interaction of interest. Such interactions are incoherent from, and thus would not necessarily affect the measurement of the interesting event in a significant way. However, occasionally an MPI or pile-up interaction may have

a momentum transfer which approaches the hard scale energy of the ‘main event’. This can produce much confusion in analysis, as it may be interpreted as *part* of the main event, indistinguishable from it, and for example, adding additional energy to the jets which are the parton candidates. These effects are not shown in figure 2.7, and are essentially indistinguishable from proton remnant fragmentation or ISR. Although vertex detection can identify that pile-up has occurred, frequently this is too difficult for calorimetry to differentiate.

## 2.5 Monte Carlo Modelling

A Monte Carlo (MC) simulation is used to make predictions from, and comparisons to, theoretical models. It also helps to determine or predict the behaviour of the detector. An MC is a computer program that uses random number generated integration to calculate cross-sections of specific physical processes within a particular kinematic region [25]. Although the matrix elements of many Monte Carlos are calculated at tree level, all higher orders are approximated, via the models specified in the following subsections. The program keeps track of particle information such as 4-vectors and particle ID for all states including detectable final state, parton level, and decay products in between.

In general there are four major steps involved in MC generation, although background effects such as MPI may also be added:

- The *elementary hard interaction* is calculated using full perturbative-QCD (at first or higher orders) and referred to as the matrix element (M.E.) in figure 2.7.
- *Parton showers* are added in the initial and final states (before and

after the hard interaction) using longitudinal pQCD (i.e. strict theory is applied only in the parton direction).

- *Decay* of massive particles, such as the top quark and W/Z bosons.
- *Hadronisation* of coloured particles followed by decays of unstable hadrons.

At ‘generator’ level, the information is the *truth* - that is to say the direct output of the MC in the final state. The next step is ‘simulated’ data in which the generator level MC is passed through a detector simulator such as ‘Géant’ [26]. This simulates the detector response to the final state particles. The ‘reconstructed’ stage assembles the simulated output into a useful form which allows predictions of what may be *observed*. These results are directly comparable to collision data from the detector, and in fact the data reconstruction software is identical.

There are several popular MCs including PYTHIA [27], HERWIG [28], SHERPA [29] and ALPGEN [30]. It is common to compare to several of these for a particular data analysis, as they represent a combination of different sub-models.

MC is also used to test reconstruction software and thus calculate experimental efficiencies. Generally in this case, the precise model used is less important, so long as the final state is a good approximation.

### 2.5.1 Parton Showering

Following the hard interaction, parton showers are statistically simulated on initial and final partons, assuming multiple splittings, although varying methods are used.

One such approach is via the DGLAP equations, which describe the PDF evolution with  $Q^2$  and  $x$ , and hence the cross-section for the various processes

involved in parton showering.

The longitudinal momentum of the showered parton is calculated via perturbative QCD (pQCD). But this cannot determine the transverse component. Instead, an effective mass,  $M$  is assigned to the new particle from the Sudakov Form Factors [27]. These distributions run between zero and the hard scale energy,  $\hat{p}_T$ , or effective mass of their parent parton. Knowing effective masses of the parent and child partons allows the calculation of the opening angle and hence transverse momentum of the showered particles. In this way it is possible to mimic the process in the absence of rigorous theory.

Subsequent splittings are calculated similarly, until a specified cut-off value,  $Q_0$ , is reached.  $Q_0$  is set by the user and has no theoretically determined value. This parton showering method is common to many MCs including both HERWIG and PYTHIA version 6.

Another variant is the colour-dipole model [31], in which the gluon field behaves in a similar way to an electromagnetic dipole field. The latest version, PYTHIA8, deals with some of the final state radiation in this way, stretching a dipole between a final-state parton and the ‘hole’ left by an initial-state parton [32]. It also continues to use the  $p_T$ -ordering common to PYTHIA since version 6.3, but not the mass-ordering favoured previously.

## 2.5.2 Hadronisation

Various models for hadronisation are used by different generators. PYTHIA uses the ‘Lund string’ model [33], in which a coloured gluon ‘string’ stretches between each  $q\bar{q}$  pair, mimicking the non-abelian nature of the strong force. When the string snaps, hadrons are formed along its length. HERWIG uses the cluster hadronisation model [34] to project colour–anticolour pairs onto singlet states called clusters, which decay to hadrons [28].

All hadronisation models are guided by theory but are not directly calculable from first principles. Instead they have a more empirical nature, and rely upon experimental observations for validation.

### 2.5.3 Pythia

Various MC programs exist which allow the modelling of hard interactions and associated UE, and return the full event data. These can be run to generate only particular sub-processes, for example the rare creation and decay of a top and anti-top or standard QCD interactions, so as to form a sample of events ‘tailor made’ for the user’s analysis. Calculations are based on a combination of analytical results and QCD-based models [27] as described above.

One such program is PYTHIA. PYTHIA is a widely-used general purpose MC generator, having particular emphasis on interactions involving the strong force. In this thesis PYTHIA is used to model the processes shown in figure 1.3. These are the dominant QCD processes in hadron-hadron collisions. As well as reflecting the latest measurements associated with the Standard Model and non-perturbative ‘minimum bias’ (low  $p_T$ ) data, PYTHIA is also capable of modelling the creation and decay of hypothetical particles belonging to such (as yet, unobserved) theories as supersymmetry. Experimentalists use the full event data to aid their search for these new particles, the results of which would confirm or deny the validity of such theories.

Constraints may be applied on the generation of events including restriction on the range of transferred momentum involved in the hard process. It is also possible to turn on and off effects such as the Multiple Parton Interaction. In fact many aspects may be altered by the user; a specific set of parameters is called a ‘tune’. There exist ‘ATLAS tunes’ and ‘Tevatron

tunes’, which are specific to those experiments, and are generally used as a standard basis for all MC generation associated with that experiment. The ATLAS tune designates, amongst other things, the format of output with respect to the event record, how software errors are reported, particle properties such as mass and lifetime, and the PDF set. These are done in a way specific to the requirements of ATLAS users, and up to date with the latest knowledge.

The MC output is ‘tuned’ to closely resemble the most up to date experimental results, whilst still remaining constrained by accepted theory. Table 2.3 gives examples of parameters used commonly within PYTHIA. A complete list with descriptions and usage may be found in [27].

Param	Argument(s)	Description
MSUB()	11,12,13,28,53,68	Switches on/off QCD subprocesses
CKIN()	3,4	Sets minimum, maximum momentum transfer (GeV)
PMAS()	6,23,24	Sets properties (mass, width) of top quark, Z, W
MSTP()	61,71,81,111	Switches on/off underlying event e.g. MPI

Table 2.3: Examples of user-set parameters within PYTHIA.

The current standard is PYTHIA Version 6.4, although Version 8 is in the process of validation and is quickly gaining prominence for certain processes.

PYTHIA has access to a wide range of PDF options. The default PDF set in PYTHIA 6.4 is CTEQ 5L [27], and this was used for the development stage of the TRAPS algorithm.



### 2.5.4 Rivet

Rivet is an acronym for Robust Independent Validation of Experiment and Theory [35]. It is a Monte Carlo interface consisting of a library written in C++ and produces simulated distributions for particle collisions at generator level. These results are output in a ROOT [36] file, which is a convenient format for analysis.

A summary of cross-sections for individual sub-processes is printed following every run, to indicate to the user the approximate quantity of real data needed to produce the dataset they have just simulated.

The TRAPS algorithm (documented in chapter 6) was developed in C++ code linked to ROOT libraries, and tested using generator level output from an old version of Rivet running PYTHIA version 6.418. Later, AGILe [37], an interface for a variety of MC event generators, was used when PYTHIA version 6.420 was required.

### 2.5.5 Athena

Athena [38] is the ATLAS software framework for analysis. It provides easy access to ATLAS data, whether real or Monte Carlo, and the ATLAS conditions database, which records the physical and electronic setup of ATLAS whenever data is taken. Athena is compatible with the Grid, a worldwide computing network [39]. Within this thesis, Athena has been utilised for most of the timing studies in chapter 4. The TRAPS algorithm was also adapted for Athena after its initial development. This allows usage of the ATLAS MC datasets, and eventually, collision data. The adaptation for, and usage of TRAPS in Athena is detailed in chapter 7.

## 2.6 Final State Particles

A final state particle is typically defined as one which lives long enough to be detected. The range of particles ‘seen’ by detectors is limited. Many particles produced at high energy are too short-lived to even reach the detector, despite their great velocity. Their existence has to be inferred by observable decay products. Neutrinos from weak decays overwhelmingly pass through the entire detector unobserved. *Their* presence is only inferred by missing transverse energy.

Other particles exist on the boundary of detection methods. For example charged kaons, mesons which are composed of a strange or anti-strange quark and other light quark, decay close to the beampipe, and therefore some are observed whereas others just leave evidence of their decay products in the detector. Sometimes only particles with electric charge are used within an analysis, as their momentum is easily and accurately ascertained from the curvature of their tracks in a strong magnetic field.

Within PYTHIA the default minimum flight distance to qualify as a final state particle is 10 mm. Thus decays of  $K_S^0$ ,  $\Lambda$ ,  $\Sigma^-$ ,  $\Sigma^+$ ,  $\Xi^-$ ,  $\Xi^0$  and  $\Omega^-$  are all ‘switched off’, but charm and bottom quarks still decay [27].

Reconstruction of an event is not a simple process, as typically many hundreds of final state particles are created. Roughly half of these are positively or negatively charged pions, and another quarter are photons from neutral pions which decay to two photons in  $\sim 10^{-16}$  s, well before detection. The rest are largely hadrons like kaons, protons and neutrons or leptons such as electrons and muons. The composition of jets at ATLAS has been calculated for a range of jet transverse energies, and appears in figure 2.8 [40].

The two or more final state partons which are produced in the majority of high- $p_T$  events, create collimated clusters of mostly hadrons and photons.

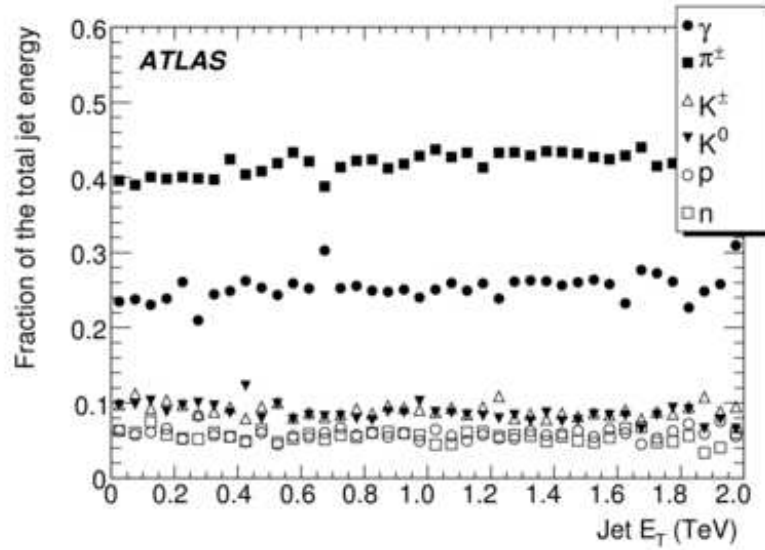


Figure 2.8: Fractional energy of jets by particle type as a function of jet transverse energy, calculated by the ATLAS experiment [40].

These are termed ‘jets’, although these objects require a strict definition of their parametrisation to be meaningful. Discussion of this is given in chapter 5. Analysis of the shape and momentum of jets gives information about the original partons. Naturally, the many background effects mentioned previously in this chapter, also appear in the final state.

# Chapter 3

## LHC and ATLAS

### 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a particle accelerator at the Organisation Européenne pour la Recherche Nucléaire (CERN). It accelerates protons to energies an order of magnitude greater than any previous device - a centre of mass (CMS) energy of 14 TeV - but is currently fixed at half design-CMS energy of 7 TeV due to the decision to run the machine far from its limits in the initial stage. The LHC also makes ‘heavy ion’ Pb-Pb collisions with an energy of 2.76 TeV/nucleon, but these are not the concern of this thesis. The accelerator is housed in the same circular tunnel previously occupied by the LEP collider<sup>1</sup>, which is 27 km in circumference and situated approximately 100 m below ground across the Swiss-French border near Geneva [1].

Having been pre-accelerated in smaller rings to an energy of 450 GeV, protons are injected into the two beam-pipes. The protons travel clockwise in one beampipe and anti-clockwise on the other, whilst being accelerated by synchronised electromagnetic fields. They are guided in the correct tra-

---

<sup>1</sup>The Large Electron-Positron Collider (LEP) operated between the years 1989–2000.

jectory by increasing magnetic fields of up to 8.33 Tesla, which are created by superconducting electromagnets that must be cooled to 1.9 K [1].

The LHC is the world's largest and highest energy particle accelerator. High luminosity rates are important in order to create the statistics required for rare processes. The LHC has been running at a luminosity of  $10^{32} - 3 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  through 2011 [41], and should increase to the intended  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$  in 2012. A year-long period of maintenance will follow<sup>2</sup>. Current estimates predict the LHC will output an integrated luminosity of 7000  $fb^{-1}$  over its lifetime, assuming major upgrades are made as planned.

At maximum energy and luminosity, the beam current is 0.584 A. This corresponds to a stored energy of 362 MJ, although both will tend to reduce over the lifetime of a physics run. The beam intensity decreases over time, largely due to proton loss from physics and beam-gas collisions. Approximately 600 MJ is stored by the magnet system during running [1].

At points around the ring, protons are collided and particle detectors measure the debris. There are four major detectors and two smaller detectors: ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are general purpose particle detectors; ALICE (A Large Ion Collider Experiment) specialises in heavy ion collisions; LHCb (Large Hadron Collider beauty) focuses specifically on B-physics; the smaller ones are TOTEM (Total Cross Section, Elastic Scattering and Diffraction Dissociation) and LHCf (Large Hadron Collider forward). The relative positions of the detectors on the LHC ring are illustrated in figure 3.1.

Many technical challenges characterise the operation of the LHC. To be brief, these include the containment of large amounts of liquid Helium (130

---

<sup>2</sup>The LHC will not function during 2013 while essential maintenance takes place. This will allow the LHC to safely achieve full design energy collisions.

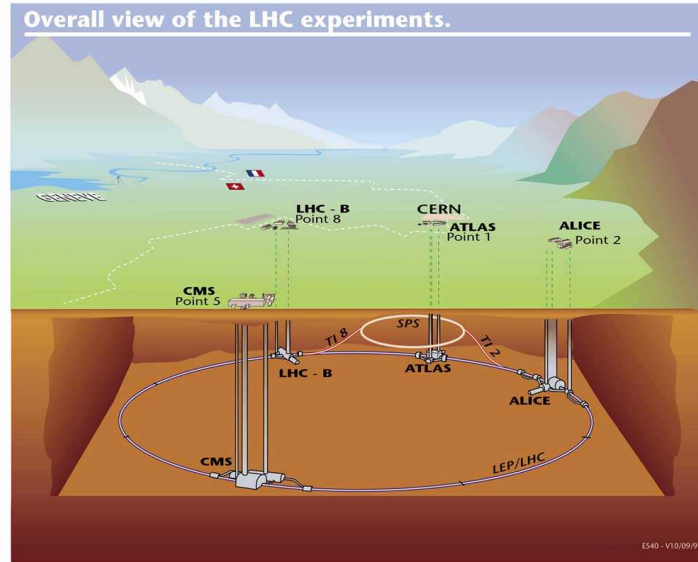


Figure 3.1: Cartoon of LHC, indicating the positions of the main detectors.

tonnes) and liquid Nitrogen (10,000 tonnes), the maintenance of beampipe vacuum at approximately  $10^{-13}$  atmospheres [42], the generation of opposing magnetic fields in beampipes brought close together and the ability to control the high energy beams and safely dissipate that energy when necessary.

As for the detectors, they must be protected from high doses of radiation [19] and kept to the correct temperature. In addition the whole structure must be supported physically in a way that maintains alignment whilst not adding large amounts of ‘dead’ (non-detecting) material. Some parts must also be mobile for access.

### 3.1.1 LHC in Numbers

Tables 3.1 and 3.2 contain a summary of quantities relating to the LHC tunnel and beam [1][43][44].

Tunnel characteristic	
Length	26.659 km
Width	3.5 m
Depth	50 - 175 m
Number of dipole magnets	1232
Length of dipole magnets	14.3 m

Table 3.1: LHC Tunnel Characteristics.

Beam characteristic	Injection	2010-12 data	2014 and beyond
Proton energy (GeV)	450	3500	7000 <sup>†</sup>
Relativistic gamma	479.6	3731	7461 <sup>†</sup>
Number of particles per bunch	As data	$1.3 \times 10^{11*}$	Unknown
Number of bunches		1380*	2808 <sup>†</sup>
Stored energy per beam (MJ)	23.3	181	362
Peak luminosity ( $\text{cm}^{-2}\text{s}^{-1}$ )	-	$10^{34}$	$10^{35}$
Inelastic cross section (mb)	-	60	70
Total cross section (mb)	-	100	110
Events per bunch crossing	-	19	23
RMS bunch length (cm)	11.24	7.55	
Revolution frequency (kHz)	11.245		

Table 3.2: LHC proton beam characteristics (maximum capacity). Note that some quantities are liable to change. In some cases, such as the number of particles per bunch, quantities have already exceeded the design value at the time of writing.

\* By end of running 2011 [45].

<sup>†</sup> Design quantity.

## 3.2 Brief Description of ATLAS

ATLAS is an acronym for A Toroidal LHC ApparatuS [19]. It is the largest volume particle detector at the LHC. Figure 3.2 shows a scale drawing of ATLAS. It is a general purpose experiment, with many aims involving verification of the Standard Model (SM) in the newly accessible energy region, including search for evidence of the Higgs boson, and more exotic discovery beyond. The experiment is an international collaboration, involving more than 3000 experimental physicists from around 40 countries [46].

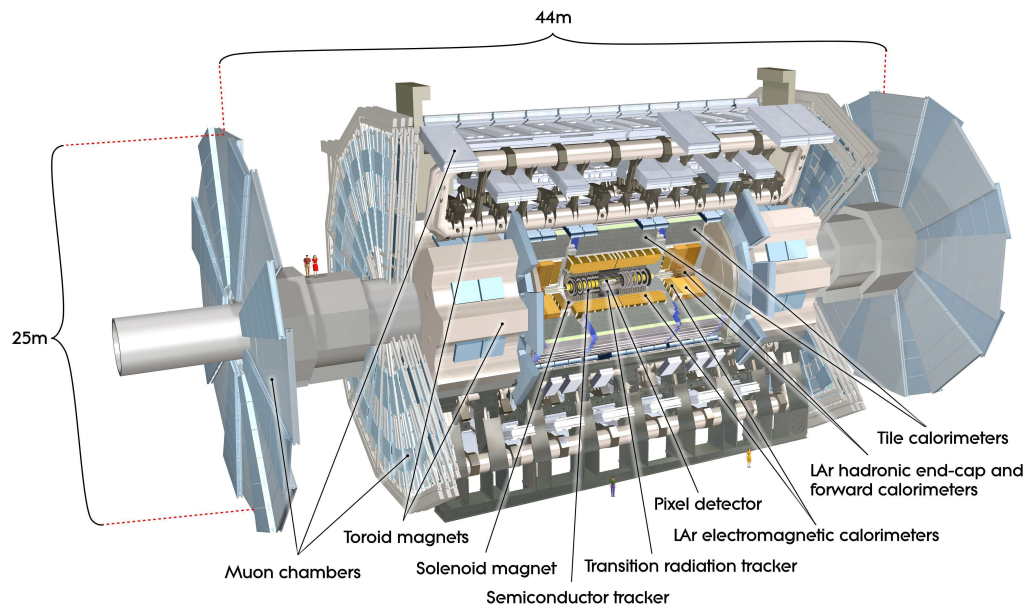


Figure 3.2: Drawing of the ATLAS detector, with sections removed so as to view the inner detector



### 3.2.1 ATLAS in Numbers

Table 3.3 contains a summary of quantities relating to the ATLAS detector [19].

Dimensions: diameter, $z$ (m)	25, 44
Mass (kg)	7,000,000
Magnetic field strength, inner detector (T)	2
Number of trigger levels	3
Bunch crossing rate ( $MHz$ )	40

Table 3.3: ATLAS characteristics at design luminosity.

### 3.2.2 Co-ordinate System

The co-ordinate system in ATLAS is defined as follows [19]: the origin is at the centre of the detector (the nominal interaction point). The  $z$ -axis is along the beampipe, with the  $x$  and  $y$  axes perpendicular to  $z$ . The positive  $x$  direction is towards the centre of the LHC ring, and positive  $y$  is in the upward direction. It is a right-handed system and positive  $z$  defines the ‘ $A$ ’-side of ATLAS.

The azimuthal angle,  $\phi$ , is measured anti-clockwise from the  $x$ -axis in the  $x - y$  plane and  $\theta$  is the polar angle between the track and the  $z$ -axis, although pseudorapidity,  $\eta$  is more typically used, as defined previously in equation 2.6 [19]. The detector is divided into two halves about  $\eta = 0$ . These are known as the  $A$ -side and  $C$ -side for positive and negative  $\eta$  respectively.

### 3.2.3 Detection Concepts

The basic principles of particle detection in ATLAS are common to many experiments, having been used throughout the history of particle physics. Some general concepts will be described here.

#### Motion of a Charged Particle in a Magnetic Field

The motion of a charged particle moving in an electromagnetic field is described by the Lorentz Force:

$$\vec{F} = q(\vec{E} + (\vec{v} \times \vec{B})) \quad (3.1)$$

where  $q$  is the charge carried by the particle,  $\vec{v}$  is the particle velocity, and  $\vec{E}$  and  $\vec{B}$  are the electric and magnetic fields respectively.

A particle moving perpendicular to a uniform  $\vec{B}$ -field has a radius of curvature given by:

$$\rho = \frac{mv}{qB} \quad (3.2)$$

#### Radiation and Interaction Lengths

A ‘radiation length’,  $\chi_0$  is associated with electromagnetic particles. It is the path length in which an electron will, on average, radiate all but  $1/e$  of its energy via a photon. Likewise a photon will create an  $e^+e^-$  pair after  $7/9 \chi_0$  [17]. Hence the original single electron creates a ‘shower’ of electrons and photons. The process continues until a critical energy, at approximately 600 MeV per nuclear charge of the medium, at which point losses due to radiation equal those of ionisation [14]. The exact distance depends on the medium through which the object is travelling.

Hadronic particle showers contain many different contributing processes but have a similar ‘interaction length’,  $\lambda$ , sometimes called ‘absorption length’, which is related in a similar way to the probability of a collision or absorption. However, some contributions to the total energy may not be counted, due to the possibilities of nuclear excitation and ‘punch-through’ (leakage) of secondary muons and neutrinos, which are not detected in the calorimeter, and must be compensated [14]. As  $\chi_0$  is typically smaller than  $\lambda$ , hadronic calorimeters are built thicker than electromagnetic so as to absorb the greatest energy.

Showers resulting from hadronic objects are significantly broader than those from electrons, and thus the shape can also be used to determine the particle type, even if the shower occurs solely in the EM calorimeter.

### 3.2.4 Detector Components

The ATLAS detector is made up of three major detecting components which form forward-backward symmetric, and coaxial cylinders around the beampipe. These are the inner detector, the calorimeters and the muon system. Figure 3.3 shows the relative radii of each. There are also several associated forward detectors. Each measures properties of charged and neutral particles which pass into them, excluding neutrinos. Figure 3.4 is a diagram indicating in which detectors specific particles may be observed. The varying methods of detection are appropriate to a wide range of physics signatures and radiation conditions in different parts of the detector [19].

The magnet system creates a 2 Tesla homogeneous magnetic field in the inner detector and a toroidal field in the muon spectrometer, to cause the trajectories of charged particles to curve respectively in  $\phi$  and  $\eta$ , thus allowing precise momentum measurements.

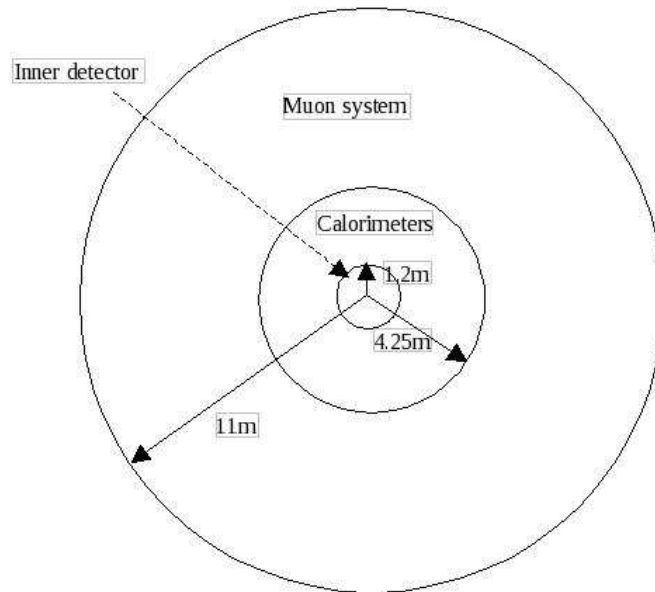


Figure 3.3: Diagram indicating relative radii of the detector layers, the remaining 3 m of height being composed of structural supports.

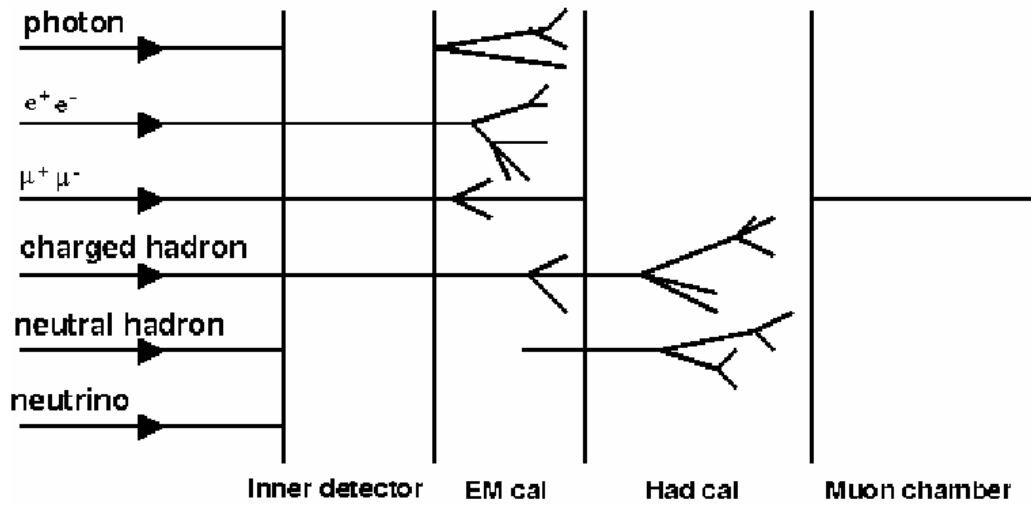


Figure 3.4: Diagram illustrating where, in the detector, particle types are observed.

## Inner Detector

The Inner Detector (ID) tracks charged particles through three different levels, from a few centimetres from the beam axis to a radius of 1.2m and within  $|\eta| < 2.5$ . It is able to establish, from the reconstructed tracks left in the detector material, the interaction point (IP), and the charge and momentum of most charged particles with an energy above 200 MeV. This knowledge assists in the identification of particles.

The first and second levels of the inner detector use silicon as the detecting material. The innermost level is called the Pixel Detector, and has the highest resolution to provide precision tracking when the particles are closest to their origin and are most energetic. The pixel size is  $50 \times 400 \mu\text{m}$  and the detector contains 82 million pixels in three layers [19].

The Semi-Conductor Tracker (SCT) differs from the Pixel Detector in that it is a 1-Dimensional silicon strip detector with the strips sized  $80 \mu\text{m} \times 12.6 \text{ cm}$ . These are arranged in four double layers. The area to be covered increases with radial distance, hence the detecting elements further out have to be larger to be practical.

The outermost layer of the inner detector is the Transition Radiation Tracker (TRT) which contains a total of 351,000 ‘straws’ [19] filled with gas. This gas has a contrasting refractive index compared with between the straws. As relativistic charged particles pass through the medium, they lose energy via transition radiation photons. The charged particles ionise the gas inside the straws; the accompanying photons increase the number of freed electrons. These free electrons accelerate toward, and then produce a current in a high-voltage wire that runs through the centre of the straw. The response of the TRT to a light charged particle, for example an electron, is different to that of a heavy hadronic charged particle, for example a charged

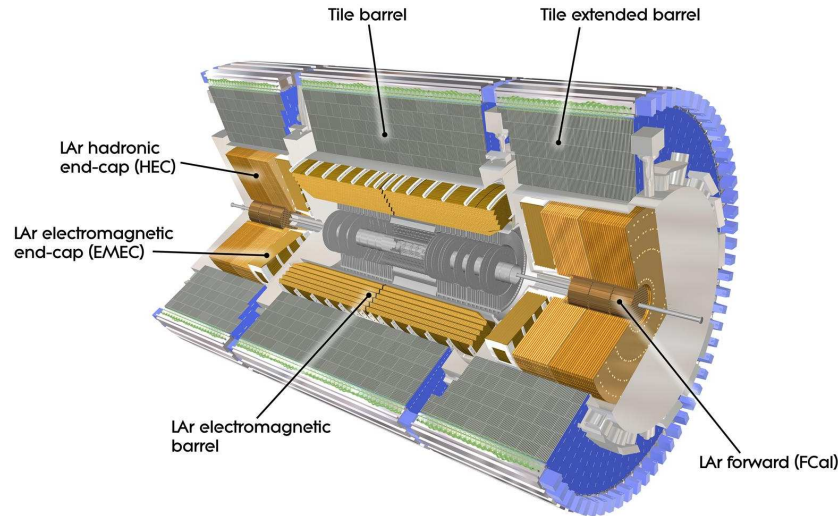


Figure 3.5: Computer generated image of the full calorimeter with partition labels.

pion, as an electron produces more transition radiation. Hence it is possible to make a distinction between them. Again, for reasons previously specified, the resolution is much lower than that in previous layers.

### Calorimeters

There are two types of calorimeter which sit radially beyond the inner detector and the solenoid magnet. The electromagnetic (EM) is on the inside and the hadronic on the outside, as indicated in figure 3.5. Each provide hermetic coverage over a large  $\eta$  range. They are both sampling calorimeters, which require alternating layers of ‘absorber’ and ‘sampler’. The particles hit the high-density absorber (commonly a metal) and create showers of particles. These showers are sampled in a detecting medium and the shape and amplitude of the signal estimates the number of produced secondaries and thus reveal the energy. They indicate whether the particle was electromagnetic, hadronic or weakly interacting.

The EM calorimeter is designed to measure particles that interact via the electromagnetic force: all particles with charge, and photons. It has total thickness greater than 22 radiation lengths,  $\chi_0$ , in the barrel and 24 in the endcaps [19]. The total number of interaction lengths,  $\lambda$ , in the calorimetry equates to approximately 9.7 and 10  $\lambda$  in the barrel and endcaps respectively [19]. Approximately two thirds of hadronic particles start their showers in the EM calorimetry. The remaining high energy hadrons continue their shower into the hadronic calorimeter, which measures only particles that interact via the strong force.

The central barrel region of the hadronic calorimeter is known as ‘TileCal’. It contains layers of steel and scintillating tiles. The rest of the calorimetry comprising the EM barrel, and both EM and hadronic endcaps and forward calorimeters (FCAL) is referred to as ‘Liquid Argon’ (LAr). This comprises alternating layers of lead and LAr. In the EM barrel an ‘accordion’ structure provides complete azimuthal coverage in three layers for high-precision measurement of electrons and photons.

The EM calorimeter is the more precise for both position and energy measurement. Its cell size is  $\Delta\eta \times \Delta\phi \simeq (0.003 \times 0.1)$  and  $(0.003-0.006 \times 0.1)$  for the first layer in barrel and endcap respectively,  $(0.025 \times 0.025)$  for the second layer, and  $(0.050 \times 0.025)$  for the third layer [19].

The smallest EM calorimeter barrel position resolution is measured as  $1.5 \times 10^{-4}$  by  $50-60(\text{mrad}) / \sqrt{E(\text{GeV})}$  in  $\eta$  and  $\phi$  respectively [47], although this value varies over the coverage of  $\eta$  and between layers. In terms of energy resolutions, these are given as  $10\% / \sqrt{E(\text{GeV})} \oplus 0.17\%$  [48][19] in the EM and  $50\% / \sqrt{E(\text{GeV})} \oplus 3\%$  [19] in the hadronic calorimeters.

Calorimeter response to electrons and photons is different to that of hadrons, and this has to be taken into account. The full energy of hadrons

is not measured and this must be compensated. ATLAS is calibrated to the ‘EM Energy Scale’, and offline corrections are applied to hadronic objects. These correction factors are known as the Jet Energy Scale (JES) [49]. Shower shape is used to differentiate and thus compensate for hadronic deposits in the EM calorimeter.

The calorimeters are split into partitions; more detail on these can be found in chapter 4. In summary, the region central in  $\eta$  is the barrel, the ends are closed by endcaps and the forward calorimeters (FCAL) sit close to the beampipe as shown in figure 3.5 [19].

Efficiency is often lower at the boundaries between partitions. For example, at the EM barrel/endcap boundary there is less detecting material and so less energy is deposited. These are known as the overlap or ‘crack’ regions.

### **Muon System**

Muons are among the most penetrating of particles. Hence the muon spectrometer sits as the outermost detecting component. There are three layers of detecting chambers. In the barrel, these are arranged in concentric cylinders and in the endcaps they are in planes perpendicular to the beam [19].

The muons are deflected from their straight-line trajectories by the outer toroidal magnetic field. Various methods of detection are used: the Monitored Drift Tubes (MDTs) and Cathode Strip Chambers (CSCs) are used for precision tracking and the Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) trigger on the tracks.

These units allow muon tracking and triggering within  $|\eta| < 2.7$ , although the RPC coverage extends only to  $|\eta| < 1.05$  [19].

A good muon identification requires independent reconstruction in the inner detector and the muon spectrometer, with a significant number of hits



in each. The two collections of tracks must match each other and point back to the reconstructed primary event vertex, so as to reduce cosmic ray and other backgrounds.

### **Magnet System**

Moving charged particles in a magnetic field are affected by the Lorentz force which curves their trajectories; the direction and level of curvature allow direct inference of the particle momentum and charge.

The inner solenoid of the magnet system surrounds the inner detector and provides a 2 Tesla homogeneous field within that sub-detector. The outer toroidal magnetic field is situated within the muon spectrometer but cannot provide a uniform field for the muon detector. To do so requires a solenoid and was prohibited due to the expense; however, knowledge of the field strength in different regions allows this problem to be overcome in the particle reconstruction. A toroidal field is also better for low angle muons, as they fly perpendicular to the field lines.

### **Forward Detectors**

In the far-forward region outside of ATLAS there are three more detector systems. LUCID (LUminosity measurement using Cerenkov Integrating Detector) and ALFA (Absolute Luminosity for ATLAS) are situated at  $\pm 17$  m and  $\pm 240$  m respectively from the IP. Their principal purpose is to evaluate the luminosity delivered to ATLAS. The ZDC (Zero-Degree Calorimeter) measures the energy of neutral particles with  $|\eta| \geq 8.2$  at  $\pm 140$  m.

## Trigger System

The data rate emerging from LHC is overwhelmingly large, and it is important to measure and store only the most useful events, which make up a small fraction of the total. At design luminosity and a bunch-crossing frequency of 40 MHz, an event rate of approximately 1 GHz is expected in ATLAS. However a large proportion of this will be low  $Q^2$ , ‘diffractive’ or elastic collisions, which are unlikely to contain new physics and are of little interest other than to assess the background and make detector calibrations.

The ‘trigger’ system, for which there are three levels, had been designed with a maximum event recording rate of 200 Hz [19]. However by the end of 2011 this recording rate has doubled. Nevertheless, this capacity allows only one in  $10^5$  bunch crossings to be kept for analysis. The ATLAS trigger system is described more fully in chapter 4.

### 3.2.5 Accessing ATLAS Data

The ATLAS analysis framework is called Athena [38]. Running an analysis through Athena is often the most convenient way of accessing data from ATLAS, whether you want to simulate MC, do a physics analysis on collision data or make a sub-detector calibration. Various forms of the data are available, with varying degrees of information. There are also many ‘streams’ which contain only data from certain triggers.

# Chapter 4

## L1Calo Timing Studies

This chapter summarises work carried out by the author for the Level-1 Calorimeter Trigger (henceforth known as ‘L1Calo’) subsystem on the ATLAS experiment. Largely it was the electronic timing that was studied. This includes verification of the trigger tower timing in early (pre-beam) running using cosmic rays, and monitoring of the trigger tower timing during collision running.

### 4.1 The ATLAS Trigger

The ATLAS trigger consists of three levels designed to reduce the huge amount of data detected by the apparatus to a recordable quantity and in doing so, select the most interesting events. At Level-1 the rate of event data is reduced from the bunch-crossing frequency (40 MHz with  $\sim 23$  events per bunch crossing at luminosity of  $10^{34}\text{cm}^{-2}\text{s}^{-1}$ ) to a maximum of 75 kHz. This is achieved within a processing/transmission latency of less than  $2.5\ \mu\text{s}$  using custom built hardware based on ASICs and FPGAs. Level-2 (L2) runs on parallel CPUs and reduces the pass-through rate to around 5 kHz with an

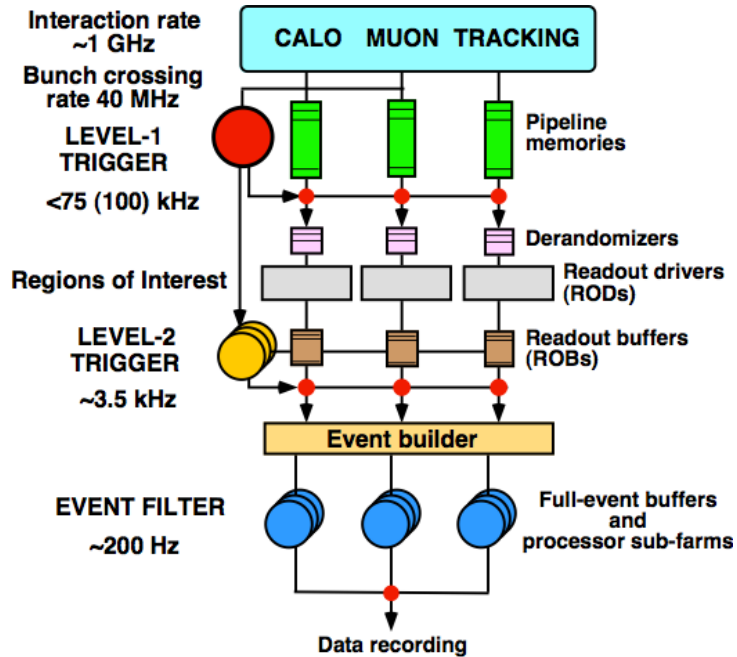


Figure 4.1: Data flow within the ATLAS trigger.

average decision-making time of 40 ms. The third level is called the Event Filter (EF) and uses CPUs to analyse the full event using offline algorithms. It must reduce the event rate to 400 Hz - the maximum recording rate in 2011 [50]. Events are recorded to disk in various data streams for monitoring and physics analysis with an average processing latency of 1 s. L2 and the EF collectively are known as the High Level Trigger (HLT). The diagram in figure 4.1 [51] shows schematically the progression of the data from the calorimeters, muon chambers and tracker through the trigger system.

The ‘trigger menu’ defines which triggers are active within the current run. In order to progress up the ‘trigger chain’ the objects in an event must pass the thresholds set at each of the three trigger levels. When the rate of a particular trigger is deemed too high (i.e. it is taking too much bandwidth), it can be removed from the menu or *prescaled*. The prescale

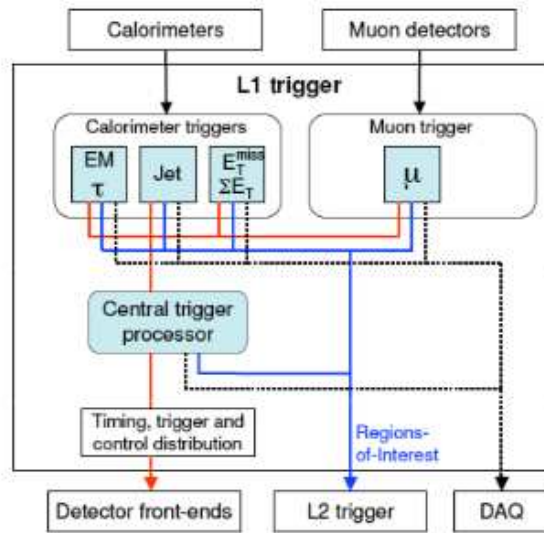


Figure 4.2: Block diagram of the Level-1 Trigger.

alters the proportion of triggered events that are accepted to the next level, and all but the highest energy triggers are usually heavily prescaled.

## 4.2 Brief Description of L1Calo

The Level-1 (L1) trigger system for the calorimeters is a major component of the L1 trigger, the other major part being the Level-1 Muon trigger subsystem which will not be discussed further here. More information about L1Muon may be found in [19]. The decision of whether to accept an event at L1 ('L1A') rests with the Central Trigger Processor (CTP), which collates information from all L1 inputs.

An area of activity above a given threshold in an event which receives an L1A is called a Region of Interest (RoI). These RoIs are passed to L2 where the event is analysed with full-granularity within the RoI. This sequence is illustrated in a basic form by the block diagram in figure 4.2 [51].

At L1 events are selected using exclusive signatures which attempt to identify high transverse-energy particles, jets and missing transverse energy. L1Calo receives input from the electromagnetic and hadronic calorimeters which measure the energy deposits of the incident particles. It uses reduced-granularity summed data to make a quick decision on each event within a period of  $2 \mu\text{s}$ , half of which is taken by transmission delays due to the distance between the detector and associated electronics.

L1Calo consists more than seven thousand ‘trigger towers’ split evenly between the EM and hadronic layers and projected towards the interaction point of ATLAS. Each tower sums the signals from a number of calorimeter cells in a local region, using the full depth in both cases.

L1Calo is composed of three main subsystems. The PreProcessor (PPr) converts analogue pulses from trigger towers to digital signals with a strobe period of 25 ns and finds an equivalent transverse energy value using a digital finite-impulse-response (FIR) filter followed by a Look-Up Table (LUT). The other two subsystems work in parallel and take digital signals from the PPr as inputs. The Cluster Processor (CP) implements the algorithm for identifying electrons, photons and taus in the region  $|\eta| < 2.5$ . The Jet/Energy-sum Processor (JEP) runs algorithms for jets and sums both total and missing transverse energy in the region  $|\eta| < 4.9$ . Note that jets with  $3.2 < |\eta| < 4.9$  are measured in the FCAL, which is seen as a different detector by the trigger system.

The maximum size of a pulse is 1023 Analogue-to-Digital-Conversion (ADC) counts. Signals larger than this value are termed ‘saturated’. Each ADC count is calibrated to an energy of 0.25 GeV, making saturation equivalent to about 255 GeV.

Each pulse is followed by an ‘undershoot’ where the signal height goes

below the pedestal. This is an effect typical of electronic devices. The size and duration of the undershoot depends on both the size of the pulse and the partition involved. The recovery (or dead) time for a trigger tower is typically five bunch crossings, during which it cannot trigger again. A full description of L1Calo can be found in [51].

### 4.2.1 Physical Layout

L1Calo is built mainly of custom electronics which sit both on the calorimeters and in the electronics cavern USA-15 which is situated adjacent to the ATLAS cavern. The trigger towers cover the whole calorimetric range on both the EM and hadronic calorimeters. They are each split into partitions which (for the purposes of the trigger) are divided as shown in table 4.1, with the positive and negative halves in  $\eta$  known as the A-side and C-side respectively.

Trigger towers are more coarse than the calorimeter cells, and each take signals from around 60 cells in the EM and 5 cells in the hadronic, summing over the full depth of either the EM or hadronic layers within a local  $\eta - \phi$  region. The variable granularity of the trigger towers is shown in figure 4.3 [51] for one half of the detector in  $\eta$  coverage and one quarter in  $\phi$ . The other sectors are identical in  $\phi$  and mirrored in  $\eta$ .

The outputs from the trigger towers are summed using a sliding window algorithm to determine regions of interest for different types of candidate physical objects. An EM element is a single trigger tower which is approximately  $0.1 \times 0.1$  in  $\eta - \phi$  (applies within  $|\eta| < 2.5$  only) whereas a jet element is  $0.2 \times 0.2$  in  $\eta - \phi$ . The actual window size is  $2 \times 2$  and  $8 \times 8$  trigger towers for CP and JEP objects respectively. The window slides by the size of the element in  $\eta$  and  $\phi$  in each case. To identify particular signatures the

Partition name	Abbreviation	$\eta$ range
<b>Electromagnetic calorimeters</b>		
Barrel	EMB $\frac{A}{C}$	$0 <  \eta  < 1.4$
Barrel/Endcap overlap	Overlap $\frac{A}{C}$	$1.4 <  \eta  < 1.5$
Endcap	EMEC $\frac{A}{C}$	$1.5 <  \eta  < 3.2$
Forward calorimeter	EM FCAL $\frac{A}{C}$	$3.2 <  \eta  < 4.9$
<b>Hadronic calorimeters</b>		
Long Barrel	Tile $\frac{A}{C}$	$0 <  \eta  < 0.9$
Extended Barrel	Tile $\frac{A}{C}$	$0.9 <  \eta  < 1.5$
Endcap	HEC $\frac{A}{C}$	$1.5 <  \eta  < 3.2$
Forward calorimeter	Had FCAL $\frac{A}{C}$	$3.2 <  \eta  < 4.9$

Table 4.1: Electromagnetic and hadronic calorimeter partition information including the barrel/endcap overlap, also known as the ‘crack’ region.

CP algorithms require a threshold to be exceeded by sums of trigger towers of two adjacent or four in a square for electron/ $\gamma$  and  $\tau$ /hadron algorithms respectively as well as various isolation conditions to be satisfied. Figure 4.4 [51] illustrates the elements used for the different algorithms.

### 4.2.2 L1Calo Timing

A signal in an L1Calo trigger tower is an analogue pulse with amplitude proportional to the energy deposited. The pulse is sampled every bunch-crossing (25 ns) and an extended (15 slice) example with fit can be seen in figure 4.5 [52]. The shape of the pulse varies slightly depending on the partition from which it came. The peak of the pulse must be in the centre 25 ns bunch crossing sample (slice) of the read-out, which is typically five slices



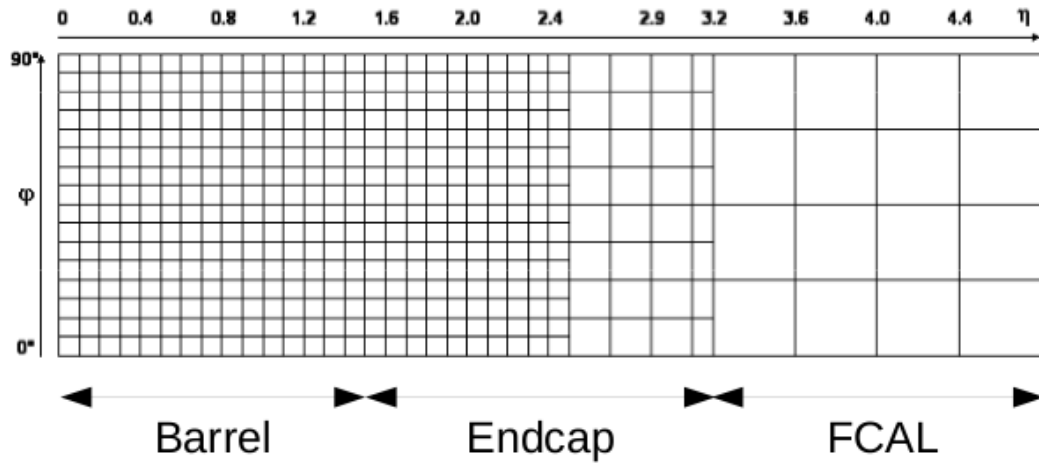


Figure 4.3: Trigger tower granularity for  $\eta > 0$  and one quadrant in  $\phi$ .

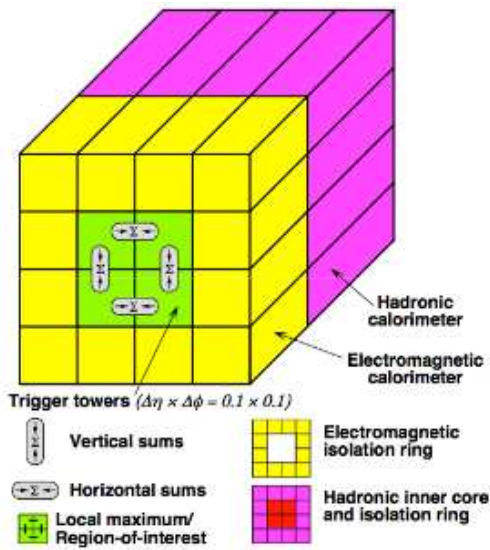


Figure 4.4: Elements used for the  $e/\gamma$  and  $\tau$ /hadron algorithms

wide during collision data-taking<sup>1</sup>. Five slices is the approximate width of the pulse, and is the minimum required for optimum energy resolution in the calorimeter readout. If the peak is not in the centre slice, then the pulse will not have an associated LUT energy value in the correct bunch crossing and hence not be summed within the correct event.

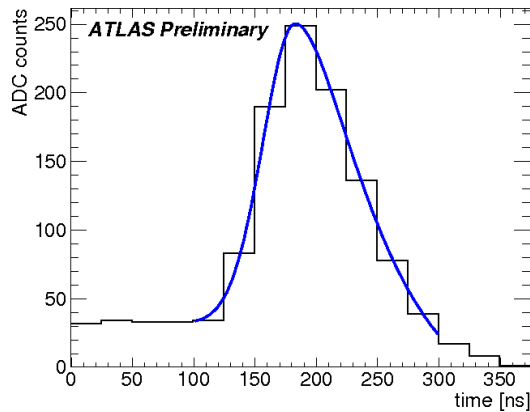


Figure 4.5: A digitised LAr pulse with functional fit, showing 15 slices around the peak.

It is essential that the sub-system is well timed in order to achieve the best energy resolution for triggers - hence L1Calo must sample at the peak of the pulse. Improved energy resolution implies sharper turn-on curves. The efficiency of Bunch Crossing Identification (BCID) is important even for small pulses, as these are used in isolation cuts which may distinguish between a jet and an electron, for example.

Timing calibration of a trigger tower so that the peak is in the centre slice is known as good ‘coarse’ timing. However, to trigger accurately, the peak of the analogue pulse must be in the middle of the centre slice - this is known as good ‘fine’ timing. The aim is to calibrate and maintain every

---

<sup>1</sup>Seven and 15 slice sample sizes were also used during early calibration so as to investigate very badly timed trigger towers.

tower within 2 ns of this centre. The smallest increment of change on the timing is 1 ns.

### 4.3 Timing Verification Using Cosmic Rays

Pre-beam timing calibration was largely performed using an electronic pulser which injects charge into the on-detector front-end electronics of each trigger tower and measures the resulting signal. This provides a pulse with a well-defined shape similar to a physics pulse, at a known energy, and at a fixed time. In most cases a good fit is made to the pulse shape, from which is derived a peak position. Hence a calibration is made so as to centre the peak. However there are potential issues with this method.

Firstly, calibration pulses may not be perfect representations of ‘physics’ (collision or cosmic) pulses. Secondly, due to the design of the front-end electronics, different calorimeter partitions register differently-shaped pulses, for example with larger undershoot or broader width. As a consequence the fitting function used for this process is variable in its performance across the calorimeters. Thirdly, the relative timing between calorimeter partitions when using the calibration pulser is different to that in physics data taking.

This third problem was what makes verification of timing using cosmics particularly important in the months (and years) of L1Calo calibration before the LHC was able to provide collisions. Cosmics and collisions have the same relative timing for neighbouring towers between partitions and hence cosmic particles can be used to check that each partition is recording peaks in synchronisation with the others.

As soon as LHC beam arrived it was possible to implement an improved

calibration using firstly ‘splash’<sup>2</sup> events [52] and then collision data. But prior to this, L1Calo was able to make initial timing calibrations, correct faulty hardware and find software bugs.

### 4.3.1 Method I

When ‘cosmic rays’, which are generally muons, hit the ATLAS detector they can generate a shower of particles. These may enter the calorimeters and some are energetic enough to register hits, with a collective rate of the order of 5 Hz for events with energy greater than 3 GeV and 5 GeV for electrons and hadrons respectively [53]. An example of a cosmic particle passing through the detector is given in figure 4.6 [54][55].

These showered particles are frequently close together in space, and are considered simultaneous, so can therefore be used to compare the response of adjacent trigger towers. Also, since towers are projective, multiple towers are commonly traversed by a single particle. The pulses are not ‘in time’ with the LHC global timing, but are comparable relative to each other. Hence cosmic pulses are only useful (for timing calibration) when multiple towers in the same region are hit, and in particular on either side of a partition boundary. In this method, timings are only measured with respect to others around them with pulses in the same ‘event’.

To find the peak of the pulse from digital readout, a fit function incorporating a Gaussian rise with a Landau fall shape is used. An example of such a fit on a five slice digital signal is shown in figure 4.7 [53] and the fit

---

<sup>2</sup>‘Splash’ data is that produced when the collision beam is allowed to impact upon one of the tertiary collimators near ATLAS. This produces messy sprays of debris particles which hit all parts of the detector. Knowing the geometry of the detector allows calibration of every channel from a small number of splashes.

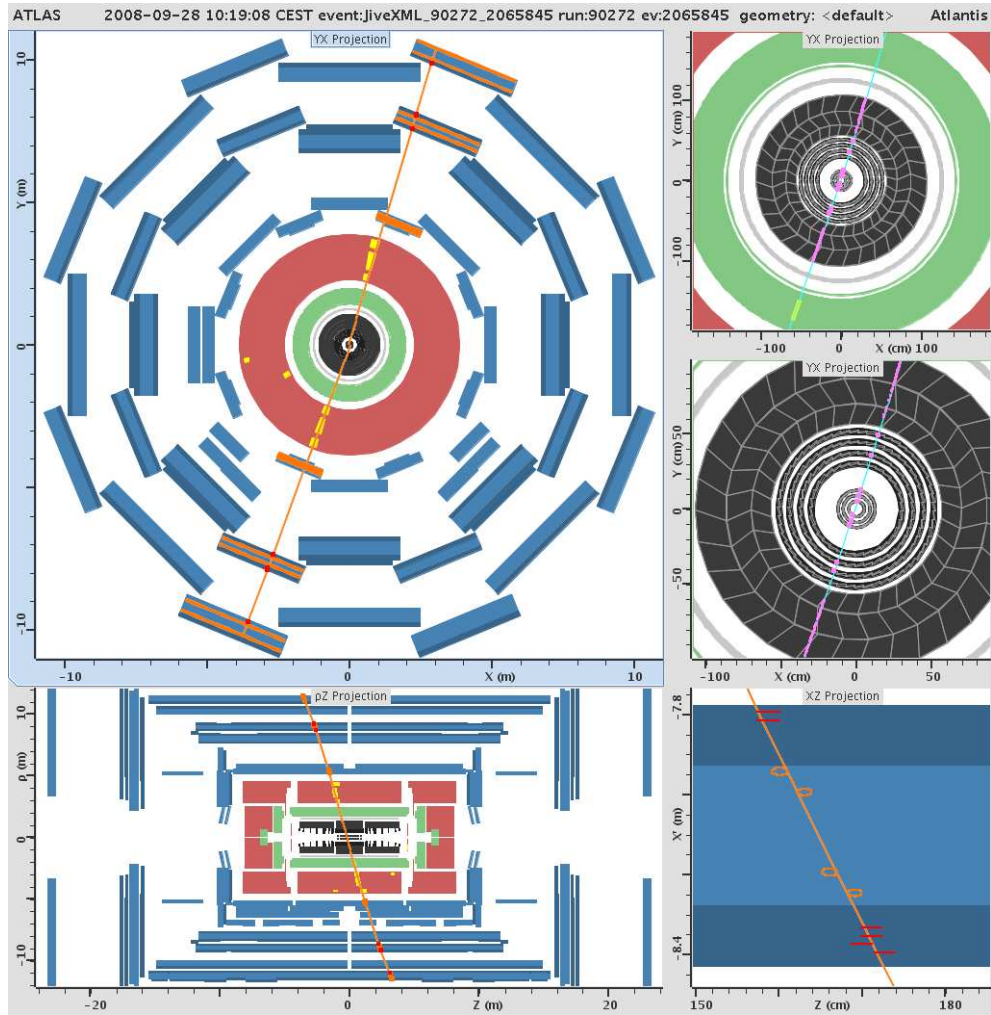


Figure 4.6: An Atlantis event display of a real cosmic particle recorded by the ATLAS detector in September 2008.

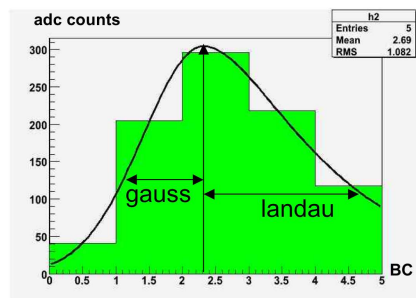


Figure 4.7: Gauss-Landau fit on a digitised pulse.

expression is given by

$$f(t \leq t_{peak}) = Ae^{-\frac{(t-t_{peak})^2}{2\sigma_g^2}-0.5} + ped \quad (4.1)$$

$$f(t > t_{peak}) = Ae^{-\frac{1}{2}\left(\frac{t-t_{peak}}{\sigma_l} + e^{-\frac{(t-t_{peak})}{\sigma_l}}\right)} + ped \quad (4.2)$$

where  $A$  is proportional to the signal amplitude,  $t_{peak}$  is the  $x$ -axis (time) value at the peak,  $t$  is the time at different points along the  $x$ -axis,  $\sigma_g$  and  $\sigma_l$  are the Gauss and Landau fit parameters respectively<sup>3</sup>, and  $ped$  is the height of the pedestal.  $A$  and  $t_{peak}$  are the free variables the fit is aiming to determine. The values of the fixed parameters used for this period are:

$$ped = 40 \text{ (ADCcounts)} \quad (4.3)$$

$$\sigma_g = 0.9 \text{ (slice)} \quad (4.4)$$

$$\sigma_l = 0.65 \text{ (slice)} \quad (4.5)$$

$$(4.6)$$

---

<sup>3</sup>These fit variables were fixed following study of pulser signal shapes, so as to achieve a successful fit for a greater proportion of pulses.

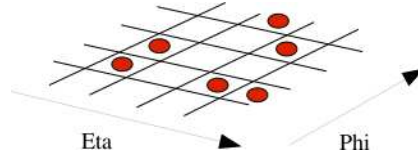


Figure 4.8: Three possible combinations for adjacent tower pairs.

### Pulse Selection Criteria

To be used in the analysis there must be a minimum of *two* towers in the same event within a  $\Delta R$  separation of 0.15 of their centres. As the region studied has a granularity of approximately  $0.1 \times 0.1$  this forms a definition of adjacent towers, and the various possibilities are illustrated in figure 4.8.

The pulses from both towers must satisfy the following conditions:

- not be disabled,
- be within  $|\eta| < 2.5$ ,
- have a pulse sample  $>$  threshold parameter (for example 60 counts).

### 4.3.2 Results I - Interpartition Timing

The analysis was run on cosmic datasets of significant duration from autumn 2008 until the start of LHC collisions in late 2009. During the early part of this period very few useful results were obtained. Strange pulse shapes were observed which could not be fitted and below-threshold pulses were discounted as unreliable for assessment of timing.

The criteria to have *two or more* adjacent towers above the threshold *and also* to see these appear across partition boundaries was a high expectation for the short and relatively low-rate cosmic runs made at the time. For example, dataset 90793 contains 27k EM triggered cosmic events, but only

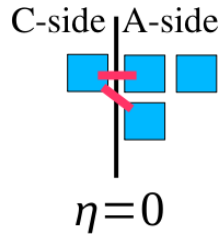


Figure 4.9: Diagram illustrating two possible combinations (pink lines) of adjacent towers across the  $\eta = 0$  boundary.

one pair of adjacent towers straddled the  $\eta = 0$  boundary with an amplitude of 20 ADC counts above pedestal. Although the two towers showed not too different derived peak times, this was hardly conclusive considering the lack of measurement reliability in many of the other towers.

In the hadronic partition, the situation was slightly more revealing as the TileCal is the most likely part of L1Calo to be hit by cosmics<sup>4</sup> as it is larger and sits outside of the EM layer. Also, cosmic muons are more likely to shower in the TileCal, due to its higher density. Even so, only 15-20 (out of a possible 64) distinct tower pairs of 20 ADC counts above pedestal were observed across the central boundary.

Following the 2008/9 winter shutdown, more success was had with pulse fitting and, through the various methods, timing in general gradually improved. It was then possible to observe cross-boundary timing issues.

In a cosmic data run from spring 2009, there were 146 and 15 pairs of adjacent towers across the  $\eta = 0$  boundary in the hadronic and EM calorimeters respectively. This boundary is illustrated schematically in figure 4.9.

The difference in timing from the mean of the cluster is shown in figure 4.10 for the EM and 4.11 for the hadronic. For each, the towers on the C-side

---

<sup>4</sup>The endcap and FCAL regions do not receive substantial numbers of cosmic hits due to their shape and orientation.



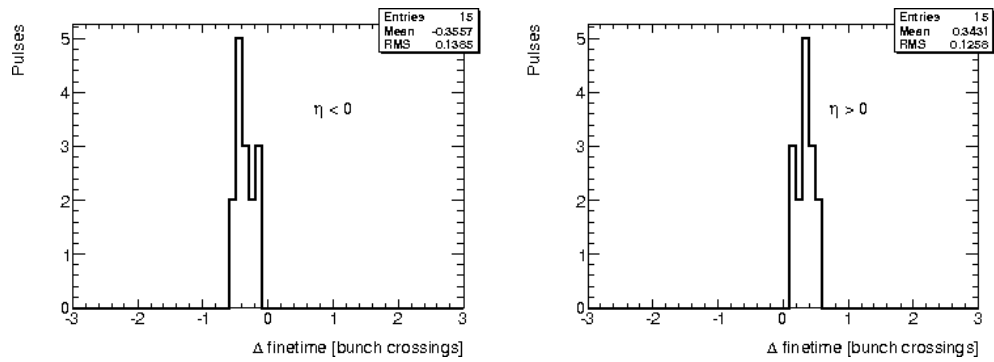


Figure 4.10: Histogram of the deviation in timing from the mean of the pair/cluster for EM towers across the  $\eta = 0$  boundary with C-side (left), A-side (right).

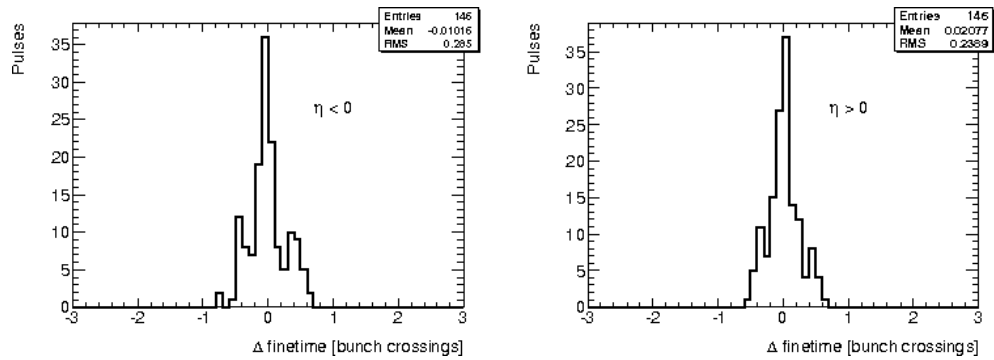


Figure 4.11: Histogram of the deviation in timing from the mean of the pair/cluster for hadronic towers across the  $\eta = 0$  boundary with C-side (left), A-side (right).

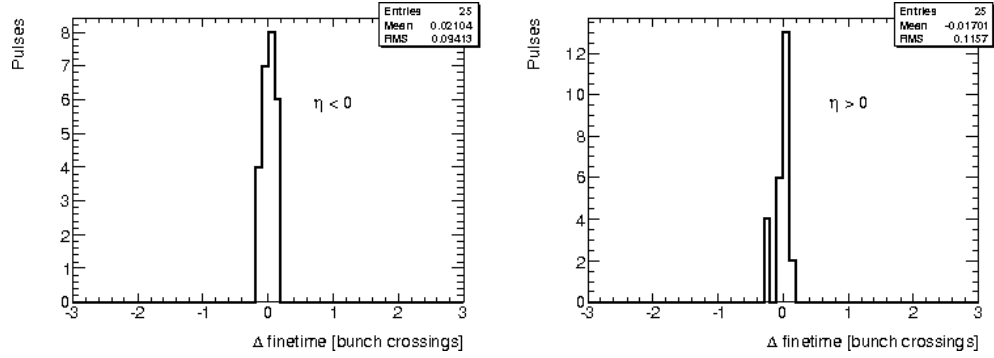


Figure 4.12: Histogram of the deviation in timing from the mean of the pair/cluster for EM towers across the  $\eta = 0$  boundary with C-side (left), A-side (right) following timing adjustments.

are plotted on the left-hand plot, and on the right-hand plot for the A-side.

In figure 4.10 it is clear from the similar shapes between the two that many of the clusters were pairs in this case. Entries on the A-side are positively timed compared to the mean, whereas on the C-side they are negatively timed. These results indicated a systematic mean difference between the timing of the A and C sides in the EM barrel of  $0.7 \pm 0.07$  bunch crossings, equivalent to  $17 \pm 2$  ns.

This problem was later corrected by applying new timing delays in the Pre-Processor Modules (PPMs). Following this, a new cosmic dataset was analysed and the new results in figure 4.12 show that the correction was successful. Now the systematic difference was measured as  $0.03 \pm 0.05$  bunch crossings (less than  $1 \pm 1$  ns, and hence within the desired timing precision).

In the TileCal the A and C sides appeared not to have any systematic shift, with a mean difference of only 0.03 bunch crossings ( $< 1$  ns) between them. However, to a greater degree than the EM, the histogram structure indicated that some individual towers were not synchronous with the rest, with several being measured half a bunch crossing or more out of time with

their neighbours on the opposite side.

## 4.4 Timing Monitoring Using Collision Data

It was always considered possible that L1Calo timing would ‘drift’ from its ideal, well-calibrated state. But it was not known which effects would be dominant, which calorimeter partitions may be most affected or whether individual channels would be correlated with their neighbours. Therefore it is important to carefully monitor the timing, particularly in the early-running stages, so as to keep the system running efficiently.

There are several factors which *may* contribute to significant timing drift. These include random drift due to electronic component variability, problems with signal transmission and other problems in the hardware.

The general aim of the monitoring is to fit a large number of proton-proton collision pulses for each trigger tower and derive the peak amplitude position of each. The difference between this value and ‘perfect’ timing (i.e. in the centre of the middle bin of the sample, as described in the previous section) is calculated. A mean and RMS are found for each trigger tower. Plotting this mean and associated error as a function of time reveals general trends on drift. This was done for both some individual towers as well as entire partitions. If any tower or region is seen to be drifting then it can be listed for re-calibration at an appropriate time.

The RMS plotted on an  $\eta - \phi$  map can quickly reveal which trigger towers may be inconsistent in their timings - with their pulse peaks occurring at a wider range of values than is acceptable for precision measurements. These channels can then be prioritised for repair or replacement.

L1Calo is a large system of around 7200 channels and problems of an

electronic nature are to be expected. However, the major influence on timing that was found during the first few months of running the LHC, was the ambient temperature. Cables and optical fibres are sensitive to temperature changes, and will increase in length when at a higher temperature. As winter turned to summer, differences between the LHC clock and the arrival of beam particles measured by the BPTX (Beam Pick-ups) [56] were observed by the Central Trigger Processor (CTP) team, the probable cause being optical fibre-length change over the several kilometres between the LHC Control Centre and the ATLAS Control Room. It was decided that L1Calo should independently monitor this effect, as well as any other timing drifts which may occur.

#### 4.4.1 Method II - ‘Simple’ Method

Although the fitting function described in Method I accurately finds the peak of a well-shaped pulse, there are occasions where no peak is found because the fit has failed. Therefore it was decided also to use a more simple, robust method which is able to reliably produce a peak value even if the pulse shape is not ideal. This method would not necessarily reach the same absolute precision as that described in Method I, as it uses only three data points and assumes a symmetric pulse shape. Nevertheless it provides a useful cross-check of that method, even for low statistics and regardless of partition. In addition it is self-consistent and hence used to make comparisons between towers within the same partition and reveal variations over time.

##### Calculation of ‘finetime’

In the ‘simple’ method, finetime is calculated by equation 4.7, where  $a$ ,  $b$  and  $c$  refer to the height of the beam crossing time slices indicated in figure 4.13.

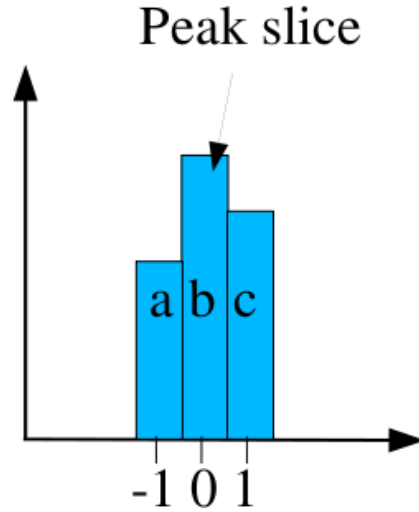


Figure 4.13: Diagram indicating the bunch crossings/slices used by the simple method to calculate finetime.

This equation gives the deviation of the peak from the centre of the peak slice as a fraction of a bunch crossing. It is then used to calculate a  $\Delta finetime$  which is the time difference in nanoseconds from the middle of the *centre* slice. Knowledge of which bunch crossing contained the peak (*peakSlice*, or *b*) is also essential for the full calculation, as shown in equation 4.8.

$$finetime = \frac{c - a}{2(2b - c - a)} \quad (4.7)$$

$$\Delta finetime = (finetime + (peakSlice - middleSlice)) * 25 \quad (4.8)$$

Here, *middleSlice* is the bunch crossing that *should* contain the peak (i.e. the centre slice).

### Pulse Shape Selection Criteria

Not all data pulses are analysed by the method. Letters A-F in the descriptions below refer to the labelled diagrams in figure 4.14. A and B are examples of ‘good’ pulses which would be analysed, whereas examples C-F are rejected based on experience of common problems. A finetime is *not* calculated if the peak

- is less than the peak threshold (example C),
- is in the first or last sampled slice (example D),
- has adjacent slices smaller than some secondary threshold (above pedestal noise) (example E),
- has adjacent slices equal in size to the peak (example F).

#### 4.4.2 Method Comparison

Figures 4.15 and 4.16 show a tower-by-tower comparison of the two methods described for pulse peak-finding, for the EM and hadronic layers respectively. The partitions of the barrel and endcaps are distinguishable in  $\eta$  slices, as presented in table 4.1, and so are the larger towers of the FCALs. The strips of badly timed towers in the EM layer at  $|\eta| = 1.4$  are associated with the ‘overlap’ region, which will be examined later in this chapter.

Runs 178044 and 178109 were analysed and combined by the author using Method II. The same datasets were analysed by Valerie Lang during the course of her timing calibration runs, using a variation of Method I<sup>5</sup>. The difference is calculated as Method II - Method I.

---

<sup>5</sup>It was found that a Landau-Landau shape better suited pulses from some partitions. Also, an additional parameter was added to better describe the undershoot.

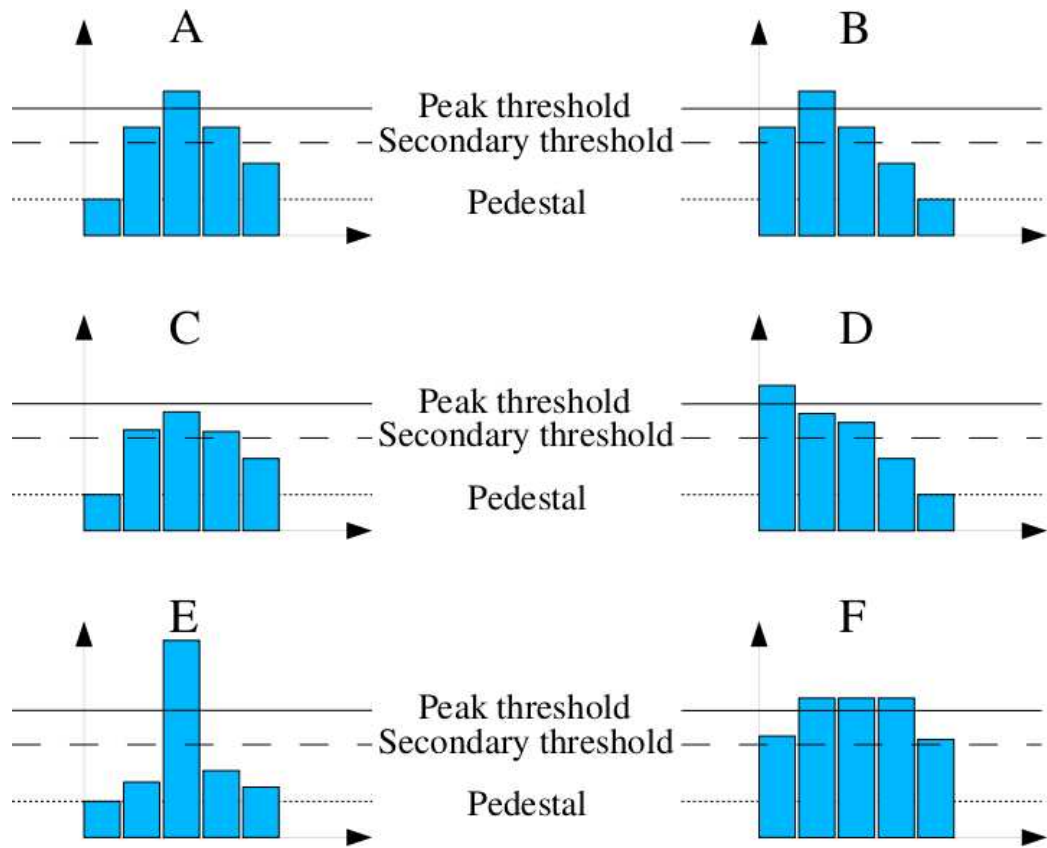


Figure 4.14: Schematic examples of L1Calo pulses. A and B are accepted whereas C-F are rejected for analysis.

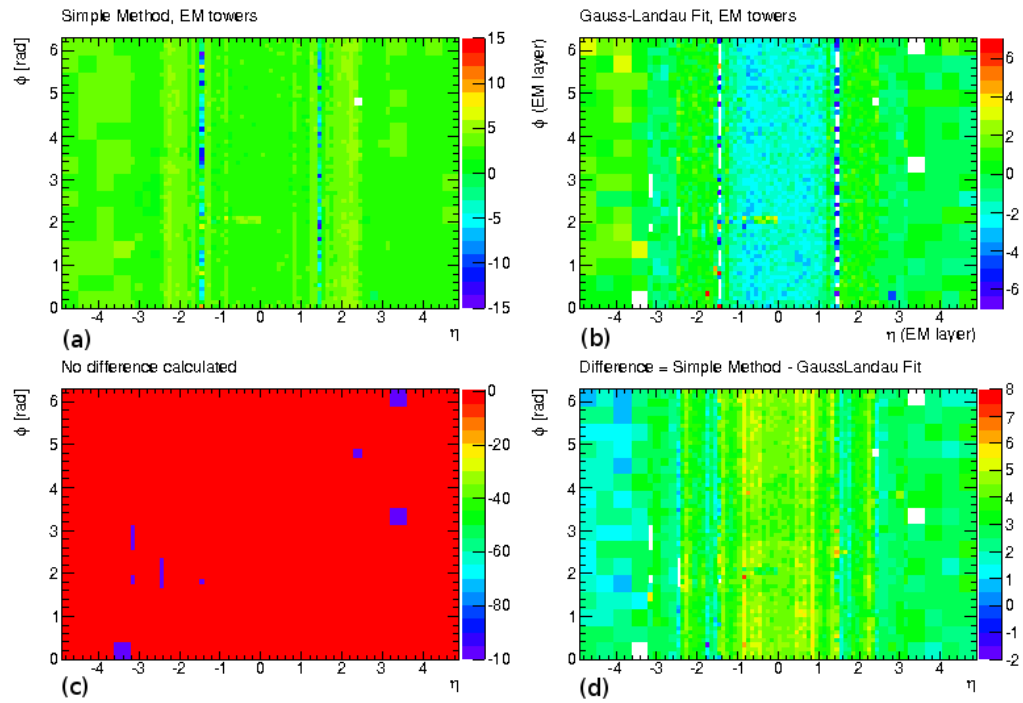


Figure 4.15: EM  $\eta$ – $\phi$  maps illustrating differences between the mean timing result of methods I and II. Plots (a) and (b) give the timing deviation measured by the two approaches. Plot (c) indicates where no difference could be calculated and plot (d) shows the difference between the two methods (= Simple Method - Gauss-Landau or Landau-Landau Fit) in units of ns.



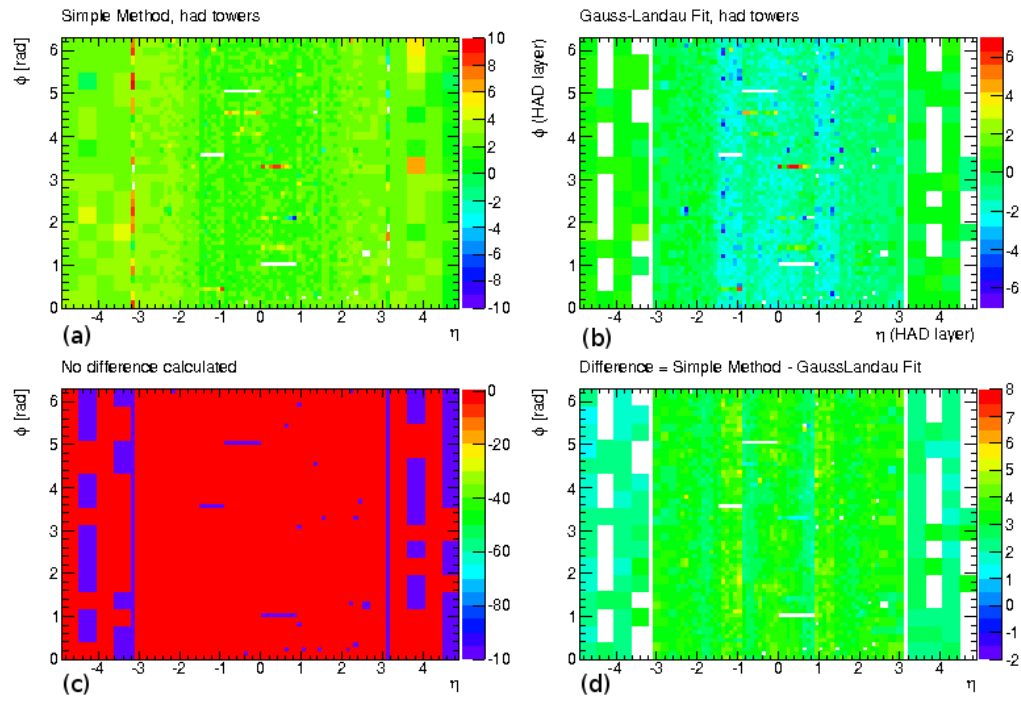


Figure 4.16: Hadronic  $\eta - \phi$  map illustrating differences between the mean timing result of methods I and II. Plots (a) and (b) give the timing deviation measured by the two approaches. Plot (c) indicates where no difference could be calculated and plot (d) shows the difference between the two methods (= Simple Method - Gauss-Landau or Landau-Landau Fit) in units of ns.

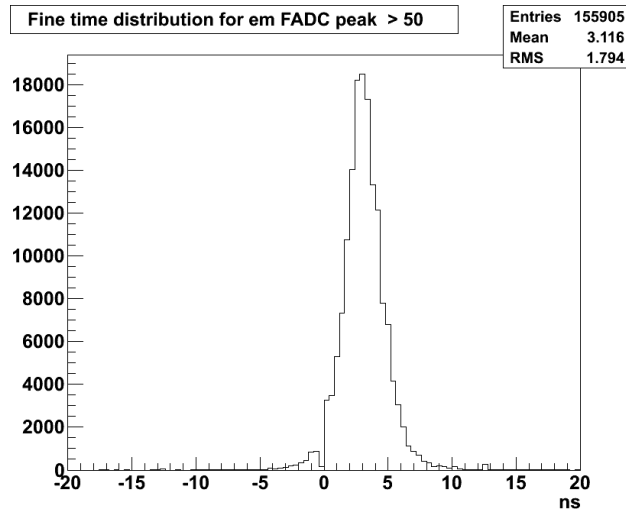


Figure 4.17: Online monitoring plot of  $\Delta finetime$  for the EM partition during a run taken on 29/07/11.

The ‘difference’ plots (d) in each of figures 4.15 and 4.16 reveal that the simple method tends to return a timing slightly ahead of the Gauss-Landau/Landau-Landau fit. The simple method measures timing deviations to be 3-6 ns ahead in the EM barrel, and slightly less in the Had barrel and endcap. This value is 1-3 ns in the EM endcap and both FCALs. However, even 6 ns is less than one quarter of a bunch crossing, and is thus a good approximation for the full fit method. Therefore it may be used for monitoring, as outlined in the previous section.

Figure 4.17 shows an example of this method being used in the ATLAS online monitoring. This plot was generated during a collisions run in July 2011 and shows clearly the timing deviations described.

### 4.4.3 Results II - Timing Monitoring

Monitoring plots are made at regular intervals, but in particular around L1Calo timing changes, CTP timing changes, unusual conditions and follow-

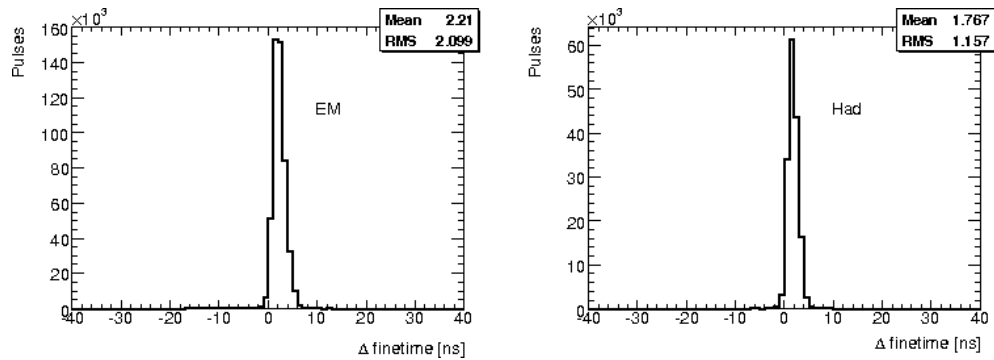


Figure 4.18: Difference from ideal in finetime for all analysed pulses in runs 177924 and 177929 combined, left EM towers, right Hadronic towers.

ing detector shut down periods. In figure 4.18 the finetime difference from ideal is shown for a large number of pulses in both the EM and hadronic layers from all partitions. These pulses came from collision data runs 177924 and 177929 (March 2011), and are selected according to the criteria detailed above with a peak threshold of 38 ADC counts and the secondary threshold at 8 ADC counts (both above pedestal<sup>6</sup>).

The peak-finding method used here is known not to be precise, since it assumes a symmetric pulse shape, which does not describe the true response of the trigger towers. Therefore this information cannot be used for timing calibration, without precise knowledge of the disparity at least; however the *changes* over time between the finetimes calculated here do provide useful information about the system as can be seen in the following.

The variation on the measurement is assessed by calculating the standard deviation on the timing of each tower over one or more data runs. Figure 4.19 shows four  $\eta - \phi$  maps each representing the trigger towers in the hadronic layer for the datasets mentioned above. 4.19(a) is simply the number of pulses used and (b) is the mean average difference in finetime from the ideal. The

<sup>6</sup>The pedestal had a value of 32 ADC counts at the time of this data taking.

standard deviation about that mean is shown in (c) and (d), with different  $z$ -axis scales.

The hadronic layer is shown to be in a good state with the mean timing of the vast majority of towers within a few ns of each other and standard deviations of around 0-2 ns. This is close to the ideal calibration, which puts all towers to within 1-2 ns of each other. One ‘warm’ tower is seen, triggering slightly more regularly than its neighbours, although three TileCal ‘drawers’<sup>7</sup> are dead at  $\eta - \phi \sim (-1.3, 3.6)$ ,  $(-0.5, 5.1)$  and  $(0.5, 1.1)$ . No signals are measured in these regions, which can only be recovered during a period of ATLAS shutdown. As a consequence, the hadronic layer trigger efficiency is reduced.

These plots can be used for spotting problem areas, as can be clearly seen in the C-side barrel of the EM calorimeter in figure 4.20. This appears to be a noise burst of probably mis-shapen pulses, which has caused a larger standard deviation, but seemingly not affected the mean timing measurement in this case. Away from the noise burst, and ignoring the ‘overlap’ region which will be assessed later, a similar picture to the hadronic maps is seen, with almost all towers within a couple of ns of others within the same partition.

### Variability Over Short and Long Periods

Although some variation in the position of the trigger tower pulse peak is inevitable, this should be kept to a minimum wherever possible. Continual monitoring is essential to the running of L1Calo. The example shown in figure 4.21 illustrates the variability of a single tower throughout a long data-taking run. This tower was chosen at random (along with several others) to

---

<sup>7</sup>The front-end readout electronics for one TileCal module receiving input from a group of TileCal cells.

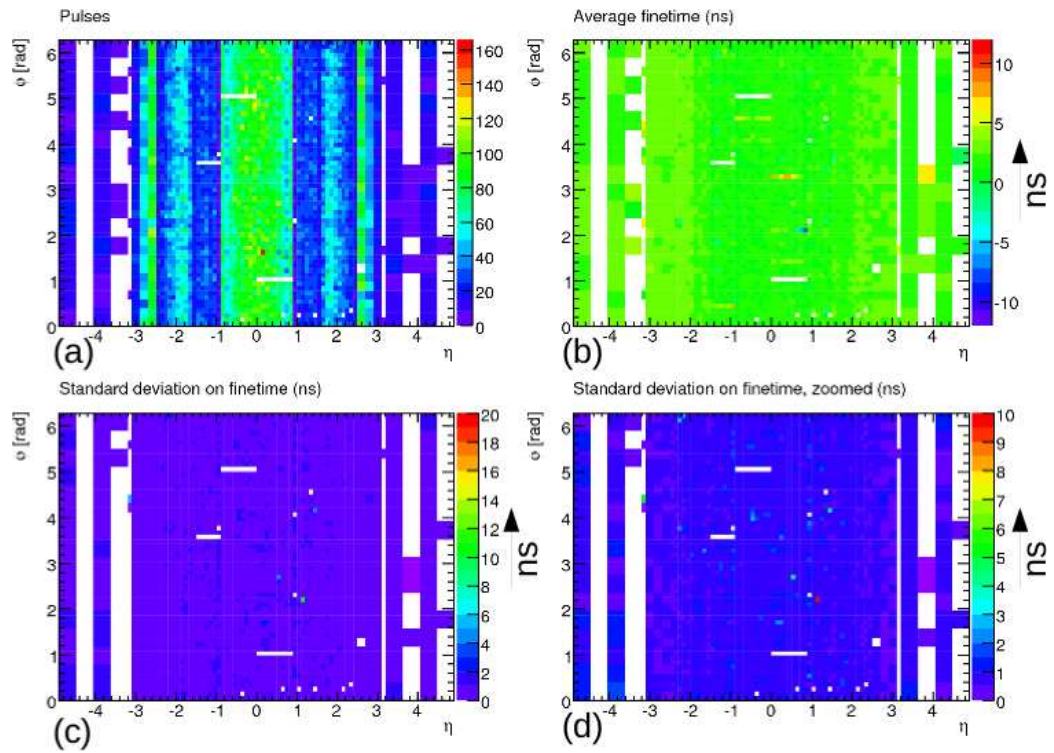


Figure 4.19: Hadronic  $\eta$ - $\phi$  map for number of contributing pulses (a), mean timing in ns (b) and standard deviation in ns with different scales (c) and (d) in runs 177924 and 177929 combined.

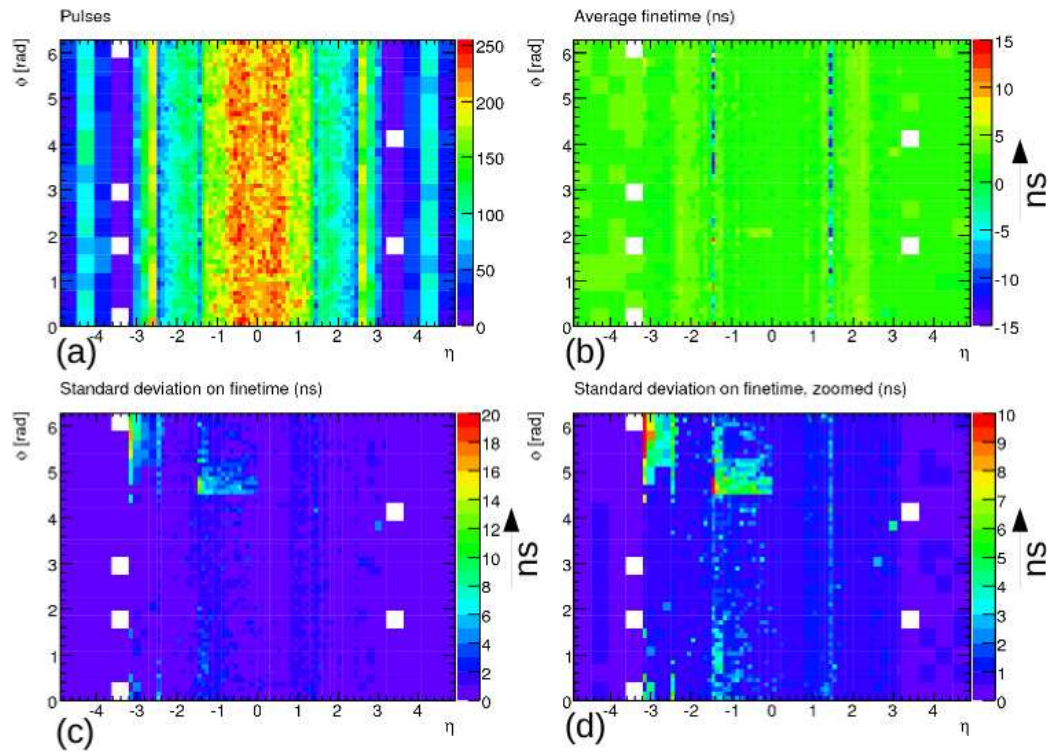


Figure 4.20: EM  $\eta - \phi$  map for number of contributing pulses (a), mean timing in ns (b) and standard deviation in ns with different scales (c) and (d) in run numbers 177924 and 177929 combined.

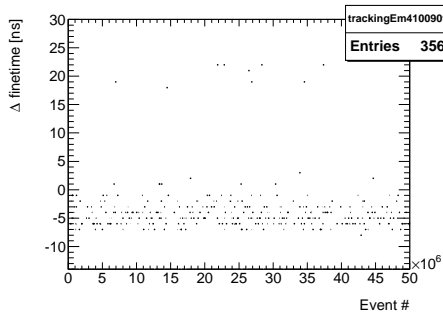


Figure 4.21:  $\Delta finetime$  tracking of EM tower 0x4100901 through physics run 153565 (23-25 April 2010).

be monitored and displays some of the problems that occur: Many of the entries are clustered around  $\Delta finetime = -4$ , with a spread of  $\pm 3$  ns. There is also a small population of entries around  $+20$  ns, which is almost one bunch crossing away from ideal. These entries may represent pulses triggered by other mis-timed triggers.

From examining this type of plot and others, no timing drift has been observed during the relatively short period of a data-taking run.

It would be very time-consuming to monitor every one of the 7200 trigger towers individually across many data-taking runs, and probably unnecessary since we would see these problems in the standard deviation  $\eta - \phi$  maps shown in figure 4.20. However the average of each *partition* can be plotted in a visually friendly graph, clearly showing the trends of each partition over many data-taking runs. This is useful to see global effects on the system and verify that timing adjustments have taken place in the desired manner.

Figures 4.22 and 4.23 show how the average timing of partitions in the EM and hadronic layers respectively varied over a number of collision datasets between June 2010 and March 2011. There are various features to understand: the downward trends which appear in summer of 2010 are thought to

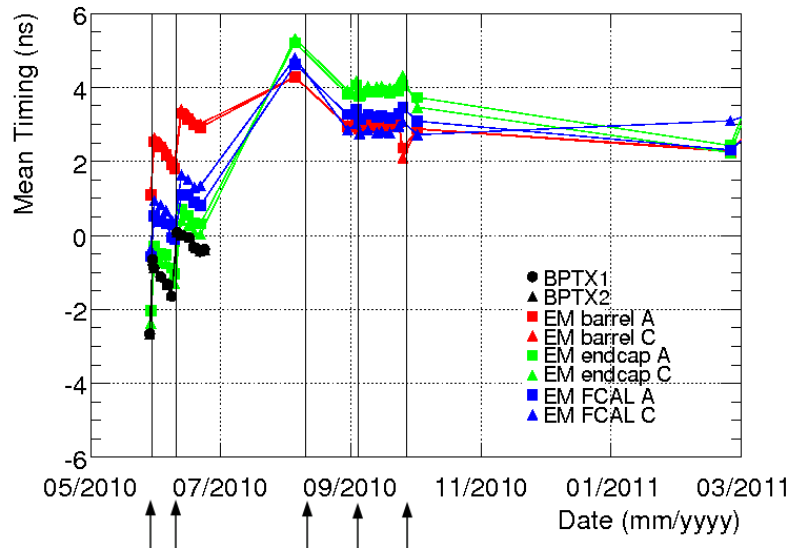


Figure 4.22: EM layer  $\Delta finetime$  tracking for period June 2010 - March 2011. The solid vertical black lines indicated by arrows represent timing changes made by the CTP or L1Calo.

be caused by the increase in ambient temperature at that time; each black vertical line indicated by an arrow represents a global change to the timing with respect to the ATLAS clock (made by the CTP or L1Calo); the lack of points between October 2010 and March 2011 is due to the heavy ion running period and the winter shutdown; the upward trends which appear around September and October presumably reflect a cooling of the ambient temperature. These upward trends are very much shorter due to this becoming a known effect which was by then closely monitored and quickly adjusted.

Monitoring was continued so as to verify various timing changes in 2011. It was discovered that by March, statistics were high enough to observe timing changes in the manner shown in figures 4.22 and 4.23 for collision periods as short as ten minutes!



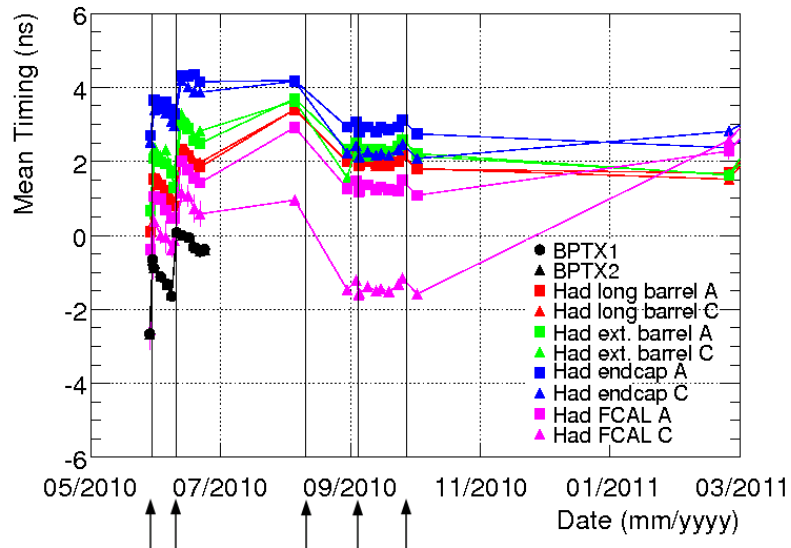


Figure 4.23: Hadronic layer  $\Delta finetime$  tracking for period June 2010 - March 2011. The solid vertical black lines indicated by arrows represent timing changes made by the CTP or L1Calo.

## 4.5 Timing Problems in the Overlap

The ‘overlap’ in the electromagnetic calorimeters is a region of transition between the EM barrel and EM endcap as illustrated in figure 4.24. Strictly speaking, it exists at  $1.37 < |\eta| < 1.52$ , although is only noticeable for L1Calo in the trigger towers defined at  $1.4 < |\eta| < 1.5$ . These towers are clearly visible in 4.20*b* as two stripes of poorly timed towers.

The fact that the trigger towers in the overlap receive a combination of signals from layers in *two* partitions, which are independently timed and each with slightly different pulse shapes, means that making a good calibration is not possible in the normal sense. The summed results compiled by the trigger tower were composed of badly timed pulses which varied in size depending on the distribution of energy deposits within the layers for a particular hit.

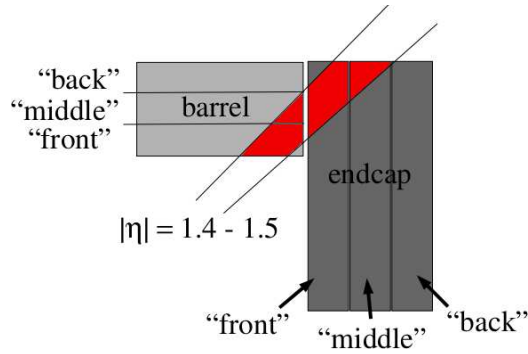


Figure 4.24: Diagrammatic representation of the ‘overlap’ region in L1Calo (EM), indicating in red the layers of the barrel and endcap which make up the trigger towers in this region. The presampler layer has been omitted for clarity.

The only way to properly analyse these pulses was to vary the conditions during collision running by turning off parts of the calorimeter.

For a few days in March 2011, ATLAS was unable to take physics data due to problems with its magnet system; however, collisions continued as normal. The opportunity was taken to do dedicated tests involving switching off some layers of EM calorimeter in various combinations. This allowed a measurement of the difference in timing between the barrel and endcap parts of the overlap. The datasets involved are numbered 177904 and 177912.

As only the *differences* in timing were being studied it was appropriate to use Method II, as documented earlier in this chapter.

#### 4.5.1 Results III - Calibration of Overlap Regions

The results of average timing differences from the analysis for the two special runs designed to observe separate layers in the overlap are given in figure 4.25. In figure 4.25 (a) the endcap region of the overlap is active with the barrel

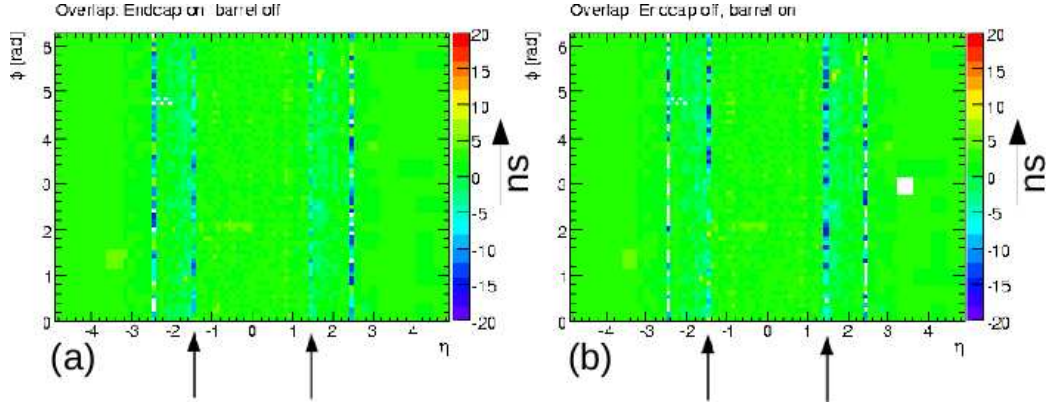


Figure 4.25:  $\eta - \phi$  maps of the EM trigger towers average timing in ns with different layers active and non-active. Overlap: endcap layer on (a), barrel layer on (b). Towers coloured white are those with no entries.

region non-active. Figure 4.25 (b) has only the barrel layer on and the endcap off. Hence the inter-layer timing difference can be calculated. This is shown in figure 4.26. It is quite clear that the largest differences are in the overlap region at  $1.4 < |\eta| < 1.5$ .

The timing differences of the 64 towers of each the A and C sides are graphed in figure 4.27 labelled ‘L1Calo’ in the legend. Timing differences were also verified by the Liquid Argon detector group and these are plotted in the same figure with the label ‘LAr’. The two results agreed well, and the prescribed corrections were applied via the Tower Builder Boards (TBBs).

To date no additional single layer data runs have been taken to directly examine the effectiveness of the corrections, but standard (all layers on) runs reveal that a significant improvement has been made. Figure 4.28 shows a timing analysis in the style shown previously, on a dataset taken following the overlap region correction. Figure 4.28 b shows that the *average* timings in all of the overlap trigger towers are commensurate with the surrounding areas.

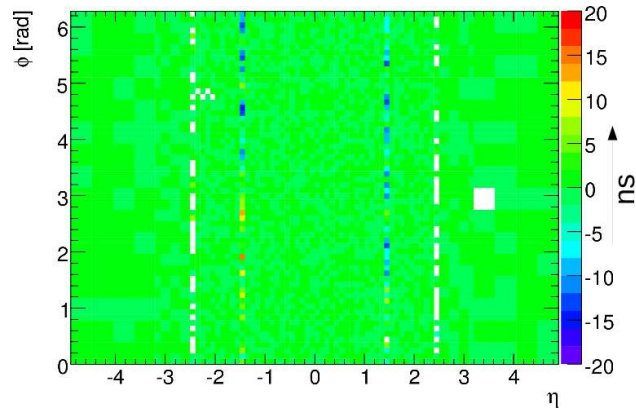


Figure 4.26:  $\eta - \phi$  maps of the difference in mean average timing in ns of the EM trigger towers ( $t(\text{barrel on}) - t(\text{endcap on})$ ).

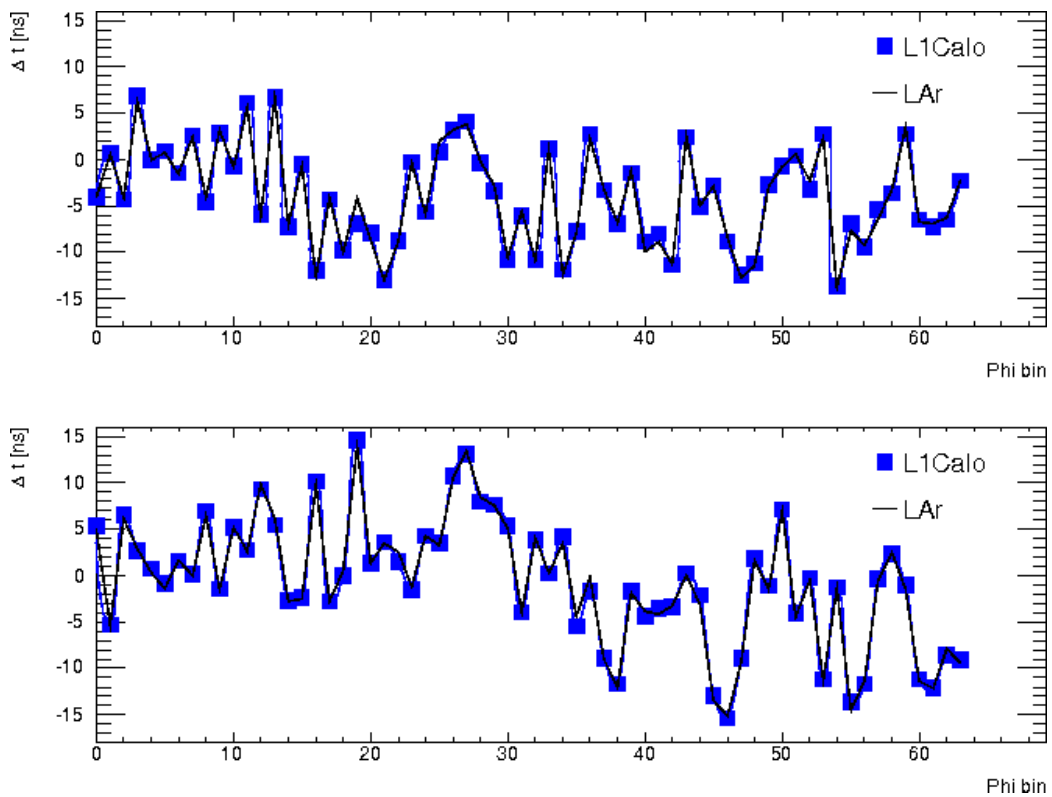


Figure 4.27: Deviation in timing as measured by Method II (L1Calo) and quadratic method (LAr) for the  $\phi$  bins of the overlap region, *A*-side (top) and *C*-side (bottom).

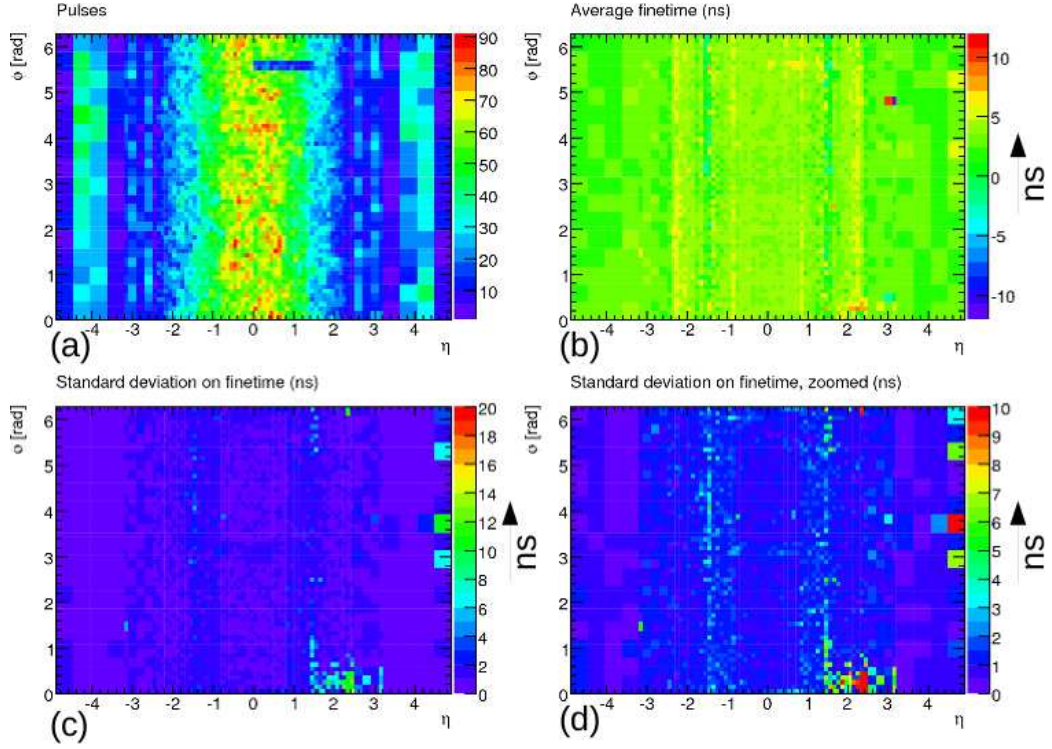


Figure 4.28: Results from timing analysis on dataset 182161 following the timing corrections in the overlap region. (a) number of contributing pulses, (b) mean timing difference in ns, (c) and (d) standard deviation in ns with different scales.

Note that the comparatively low number of pulses in the area of the barrel coloured blue at  $\phi \sim 5.6$  is due to the front-end electronics for the middle and back layers temporarily not working for a number of towers. Also, a noise burst is seen at  $(\eta - \phi) \sim (2.5, 0)$ .

However, figure 4.28 *c* and *d* reveal that the standard deviations remain larger in the overlap than in much of the rest of the barrel and endcap. Unfortunately these measurements are unlikely to be improved much further. Variance in the pulses reported by the trigger towers in the overlap region is inevitable as each pulse is a combination of barrel and endcap endcap

calorimeter signals. The pulse shapes associated with those two partitions are slightly different, and varying contributions from them will apply for each pulse depending on how the energy was deposited in the layers.

## 4.6 Summary

Methods for electronic timing calibration and monitoring of the ATLAS experiment Level-1 Calorimeter subsystem have been devised and implemented. Signals in the apparatus were produced by either cosmic or collision particles. The resulting digitised pulses were fitted using one of two methods, as described in this chapter.

The inter-partition timing was studied using pre-collision cosmic data, so as to make synchronous independently calibrated sections of the subsystem. Corrections amounting to less than one bunch crossing (25 ns) were applied to the EM barrel so as to align the A and C sides.

Monitoring of timing ‘drift’ on short- and long-term bases was carried out over a period of one year. The main source of variation was attributed to the change in ambient temperature and resulting effect on the transmission of the LHC clock signal. Corrective adjustments are now made regularly by the CTP subgroup.

Using special collision data runs, the EM ‘overlap’ region at the join between the barrel and endcap was calibrated. This presented a particular problem because the signals coming from these regions are a combination of pulses from two different partitions, and hence have different shapes. By calibrating the two partitions individually, an improvement is made so as to see the disparity between the overlap and surrounding areas reduce from 10-15 ns to <5 ns.

# Chapter 5

## Jet-finding Algorithms

A ‘jet’ is a collection of collimated particles, which may be considered as the result of the hadronisation of an individual parton. A jet algorithm is a general method for obtaining defined objects, jets, from these particles. In order to be relevant and comparable, a jet must be defined precisely, by factors such as the measurement conditions (for example, description of calibration), jet-finding algorithm used and associated parameters. All of these details will change the number and size of the jets found, in all but the simplest of cases. Qualification must be made of the possible jet constituents, i.e those particles that may be considered for inclusion in a jet. The rapidity region used and minimum  $p_T$  sensitivity should be quoted, as well as the treatment of background effects. The comparison to MC must be stated as the scattered or showered parton-level, or final state stable hadrons, in which case the minimum particle lifetime should be given. In experimental data a more detailed description of the calibration scheme, detector response and trigger menu will be required. Additionally the summation scheme should be stipulated. This could be either the 4-momentum additive ‘E-scheme’ or the transverse energy component  $E_T$ -weighted average of  $\eta$  and  $\phi$ , although

it is the E-scheme which is more highly recommended in recent years [57].

A jet-radius or ‘distance parameter’,  $R$  is the main user-input variable of many algorithms and will be described for individual algorithms later in this chapter. This value has an effect on the number of jets found as well as their 4-momenta, whichever algorithm is used, as can be seen in [58]. The chosen size may depend on the nature of the analysis involved. For example, the inclusive QCD jet cross-section measurement typically uses a larger size so as to capture, safely, all radiation from the scattered parton, although too large an area may include increased levels of background. The search for particular processes producing multi-jet states may require a smaller sized jet-radius in order to identify different elements of the event [40]. Sometimes analyses are run with more than one jet-finding algorithm and various jet-radii for comparison.

This chapter summarises jet-finding techniques currently used at hadron colliders. These can be divided into the categories ‘sliding window’, ‘cone’ and ‘recombinatorial’. They will then be compared for relative advantages. It is worth noting that the principal fragmentation function study of  $e^+e^-$  annihilation and DIS involved no jet-finders, instead relying on hemispherical phase space for  $e^+e^-$  or the Breit frame [3] for DIS. Of course these collision events benefit greatly from a noise-free environment and hence do not *require* complex jet-finding techniques. Jet-finding for  $e^+e^-$  and DIS will not be discussed here.

## 5.1 Infrared and Collinear Safety

Two important characteristics of a jet-finder are those of infrared and collinear safety. Only algorithms that are both infrared and collinear (IRC) safe are



recommended for use in analysis. IRC safety is particularly essential at the LHC due to the pervasion of multi-jet states. IRC unsafety breaks the relation between the parton and corresponding, observed, jet. It also makes it impossible to make finite, perturbative calculations. A good jet-finder should produce the same collection of jets whether analysed at parton-, particle- or detector-level [40], or whether calculated at indefinitely higher orders. Within pQCD there are cross-sectional calculation singularities for collinear or low-energy radiation. In addition, soft tracks are produced by underlying event or pile-up.

### **Infrared Safety**

An algorithm is infrared unsafe if the jet configuration for a particular event depends critically on a single soft particle being above the measurement threshold, as this may lead to the combination of jets which would otherwise be separate. Also, if all particles act as jet-seeds, or there is a low seed threshold applied, then it is possible that the inclusion of a relatively soft particle may produce a new jet. When this happens, an algorithm is IR unsafe, as the final result depends on whether the soft particle formed a jet, combined two jets or otherwise changed the final jet state.

In figure 5.1, the addition of a soft particle between particles belonging to the two jets has affected the final combination of jets. In this case the algorithm is IR unsafe. The additional particle may, for example, have originated from calorimeter noise. It is therefore not necessarily part of the hard event and hence should not alter the final jet state.

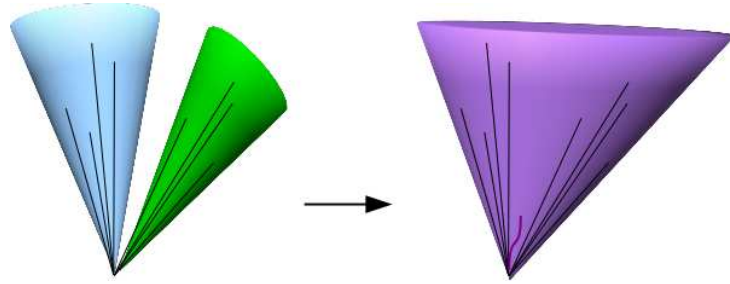


Figure 5.1: Diagram illustrating the problem of infrared unsafety, where two different final jet states are found, subject to an additional soft particle.

### Collinear Safety

A jet algorithm is collinear unsafe if exchanging one particle with momentum  $2x$  with two collinear particles each with momentum  $\sim x$ , as illustrated in figure 5.2, alters the results of the algorithm. The total 4-momentum is the same in each case and therefore the final result should be identical. If the algorithm is seeded above a certain threshold, the result may alter when varying that threshold, particularly if there is any  $p_T$ -ordering involved.

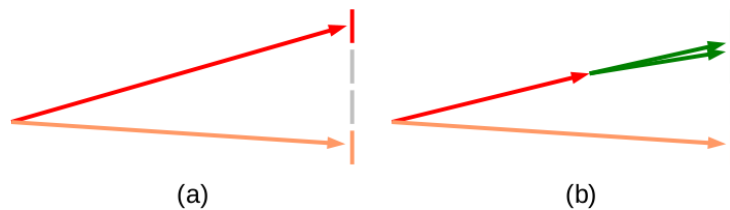


Figure 5.2: Diagram showing two possible situations, where a hard particle (red) is measured as one hard (a) or two softer particles shown in green (b). In case (a) the hardest particle is chosen as such, but in (b) the less-hard (orange) particle is selected as the hardest.

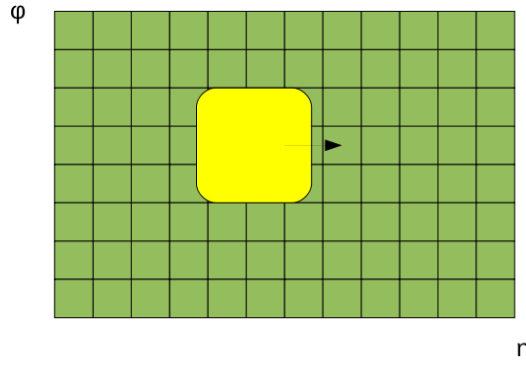


Figure 5.3: Diagram illustrating the sliding window principle, with the window in yellow moving over calorimeter cells.

## 5.2 Sliding Window

A sliding window algorithm is a simple generic jet-finder. A square or rectangular ‘window’ of fixed, assigned size, but usually several calorimeter cell widths wide, slides across the Lorentz-invariant intervals of  $\eta - \phi$  space, as illustrated by the diagram in figure 5.3. It sums the energy deposits within the window, and then moves on by one unit in either of the two dimensions, so as to identify areas of activity above some threshold. This is a fast and robust technique that is still used at the most basic level of trigger systems, for example on the ATLAS Level-1 Calorimeter trigger. Of course the jet is unlikely to be well-represented by this search shape, as a combination of radiation and boosts produce, on average, conical jets in  $\eta - \phi$  space.

## 5.3 Cone-type

The simple cone jet algorithms take final state particles as ‘seeds’ in an iterative search for stable cones with a predefined circular size. Cone-type algorithms are known as ‘top-down’. High- $p_T$  particles above a certain threshold

are labelled as seeds for initial cone directions. Each cone has a fixed circle radius in  $\eta - \phi$  space dictated by the user, known as  $\Delta R$ . This is defined via equation 5.1, where  $\Delta\eta$  and  $\Delta\phi$  are the distances from the circle centre.  $\Delta R$  is assigned values such as 0.4 or 0.7 depending on the analysis undertaken.

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (5.1)$$

The search cone is defined in  $\eta - \phi$  space such that the jet is invariant to, and hence independent of, boosts in the beam direction. Particles are included in the jet if within the radius  $\Delta R$  of the initial cone directions.

Equal coefficients of  $\Delta\eta$  and  $\Delta\phi$  (i.e. a circle) are justified by the argument that for an independent process, radiation is emitted azimuthally with equal probability about the parton's direction.

In an  $x - y$  plane, the  $k_T$  of a radiated particle from a parton is given, for small  $\delta\phi$ , by

$$k_T \sim p_T \delta\phi \quad (5.2)$$

and similarly in the  $z$ -transverse plane by

$$k_T \sim p \delta\theta. \quad (5.3)$$

As  $\theta$  is not a Lorentz invariant quantity, it must be replaced. This is done by differentiating  $\eta$ , as defined in equation 2.6, with respect to  $\theta$ .

$$\frac{d\eta}{d\theta} = -\frac{1}{\tan\left(\frac{\theta}{2}\right)} \frac{d}{d\theta} \tan\left(\frac{\theta}{2}\right) = -\frac{1}{\sin\theta} \quad (5.4)$$

Hence,  $\delta\theta \sim -\sin\theta\delta\eta$ .

By substitution, and equating equations 5.2 and 5.3, and given that radiation in either direction is equally likely, it is shown that the magnitudes

of  $\eta$  and  $\phi$  are equivalent, and by this reasoning, the parton is equally likely to radiate into an element of  $\delta\eta$  or  $\delta\phi$ .

$$\delta\eta \sim \delta\phi \tag{5.5}$$

Of course the circular shape is only created on average, and radiation may not be totally independent, e.g. there may be influence from the colour field between the proton remnant and the parton.

In principle, the 4-momenta of all particles found within each cone is summed and the result becomes the direction of a new cone. This process is iterated until stable, when the 4-momentum of the new sum is the same as within the previous cone.

The basic principle is illustrated in figure 5.4, where lines represent the paths of particles from the interaction point. In the first instance, cones are centred on the seed particles, represented by arrows. In the second diagram, which follows the first iteration of the process, the cone centres have moved slightly, and in the case of the cone with the arrow pointing up, incorporated an additional particle, precipitating a further iteration.

When applied to a more complex state than the one shown in figure 5.4, the process may identify more than two stable cones. These could overlap and some particles may be members of more than one jet, which is clearly problematic. Two methods exist to combat this problem [57]. One involves ‘progressive-removal’ - the exclusion of particles once a jet has been defined, so they are not available for inclusion in other jets. Hence the order of jet selection becomes important. This is done by  $p_T$ - or mass-ordering, where the first cone is iterated from the hardest  $p_T$  or highest mass particle. The second cone is iterated from the hardest  $p_T$  particle remaining, and so on.

The other possibility is the split-merge method. Two cones are amalga-

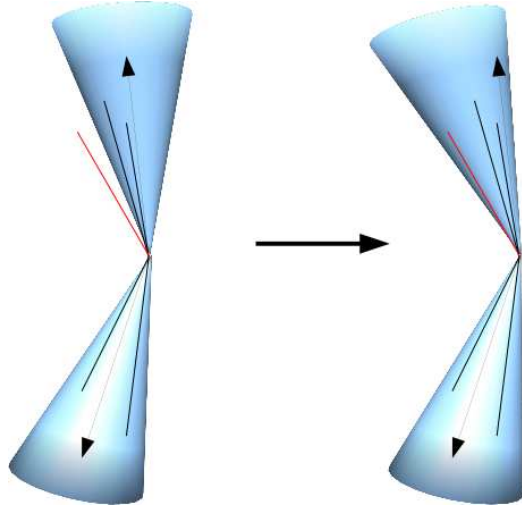


Figure 5.4: Diagram illustrating the principle of the cone jet algorithm, where the lines represent particles originating at the interaction point.

mated if greater than fraction  $f$  of the softer cone's  $p_T$  is in common with the harder cone. If the fraction is less than  $f$  then the shared particles are attributed to the closest cone. Naturally, this can cause jets to be non-circular in  $\eta - \phi$  space. Typically  $f$  is given a value of 0.5 or 0.75, although 0.75 is considered safer, as smaller values can overly merge several jets, creating ‘monster-jets’ [57].

There can exist a theoretical stable cone with hard particles on opposite sides but nothing in between. Hence two stable cones are found centred on each of the hard particles. But if any soft particle between them acts as a seed, then the third cone containing all three particles is also found, and the method is IR unsafe. This problem was partially overcome by adding artificial ‘midpoint’ particle seeds between stable cones separated by less than  $2R$  [59]. Then the algorithm would search for new stable cones. However, this does not solve the similar scenario of three particles in a triangular configuration, and so the problem is postponed to higher orders in the coupling.

### 5.3.1 SIScone

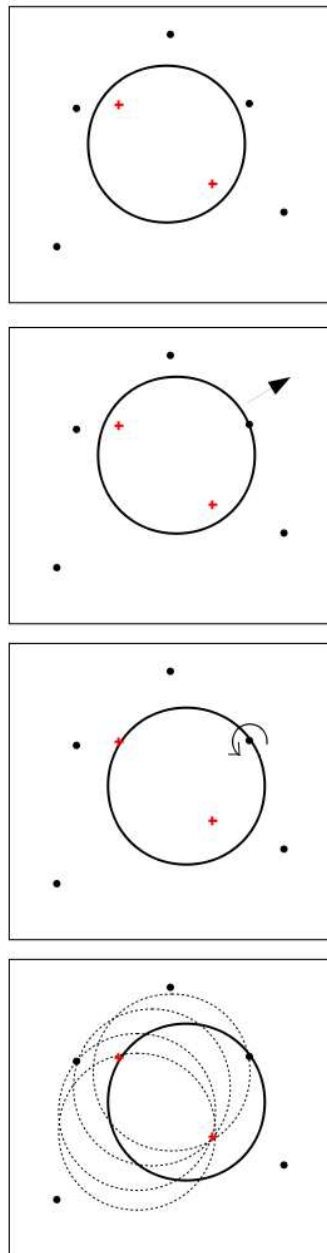
An exact seedless cone jet algorithm is a cone algorithm that can identify *all* stable cones independently of seeds. It defines *every* particle in the event as a seed. A cone of fixed radius  $R$ , is moved throughout  $\eta - \phi$  space until it finds a particle. The cone is then iterated until stable in the manner described previously. Other cones are constructed in the same way, and they are then split/merged by the method described previously.

The time taken to execute this exact method scales as  $\mathcal{O}(N^2N)$ , where  $N$  is the number of particles in the event. For typical jet events, this equates to an unacceptable duration. The Seedless Infrared-Safe Cone (SIScone) jet algorithm [59] has reduced this to  $N^2 \ln N$ .

The procedure involved is complex, but the general approach is described pictorially in figure 5.5.

## 5.4 Recombinatorial

The basis of sequential recombination algorithms is to define a ‘distance’ between particles or the energy content of calorimeter towers and iteratively recombine the ‘closest’ pair of objects up to some stopping limit, known as the ‘distance parameter’ (akin to the jet radius in the cone algorithm). This is usually a dynamic, rather than angular variable, such as  $k_T$  and is described as a ‘bottom-up’ approach. However this method takes a time  $\mathcal{O}(N^2N)$ , among  $N$  particles or towers. Hence in practice it is computationally intensive for complex jet events, and the time required to carry out may be prohibitive. Therefore in usage, particles are often pre-clustered before the recombinatorial technique begins.



1. A circle with radius  $R$  is placed in an initial position on the rapidity- $\phi$  plane.

2. The circle moves in a random direction until it touches an external or internal particle.

3. The circle pivots about the point touching the circle edge until it touches a second particle.

4. Each pair of particles within distance  $2R$  of one another have two associated radius  $R$  circles for which they lie on the circle's edge.

Figure 5.5: General approach to the SIScone method. The red crosses represent the initially enclosed particles and black dots are other particles in the region.



### 5.4.1 $k_T$ algorithm

The  $k_T$  algorithm takes all final state hadrons or, experimentally, calorimeter cell deposits, as input. If computing load is an issue, it forms proto-jets from cones with a nominal radius out of seed towers, where towers must record a transverse energy greater than some threshold [60]. It then continues with a sequential recombination, clustering these proto-jets into full jet objects.

The stages of the inclusive  $k_T$  algorithm are illustrated by the flowchart in figure 5.6, where  $E_{T,i}$  is the transverse energy associated with object  $i$ ,  $\Delta R_{ij}$  is the  $\eta - \phi$  distance between objects  $i$  and  $j$ , and  $R$  is the jet-parameter. All two-object combinations are considered for each iteration, before one pair are either combined to form a new object, or one or both objects forms a final state jet.

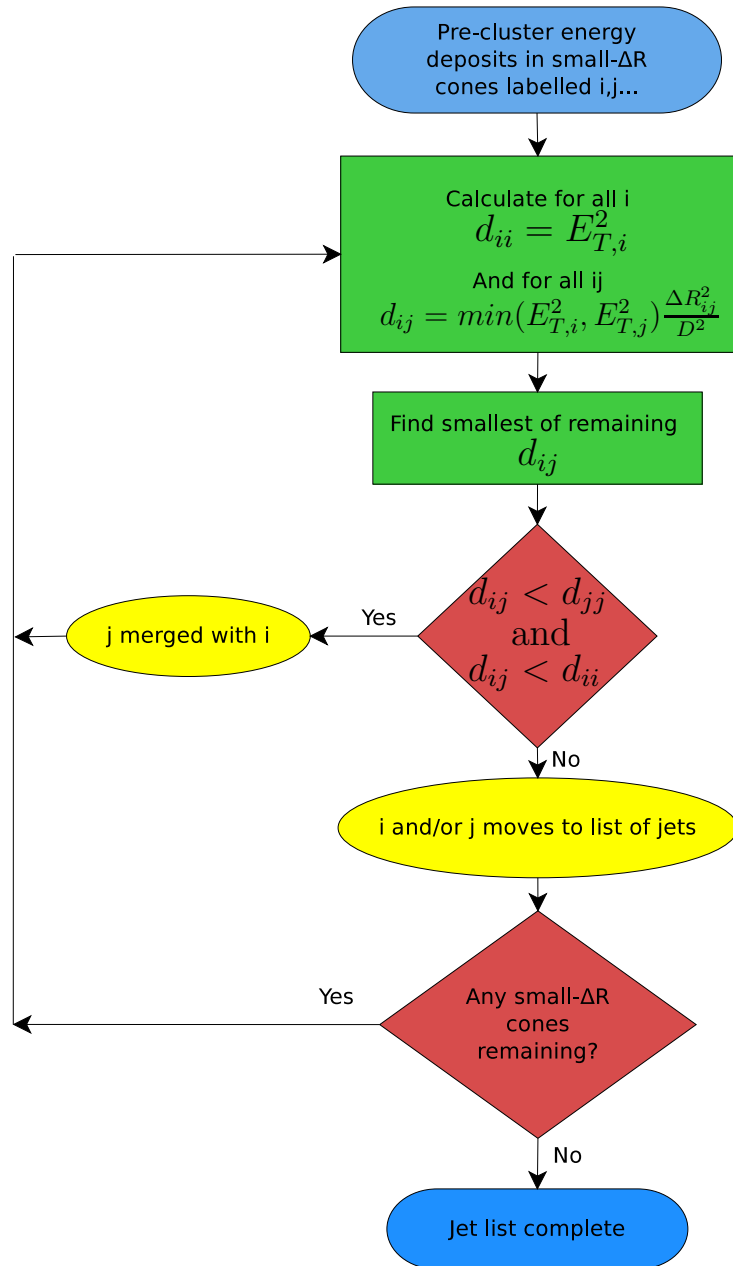
The *exclusive*  $k_T$  algorithm has a similar basis, but outputs either a user-specified number of jets, or all jets above some energy threshold,  $d_{cut}$ . This is done by merging objects close to the beam with a ‘beam jet’.

Potential differences in jet area when using the  $k_T$  algorithm compared with a cone are shown schematically in figure 5.7. The cone may exclude some of the energy in the proto-jets, whereas the  $k_T$  algorithm includes everything in the vicinity.

### 5.4.2 Variations on the $k_T$ algorithm

There are two variants on the  $k_T$  algorithm, known as anti- $k_T$  and Cambridge-Aachen algorithms. The method is broadly the same in each, as characterised in figure 5.6. The difference lies in the exponent of the transverse energy,  $E_{T,i}^{2p}$ , as used in figure 5.6, with:

- $p = 1$  :  $k_T$

Figure 5.6: Flowchart describing the key steps of the inclusive  $k_T$  algorithm.

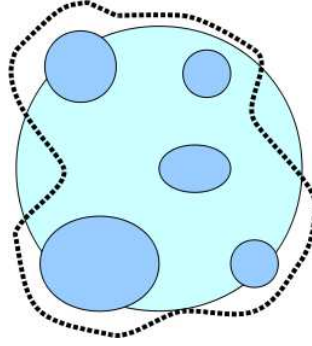


Figure 5.7: Diagram in  $\eta - \phi$  space, showing the different jet areas when using a cone (light blue) and the  $k_T$  algorithm (dotted line). The darker blue areas represent significant transverse energy deposits.

- $p = 0$  : Cambridge-Aachen
- $p = -1$  : anti- $k_T$

In the case of Cambridge-Aachen, only the  $\eta - \phi$  distance between two objects is used. The main difference in practice between  $k_T$  and anti- $k_T$  is the ordering of the recombination. Where  $k_T$  tends to combine small objects in the vicinity of each other together first, anti- $k_T$  combines small objects with nearby large objects first. The results of *these* two approaches are often similar or identical final jet states, and the differences will be discussed in the following section.

## 5.5 Comparison of Jet-finding Algorithms

The most simple method to understand is the sliding-window algorithm. But while this is fast and robust, it is also a very crude method, which can only sum the deposits within a fixed area. This is useful for a ‘first-look’ at an event, but cannot identify any complicated structure. Hence it is used only

in low-level triggers.

The cone algorithm is more sophisticated, but either uses a fixed area for its stable cone search or achieves even greater levels of complexity and parametrisation if cones must be split or merged. Results depend on the size of the search area and the seed energy threshold. Reliance on seed particles can also cause problems in certain particle configurations, when a low energy deposit may entirely change the list of jets produced, as seen in the example in figure 5.1. Hence the basic cone algorithm is not infrared safe. The low energy deposit may or may not have occurred or been detected, or it may even have been caused by electronic noise. These arbitrary factors that cause potentially different final jet results are dangerous features of jet-finding.

The  $p_T$ -ordered progressive removal technique can create a collinear unsafe situation, if the hardest particle splits into an approximately collinear pair, and causes an alternative particle to be the seed for the first jet. In this case, a different final state may occur.

The use of a seedless cone, albeit one necessarily ‘seeded’ by the calorimeter’s minimum threshold, and essentially making all particles into potential seeds, is considered IRC safe, but is generally a slow process due to the large number of cone combinations that must be considered. The SISCone algorithm has been shown to be an IRC safe cone algorithm. Performance of this code is considerably more efficient, although not as fast as the sequential recombination-type of algorithm.

As they are seedless, and soft radiation does not affect the jet-boundary, sequential recombination algorithms are IRC safe. The  $k_T$  algorithm frequently produces non-regularly shaped jets, which is considered non-ideal. Although in reality jets are *not* circular, knowing the area of the jet search is preferable when considering a background subtraction. All jets found by

the anti- $k_T$  algorithm *are* circular, excepting those that may have been split or merged, with a radius equal to the distance parameter. Hence cone-like jets are produced without the use of a fixed cone size.

The amount of time taken to run an algorithm contributes to its practicality in use. Recombination algorithms typically take  $\mathcal{O}(N^3)$ , which translates to approximately 1 s for  $N = 1000$  and is regarded as too slow. However, this can be greatly improved upon by engaging the FASTJET [61] software. FASTJET is a tool for fast implementation of recombination-type jet finders and is a uniform interface to others such as SISCone. It includes tools for calculating jet areas and background. Table 5.1 [62] summarises the scaling of commonly used jet algorithms within FASTJET, and demonstrates with examples how that scaling affects the run time for different numbers of  $N$  particles.

Algorithm	Formula	t(N = 100)	t(N = 500)	t(N = 1000)
Recombinatorial	$N \ln N$	1	7	15
SISCone	$N^2 \ln N$	100	3400	15,000

Table 5.1: Relative timings of commonly used jet-finding algorithms using FASTJET, normalised to the recombinatorial-type with  $N = 100$ .

Anti- $k_T$  is the current default algorithm in the ATLAS experiment. The reconstructed jets best match the truth in Monte Carlo, having a better efficiency and purity across a wide transverse energy scale [63]. In addition, the preference for circular jets is strong, as these are less prone to pile-up.

# Chapter 6

## Introducing TRAPS

### 6.1 Motivation: Why Develop TRAPS?

As summarised in the previous chapter, there are many existing jet-finders which are commonly used and accepted by the particle physics community. However there is no ‘perfect’ jet definition and with the resolution on jet energy being the dominant systematic error on many measurements, there is always room for improvement, even if that improvement is seen only for specific circumstances or in a particular energy regime. Therefore it may be necessary to specify a different goal to previous work, or to make significant progress in efficiency in order to cope with the huge output of high luminosity experiments.

TRAPS is a different type of jet-finding algorithm to those discussed in chapter 5, in that it seeks to produce exactly *two* jets per event, and without a pre-defined size or other scaling variable. Nor does it wish to have to cut out events because the energy deposits are not azimuthally back-to-back, or do not balance in transverse momentum, for example due to initial state radiation. This would restrict the sample and may create a bias in the calcu-

lations of fragmentation functions, structure functions and other important results essential for creating better models in QCD.

TRAPS is an algorithm that attempts to reproduce properties of partons exiting the hard interaction. Currently there is no interest in separating out multi-jets, the aim being to assess the energy deposits resulting from the final state partons outgoing directly from the hard scatter, and subtract the relevant underlying event (UE).

The ethos is:

- to produce a result for every event with few or no event cuts and hence use the greatest possible sample of high-energy QCD events,
- to avoid parameters where possible and show them to be arbitrary and thus insensitive if not,
- to avoid dictating in advance the size or shape of the ‘jet’,
- to write a simple, fast and robust algorithm.

The main assumption is that each high- $p_T$  event is dominated by a  $2 \rightarrow 2$  parton scatter; therefore the algorithm searches for two ‘TRAPS-objects’, which are considered *scattered parton candidates*. In doing so it intrinsically searches at a resolution at the hard scale of the interaction. However, the algorithm does not attempt to work with only  $2 \rightarrow 2$ -type events. However, these are assumed to make up a significant proportion of the total of all jet events.

TRAPS is specifically designed to work at LHC energies and thus explore the high momentum transfer at high luminosity that has only recently become available. Therefore the scope of TRAPS is to analyse events with a hard scale greater than  $\sim 100$  GeV, where the potential for new physics is greatest and

existing  $e^+e^-$  and DIS data ends. These also tend to be the cleanest QCD events produced at the LHC.

The TRAPS algorithm includes an event-by-event underlying event subtraction. While the idea of this is to subtract UE energy deposits (ISR, MPI, etc. as previously discussed) which may lie within the TRAPS-object search area, but keep anything coming from the hard scattered partons (e.g. FSR), the developers of TRAPS realise that it is intrinsically impossible to know *which* deposits are which without the insight of the MC model. The method used here attempts to assess the UE energy density in non-TRAPS-object regions and then calculate the subtracted value based on the area of the search size. In this way it is hoped that a good energy measurement is made *on average* whilst still enabling the UE assessment on the very same events as the signal.

## 6.2 Development Techniques

TRAPS was developed using PYTHIA Version 6.418 Monte Carlo in conjunction with Rivet Version 1.1.1. Rivet can produce ROOT output files containing final state particle 4-vectors at ‘generator’ level. Thus the particles have not undergone any smearing or other detector effects associated with ‘simulated’ level data. Various cuts on minimum momentum transfer ( $\hat{p}_{T_{min}}$ ) values were used, with the main focus during development on 100, 200 and 500 GeV, where the LHC will provide significant quantities of new information.

There are major advantages to developing TRAPS using MC. The samples of events are easily controlled, so that  $p_{T_{min}}$  may be varied and UE such as Multiple Parton Interactions turned on and off so as to examine the relative effects.



In order to analyse fully reconstructed Monte Carlo and ATLAS data, as well as to easily switch between various MC models for more rigorous comparison, TRAPS had to be incorporated into Athena [38] - the ATLAS software. This has the capacity to add the effects of the detector to MC output and also permits access to ATLAS datasets so that theory and experiment may be compared directly. The usage of TRAPS in Athena will be described in chapter 7.

### 6.2.1 Algorithm Testing

Feynman diagrams show partons as both ‘initial state’ and ‘final state’ particles. This reflects the philosophy of the TRAPS algorithm to regard the initial and final partons taking part in high energy QCD processes as measurable entities and hence physically existent. These final state partons are ‘off mass shell’, i.e. when  $E^2 - p^2 > m^2$ , in a time-like manner, but are considered by TRAPS as *real*, i.e. as a collapsed wavefunction.

However, the partons generated by the Monte Carlo do not necessarily reflect reality - their behaviour is model dependent. Hence to consider these partons as ‘real’ is clearly a simplified assumption. However, it is necessary when studying quantum mechanical processes at the human scale, to take a pragmatic approach ‘for all practical purposes’. Therefore during the development stage, as a working approximation, the TRAPS-objects will be compared with the partons as supplied by the MC. This technique, described in section 6.6, is necessary, but not necessarily sufficient.

For the final assessment, and in order to compare in an appropriate manner with existing jet-finders, TRAPS is applied to the decay products of a sample of hypothetical excited quarks, generated in MC with a high mass and narrow width. The distribution of invariant masses of the TRAPS-objects

is compared to that of the partons being produced in MC by the excited quark. The distribution width is assessed and compared with results from top-of-the-range jet-finding algorithms. Results are shown in section 8.1.

### 6.3 Input Objects

The input objects to the TRAPS algorithm are 4-momenta associated with final state hadrons generated by MC, calorimeter cell deposits, or topoclusters in real or simulated data. However, due to the summation techniques involved, the algorithm is insensitive to the choice of object type. Only particles within the rapidity region  $|\eta| < 4.9$  are considered since that is the limit of the ATLAS calorimetry.

## 6.4 Technical Description of TRAPS

This section will give a complete description of all steps making up the TRAPS algorithm. Some discussion of alternative subroutines that were investigated and rejected will also be given. Comparisons with the parton information will follow.

### 6.4.1 Preparation - Setting $\eta_{range}$

When a suitable dataset has been created or acquired according to some Data Analysis Model (DAM) such as that outlined in section 9.1, TRAPS calculates the possible limits of pseudorapidity that may be reached by the hard scattered partons in a sample with particular minimum momentum transfer,  $\hat{p}_{T_{min}}$ . This range is calculated on a sample-by-sample basis, via equation 6.1, assuming a maximum possible  $x = 1.0$ . Partons from the vast

majority of events comply, nevertheless, this range *can* be violated with the boost of a suitably large ISR or by a rare  $s$ -channel process.

$$\eta_{range} = \pm \ln\left(\tan\left(\frac{1}{2} \arcsin\left(\frac{\hat{p}_{Tmin}}{p_{Beam}}\right)\right)\right) \quad (6.1)$$

These  $\eta$  define the range over which TRAPS will look for the scattered partons and immediately excludes regions of rapidity containing, in most cases, only proton remnant and other UE. As an example, for a sample of events with  $\hat{p}_{Tmin}$  of 500 GeV,  $\eta_{range}$  is set to  $\pm 2.63$ . Naturally, if this range exceeds the absolute limits of the calorimetry, then it is set to that limit.

### Momentum Transfer

The  $Q^2$  (momentum transfer squared) parameter of deep inelastic scattering is a measure of the probing strength executed in a single collision. The amount of momentum transferred in a proton-proton collision,  $\hat{p}_T$  is

$$\hat{p}_T = p^* \sin\theta^* \quad (6.2)$$

where  $p^*$  is the momentum of any of the colliding partons in their CMS frame, and  $\theta^*$  is the angle of scatter in the same frame.

This can be roughly approximated by a scalar sum of the transverse momentum,  $p_T$ , of all of the particles produced by the hard scatter. The value of this summation has great variation, even for events at the same CMS energy since the range of momentum taken by either interacting parton may vary between essentially zero (or threshold  $p_T$  of a given trigger) and the full beam energy. This is discussed in chapter 2.

### 6.4.2 Step I - Splitting the Event in Two

Taking all of the particles within the theoretical kinematic acceptance of the energy range, TRAPS calculates the angle of the ‘plane’ ( $\Phi$ ) for the event. This plane is defined about the  $z$ -axis, as that which has the least event momentum outside of it. During the course of this, acoplanarity, a measure of the minimised momentum outside of the plane, is also calculated. The event is then divided somewhat arbitrarily into two halves in preparation for Step II. Mathematical methods and further details are given below.

#### Acoplanarity and Event Plane

Whether assuming every event to be a dijet-type event or not, a plane can be defined that contains the largest projection of event transverse momentum. This plane has a different angle with respect to the  $x$ - $y$  plane for each event, but always uses the  $z$ -axis as its axis of rotation. Hence it is represented by an azimuthal angle  $\Phi$  in the  $x - y$  plane. This angle is calculated by equation 6.10, which is derived from acoplanarity via equations 6.3 - 6.9.

Due to ISR and FSR, which are distributed evenly in the azimuthal angle about the beam and scattered parton respectively, and even the manner in which final state particles are created, it is to be expected that none of the momentum is precisely on that plane. The algorithm calculates the acoplanarity,  $A$ , via equation 6.3 which represents the proportion of momentum squared out of the plane; an acoplanarity of zero represents a perfectly planar event and  $A = 0.5$  is an isotropic event. A new coordinate system facilitates this calculation, with  $z' = z$ ,  $x'$  parallel to the plane ( $\Phi$  from  $x$  in the  $x - y$  plane) and  $y'$  orthogonal to  $x'$  and  $z'$  as illustrated by the diagram in figure 6.1.

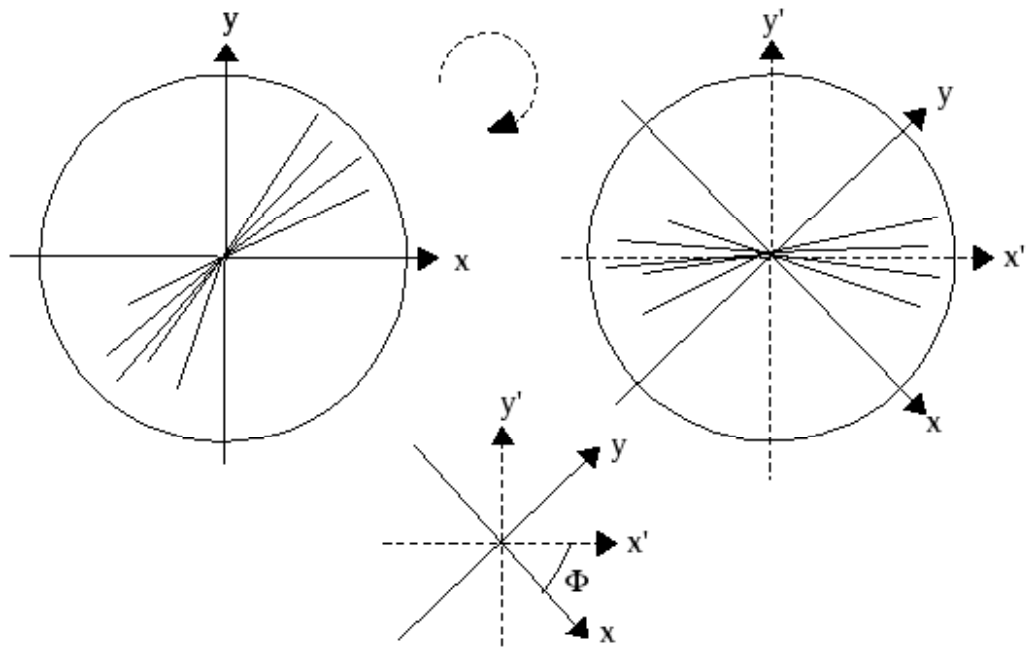


Figure 6.1: Diagram of a simple 2-jet event shown in the standard ATLAS coordinate system (left) and in the rotated coordinate system (right).

$$A = \frac{\sum_i p_{y_i}'^2}{\sum_i p_{T_i}^2} \quad (6.3)$$

where

$$p_{y_i}' = p_{y_i} \cos \Phi - p_{x_i} \sin \Phi \quad (6.4)$$

By minimising the acoplanarity with respect to  $\Phi$  it becomes possible to *calculate*  $\Phi$  and hence  $A$  in terms of only sums including the particle's  $x$  and  $y$  momentum components:

$$\frac{dA}{d\Phi} = \frac{1}{\sum_i p_{T_i}^2} \sum_i 2(p_{y_i} \cos \Phi - p_{x_i} \sin \Phi)(-p_{y_i} \sin \Phi - p_{x_i} \cos \Phi) \quad (6.5)$$

$$= \frac{1}{\sum_i p_{T_i}^2} \sum_i 2(-p_{y_i}^2 \sin \Phi \cos \Phi + p_{x_i}^2 \sin \Phi \cos \Phi + p_{x_i} p_{y_i} \sin^2 \Phi - p_{x_i} p_{y_i} \cos^2 \Phi) \quad (6.6)$$

$$= \frac{1}{\sum_i p_{T_i}^2} \sum_i (2 \sin \Phi \cos \Phi (p_{x_i}^2 - p_{y_i}^2) + 2 p_{x_i} p_{y_i} (\sin^2 \Phi - \cos^2 \Phi)) \quad (6.7)$$

$$= \frac{1}{\sum_i p_{T_i}^2} (\sin 2\Phi \sum_i (p_{x_i}^2 - p_{y_i}^2) - 2 \cos 2\Phi \sum_i p_{x_i} p_{y_i}) \quad (6.8)$$

When  $\frac{dA}{d\Phi}$  is equal to zero,

$$\sin 2\Phi \sum_i (p_{x_i}^2 - p_{y_i}^2) = 2 \cos 2\Phi \sum_i p_{x_i} p_{y_i} \quad (6.9)$$

Hence

$$\Phi = \frac{1}{2} \arctan\left(2 \left(\frac{\sum_i p_{x_i} p_{y_i}}{\sum_i (p_{x_i}^2 - p_{y_i}^2)}\right)\right) \quad (6.10)$$

Figure 6.2 is a histogram showing the distribution of acoplanarity over 100,000 PYTHIA events for generated  $\hat{p}_{T_{min}}$  of 2.5, 20, 50, 100 and 200 GeV. With low minimum momentum transfer a large proportion of events have a high acoplanarity. However at  $\hat{p}_{T_{min}} = 200$  GeV, the majority of events

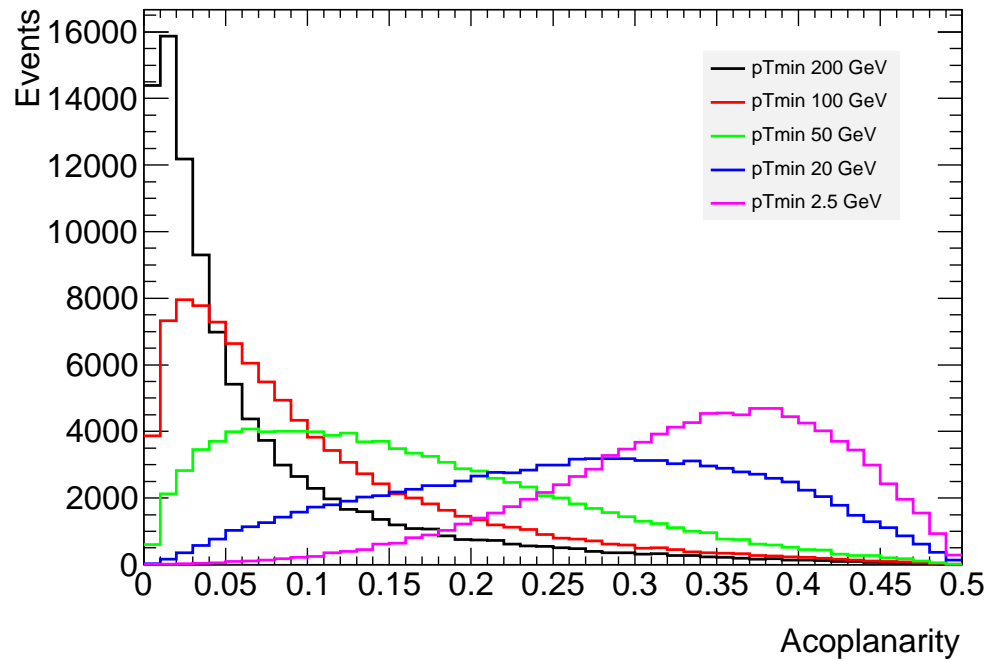


Figure 6.2: Histogram of acoplanarity of 100,000 PYTHIA events with CMS energy = 14 TeV, demonstrating average acoplanarity falling sharply with increasing generated minimum momentum transfer.

are highly planar. This implies that at high momentum transfer, the jets of particles become narrower and more well-defined, and also that the UE effects are less significant as phase space effects begin to limit ISR production.

### Defining Semicylinders

Consider splitting the cylindrical detector into two semicylinders, where the division takes place at a right-angle from the event plane ( $\Phi$ ) in azimuth. The semicylinder containing the positive  $x'$ -axis is labelled ‘positive’ and the other ‘negative’ as shown in figure 6.3. This defines the physical regions where Step II will begin to assess the momentum distribution of the event.

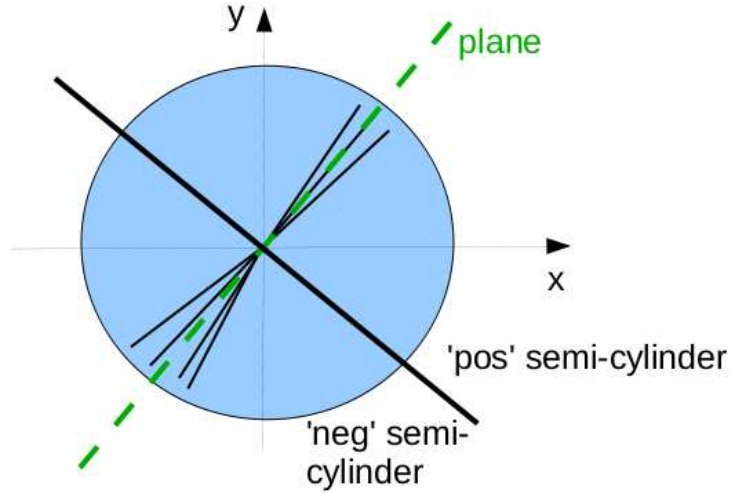


Figure 6.3: Diagram illustrating definition of ‘positive’ and ‘negative’ semi-cylinders. The event plane (dashed green line) lies along the  $x'$ -axis as defined in fig. 6.1

### 6.4.3 Step II - Floating Boundary Mean

Once the two semicylinders have been defined in Step I, TRAPS assesses the mean distribution of transverse momentum in  $\eta$  and  $\phi$  for each half of the defined phase space. This is a coarse method for quickly establishing two starting search-points. In clean events where the collimated clusters are narrow and the UE is minimal, these  $\eta$ - $\phi$  points can be accurate at assessing the initial direction of the partons.

This is well-illustrated by the negative semicylinder when running PYTHIA for one event with a minimum momentum transfer of 500 GeV. This example is shown in the schematic diagram in figure 6.4. The  $p_T$ -weighted energy deposits are represented by black dots, and the calculated means in each dimension are indicated by blue lines.

However, this step alone may not be so accurate for the positive semi-



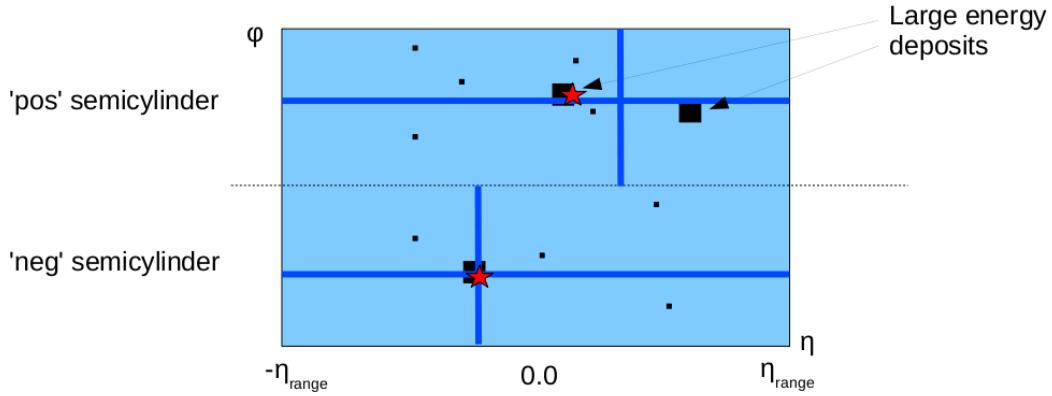


Figure 6.4: Schematic of a PYTHIA event on an  $\eta$ - $\phi$  map indicating (in black) the  $p_T$ -weighted energy deposits, position of the two semicylinders, the mean  $\eta$ - $\phi$  positions (blue lines) and true final state parton positions (red stars).

cylinder in the same event, where there are two significant  $p_T$  deposits, and the mean lands somewhere between them.

In figure 6.4 the direction of the final state partons is indicated by red stars, coinciding with one of the positive semicylinder deposits, and perhaps suggesting that the other energy deposit was caused by a large radiated particle. The appearance of two distinct energy deposits in one semicylinder is observed frequently. However, without the benefit of the MC, it is not possible to know which was the radiated particle. The algorithm can not differentiate between ISR and FSR, and although it seeks to incorporate all FSR and ignore ISR, it can only rely on the statistical nature of radiated partons, under the common definition that FSR are more highly correlated with the parton of interest and the ISR with the beam direction.

### Changing Boundaries

Following the determination of an initial set of mean values, the iterative ‘floating boundary’ process begins. This allows some removal of UE effects

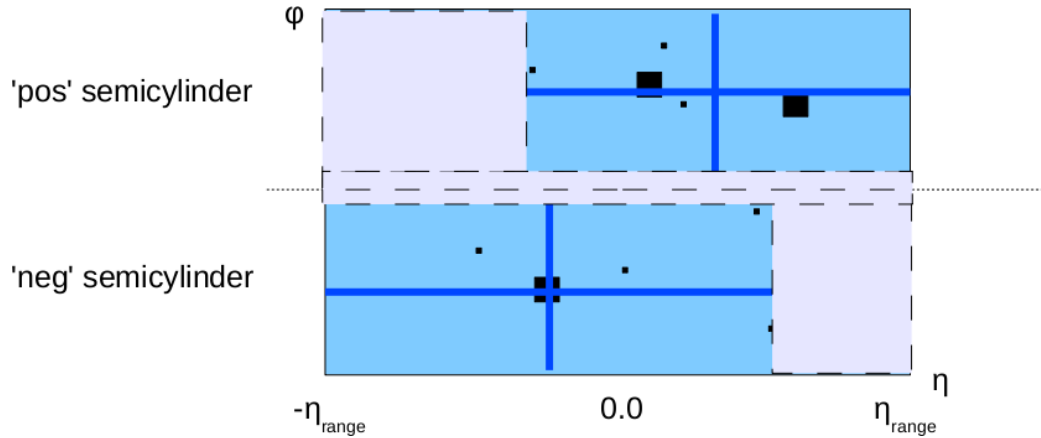


Figure 6.5: Schematic of a PYTHIA event on an  $\eta$ - $\phi$  map demonstrating the floating boundary method, where the grey areas are excluded for the following iteration.

in certain events.

The  $\eta$  and  $\phi$  boundaries are redefined so as to position the means in their centre, as shown in figure 6.5. The mean values are then recalculated using only energy deposits found within the new boundaries. This process is iterated until each of the mean positions is stable to a change of less than 0.005 (effectively zero) between iterations, and while the number of iterations is less than 25. This stopping point is merely a safety device, as this limit is virtually never reached. All four boundaries of each semicylinder are free, in practice, to move in either direction so as to centre the means, although they will not expand outside of their original ranges and cannot overlap. For each iteration, the new boundary will always be inside of the old boundary, since only deposits inside of the boundary are considered.

### 6.4.4 Step III - Calculation of Standard Deviation

Once the floating boundary mean is established, TRAPS assesses the spread of energy in each semicylinder, but only of those particles within the final boundaries set in Step II. The standard deviations about each mean are calculated for  $\eta$  and  $\phi$ . Events have a wide range of shapes and sizes, and this is reflected in figure 6.6. In this plot, the standard deviations in  $\phi$  vs  $\eta$  are plotted for 100,000 PYTHIA events with two regions (and hence entries) per event. Clearly the spread is greater in  $\eta$  units than in  $\phi$ , due to simple phase space effects.  $\sigma_\phi$  is limited to  $\sim \pm\frac{\pi}{2}$  about the mean whereas  $\sigma_\eta$  may cover the whole acceptance of  $\pm\eta_{range}$ .

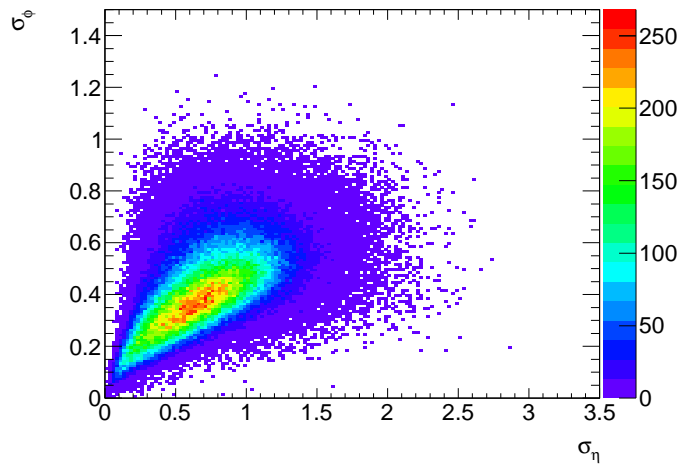


Figure 6.6: Scatter plot illustrating the standard deviations in  $\eta$  and  $\phi$  from a sample of PYTHIA events with  $\hat{p}_T = 140 - 280$  GeV.

### 6.4.5 Step IV - Ellipse Method

The investigations into standard deviation, and energy distributions such as that seen in figure 6.4, reveal that a circular search area is not necessarily

optimal. The relative spreads in  $\eta$  and  $\phi$  are more often different than the same. This information may be effectively used by adopting an elliptic search region.

The length of the two axes of the ellipse are set proportional to the  $\eta$  and  $\phi$  standard deviations,  $\sigma_\eta$  and  $\sigma_\phi$ , found in that semicylinder - hence both *size and shape of the search areas are appropriate to each semicylinder for individual events*. This removes the need to predefine the size of the jet-cone, whilst attempting to minimise the quantity of UE unassociated with the final state partons.

An ellipse is centred on the mean positions in each semicylinder, defined by the final iteration of Step II, and as indicated by figure 6.7. The length of the ellipse axes is defined in equations 6.11 and 6.12, where  $n$  can take a range of values. The default value of  $n$  is 2.0, as in a Gaussian distribution this would include 95% of the energy content. Within the ethos of TRAPS parameters are discouraged;  $n$  is recognised as a parameter, but is merely a scaling factor rather than an absolute, dimensionful parameter. In any case, the dependence on  $n$  is weak within a sensible range, as demonstrated in section 7.3.

$$axis_\eta = n\sigma_\eta \tag{6.11}$$

$$axis_\phi = n\sigma_\phi \tag{6.12}$$

The 3-momenta of particles found within the ellipses are vector summed, including any particles which land outside of the previously defined boundaries. The maximum limits are taken as the edge of the ATLAS calorimetry at  $\eta = \pm 4.9$ , and the boundary of the semicylinder in  $\phi$ . Since the previous limit,  $\eta_{range}$ , is well inside this boundary at the hard scales considered, good efficiency is expected, even close to the boundary. Scalar  $p_T$  is also summed

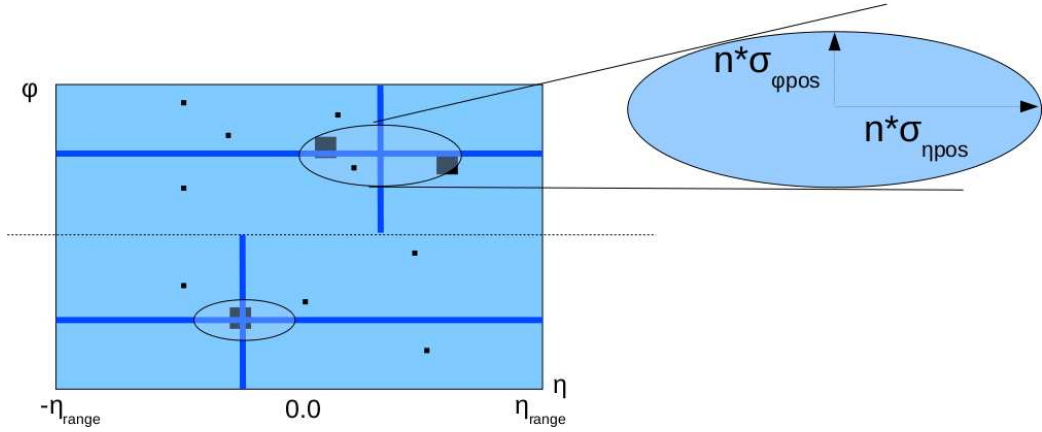


Figure 6.7: Diagram of the two TRAPS-object ellipses placed on the event, with a close-up example of an ellipse and its dimensions.

within the same region, for later use.

Using the summed momenta in each ellipse, two new  $\eta - \phi$  values are calculated. The ellipse is repositioned as shown in figure 6.8 although maintains the same dimensions, and this process is iterated until stable in the manner of several other cone-type jet algorithms. Each ellipse is iterated individually while the expression given in equation 6.13 is true and the number of iterations is less than 25.

$$\Delta R^2 < \Delta R_{old}^2 \quad (6.13)$$

where  $\Delta R$  is the shift in position from the previous centre point, as given by equation 6.14

$$\Delta R^2 = (\eta - \eta_{old})^2 + (\phi - \phi_{old})^2 \quad (6.14)$$

Typically for a  $\hat{p}_T$  range of 140 - 280 GeV the number of iterations is 5, but for harder scatters at 560 - 1120 GeV fewer iterations are typical, as shown in figure 6.9. This iterative process is curtailed after 24 iterations, but

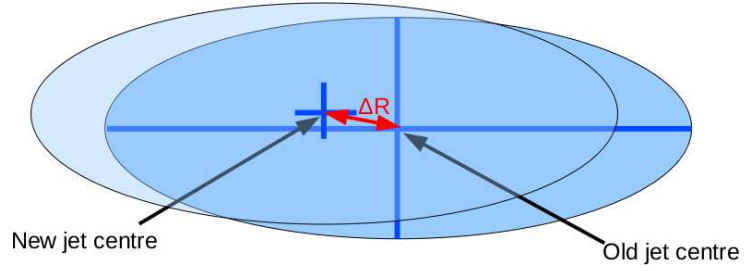


Figure 6.8: Diagram to illustrate the movement of the ellipse during the iterative process.

at generator-level this limit is rarely reached.

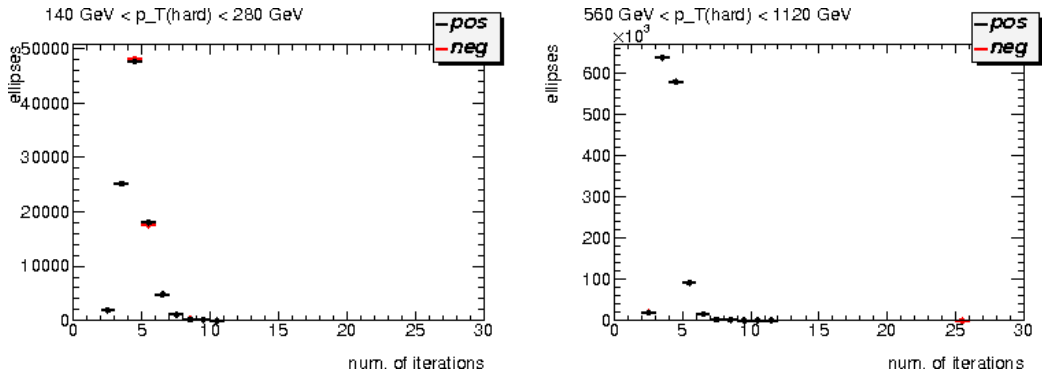


Figure 6.9: Histogram of the number of iterations taken to converge for a large number of PYTHIA events in two  $p_T$  ranges.

Once both ellipses have converged, the final  $\eta$  and  $\phi$  calculated define the absolute position of the TRAPS-objects. The last summations of  $p_x$  and  $p_y$  are used to calculate a starting value for the TRAPS-object  $p_T$ , and the constituent hadrons or topoclusters are saved to a vector.

In the case when no energy is found in an ellipse, the algorithm would default to the previously calculated  $\bar{\eta}$  and  $\bar{\phi}$ , and takes the entire bound area at the end of Step II for the  $p_T$  sum. Note that this has never been observed.

### 6.4.6 Step V - Underlying Event Subtraction

For each of the TRAPS-objects, two underlying event areas at  $(\eta_{UE}, \phi_{UE_{1,2}})$  are established in positions away from both objects. The jet underlying event has a small  $\eta$ -dependence, which in turn is dependent on the  $Q^2$  of the event. As far as possible these regions are chosen to be at the same  $|\eta|$  as the TRAPS-object, unless this is compromised by the nearby presence of either TRAPS-object. The positions are determined by the following criteria and the various combinations are illustrated in figures 6.10 and 6.11, in which the ellipses represent the jets and the squares are the underlying event positions. The subscript ‘jet’ is used as short-hand for TRAPS-object where convenient.

**In  $\phi$**

- + and -  $\pi/2$  from each TRAPS-object.

**In  $\eta$**

If both TRAPS-objects have the same sign in  $\eta \dots$

- $\eta_{UE} = -\eta_{jet}$ 
  - Unless  $\eta_{jet}$  is ‘close’ to  $\eta = 0 \dots$

- $\eta_{UE} = -sign(\eta_{jet}) \cdot \eta_{mid}$

If the TRAPS-objects have opposite sign in  $\eta \dots$

- $\eta_{UE} = \eta_{jetMid}$ 
  - Unless  $\eta_{jet}$  is ‘close’ to  $\eta = 0 \dots$

- $\eta_{UE} = -sign(\eta_{jet}) \cdot \eta_{mid}$

where  $\eta_{mid}$  is the halfway point between 0 and  $\eta_{range}$ ;  $\eta_{jetMid}$  is the halfway point between the two TRAPS-objects; and ‘close’ is defined as  $|\eta_{jet}|$  less than the ellipse  $\eta$ -axis. All of these are thus topology dependent, but otherwise insensitive parameters. Under this scheme, a large proportion of TRAPS-objects have their UE assessed at their oppositely signed  $\eta_{mid}$ . However the  $p_T$  summed per UE region has a similar distribution whichever positional option is selected.

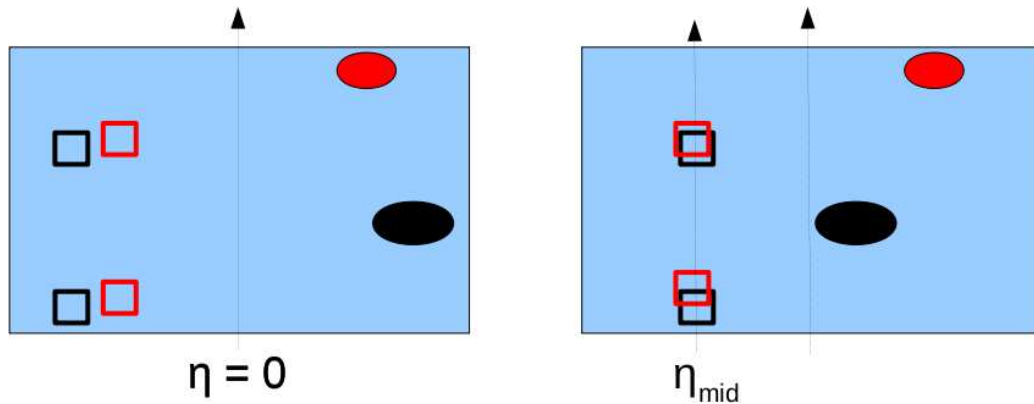


Figure 6.10: Diagram illustrating UE assessment positions for when the TRAPS-objects have the same  $\eta$  sign.

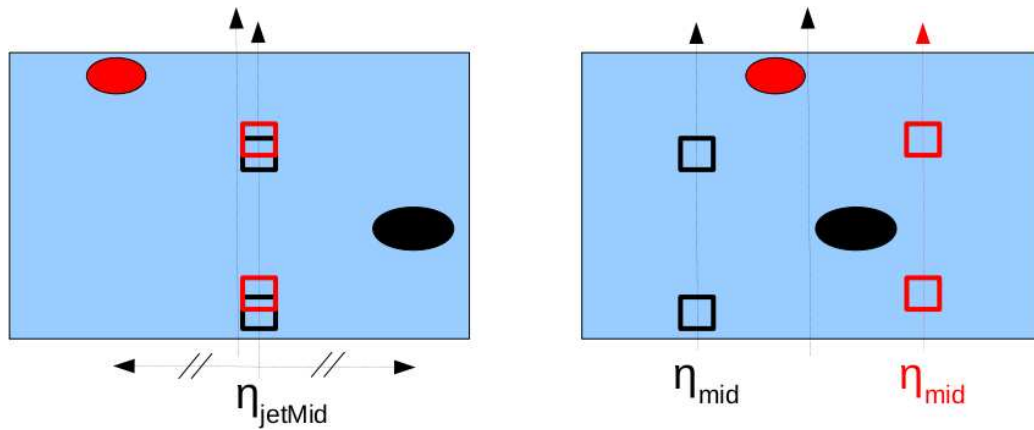


Figure 6.11: Diagram illustrating UE assessment positions for when the TRAPS-objects have opposite  $\eta$  sign.



The size of the UE regions is given a working value of a unit square in  $\eta-\phi$  space. This area was decided upon following a quick test that established that in a wide range about this value the mean  $\sum p_T$  per unit area was independent of this value. Table 6.1 shows the mean value of the UE  $p_T$  per unit area for a range of sizes. A more detailed study, taking into account the variability in the UE  $p_T$  assessment, can be found in the following chapter.

UE region size	Mean $p_T$ /unit area, GeV
0.6 x 0.6	2.56
1.0 x 1.0	2.58
1.2 x 1.2	2.57
1.5 x 1.5	2.55

Table 6.1: Table of mean UE scalar  $p_T$  summed, per unit area, found in UE regions of different sizes in the  $\hat{p}_{T_{min}} = 100$  GeV events.

The mean average energy density in the associated two regions is multiplied by the TRAPS-object ellipse area and subtracted from the TRAPS-object 4-momentum.

### 6.4.7 Step VI - Virtuality Correction

A final state on mass shell parton would be massless, but in order to closely mimic infinitely higher orders some MC give only the Sudakov-modified off mass shell value. TRAPS-objects intrinsically have mass. To force the TRAPS-object to represent the MC parton in a massless way,  $m^2 = E^2 - p^2$  must be equal to zero. This is done by making  $E$  and  $p$  the same value<sup>1</sup>.

---

<sup>1</sup>This is the reverse of a process adopted by some MCs when allocating a mass to a matrix-element generated parton.

Given perfect measurement of the final state hadrons, the actual parton momentum should lie somewhere between the scalar and vector momentum quantities. Somewhat arbitrarily, an average is taken of the two, in an attempt to correct for this ‘virtual’ mass adopted by the TRAPS-object. The previously calculated scalar and vector  $p_T$  sums are taken as  $E_T$  and  $p_T$  respectively.

This is typically a small mean average correction of the order of  $\sim 2\%$  to the TRAPS-object momentum. Further analysis is not justified at this stage. This completes the algorithm description.

### 6.4.8 Visual Summary

A flowchart summarising the steps of the TRAPS algorithm, as detailed in the preceding section, can be seen in figure 6.12.

## 6.5 Running TRAPS

TRAPS is a stand-alone algorithm written in C++. To date it has been run in conjunction with compiled ROOT code and within the ATLAS Athena software (see section 7.1 for information). The only compulsory inputs are a vector of TLorentz vectors (4-vectors). Other, variable, inputs are  $p_{beam}$ , the beam energy;  $\hat{p}_{T_{min}}$ , the minimum transferred momentum of the sample that is required for this run; the  $n$ -parameter, as described in section 7.3; and  $\phi$  half side and  $\eta$  half side, the half side dimensions of the UE regions.

Table 6.2 specifies the default settings for these parameters, for the case when only 4-vectors are input. Naturally  $p_{beam}$  and  $\hat{p}_{T_{min}}$  must be changed appropriately in order to set correctly the initial pseudorapidity range,  $\eta_{range}$ . Altering the other variables is for future algorithm experimentation only.

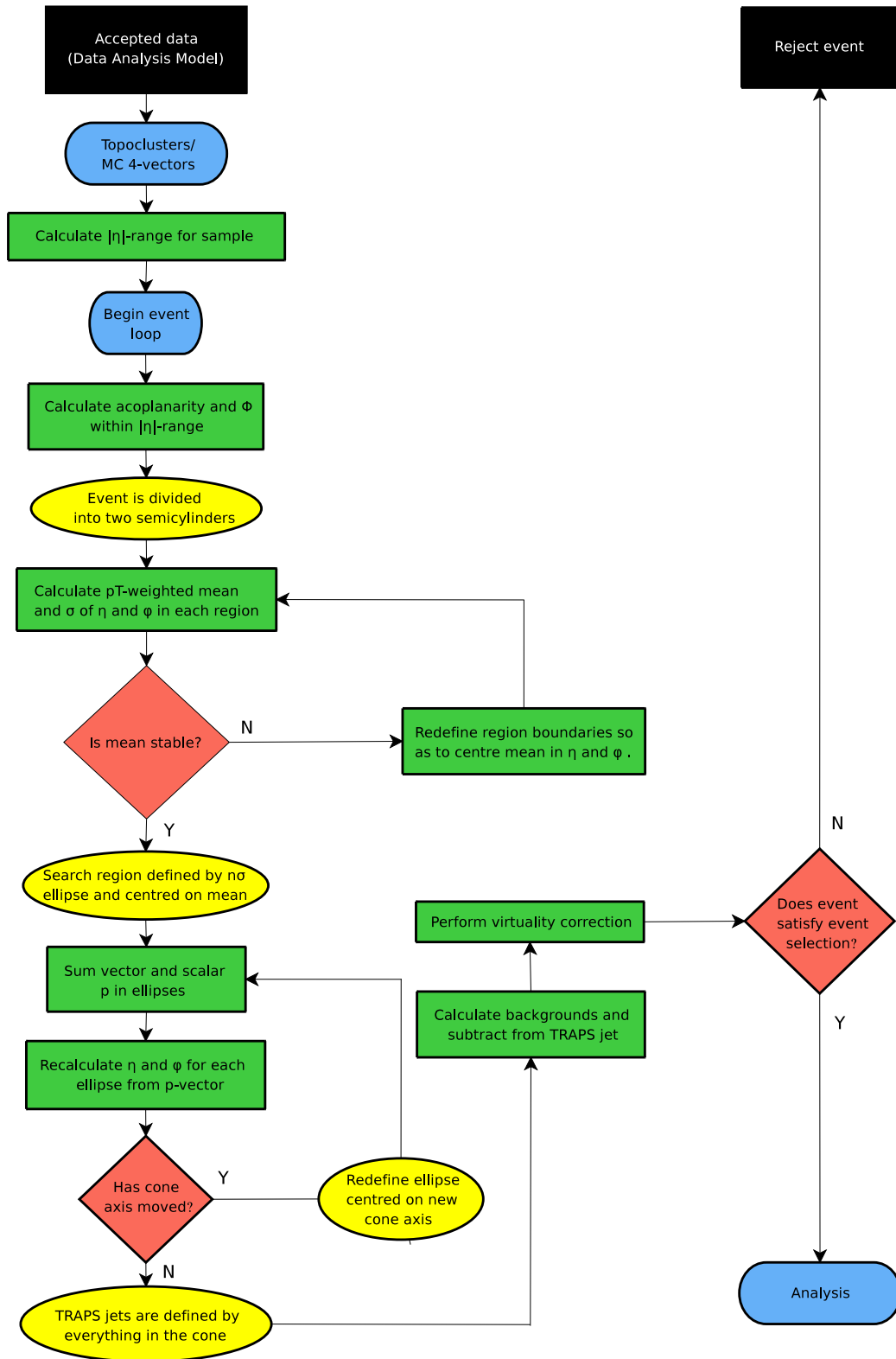


Figure 6.12: Flowchart of TRAPS steps.

Parameter	Default setting	Unit
$p_{beam}$	3500	GeV
$\hat{p}_{T_{min}}$	100	GeV
$n$	2.0	-
$\phi_{halfSide}$	0.5	rad
$\eta_{halfSide}$	0.5	-
virtuality correction	$E' = \frac{p+E}{2}, p' = \frac{p-E}{2}$	-

Table 6.2: List of TRAPS input parameters.

### 6.5.1 Event Flags

A selection of event ‘flags’ to indicate particular event or TRAPS-object properties is provided with descriptions in table 6.3. No events are excluded automatically by TRAPS, merely flagged for user benefit.

Flag	Description
flagAco	large acoplanarity, $A > 0.4$
flagEtaJet	$ \eta_{jet}  > 4.0$
flagNegPT	$p_{T_{jet}}$ (after UE sub) $< 0$
flagLowPT	$p_{T_{jet}} < p_{T_{min}}$

Table 6.3: List of TRAPS event flags.

A TRAPS-object may have a negative  $p_T$  when, by chance, the UE region falls upon an energy-dense ISR, or other deposit whether UE or signal, in the latter case assuming a failure of the algorithm to properly locate the scattered parton. If this is coupled with a particularly large TRAPS-object search ellipse, then the UE  $p_T$  scaled by the ellipse area becomes larger than

the TRAPS-object  $p_T$ .

The flagEtaJet indicator is set if the TRAPS-object pseudorapidity is greater than 4.0. At large  $|\eta|$  a TRAPS-object is in danger of ‘leaking’ some of its energy outside of the calorimeter limit,  $|\eta| = 4.9$ , and hence a confinement of  $|\eta_{jet}|$  is considered the safest option. The overwhelming proportion of partons *do* land within the kinematic limits, as defined in section 6.4.1, but a radiated particle with an energy approaching the hard scale of the interaction, may occasionally cause partons to have a pseudorapidity outside of this range. TRAPS will often mis-represent the parton in this case, and cannot know if the ISR boost caused it to escape down the beampipe. Hence, this flag appears extremely rarely at the momentum transfers analysed, as TRAPS is designed to search broadly within  $\eta_{range}$ , and only collect additional energy from outside of this range if necessary.

### 6.5.2 Primary TRAPS outputs

A list of associated TRAPS member functions is given in table 6.4. Energy units in TRAPS are output in MeV and  $\phi$  is given in radians in the range  $[-\pi, \pi]$ .

## 6.6 Development Tests

Working within the pragmatic approach, as described in section 6.2.1, comparisons are made in Monte Carlo between the TRAPS-object and the parton that is presumed to have created it, of the two coming from the hard interaction. In this manner the success of TRAPS was assessed during development.

To decide which parton/TRAPS-object pair should be compared, a simple  $\Delta R$  calculation is made - the parton/TRAPS-object with the smallest  $\Delta R$

Member name	Brief description
Event info	
$\min_{\max}$ EtaDetectorLimit()	$\eta$ -limits of calorimeter detection
$\min_{\max}$ EtaLimit()	$\eta$ -limits of calculated phase space for sample
eventPlane()	Azimuthal angle of the event plane ( $\Phi$ )
acoplanarity()	Acoplanarity of the event
floatNItr()	# iterations to stabilise floating boundaries
$\text{pos}_{\text{neg}}$ JetNItr()	# iterations per TRAPS-object, ellipse method
TRAPS-object output	
$\text{pos}_{\text{neg}}$ JetConstituents()	Vector of particles designated as TRAPS-object
$\text{pos}_{\text{neg}}$ JetEta()	$\eta$ of TRAPS-object
$\text{pos}_{\text{neg}}$ JetPhi()	$\phi$ of TRAPS-object
$\text{pos}_{\text{neg}}$ JetEt()	$E_T$ of TRAPS-object
$\text{pos}_{\text{neg}}$ JetPt()	$p_T$ of TRAPS-object
$\text{pos}_{\text{neg}}$ JetScalarPt()	Scalar summed $p_T$ of TRAPS-object before UE sub
$\text{pos}_{\text{neg}}$ JetVectorPt()	Vector summed $p_T$ of TRAPS-object before UE sub
UE	
$\text{pos}_{\text{neg}}$ BgEta()	$\eta$ of UE search region
$\text{pos}_{\text{neg}}$ BgPhi <sub>2</sub> <sup>1</sup> ()	$\phi$ of UE search region
$\text{pos}_{\text{neg}}$ BgPt()	$p_T$ used in UE subtraction
TRAPS-object internal workings	
$\text{pos}_{\text{neg}}$ SemicylinderPt()	Scalar sum $p_T$ in entire semicylinder before UE sub
sdEta $_{\text{neg}}^{\text{pos}}$ ()	$\eta$ standard deviation within final floating boundaries
sdPhi $_{\text{neg}}^{\text{pos}}$ ()	$\phi$ standard deviation within final floating boundaries
eta $_{\text{neg}}^{\text{pos}}$ Axis()	Length of the $\eta$ axis of TRAPS-object search ellipse
phi $_{\text{neg}}^{\text{pos}}$ Axis()	Length of the $\phi$ axis of TRAPS-object search ellipse
areaEllipse $_{\text{neg}}^{\text{pos}}$ ()	Area of TRAPS-object search ellipse

Table 6.4: List of member functions of the TRAPS class. ‘pos’ and ‘neg’ of TRAPS-objects refer to the positive and negative semicylinders, as defined in section 6.4.2. UE is subtracted unless specified.

separation become one associated pair and the other combination become the second pair. Kinematic variables can then be evaluated for algorithmic method (separate from any experimental) resolution.

The most fundamental variables, i.e.  $\phi$ ,  $\eta$  and  $p_T$  were used throughout the development process in order to determine when improvements were made and to remove sub-algorithms that had no significant effect.

Histograms demonstrating the differences between a TRAPS-object and a parton over a large number of events, are termed the *delta plots*. The regularly plotted variables are defined in equations 6.15 - 6.19, in which the word ‘jet’ is used as short-hand for ‘TRAPS-object’.

$$\Delta\phi = \phi_{parton} - \phi_{jet} \quad (6.15)$$

$$\Delta\eta = \eta_{parton} - \eta_{jet} \quad (6.16)$$

$$\Delta\eta_{signed} = \frac{\eta_{parton}}{|\eta_{parton}|} (\eta_{parton} - \eta_{jet}) \quad (6.17)$$

$$\frac{\Delta p_T}{p_T} = \frac{p_{T_{parton}} - p_{T_{jet}}}{p_{T_{parton}}} \quad (6.18)$$

$$A_{p_T} = \frac{p_{T_{parton}} - p_{T_{jet}}}{p_{T_{parton}} + p_{T_{jet}}} \quad (6.19)$$

$\Delta\eta_{signed}$  is useful as it not only reveals the parton/TRAPS-object absolute  $\eta$ -difference, but also any bias toward central or beam pseudorapidity values by TRAPS.  $A_{p_T}$  has the benefit of including all entries within the range  $[-1, 1]$ , unlike  $\Delta p_T/p_T$ . Fine scale resolution is best examined with  $\Delta p_T/p_T$  whereas the far outlying results are easier to analyse using  $A_{p_T}$ .

### 6.6.1 Analysing the Delta Plots

Normalised examples of the delta plots for the most recent version of the algorithm running over the 560-1120 GeV event sample with  $n = 2$  can

be found in figure 6.13. Both linear and log scale plots are examined as an indicator of the central resolution and to look for outlying structure, at lower frequency. Symmetric distributions with a mean of zero, within errors, demonstrate that the algorithm is not biased in one way or another. Note that a mode average persistently around zero, is also a strong endorsement of the algorithm's lack of bias.

The mode, mean, RMS and Full-Width-Half-Max/2.35 (FWHM/2.35) statistics associated with each of the delta plots are observed, the latter two giving indication of the distribution width. FWHM/2.35 is calculated by an interpolation over central bins, and divided by 2.35 as this gives an equivalent to the RMS, assuming the distribution to be a Gaussian. However, the large discrepancies between them, indicate that this is not the case. The mean and RMS are calculated by the standard method in ROOT, with the error on the mean being  $RMS/\sqrt{N}$  where  $N$  is the number of entries. The mode and its error are calculated using a three-bin linear fit, as described by equations 6.20 and 6.21. This uses the height of the tallest bin ( $A$ ), and the adjacent bins ( $B$  and  $C$ ), where  $B$  has more entries than  $C$ .

$$Mode = La + \frac{w}{2} \left( 1 - \left( \frac{B}{A} \right) \left( \frac{B-C}{B+C} \right) \right) \quad (6.20)$$

$$Error_{Mode} = \frac{w}{2} \left( 1 - \left| \left( \frac{B}{A} \right) \left( \frac{B-C}{B+C} \right) \right| \right) \quad (6.21)$$

$La$  is the value of the variable at the leading edge of  $A$  and  $w$  is the bin width. On occasion, the histogram binning was made coarser before these equations were applied, as the extremely fine binning required for such variables as  $\Delta\Phi$  measuring the highest- $p_T$  sample, leads to statistical fluctuations at the peak of the rather flatter low- $p_T$  distribution.

With respect to the width assessments, the smallest possible values are



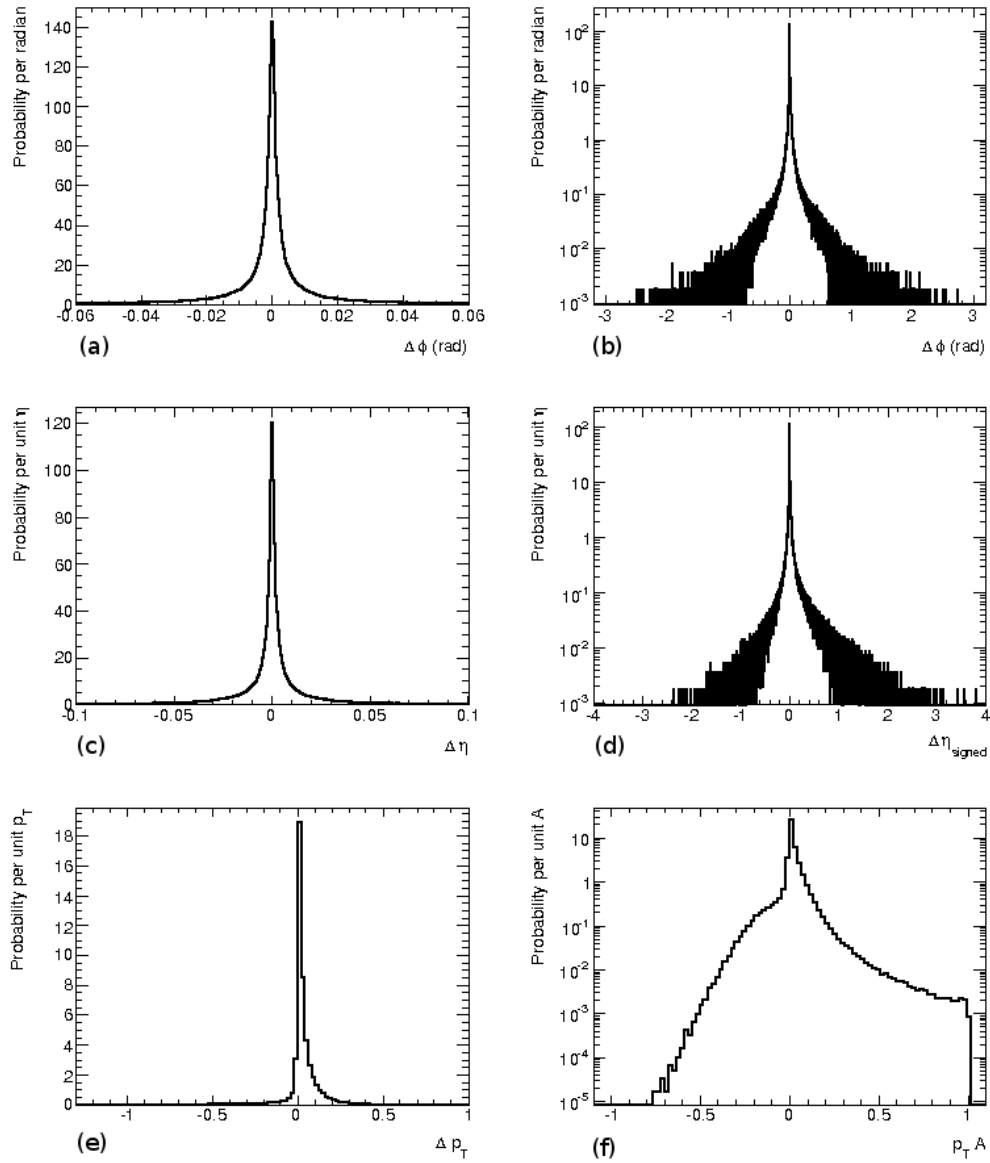


Figure 6.13: Example of normalised delta plots for 560-1120 GeV PYTHIA MC. Histograms (a) and (b) show  $\Delta\Phi$  resolution on linear and log scales. Histograms (c) and (d) show  $\Delta\eta$  and  $\Delta\eta_{signed}$  respectively. Plots (e) and (f) are  $\Delta p_T/p_T$  resolution and  $A_{p_T}$  respectively. Related statistics are shown in table 6.5.

desired, in which the vast majority of parton/TRAPS-object pairs occupy the central bins of the delta plots. However, the RMS is strongly skewed by the far outliers which represent the ‘very wrong’ results. FWHM/2.35 is not sensitive to the wings of the distribution, but gives a clearer indication of the spread of the majority.

Example statistical values, extracted from the plots in figure 6.13 are given in table 6.5. Here a positive bias is seen in the mean values of  $\eta$  and the  $p_T$  variables, while the mode is zero within errors. Full analysis and results including other momentum transfer samples can be found in chapter 7.

Histogram	Mode	Mean	RMS	FWHM/2.35
$\Delta\phi$	$(-9.7 \pm 10.0) \times 10^{-5}$	$(4 \pm 7) \times 10^{-5}$	0.11	0.001
$\Delta\eta_{signed}$	$(-9.9 \pm 10.0) \times 10^{-5}$	$(92 \pm 0.8) \times 10^{-4}$	0.13	0.001
$\Delta p_T/p_T$	$(1.2 \pm 0.9) \times 10^{-2}$	$(239 \pm 0.8) \times 10^{-4}$	0.13	0.013
$A_{p_T}$	$(1.1 \pm 1.0) \times 10^{-2}$	$(163 \pm 0.4) \times 10^{-4}$	0.07	0.012

Table 6.5: Table of resolution statistics relating to example delta plots in figure 6.13.

The general case of much smaller FWHM/2.35 than RMS values indicate a large proportion of partons are measured *very* well, whilst the RMS is strongly affected by those few that are measured *very* badly.

## 6.7 Alternative TRAPS Sub-methods



In the course of developing the TRAPS algorithm there were several dead-ends, variations and alternative sub-methods which were tried and discarded

for various reasons. Here are described the most significant.

### 6.7.1 No Floating Boundary

In the earliest incarnation of the algorithm, there was no iterative method with movable boundaries, as described in Step II. Instead, the algorithm would simply calculate the  $p_T$ -weighted means, within the initial  $\eta - \phi$  limits, i.e.  $(|\eta_{range}|, \Phi \pm \frac{\pi}{2}$  or  $(\Phi - \pi) \pm \frac{\pi}{2})$ , and then move directly onto Step III. However, it was decided that the improvements seen in the delta plot statistics for the lower momentum range when using the floating boundary iterative method, as seen in figure 6.14, outweigh a slight increase in computing time and algorithm complexity.

Figure 6.14 is colour coded for visual simplicity. In this scheme, green indicates a preferred value over the alternative method and blue indicates a similar value, within  $\pm 10\%$  of the other, and hence not a significant difference.

Pythia, n = 2.0		smaller number
		similar number

<b>J4 = 140-280 GeV</b>				<b>J4 = 140-280 GeV</b>			
<b>With floating boundary</b>		Entries = 2740378		<b>Without floating boundary</b>		Entries = 2738648	
	mean	RMS root	FWHM/2.35		mean	RMS root	FWHM/2.35
Delta Phi	0.0002	0.194	0.0066	Delta Phi	0.0002	0.192	0.0080
Delta Eta (signed)	0.0200	0.381	0.0091	Delta Eta (signed)	-0.0052	0.411	0.0149
Delta pT	-0.0349	0.243	0.0325	Delta pT	-0.0489	0.260	0.0347
A pT	-0.0055	0.111	0.0189	A pT	-0.0102	0.122	0.0238

<b>J6 = 560-1120 GeV</b>				<b>J6 = 560-1120 GeV</b>			
<b>With floating boundary</b>		Entries = 2681422		<b>Without floating boundary</b>		Entries = 2684968	
	mean	RMS root	FWHM/2.35		mean	RMS root	FWHM/2.35
Delta Phi	0.0000	0.108	0.0012	Delta Phi	0.0000	0.107	0.0012
Delta Eta (signed)	0.0092	0.129	0.0013	Delta Eta (signed)	0.0087	0.131	0.0016
Delta pT	0.0239	0.132	0.0128	Delta pT	0.0141	0.136	0.0120
A pT	0.0163	0.067	0.0115	A pT	0.0115	0.069	0.0116

Figure 6.14: Spreadsheet showing the mean and spread of delta plots for two momentum transfer ranges, when using (left) or excluding (right) the floating boundary method.

### 6.7.2 4-jet Sink

This sub-method would have taken the place of step III. The premise is to initialise the position of four ‘jet sinks’ - one on each of the mean values found in step II and one on the beam line in each direction. All of the inputs are then assigned to one of the sinks, depending on its proximity in  $\eta - \phi$  space. From these constituents, the four sink positions are re-calculated using vector summation. The particles are then released from their sink-assignment and the process iterates until the sink positions stabilise. The particles in the beamline sinks are discarded and those in the other two make up the TRAPS-objects.

Although this method produces good results, comparable with the ellipse method in terms of parton reconstruction, it was abandoned. This was because an associated UE subtraction method was not obvious since the TRAPS-objects have no distinct or quantifiable area. In addition, the arbitrary parameter used to define which object to associate each particle would be  $p_T$  dependent.

# Chapter 7

## Using TRAPS

### 7.1 Adaption for Athena

To access ATLAS data of certain types, it is beneficial to work within the Athena framework, which is mentioned in chapter 2. In the context of TRAPS there are three key reasons. Firstly, it allows usage of the official production MCs, for which there are several different generators available. Secondly, it is possible to work with MC that has been reconstructed, using the same software code applied to real data, from simulation of the response of the ATLAS detector. And thirdly, it permits access to ATLAS collision data in various formats.

Jet-finding in ATLAS is generally undertaken within the Athena framework on ESD (Event Summary Data) files, which contain information of all calorimeter cells. TRAPS in Athena may be run on RAW or (Derived)ESD files, both of which contain the calorimeter cell data.

The structure of TRAPS in Athena is based on the structure of other jet-finders utilised by the ATLAS collaboration. The software consists of

five packages, which are available in the CERN subversion repository, SVN<sup>1</sup>. The standard *JetRec* reconstruction package, which is responsible for calling all jet-finders used in ATLAS, is modified to allow TRAPS jet-finding using the same inputs, at either truth or reconstructed level, as other jet-finders. It supplies these inputs as a set of FASTJETS [61]. A FASTJET has a wide definition from a fully-formed jet, down to a single particle, calorimeter cell or topocluster<sup>2</sup>. Topoclusters are the input used by TRAPS in Athena.

The FASTJETS are passed to *TRAPS Athena*, which converts them into TLorentz vectors. TLorentzVector is a generic ROOT [36] class, which describes a four-vector either in terms of position and time  $(x, y, z, t)$  or momentum and energy  $(p_x, p_y, p_z, E)$ . These vectors are then fed to the *TRAPS Algorithm* itself, which carries out the steps described in chapter 6. The *TRAPS Algorithm* package is a standalone C++ class that is independent of the Athena framework, but requiring TLorentz vectors as input.

The data members of the TRAPS class are passed through *TRAPS Event*, and then saved to StoreGate<sup>3</sup> [64] as *TRAPS Event* POOL objects. This holds them for retrieval by other algorithms. *TRAPS Analysis* is a user analysis package that is called after all reconstruction is completed. The TRAPS-objects are accessed in much the same way as any other jet-finding results, by retrieving the *TRAPS Event* object from StoreGate. All analysis, event selection and histogram filling are completed within this package, and the output is written to a ROOT file. This progression of information is illustrated by the block diagram in figure 7.1.

---

<sup>1</sup>Path: Institutes/QMUL

<sup>2</sup>Energy contained in a local collection of calorimeter cells.

<sup>3</sup>A toolkit for client algorithms to interact with the data store to record or retrieve C++ data objects.

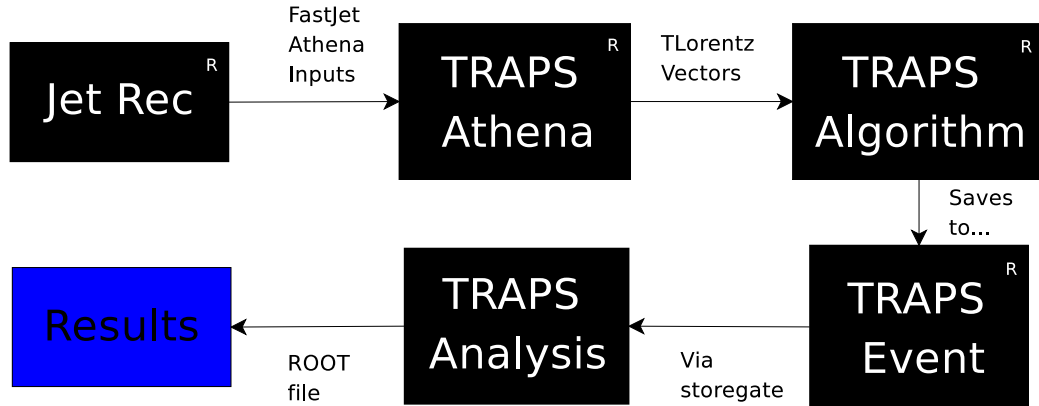


Figure 7.1: Block diagram showing simply the sequence of packages for TRAPS in Athena. Those labelled ‘R’ are part of the reconstruction.

### 7.1.1 Official Production MC

There exist various official production MC samples, which are produced centrally by ATLAS using several different generators. TRAPS is tested on several high momentum transfer ‘JX’ (jet) samples which consist of standard QCD events, as were previously generated by Rivet and described in chapter 6. The samples used in this chapter are the 2009 set, with a CMS energy of 7 TeV and are at generator level only. They all use the PYTHIA or HERWIG ‘ATLAS tune’ depending on MC for their steering parameters, including the MRST LO\* PDF set<sup>4</sup> [65].

Each sample has a fixed  $\hat{p}_T$  range, and those used in this thesis, covering more than an order of magnitude, are defined as:

- J3 : 70 - 140 GeV
- J4 : 140 - 280 GeV

---

<sup>4</sup>A PDF set to be used with leading order MC generators, tuned to better represent higher orders and ATLAS data.

- J5 : 280 - 560 GeV
- J6 : 560 - 1120 GeV

Both PYTHIA and HERWIG are used here, and ALPGEN with Jimmy is also available. The number of events and the mean cross-section of each sample are given in table 7.1. Sometimes the number of events taken in a particular analysis will differ slightly from these figures, as a small fraction of events may be lost to failed computing jobs. This is accounted for in normalisations. The HERWIG cross-sections are taken from the 2010 MC productions, as the 2009 values are unavailable, but are assumed to be similar.

Energy range	PYTHIA		HERWIG	
	#events	$\sigma$ [nb]	#events	$\sigma$ [nb]
J3	1397430	2.1936E+03	399643	2.3576E+03
J4	1397401	8.7704E+01	399091	9.4236E+01
J5	1391612	2.3498E+00	399238	2.5813E+00
J6	1347654	3.3615E-02	394376	3.9439E-02

Table 7.1: Number of events and mean cross-sections of ATLAS official production PYTHIA and HERWIG samples, as quoted from [66].

### 7.1.2 Comparison with Rivet

It is essential that TRAPS installed in Athena produces results that are commensurate with those produced by the compiled ROOT code used during development. Hence, new PYTHIA MC samples were generated via Rivet to



represent the JX samples, using identical job options, although without access to the same PDF set. Instead the PYTHIA default was used. The delta plots showed the results to be comparable within errors, allowing for the different PDFs. This confirms that the installation in Athena was successful.

## 7.2 Algorithm Resolution

The ‘algorithm resolution’ is determined in MC relative to the ‘truth’ input - in this case the final state partons, as described in chapter 6. This resolution is unrelated to the detector resolution, which is discussed briefly at the end of chapter 8. The  $n$ -parameter, which was described in section 6.4.5 and will be studied more fully in section 7.3 is set to 2.0 and the length of the underlying event square side is 1.0.

Figure 7.2 is a set of key delta plots for each of the PYTHIA JX samples specified in section 6.6. Figure 7.3 shows the same plots, but on a log scale. As expected, the best algorithm resolution is achieved for the highest momentum transfer event sample (J6), where the fragmented partons are highly collimated and background levels are suppressed.

To give an idea of relative quantities contained within the limits of the  $\Delta\phi$  and  $\Delta\eta$  plots in figures 7.2 and 7.3, the approximate percentages of the total are as follows: J6  $\sim$  90 %, J5  $\sim$  80 %, J4  $\sim$  70 %, J3  $\sim$  57 % ( $\Delta\phi$ ) and  $\sim$  44 % ( $\Delta\eta$ ).

Tables 7.2 - 7.5 detail the mode, mean, RMS and FWHM/2.35 information for each histogram, in the format described in chapter 6. As expected the spread of measurements widens with decreasing momentum transfer, although the modes remain largely at zero within errors.

The mean averages of  $\Delta\phi$  are within errors of zero - it is clear there is no

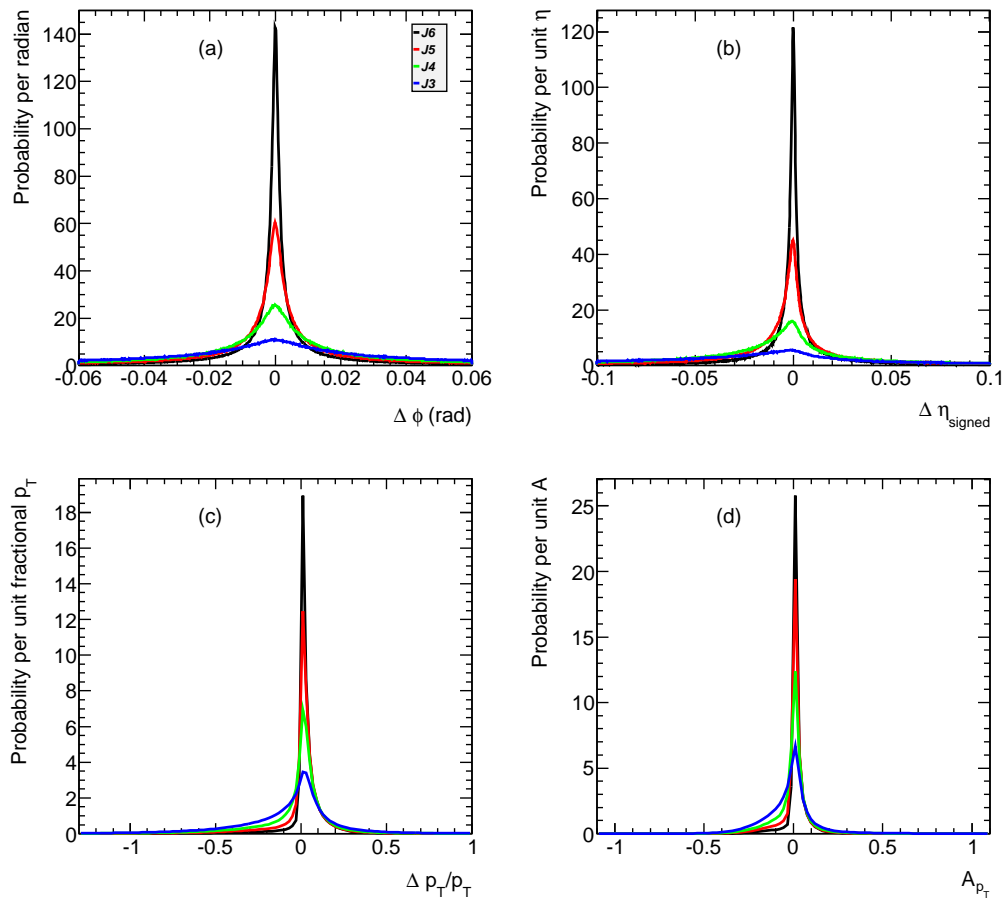


Figure 7.2: Delta plots comparing four PYTHIA JX samples. (a)  $\Delta\phi$ , (b)  $\Delta\eta_{signed}$ , (c)  $\Delta p_T/p_T$ , (d)  $A_{p_T}$ . Note the very different scales on the abscissae.

bias here, and the resolution is high. For example, a  $\text{FWHM}/2.35$  of 0.002 represents only  $0.1^\circ$  variation in the azimuth. Of course, more significant errors are expected from the calorimeter resolution.

The means of the  $\Delta\eta_{signed}$  distributions are preferentially positive in value, indicating a slight bias toward the central region in  $\eta$ . These biases exist at a level of 0.02  $\eta$ -units or less in the assessments made here.

Figure 7.4 investigates the possible cause of the asymmetry. Since a loss of FSR should be symmetric in  $\eta$ , an asymmetric error in  $\eta$  is most likely

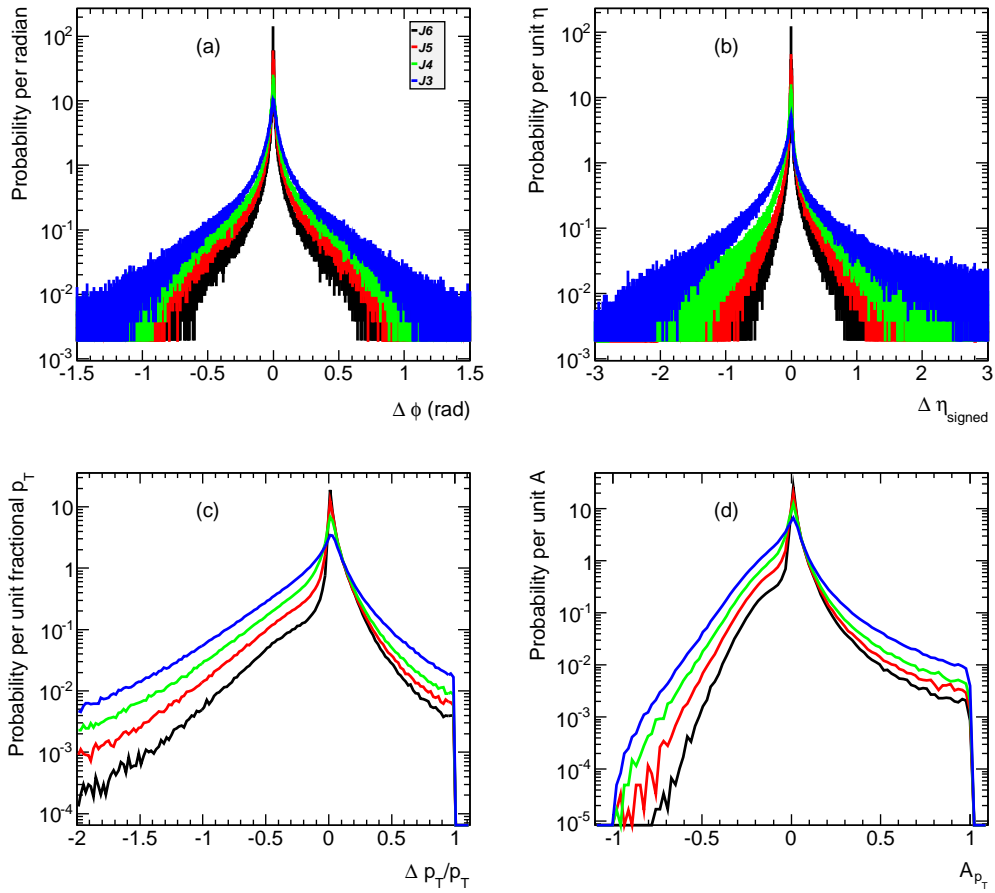


Figure 7.3: Log scale delta plots comparing four PYTHIA JX samples. (a)  $\Delta \phi$ , (b)  $\Delta \eta_{\text{signed}}$ , (c)  $\Delta p_T/p_T$ , (d)  $A_{p_T}$ .

to be caused by effects of ISR. ISRs are normally much smaller in magnitude than the parton and would be expected to be found mainly towards the beam side of a given parton. Partial inclusion would cause an average small ‘drag’ of the parton’s position towards the beam, as can be seen by the slight bulges in the dense distribution of partons in the central  $\eta$  region. Occasionally an energetic recoil ISR is mistaken for the parton. The largest phase space for this is provided by ISR or other underlying event deposit towards the opposite end of the apparatus causing a much larger  $\eta$  error of opposite sign

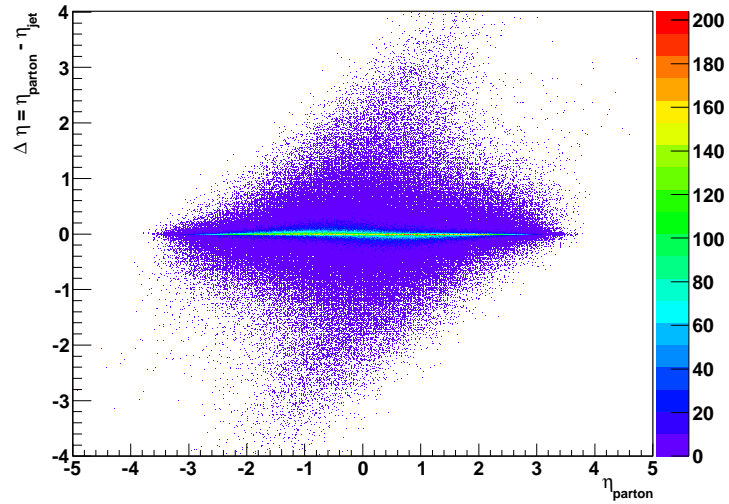


Figure 7.4:  $\Delta\eta$  as a function of  $\eta_{parton}$  for the J4 PYTHIA MC sample.

to the first effect. This feature can also be seen in figure 7.4, in the left-right asymmetry of entries with large  $|\Delta\eta|$ . The relative influence of the two effects causes the asymmetry in  $\Delta\eta_{signed}$ .

Although the  $p_T$  modes are approximately zero within errors, the means are not, although they remain small, within a few percent of zero except for the lowest  $p_T$  range at 70-140 GeV. The  $p_T$  resolution is certainly worse than in the positional measurement, although the figures cannot be compared directly. Whereas an RMS of 0.2 in  $\Delta\phi$  represents  $11^\circ$ , the equivalent value is a 20% mis-measurement in  $A_{p_T}$ . As mentioned in the previous chapter, an order of magnitude difference between RMS and FWHM/2.35 values demonstrate that the distribution is very much non-Gaussian. In fact the majority of events are measured well, as the FWHM/2.35 reveals - to within a few percent even for the J3 sample. However, the much larger RMS suggest that a small proportion of events are measured *very* badly, mainly due to inclusion or subtraction of large ISR, and this must be taken into account when

assessing errors.

Due to the nature of the distribution, a single algorithm resolution cannot be quoted, and hence both RMS and FWHM/2.35 are given throughout.

### 7.3 The $n$ -parameter

The  $n$ -parameter is the main parameter used by TRAPS, although it is in fact merely a scaling factor and not an absolute radius, whilst still permitting UE subtraction. This proportional constant for determining the length of the ellipse axes is set to a default of 2.0, but may vary widely without significant consequence to the resulting MC comparisons. This demonstrates that TRAPS is largely insensitive to  $n$ . However, as a guide, a Gaussian distribution would contain 95% of the jet on average inside  $\pm 2\sigma$ , which would be within typical jet energy scale errors.

Figure 7.5 shows the mode, mean, RMS and FWHM/2.35 of the delta plots (described in section 6.6) for a range of  $n$ . Results from both J4 and J6 PYTHIA MC samples are shown.

An absolute mean close to zero is ideal, as this indicates no strong bias of the algorithm. Note the preference for a small positive value of  $\bar{\eta}_{signed}$  at low  $p_T$ , indicating a small ‘drag’ effect on the found position of the parton.

The RMS and FWHM/2.35 are two different measures of spread about the mean. The ‘true’ resolution can be thought to lie somewhere between them. It is clear from the extreme change of scale between them, that there are a small number of partons which are measured catastrophically.

The dependence on  $n$  is particularly low for the measurement of  $p_T$  over the range 1.5-3.0, even for the less clean J4 events. This indicates that the UE subtraction is working effectively. The azimuthal variable,  $\phi$  is well

Histogram	Mode	Mean	RMS	FWHM/2.35
$\Delta\phi$	$(4.0 \pm 4.0) \times 10^{-4}$	$(3.5 \pm 1.5) \times 10^{-4}$	0.25	0.015
$\Delta\eta_{signed}$	$(-4.1 \pm 6.0) \times 10^{-4}$	$(3.1 \pm 3.9) \times 10^{-4}$	0.65	0.021
$\Delta p_T/p_T$	$(1.1 \pm 1.0) \times 10^{-2}$	$(-857 \pm 2) \times 10^{-4}$	0.32	0.065
$A_{p_T}$	$(9.6 \pm 10.0) \times 10^{-3}$	$(-218 \pm 1) \times 10^{-4}$	0.15	0.037

Table 7.2: Table of resolution statistics relating to PYTHIA J3 delta plots.

Histogram	Mode	Mean	RMS	FWHM/2.35
$\Delta\phi$	$(-1.0 \pm 1.0) \times 10^{-4}$	$(1.7 \pm 1.2) \times 10^{-4}$	0.19	0.007
$\Delta\eta_{signed}$	$(-6.0 \pm 2.0) \times 10^{-4}$	$(20.0 \pm 0.2) \times 10^{-3}$	0.38	0.009
$\Delta p_T/p_T$	$(1.2 \pm 1.0) \times 10^{-2}$	$(-35.0 \pm 0.1) \times 10^{-3}$	0.24	0.033
$A_{p_T}$	$(9.8 \pm 10.0) \times 10^{-3}$	$(-55.0 \pm 0.7) \times 10^{-4}$	0.11	0.019

Table 7.3: Table of resolution statistics relating to PYTHIA J4 delta plots.

Histogram	Mode	Mean	RMS	FWHM/2.35
$\Delta\phi$	$(1.0 \pm 1.0) \times 10^{-4}$	$(3 \pm 9) \times 10^{-5}$	0.15	0.003
$\Delta\eta_{signed}$	$(-2.0 \pm 2.0) \times 10^{-4}$	$(16.9 \pm 0.1) \times 10^{-3}$	0.23	0.003
$\Delta p_T/p_T$	$(1.2 \pm 1.0) \times 10^{-2}$	$(5.3 \pm 1.1) \times 10^{-4}$	0.19	0.019
$A_{p_T}$	$(1.0 \pm 1.0) \times 10^{-2}$	$(76.9 \pm 0.5) \times 10^{-4}$	0.09	0.014

Table 7.4: Table of resolution statistics relating to PYTHIA J5 delta plots.

Histogram	Mode	Mean	RMS	FWHM/2.35
$\Delta\phi$	$(-9.7 \pm 10.0) \times 10^{-5}$	$(4 \pm 7) \times 10^{-5}$	0.11	0.001
$\Delta\eta_{signed}$	$(-9.9 \pm 10.0) \times 10^{-5}$	$(92 \pm 0.8) \times 10^{-4}$	0.13	0.001
$\Delta p_T/p_T$	$(1.2 \pm 0.9) \times 10^{-2}$	$(239 \pm 0.8) \times 10^{-4}$	0.13	0.013
$A_{p_T}$	$(1.1 \pm 1.0) \times 10^{-2}$	$(163 \pm 0.4) \times 10^{-4}$	0.07	0.012

Table 7.5: Table of resolution statistics relating to PYTHIA J6 delta plots.

measured over the range of  $n$  and is not of particular concern. However,  $\Delta\eta_{signed}$  resolution has a slight preference towards smaller values of  $n$ .

A value of 2.0 is accepted as the default, although this is considered a compromise between a larger UE subtraction and a loss of jet energy.

## 7.4 UE Sample Region Size

The size of the underlying event region has been set in previous sections to a default of 1.0 units/radians in  $\eta$  and  $\phi$ . In this section, investigations are carried out to assess whether further optimisation can be achieved. Unlike other sections in this chapter, these tests are done using the  $\hat{p}_{T_{min}} = 100, 500$  GeV MC samples, as used in chapter 6.

The distributions of  $p_T$  per unit  $\eta - \phi$  area were measured for 10,000 events. A range of different UE square sizes were applied, between 0.5 and 3.0  $\eta - \phi$  units. A constant mean through a range of differently sized squares implies a stable level of underlying event deposits. Also, a small RMS is desired, as this should ‘smooth out’ wild variations in the amount of UE measured in each event.

Figure 7.6 shows the mean and RMS results for  $\hat{p}_{T_{min}} = 500$  GeV (left) and  $\hat{p}_{T_{min}} = 100$  GeV (right). The  $\hat{p}_{T_{min}} = 500$  GeV numbers have been divided by 5 for direct comparison with the  $\hat{p}_{T_{min}} = 100$  GeV results. The mean values start to show a dramatic rise for a side length greater than 2.5 units for the  $\hat{p}_{T_{min}} = 500$  sample and slightly higher for the  $\hat{p}_{T_{min}} = 100$ . This would suggest the UE region starting to enter the region of the jet and that the size is too large. On the other hand, the spread about the mean, as given by the RMS, shows a decrease with rising side length, as a more consistent value is found, which reduces to a minimum value and then rises again as

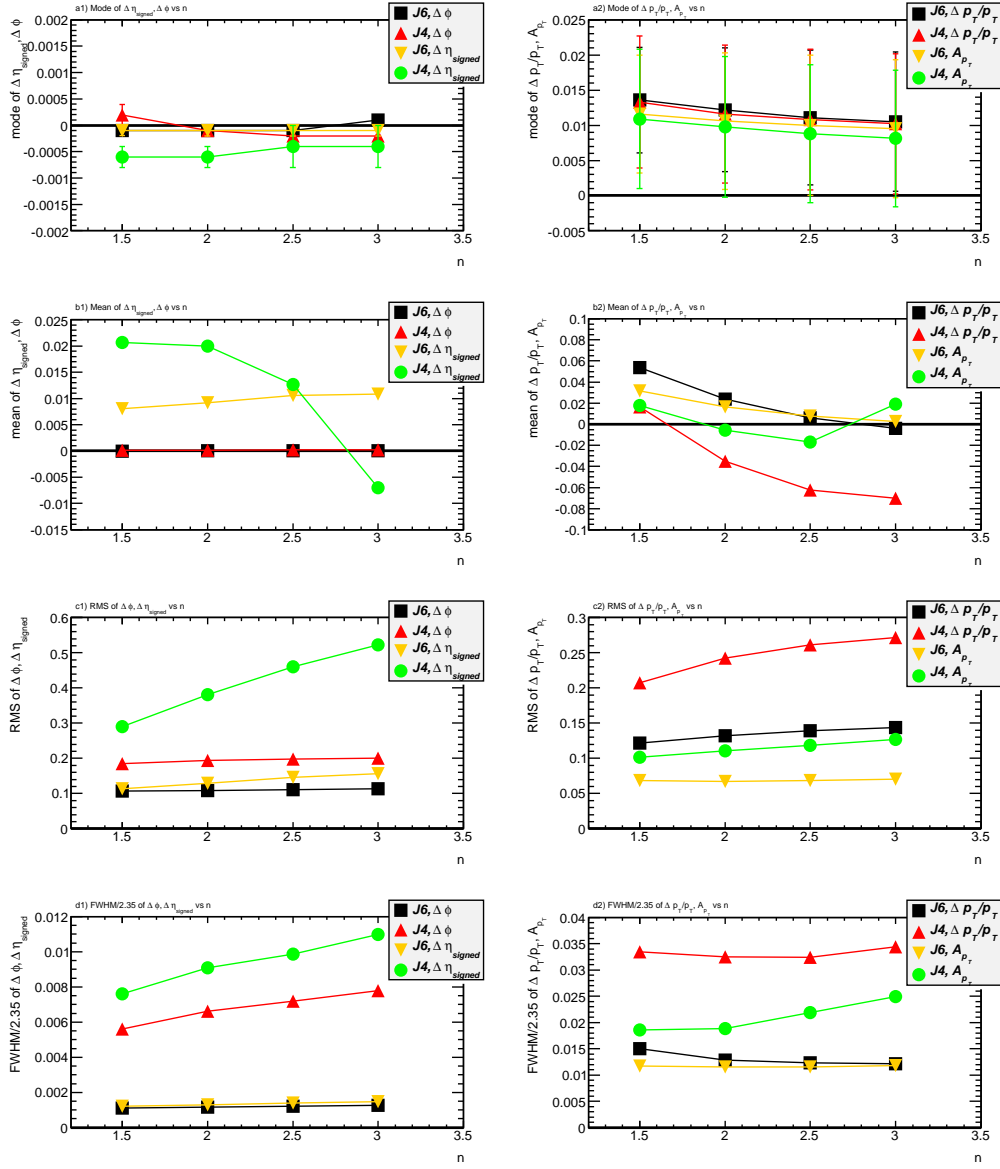


Figure 7.5: Graphs showing the  $n$ -dependence of parton resolution using the TRAPS algorithm. Mode (a), mean (b), RMS (c) and Full Width Half Max (d) values are shown for  $\Delta\phi$ ,  $\Delta\eta_{signed}$  (left) and  $\Delta p_T/p_T$ ,  $A_{p_T}$  (right). Results from both J4 and J6 PYTHIA MC are shown.



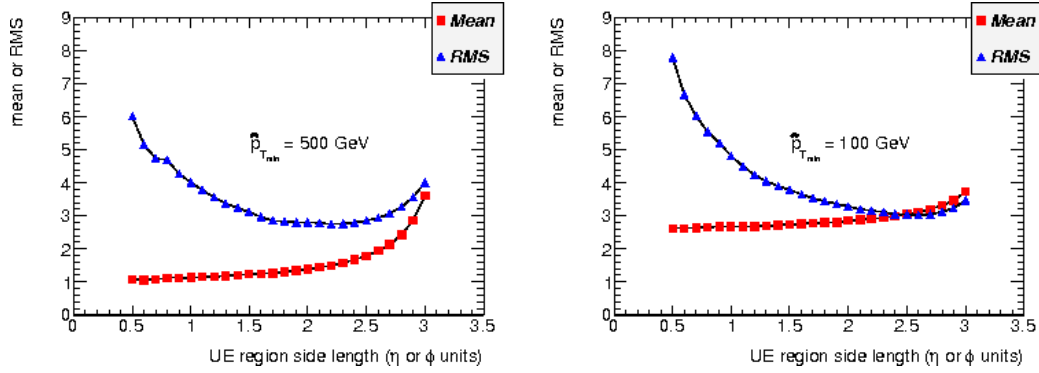


Figure 7.6: Graphs showing the mean and RMS on  $p_T$  per unit area in the UE sample regions as a function of region square size length. Left -  $\hat{p}_{T_{min}} = 500$ , right -  $\hat{p}_{T_{min}} = 100$ .

varying amounts of the jet  $p_T$  are included.

Certainly it is not desirable to include parts of the jets within the underlying event assessment. As a compromise between the results of the two plots, a value of 2.0 square side length is chosen as optimum, although the default value of 1.0 may still be used as specified.

## 7.5 Consideration of Selection Cuts

As shown in the previous section, an assessment of the resolution of TRAPS is made using key comparisons  $\Delta\phi$ ,  $\Delta\eta_{signed}$  and  $\Delta p_T/p_T$ . It is observed that both parton  $\eta$  and  $\phi$  are mostly reconstructed well, but there is greater variation in the  $p_T$  measurement. Some investigation is provided here to find out what may be done to reduce this, although cuts may introduce a potential bias. No cuts are implemented to the algorithm as a default and investigations here are indicative of use.

### 7.5.1 Acoplanarity Cut

The most obvious selection is to cut on the event acoplanarity, since this value is calculated anyway during the course of the algorithm. Events that are  $2 \rightarrow 3$ -like, or strongly affected by underlying event, are likely also to have a high acoplanarity. However, selecting events based on their acoplanarity may bias the sample for certain analyses. An example of such bias may occur in the measurement of fragmentation function, in which the cut would likely favour low multiplicity events.

Nevertheless, the application of an acoplanarity cut does reduce the width of the  $\Delta p_T/p_T$  of the J3 sample, as seen in figure 7.7, but has very little effect on the J6, since these are mostly very planar events. The reduction in both the RMS and FWHM/2.35 is indicated in table 7.6.

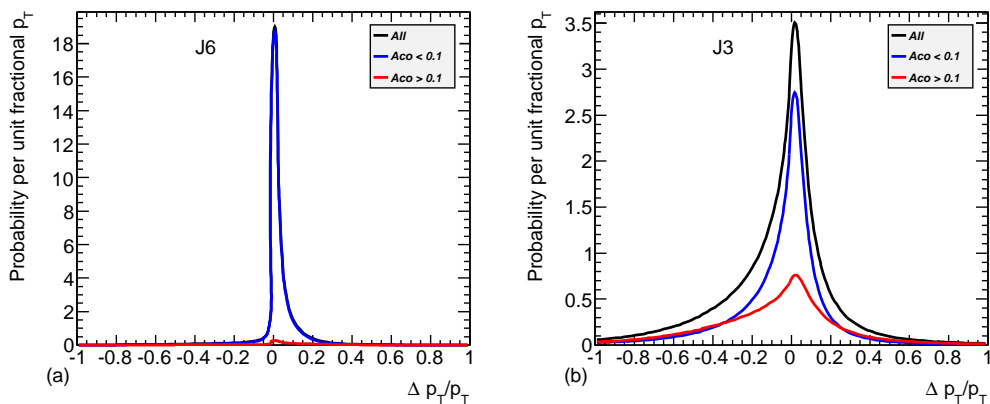


Figure 7.7:  $\Delta p_T/p_T$  for all events, and those with acoplanarity greater than or less than 0.1. (a) PYTHIA J6, (b) PYTHIA J3

### 7.5.2 Other Cut Possibilities

In this section, a range of potential event cuts are assessed for usefulness, the objective being to find one or more variables that may be used to exclude a

$\Delta p_T/p_T$	RMS	FWHM/2.35
No cut	0.32	0.065
$A_{co} < 0.1$	0.24	0.055

Table 7.6: PYTHIA J3  $\Delta p_T/p_T$  statistical information comparing all events to those with acoplanarity  $< 0.1$ .

significant quantity of badly reconstructed partons whilst not removing large fractions of those that are well reconstructed. It is assumed that such cuts would make the sample more  $2 \rightarrow 2$ -like, as is the overriding assumption of the algorithm, and simultaneously remove events that are badly measured and/or dominated by background.

The  $p_T$  asymmetry,  $A_{p_T}$  ensures all events to be within the range (-1, 1). For the purposes of these tests on the J4 PYTHIA sample, the central  $\sim 78\%$  is arbitrarily defined as ‘good’ i.e. well measured. This encloses the range  $-0.13 < A_{p_T} < 0.08$ . The rest are either ‘low’ with  $A_{p_T} < -0.13$  or ‘high’ with  $A_{p_T} > 0.08$ , each having 11% of the total. It is noted that in this convention, those events labelled ‘low’ in fact represent TRAPS-objects that are measured with greater  $p_T$  than the corresponding parton. With hindsight, it would be more logical to reverse the ‘parton - jet’ convention in the delta plot variables!

Figure 7.8 presents histograms for six example variables which are considered. Each plotted variable is calculated using only TRAPS outputs, i.e. no parton information, as would be available in data analysis. The results are split as defined above.

The variables plotted here are:

- a) background  $p_T$  as a proportion of TRAPS-object  $p_T$ ,  $\frac{p_{T_{bg}}}{p_{T_{jet}}}$

- b) TRAPS-object  $p_T$  asymmetry,  $\frac{p_{T_{jet1}} - p_{T_{jet2}}}{p_{T_{jet1}} + p_{T_{jet2}}}$
- c) TRAPS-object  $p_T$  as a proportion of full semicylinder  $p_T$ ,  $\frac{p_{T_{jet}}}{p_{T_{semicylinder}}}$
- d)  $\Delta\phi$  between the two TRAPS-objects,  $\phi_{jet1} - \phi_{jet2}$
- e) TRAPS-object  $\eta$  distribution ,  $\eta_{jet}$
- f) Number of iterations taken by the floating boundary method to converge

The results of these tests show no single, conclusive answer. Some possible cuts on the histogrammed variables are suggested in table 7.7, with the proportion of events excluded in each case from the ‘good’ and ‘bad’ collections.

It is interesting to note that the reduction effect is largely on the TRAPS-objects that are classed as ‘high’, i.e. those that contain too little  $p_T$ . The shape of the ‘low’ distributions strongly resemble the ‘good’, and it is therefore impossible to have a significant effect on them.

Perhaps the most promising of the possible selection cuts detailed here are those on TRAPS-objects with a high proportion of background  $p_T$  (figure 7.8 (a)) *or* TRAPS-objects taking a low proportion of the hemisphere  $p_T$  (figure 7.8 (c)), since these presumably cut many of the same events. In addition a stringent cut on  $\Delta\phi$  between TRAPS-objects (figure 7.8 (d)) may be effective, in removing further badly measured events not excluded by the first cut.

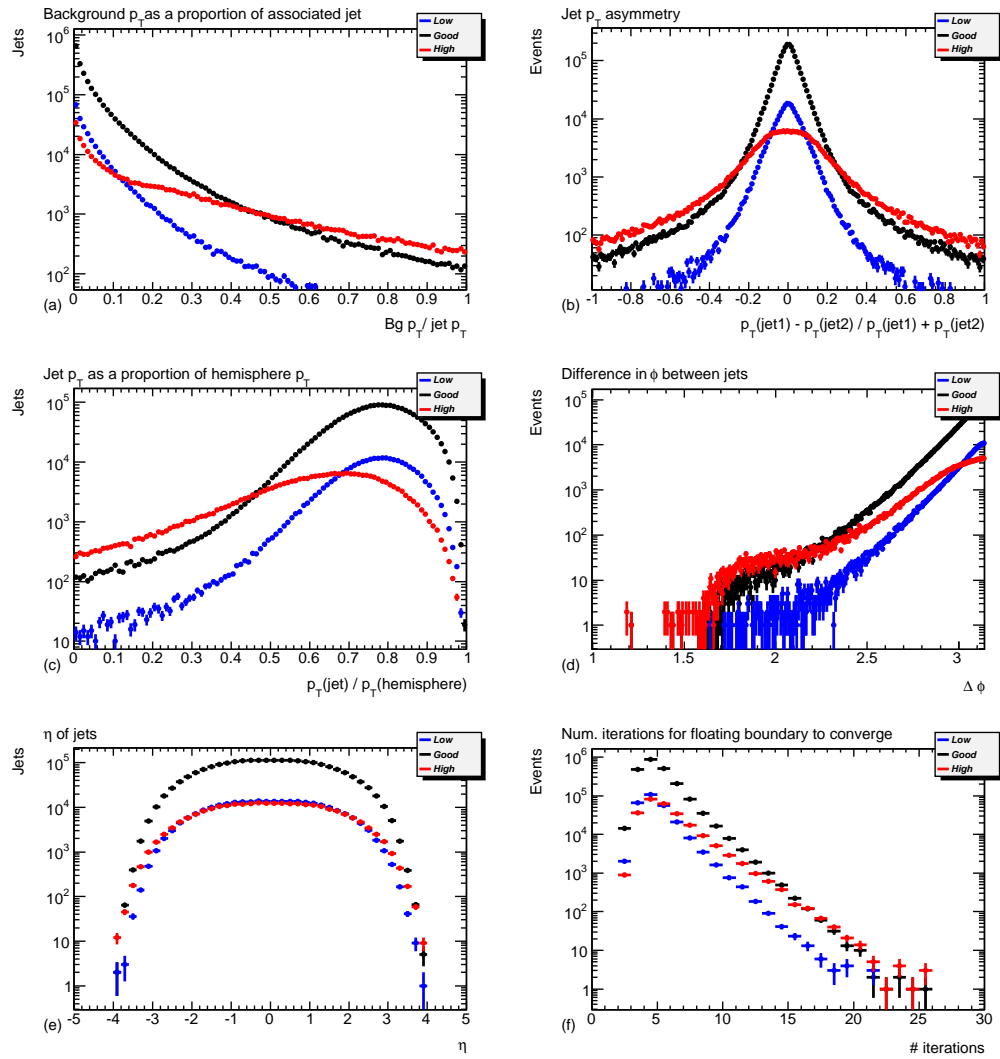


Figure 7.8: Histograms of possible selection variables as defined, for three prescribed regions of  $A_{p_T}$ : Low:  $A_{p_T} < -0.13$ ; Good:  $-0.13 < A_{p_T} < 0.08$ ; and High:  $A_{p_T} > 0.08$ .

Cut	Good events cut	Bad events cut
Histogram (a)		
0.2	5%	21%
0.1	14%	32%
0.05	27%	45%
0.03	45%	61%
Histogram (b)		
0.1	10%	7%
0.05	28%	11%
Histogram (c)		
0.5	2%	13%
0.6	7%	23%
0.7	25%	44%
0.8	66%	78%
Histogram (d)		
2.75	3%	7%
2.85	7%	13%
2.95	17%	27%
3.0	27%	38%
3.05	44%	55%
Histogram (e)		
2.0	11%	12%
1.5	25%	26%
Histogram (f)		
5	39%	44%

Table 7.7: Example cuts on the selection plots in figure 7.8 indicating the percentage of ‘good’ and ‘bad’ events excluded by the cut, where ‘bad’ events are a sum of those labelled ‘low’ and ‘high’.

# Chapter 8

## Evaluating TRAPS

### 8.1 Comparisons with Standard Jet-Finders

Naturally it is essential to test the reconstruction powers of TRAPS against some standard alternatives. However TRAPS is a specialist jet-finder, designed to find two TRAPS-objects in each high- $p_T$  parton scatter event. Therefore careful choices must be made on any comparison with standard, multi-purpose algorithms. Appropriate competition is chosen as follows:

- ‘Official’ pre-produced inclusive anti- $k_T$  jets with jet parameters of 0.4 and 0.6, in which the two highest  $p_T$  objects are used,
- Exclusive  $k_T$  jets produced by the author, in which the algorithm is instructed to give exactly two jets with the maximum possible jet parameter of  $R = \pi/2$ .

These choices are made from top-of-the-range algorithms so as to provide two jets in every event as an analogy with TRAPS. No specific tuning of any parameter is applied to any of the jet-finders used here and in the following sections. For TRAPS, the parameter  $n$  maintains its default value of 2.0, and

the underlying event sample region takes a value of  $\eta \times \phi = 2.0 \times 2.0$ , as determined in section 7.4.

To make a simple test of comparison, pre-generated samples of PYTHIA events consisting of the decay of a fictitious, excited quark,  $q^*$  with various fixed masses were used. A large proportion of the  $q^*$  decay to a quark/gluon final state. As the test is only required to evaluate QCD-type interactions, the events involving a weak process as a part of the  $q^*$  decay, and hence large proportions of missing energy, leptons and/or a multi-jet states, are removed. These represent approximately 15% of each sample. Events with the  $q^*$  decaying to a photon/quark state remain in the sample, but will be referred to as a two parton final state in this thesis for simplicity. No additional selections are made.

These excited quark events differ from the standard QCD JX events used previously, in that the  $q^*$  is a resonant,  $s$ -channel process. Although it decays in the hadronic channel to two objects, these may travel along the direction of the beam axis and hence have relatively small  $p_T$ , a phase space unlike that of parton scatters for which TRAPS was designed. Therefore  $\eta_{range}$ , given by equation 6.1, is taken to be the default limits of calorimeter detection at  $|\eta| = 4.9$  when using this sample. As the JX samples contain predominantly  $t$ -channel processes, the effect on these is minimal.

The Lorentz invariant mass, which is considered to be consistent between MC models, is calculated by equation  $m = \sqrt{E^2 - p^2}$ . It is via this *physical* kinematic quantity that a comparison between TRAPS and the other jet-finders shall be made.

Figure 8.1 shows for illustrative purposes the invariant mass of the two partons which are the decay products of the  $q^*$  in a 2 TeV/ $c^2$  sample, taken from the MC. This result shows a mode value within errors of the generated



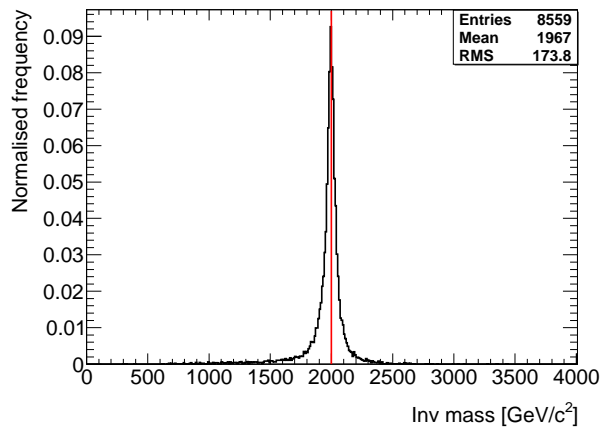


Figure 8.1: Histogram of the invariant mass of decay products at parton level from excited quarks in the 2 TeV/ $c^2$  sample. The red line indicates the generation mass.

mass. Clearly the  $q^*$  has been generated with a natural width,  $\mathcal{O}(10\%)$  of its mass. However, the shape is not symmetric as one might naïvely expect from a Breit-Wigner distribution. It has a deficit on the higher mass side, as might be produced by a greater Lorentz Invariant phase space below the generation mass, leading to a greater probability for a  $q^*$  creation on the lower side of 2 TeV/ $c^2$ . Another influence may be the constraints of the PDF, as a higher  $x$  is required for higher masses.

The equivalent histogram for the 1 TeV/ $c^2$  sample, shown in figure 8.2, does not show such variation in mean from the generation mass. This is perhaps due to the lower  $x$  requirement and available phase space being approximately equal on either side of the generated resonance. Given excellent hadronisation resolution, the jet-finders would ideally reconstruct the invariant mass with the same distribution as that shown by the partons in each case.

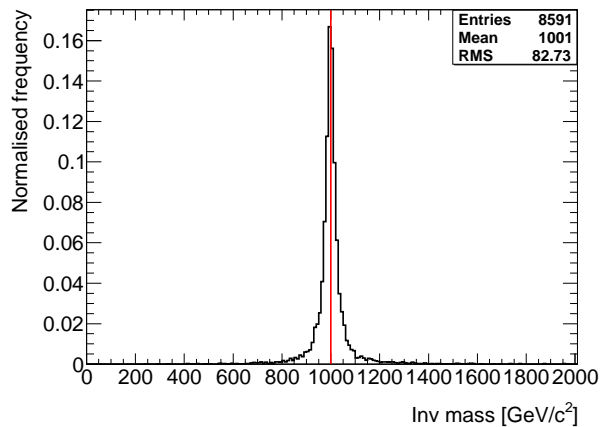


Figure 8.2: Histogram of the invariant mass of decay products at parton level from excited quarks in the  $1 \text{ TeV}/c^2$  sample. The red line indicates the generation mass.

Comparison shows that the clustering carried out by the author using the anti- $k_T$  algorithm matches the official jets closely and within errors. Following this validation, the  $k_T$  algorithm was run in exclusive mode to produce two jets. For this requirement it is only appropriate to use  $k_T$  and not anti- $k_T$  due to the nature of the two algorithms, and in particular the ordering of the clustering. The jet-parameter is set to the maximum possible value of  $R = 1.57$  ( $\pi/2$ ). This allows the maximum possible coverage in  $\phi$  and a distance of up to  $67^\circ$  in  $\theta$ , which should, as a minimum, exclude much of the underlying event close to the beam assuming fairly central jets.

Figure 8.3 shows that for the  $2 \text{ TeV}/c^2$  sample, both of the anti- $k_T$  official jets, with jet-parameters of 0.4 and 0.6, measure a mode invariant mass fairly close to that of the partons, but their means are  $\sim 10\%$  and  $\sim 8\%$  too low respectively. Similar deviations are seen for the  $1 \text{ TeV}/c^2$  sample, and full statistical data can be found in tables 8.1 and 8.2. These values imply that the jet area is more often not large enough to enclose all of the momentum

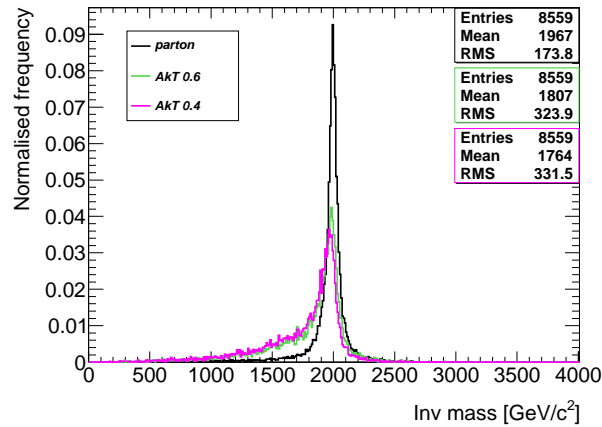


Figure 8.3: Histogram of the invariant mass of decay products from excited quarks in the 2 TeV/ $c^2$  sample, as calculated from the two highest  $p_T$  jets given by the anti- $k_T$  algorithm with jet-parameter of 0.4 and 0.6.

originating from the partons. Of the two, the larger jet-parameter obtains a better result as expected, and the 0.4 anti- $k_T$  jets will be excluded from following plots so as to provide greater clarity.

### 8.1.1 Invariant Mass Calculation Results

Figures 8.4 and 8.5 are histograms of the invariant masses, as calculated by the various algorithms described, of events from the 1 TeV/ $c^2$  and 2 TeV/ $c^2$  samples respectively. The parton invariant mass is also shown, along with a vertical black line representing the generation mass. Accompanying statistics are given in tables 8.1 and 8.2. These numbers were derived in the manner described in section 6.6.1. For FWHM/2.35, the error is approximated as half a bin width divided by 2.35. Results for other  $q^*$  masses and analysis can be found in the following section.

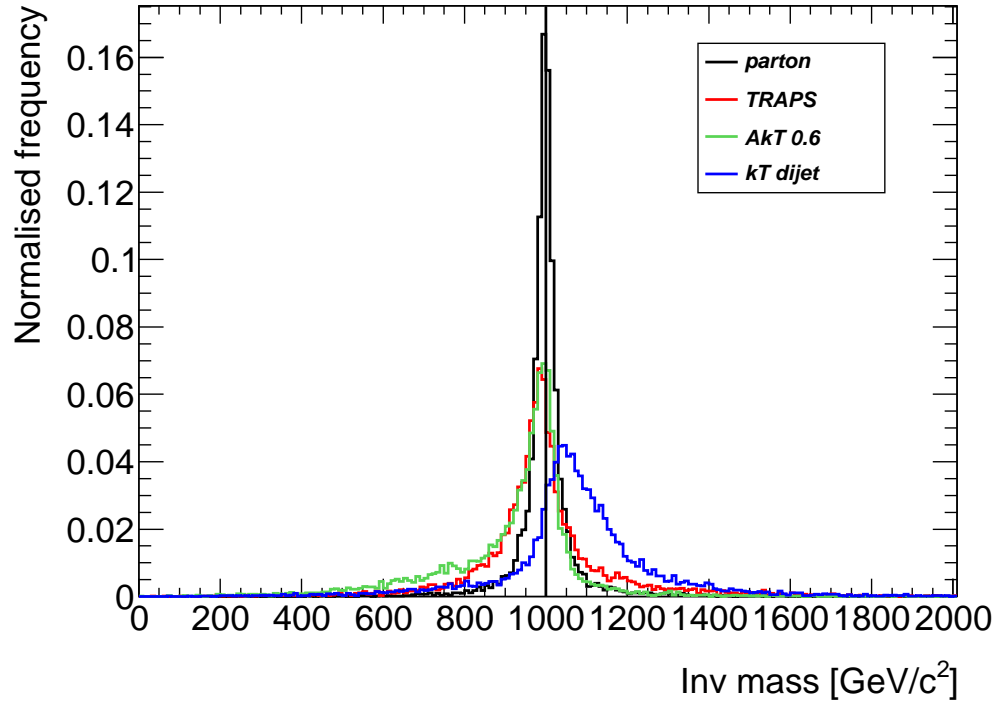


Figure 8.4: Histogram of the invariant mass of decay products from excited quarks in the  $1 \text{ TeV}/c^2$  sample, calculated as described in the text by the anti- $k_T$  0.6, the  $k_T$  dijet and the TRAPS algorithms.

Algorithm	Mean	RMS	FWHM/2.35	Mode
parton	$1001 \pm 1$	$91 \pm 2$	$18 \pm 2$	$996 \pm 5$
TRAPS	$1001 \pm 2$	$197 \pm 4$	$33 \pm 2$	$985 \pm 5$
Anti- $k_T$ 0.6	$921 \pm 2$	$179 \pm 4$	$34 \pm 2$	$995 \pm 5$
Anti- $k_T$ 0.4	$892 \pm 2$	$183 \pm 4$	$38 \pm 2$	$985 \pm 5$
$k_T$ dijet	$1081 \pm 2$	$209 \pm 5$	$66 \pm 4$	$1050 \pm 10$

Table 8.1: Statistical results from the invariant mass distributions, calculated by various algorithms using the  $1 \text{ TeV}/c^2$   $q^*$  sample. All values are in  $\text{GeV}/c^2$  and errors given are statistical only.

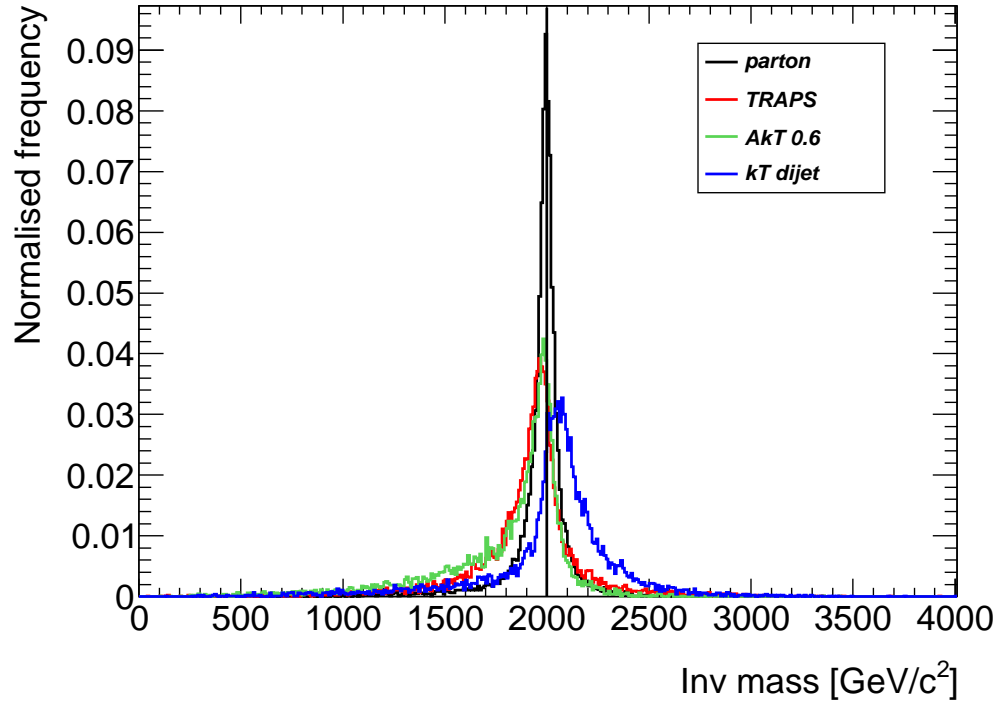


Figure 8.5: Histogram of the invariant mass of decay products from excited quarks in the 2 TeV/ $c^2$  sample, calculated as described in the text by the anti- $k_T$  0.6, the  $k_T$  dijet and the TRAPS algorithms.

Algorithm	Mean	RMS	FWHM/2.35	Mode
parton	1967 $\pm$ 2	174 $\pm$ 4	29 $\pm$ 2	1995 $\pm$ 5
TRAPS	1904 $\pm$ 3	289 $\pm$ 6	65 $\pm$ 2	1965 $\pm$ 5
Anti- $k_T$ 0.6	1807 $\pm$ 4	324 $\pm$ 7	52 $\pm$ 2	1985 $\pm$ 5
Anti- $k_T$ 0.4	1764 $\pm$ 4	332 $\pm$ 7	60 $\pm$ 4	1970 $\pm$ 10
$k_T$ dijet	2026 $\pm$ 4	336 $\pm$ 7	71 $\pm$ 4	2050 $\pm$ 10

Table 8.2: Statistical results from the invariant mass distributions as calculated by various algorithms using the 2 TeV/ $c^2$   $q^*$  sample. All values are in GeV/ $c^2$  and errors given are statistical only.

### 8.1.2 Performance as a Function of Hard Scale

Other  $q^*$  masses were added to the study to analyse how well the various jet-finders reconstruct across a range of hard scale. Graphs showing the relative performance can be seen in figure 8.6. It is clear from e.g. figure 8.1 that the excited quark is generated with a certain width, and with a mean value that is not necessarily exactly at the generation mass. Therefore the mean and mode averages in (a) and (b) are plotted as the difference between the values obtained by the algorithm and the diparton effective mass values.

Since the generated width and algorithmic resolution are independent, the diparton RMS and FWHM/2.35 in (c) and (d) are subtracted from the algorithm values in quadrature, to give a resolution for each mass scale. The resolutions are given as a percentage of the generation mass. Again, only statistical errors are represented, and combined appropriately. Results are shown in figure 8.6.

The three algorithms are comparable across the mass range studied, each having relative advantages. The  $k_T$  dijet has a comparatively small dependence on the hard scale, but shows a higher mean and mode as might be expected if ISR and other underlying event are absorbed into the jet. This problem not increasing with hard scale is possibly due to the events becoming cleaner at higher  $x$ .

The decrease in mean and mode of the anti- $k_T$  algorithm result is consistent with expectations if the use of a fixed radius excludes some FSR from the jet. TRAPS suffers from the same issue, although slightly less significantly as it gives more symmetric distributions within the range shown here. It is disappointing that the variable size search area does not compensate more reliably for this, although the scale-dependence is slightly smaller than for anti- $k_T$ .

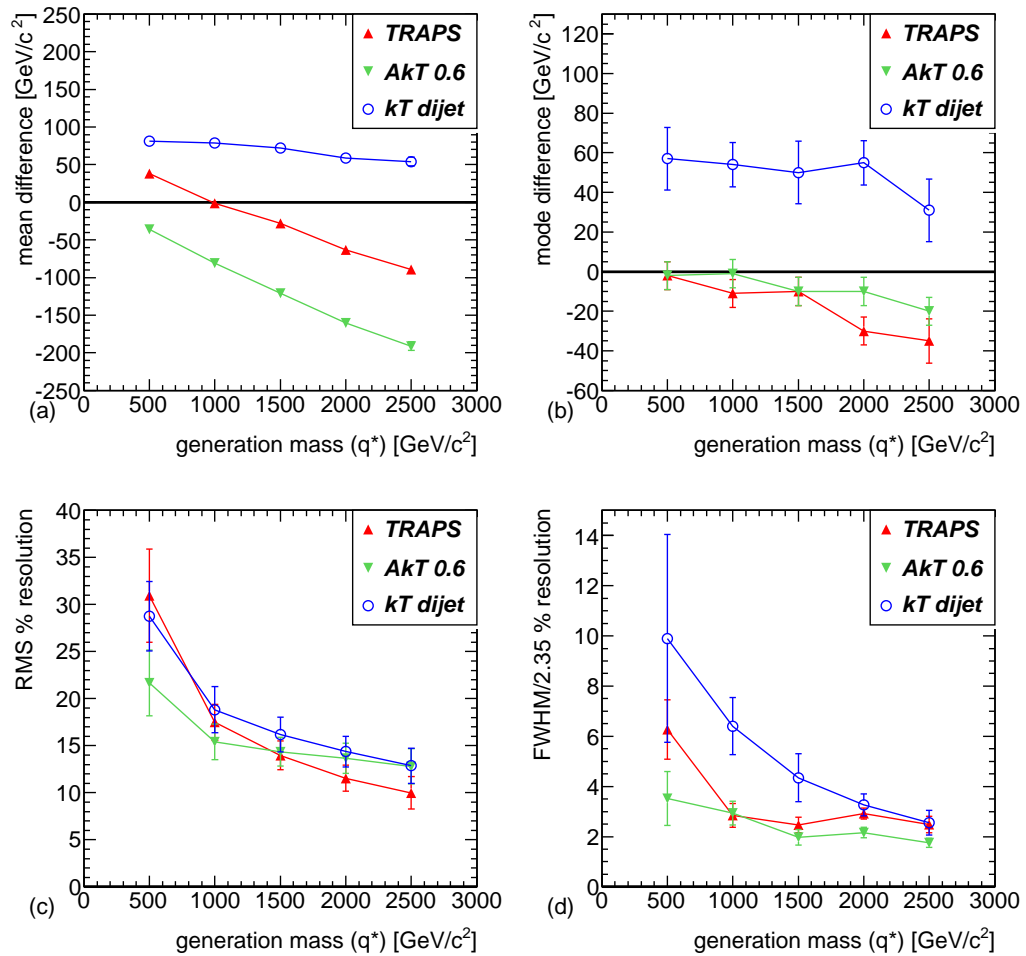


Figure 8.6: (a) mean, (b) mode, (c) RMS and (d) FWHM/2.35 relative to the natural generated values, for the  $q^*$  with varying generation mass. (a) and (b) are differences between values found by the algorithm and those calculated from the MC parton. (c) and (d) have the parton values subtracted in quadrature and are given as a percentage of generation mass, as indication of algorithm resolution.

The modes shows much less scale dependence than the mean, as would be expected by the dominance of clean events. TRAPS and anti- $k_T$  give the smaller differences compared to the diparton, although  $k_T$  dijet is within errors.

Both measures of resolution show improved performance with increasing hard scale. The differences between RMS and FWHM/2.35 values imply a non-Gaussian shape. It is perhaps consistent with a Breit-Wigner shape, although not precisely due to the discrete nature of the underlying event. This shows distinct similarities with the  $p_T$  resolutions shown in the previous chapter.

## 8.2 TRAPS and IRC Safety

As discussed in chapter 5, both collinear and infrared safety are essential characteristics of a general jet-finder. Collinear unsafety is a problem associated with jet-finders that are seeded by a high- $p_T$  particle; since TRAPS does not use seeds, this need not be studied. However, the effects of soft radiation and hence the hypothesis that TRAPS is infrared safe, are tested here.

TRAPS is expected to give the same, safe result on the vast majority of occasions. By looking for two, and only two jet-objects, TRAPS intrinsically searches with a resolution at the scale of the ‘hard’ scatter.

The test for infrared safety is based on that used in [59], in which a hard event was generated of 2-10 particles with random positions and  $p_T < 1000$  GeV/ $c$ . The algorithm was run with random jet parameters. The algorithm was run a second time, but with the addition of a soft event, consisting of 1-5 randomly positioned particles with a scale  $p_T < 10^{-100}$  GeV/ $c$ . The differences in the final jet state were then assessed.



The procedure to test TRAPS is similar, although it substitutes the fully hadronised excited quark samples employed in section 8.1 for the hard event. A single, artificial particle is positioned randomly within the event, distributed evenly, throughout  $\phi$  and within  $|\eta| = 4.9$ . Three differently sized particles are added, in turn, to both the 1 and 2 TeV/ $c^2$   $q^*$  samples; these have energy of 1 GeV, 10 MeV and 0.1 MeV. An equivalent test is applied to the inclusive anti- $k_T$  (generated by the author) and the exclusive  $k_T$  dijet algorithms as defined in section 8.1, for comparison. Here TRAPS takes the default parameters  $n = 2.0$  and an UE sample size of  $1.0 \times 1.0$ .

The difference in variables  $\eta$ ,  $\phi$ ,  $p_T$  and invariant mass are calculated, with the sign convention: (with infrared particle - without infrared particle). Figures 8.7, 8.8 and 8.9 show these differences as labelled, and tested with the 1 GeV, 10 MeV and 0.1 MeV particles respectively. Note that the  $y$ -axis is drawn with a log scale and the histogram areas are normalised to 100. Statistical data as calculated by ROOT is provided for the  $p_T$  difference plots in tables 8.3, 8.4 and 8.5. The underflows and overflows have been included as these entries are not taken into account in the calculation of RMS.

The addition of a 1 GeV particle provides a good demonstration of the method. It is evident that the anti- $k_T$  algorithm rarely has a problem, even with a particle of this size. Of course, when the additional particle happens to land *within* the jet area, it is rightly included in that jet. This is clear from the  $p_T$  difference plot - additional entries between 0 and 1 GeV/ $c$ . The shape is explained by the fact that 1 GeV is the *energy* of the added particle, whereas it is the transverse momentum which is measured here.

Overall, the TRAPS method produces results that are between the best result from the anti- $k_T$  algorithm and the rather more sensitive  $k_T$  dijet. On the whole, the performance of TRAPS is much closer to that of the infrared safe

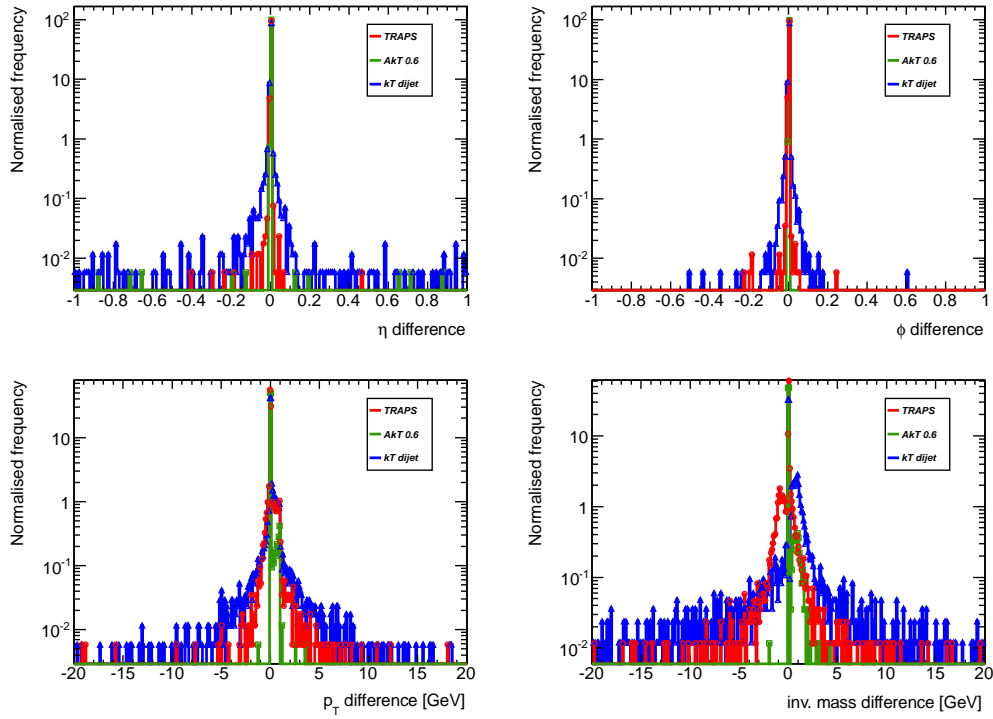


Figure 8.7: Infrared safety test on the  $1 \text{ TeV}/c^2$  sample, using  $1 \text{ GeV}$  additions distributed randomly in  $\eta$ - $\phi$  space.

Algorithm	Mean	RMS	Underflow	Overflow
TRAPS	$(3.20 \pm 0.36) \times 10^{-2}$	$(4.69 \pm 0.07) \times 10^{-1}$	0.026	0.029
Anti- $k_T$ 0.6	$(1.18 \pm 0.08) \times 10^{-2}$	$(9.95 \pm 0.15) \times 10^{-2}$	0	0
$k_T$ dijet	$(5.89 \pm 0.79) \times 10^{-2}$	$(1.03 \pm 0.02)$	0.058	0.070

Table 8.3: Statistical results on  $p_T$  from the infrared safety test on using the  $1 \text{ TeV}/c^2$   $q^*$  sample with an addition of a  $1 \text{ GeV}$  particle distributed randomly in  $\eta$ - $\phi$ . All values are in  $\text{GeV}/c$  and errors given are statistical only.

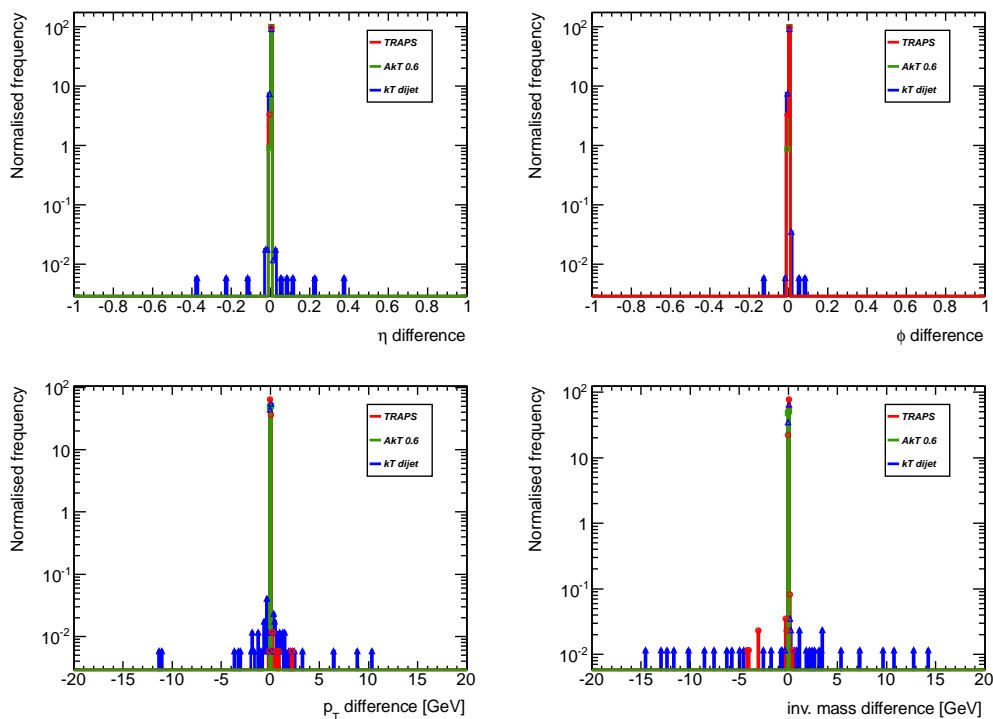


Figure 8.8: Infrared safety test on the  $1 \text{ TeV}/c^2$  sample, using  $10 \text{ MeV}$  additions distributed randomly in  $\eta$ - $\phi$  space.

Algorithm	Mean	RMS	Underflow	Overflow
TRAPS	$(5.26 \pm 1.98) \times 10^{-4}$	$(2.57 \pm 0.04) \times 10^{-2}$	0	0
Anti- $k_T$ 0.6	$(1.21 \pm 0.07) \times 10^{-4}$	$(9.51 \pm 0.14) \times 10^{-4}$	0	0
$k_T$ dijet	$(4.62 \pm 13.7) \times 10^{-4}$	$0.184 \pm 0.003$	0	0.0058

Table 8.4: Statistical results on  $p_T$  from the infrared safety test using the  $1 \text{ TeV}/c^2$   $q^*$  sample with an addition of a  $10 \text{ MeV}$  particle distributed randomly in  $\eta$ - $\phi$ . All values are in  $\text{GeV}/c$  and errors given are statistical only.

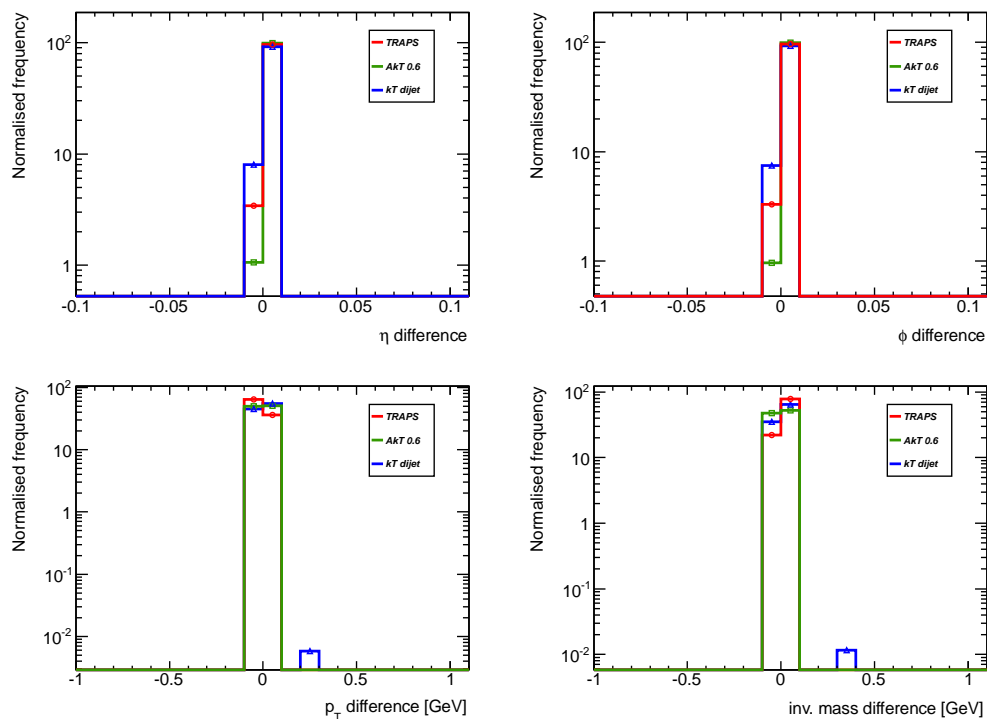


Figure 8.9: Infrared safety test on the  $1 \text{ TeV}/c^2$  sample, using  $0.1 \text{ MeV}$  additions distributed randomly in  $\eta$ - $\phi$  space.

Algorithm	Mean	RMS	Underflow	Overflow
TRAPS	$(1.13 \pm 0.15) \times 10^{-6}$	$(2.01 \pm 0.3) \times 10^{-5}$	0	0
Anti- $k_T$ 0.6	$(1.34 \pm 0.08) \times 10^{-6}$	$(9.92 \pm 0.15) \times 10^{-6}$	0	0
$k_T$ dijet	$(5.02 \pm 0.38) \times 10^{-3}$	$(4.98 \pm 0.08) \times 10^{-2}$	0	0

Table 8.5: Statistical results on  $p_T$  from the infrared safety test using the  $1 \text{ TeV}/c^2$   $q^*$  sample with an addition of a  $0.1 \text{ MeV}$  particle distributed randomly in  $\eta$ - $\phi$ . All values are in  $\text{GeV}/c$  and errors given are statistical only.

anti- $k_T$  algorithm than the less safe  $k_T$  algorithm. Certainly for the smallest particle addition of 0.1 MeV, all three algorithms appear to be intrinsically safe. However the  $p_T$  difference mean and RMS values demonstrate that the infrared safety of TRAPS approaches that of the state-of-the-art anti- $k_T$ , whereas the  $k_T$  algorithm forming a dijet state is confirmed as a less safe option.

Similar results are seen when applying the infrared safety test to the 2 TeV/ $c^2$   $q^*$  sample.

Incoherent underlying event, such as multiple parton interactions, pile-up and apparatus noise are a difficulty to be overcome in all events produced at high-luminosity hadron colliders, particularly when the UE forms a significant proportion of the total energy detected in the event. TRAPS is not completely immune to such issues, as this UE will afford a larger standard deviation and hence a greater search area. But in general, the overall effect will only be to alter slightly the direction and/or  $p_T$  of the TRAPS-object, and not completely change the final state. In addition, uncorrelated UE subtraction is an intrinsic part of the algorithm, so at least these difficulties are solved in any statistical average over events, preserving the mean. However, in MC there exist rare occasions when the UE energy approaches the hard scale of the interaction, and in these cases TRAPS will have just as much trouble in differentiating between the two as any other jet finder.

### 8.3 Speed of TRAPS

The speed of an algorithm can be an important factor in its usefulness, as mentioned in chapter 5. TRAPS uses simple for-loop summing, and on the whole, low numbers of iterations to complete its task. Some quantity of

particles are discounted at each step, thereby decreasing the number that are included in the summing. This gives the algorithm an intrinsically low computational complexity. It is hence fast and practical.

TRAPS does consider large numbers of possible input objects, but not combinations as sequential recombination or the SIScone algorithms, hence its scaling with  $N$ , the number of inputs, is of a low order.

Recombinatorial-type jet-finders have much higher complexity, and therefore take longer to execute. However, the FASTJET [61] software has greatly improved the speed of such algorithms, and coupled with increases in computing power, jet algorithms have reached a state where duration is not considered as significant as once it was.

Nevertheless, it is important to verify whether TRAPS can execute within a similar time-scale as the current fastest jet-finders. The alternate jet algorithms used in section 8.2 are also timed using the same machines for direct comparison. In figure 8.10 execution time is given as a function of  $N$  particles in the 1 TeV/ $c^2$   $q^*$  sample, and increases with increasing  $N$  as expected. TRAPS clearly takes longer than the  $k_T$  and anti- $k_T$  algorithms, although it is within the same order of magnitude, and shows a  $\sim 1$  ms minimum call time. The evident structure in the TRAPS entries result from the different numbers of iterations involved in a particular event.

One reason that TRAPS may take a longer time is the larger number of tasks it is doing. Whereas the timing of the  $k_T$ -type algorithms encloses only the clustering process itself, the timing of TRAPS includes saving the member values to Storegate, and various surplus processes providing further detail to be used during development, with an execution time of  $\sim 4$  ms/event. It was not considered worthwhile to put effort into professional coding at this stage, or reducing the algorithm down to the bare minimum - i.e. outputting only

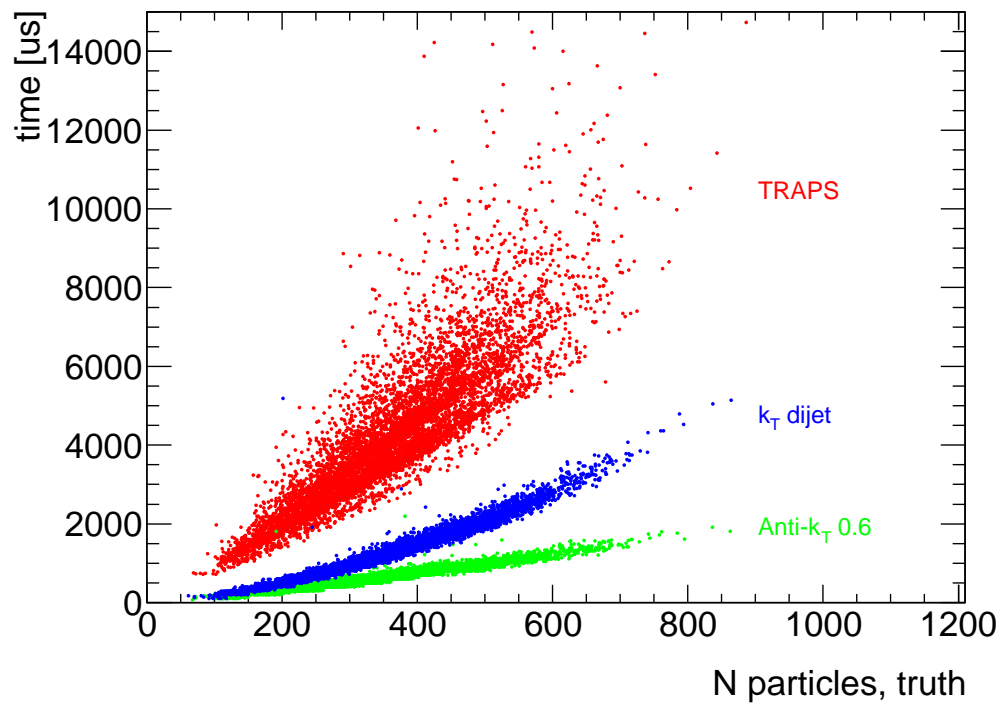


Figure 8.10: Scatter plot comparing the time taken by TRAPS as a function of  $N$ -particles with contemporary jet-finders, using FASTJET.

the  $\eta$ ,  $\phi$  and  $p_T$  of the two TRAPS objects - simply for the purpose of timing tests.

## 8.4 Sensitivity to Monte Carlo

Although they are tuned to resemble the data, different MCs will produce slightly different results, both in the final state and at the parton level. For example, it is observed that HERWIG generates more partons outside of the theoretical calculated rapidity region, known within TRAPS as  $\eta_{range}$ , than PYTHIA. These are boosted down the beampipe due to large ISB that approach the energy of the hard scale. However, as each MC represents a different model, we can use comparisons between them to estimate the error associated with choice of model. This could be described as the Monte Carlo ‘resolution’.

Here, a comparison is made between the previously given PYTHIA results and TRAPS run on HERWIG, both at generator level. The delta plots are used to assess the differences, although both the hadron and parton level effects influence their distributions. A small proportion of TRAPS-objects having the flagNegPT flag (negative  $p_T$ ), as specified in section 6.5.1, have been removed from both datasets. This amounts to approximately 2 % for the J3 and 0.3 % for the J6 samples.

Figure 8.11 shows that there are small differences between the two J6 MC samples. The positional spread is marginally larger in HERWIG, but, and it is perhaps surprising given the  $\eta - \phi$  measurement, the  $p_T$  assessment shows that HERWIG partons are just as, or more accurately reconstructed by TRAPS than those in PYTHIA.

Similarly, figure 8.12 indicates small differences at the J3 level, again



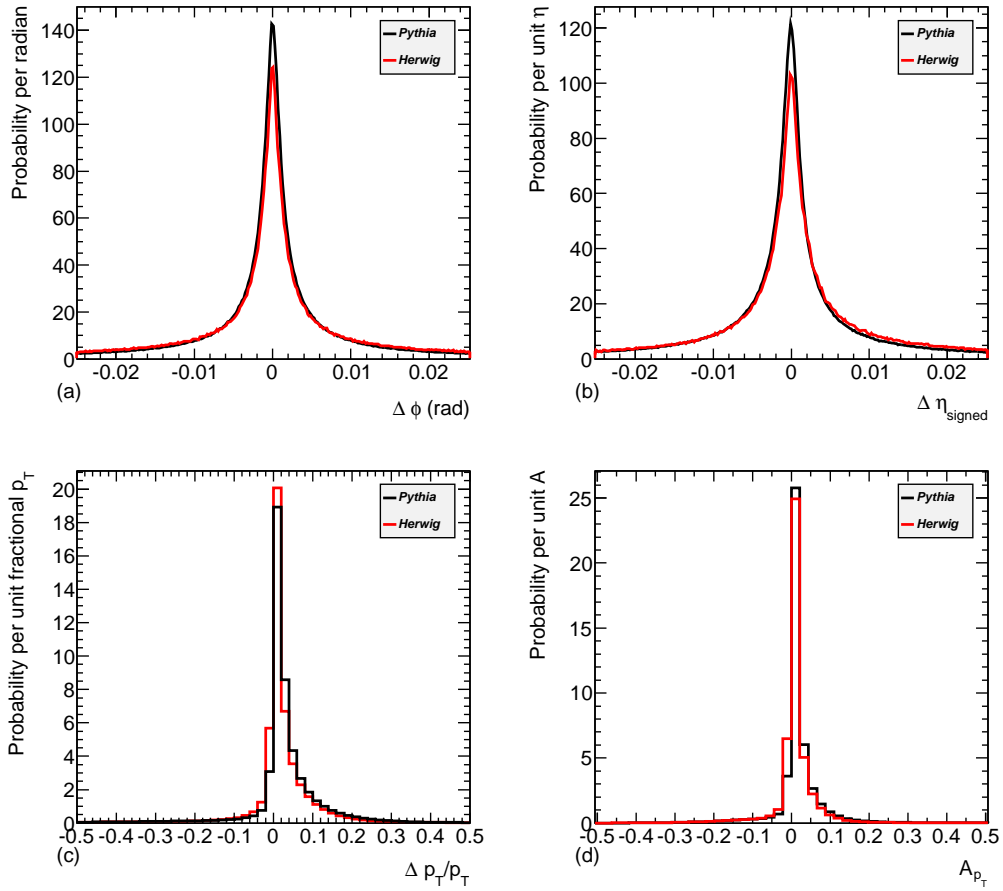


Figure 8.11: Delta plots comparing TRAPS run on HERWIG and PYTHIA, J6 sample. (a)  $\Delta\phi$ , (b)  $\Delta\eta_{signed}$ , (c)  $\Delta p_T/p_T$ , (d)  $A_{p_T}$ .

with the modes remaining resolutely at zero. As with the PYTHIA samples, the means are generally not zero but remain small compared with detector resolution. The distribution statistics are given in tables 8.6 and 8.7 for HERWIG J6 and J3 respectively. A comparison with the PYTHIA numbers given in tables 7.2 - 7.5 is made in figure 8.13. Here, the modes, means, RMS and FWHM/2.35 are graphed as a function of the hard energy scale, with this value taking the mean average  $\hat{p}_T$  of each JX sample.

To summarise, TRAPS measures each delta plot variable similarly, regard-

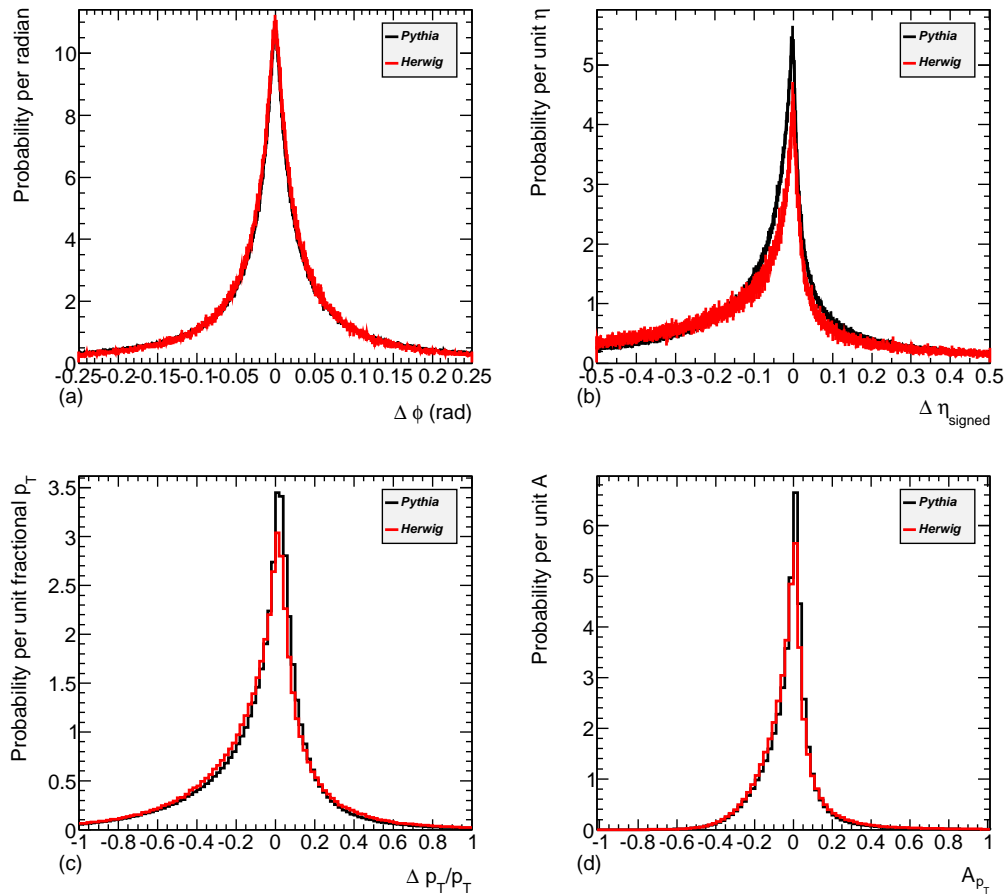


Figure 8.12: Delta plots comparing TRAPS run on HERWIG and PYTHIA, J3 sample. (a)  $\Delta\phi$ , (b)  $\Delta\eta_{\text{signed}}$ , (c)  $\Delta p_T/p_T$ , (d)  $A_{p_T}$ . Note the change in scales from figure 8.11.

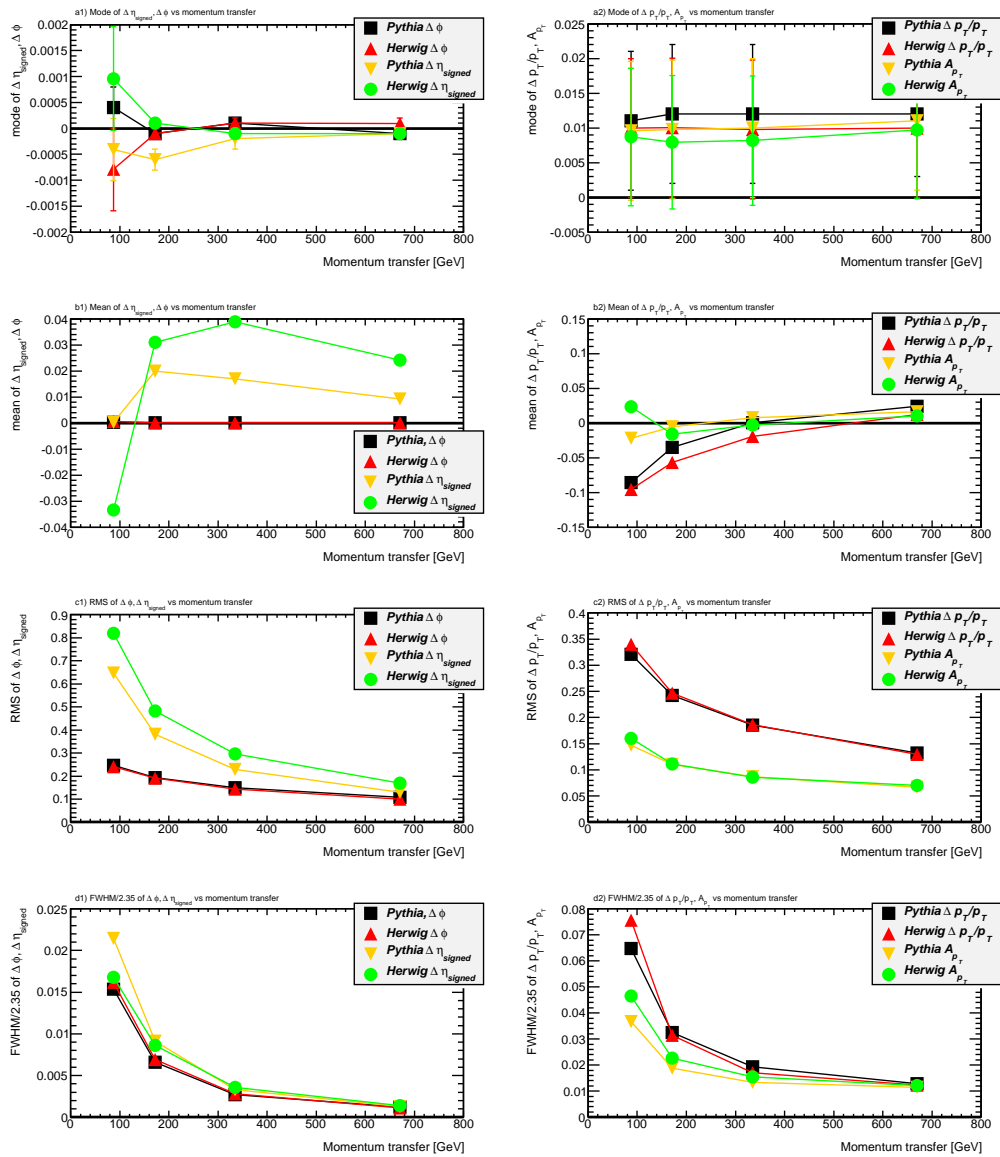


Figure 8.13: Graphs comparing PYTHIA and HERWIG delta plot statistics for the four JX samples. Mode (a), mean (b), RMS (c) and Full Width Half Max (d) values are shown for  $\Delta\phi$ ,  $\Delta\eta_{signed}$  (left) and  $\Delta p_T/p_T$ ,  $A_{p_T}$  (right).

Histogram	Mode	Mean	RMS	FWHM/2.35
$\Delta\phi$	$(9.6 \pm 10.0) \times 10^{-5}$	$(1.0 \pm 1.1) \times 10^{-4}$	0.10	0.001
$\Delta\eta_{signed}$	$(-9.7 \pm 10.0) \times 10^{-5}$	$(241 \pm 2) \times 10^{-4}$	0.17	0.001
$\Delta p_T/p_T$	$(1.0 \pm 1.0) \times 10^{-2}$	$(122 \pm 2) \times 10^{-4}$	0.13	0.012
$A_{p_T}$	$(9.7 \pm 9.9) \times 10^{-3}$	$(101 \pm 0.7) \times 10^{-4}$	0.07	0.012

Table 8.6: Table of resolution statistics relating to HERWIG J6 delta plots.

Histogram	Mode	Mean	RMS	FWHM/2.35
$\Delta\phi$	$(-7.9 \pm 8.0) \times 10^{-4}$	$(6.3 \pm 2.8) \times 10^{-4}$	0.24	0.016
$\Delta\eta_{signed}$	$(9.6 \pm 10.0) \times 10^{-4}$	$(-334 \pm 9) \times 10^{-4}$	0.82	0.017
$\Delta p_T/p_T$	$(1.0 \pm 1.0) \times 10^{-2}$	$(-956 \pm 4) \times 10^{-4}$	0.34	0.076
$A_{p_T}$	$(8.7 \pm 9.9) \times 10^{-3}$	$(-237 \pm 2) \times 10^{-4}$	0.16	0.047

Table 8.7: Table of resolution statistics relating to HERWIG J3 delta plots.

less of whether PYTHIA or HERWIG are used. There are some differences between them, most noticeably in the  $\Delta\eta_{signed}$  mean values. These discrepancies are thought to be as a result of the different ways the partons are treated in the two MCs. Hence a larger error would need to be applied to account for the model variance.

## 8.5 Exercise on $d\sigma(x)$

Each of the two partons entering into the hard scatter has an associated  $x$ , as described in chapter 2. They are known individually as  $x_1$  and  $x_2$  and related by equation 8.1, where  $\sqrt{s}$  is the CMS energy, or  $2p_{Beam}$  for equal momentum proton-proton beams.

$$\hat{s} = x_1 x_2 s \quad (8.1)$$

Briefly discussed in section 2.1.2,  $\hat{s}$ , can also be calculated by equation 8.2, where subscripts 1, 2, 3 and 4 refer to the two incoming and outgoing matrix element partons respectively. This is also the equivalent of invariant mass squared, as calculated in section 8.1.

$$\hat{s} = (\tilde{p}_1 + \tilde{p}_2)^2 = (\tilde{p}_3 + \tilde{p}_4)^2 \quad (8.2)$$

Equations 8.3 and 8.4 give  $x_1$  and  $x_2$ , where  $Y$  is the combined object rapidity given by equation 8.5, in which 3 and 4 refer to the two (reconstructed) partons. Hence  $x$  can be measured using TRAPS outputs.

$$x_1 = \sqrt{\frac{\hat{s}}{s}} e^Y \quad (8.3)$$

$$x_2 = \sqrt{\frac{\hat{s}}{s}} e^{-Y} \quad (8.4)$$

$$Y = \frac{1}{2} \ln \frac{(E_3 + E_4) + (p_{z3} + p_{z4})}{(E_3 + E_4) - (p_{z3} + p_{z4})} \quad (8.5)$$

### 8.5.1 Employment of Migration Matrix

A migration matrix gives the probability that an entry in one histogram bin has migrated from another, and hence if known, the inverse can be used to correct for such effects as algorithm and detector resolution. This is better than simple bin by bin corrections because that technique considers only *net* migration.

A correlation plot of, for example, a reconstructed vs. truth variable can be turned into a matrix by taking the number of entries from each bin. For

this trial analysis, a coarse  $9 \times 9$  matrix will be used. This method works well when the matrix contains a strong diagonal, i.e. when many of the entries are reconstructed well. Because the true value must, by definition, be correct, the badly measured reconstructed value will fall into a bin above or below the diagonal element, and generally a consecutive bin assuming the corrections will be small. This, suitably normalised by ‘truth’, will be termed the ‘*RECO*’, or reconstructed matrix.

The sum of each column is taken to be the true value for each truth bin and makes a  $1 \times 9$  vector, *TRUE*. The expected experimental result, *EXP* is found by matrix multiplication, as shown in equation 8.6.

$$(RECO)(TRUE) = (EXP) \quad (8.6)$$

However, to use the migration matrix to find the truth from experiment, it must be inverted, as in equation 8.7.

$$(TRUE) = (RECO)^{-1}(EXP) \quad (8.7)$$

The inversion of a matrix is a non-linear operation, and thus has an associated error matrix. This is calculated using a large number of random variations on the migration matrix (mean) value, within a normal distribution.

### **Inclusive differential cross-section**

Here is presented an indicative study on the inclusive differential cross-section with respect to  $x$  and mean sample  $\hat{p}_T$ . It is not intended as a full analysis, merely an exercise in using TRAPS to measure a kinematic variable.

Figure 8.14 compares the distributions of  $x$  calculated in the above manner for the J3-J6 PYTHIA-truth partons (black lines), and the parton re-

constructed by TRAPS split into nine  $x$  bins per sample (blue circles). Nine bins are chosen so as to approximately equalise the number of entries in each bin, and this is altered appropriately for the different JX samples, whilst maintaining as much commonality between them as possible for comparison.

The red dots with error bars labelled ‘MM TRAPS’ in figure 8.14 represent the PYTHIA data points unfolded for algorithmic effects by a migration matrix. These are constructed using MC, for each JX HERWIG sample. The HERWIG-truth is applied to the PYTHIA-TRAPS result, in order to assess the sensitivity of TRAPS to the choice of MC. The plotted error bars include both statistical errors and MC differences added in quadrature, although the MC contribution dominates.

The distributions are normalised by the number of events and scaled by the mean cross-sections of the PYTHIA samples, as given in table 7.1, and the acceptance. These acceptance corrections are very small, as the number of events excluded by the algorithm is small. The originate only from TRAPS-objects given either `flagNegPT` or `flagEtaJet`, and the total percentage of excluded PYTHIA events for samples J3-J6 are 2.3%, 1.1%, 0.63% and 0.29% respectively, leading to correction factors of 1.024, 1.011, 1.006 and 1.003 respectively. No other cuts have been applied to this test. The removal of up to 2% of jets would perhaps introduce a small bias to the measurement, but this shall not be investigated in the course of this exercise.

Figure 8.15 shows the same truth/TRAPS comparisons with the differential cross-section on a linear scale. In general, there is good agreement between the TRAPS-measured partons and the PYTHIA-truth, even for the more poorly resolved J3 sample. Typically it is at the extremes of  $x$  where the algorithm performs less well. It is quite possible that partons appearing in both regions belong to the *same* events, since generally one parton with

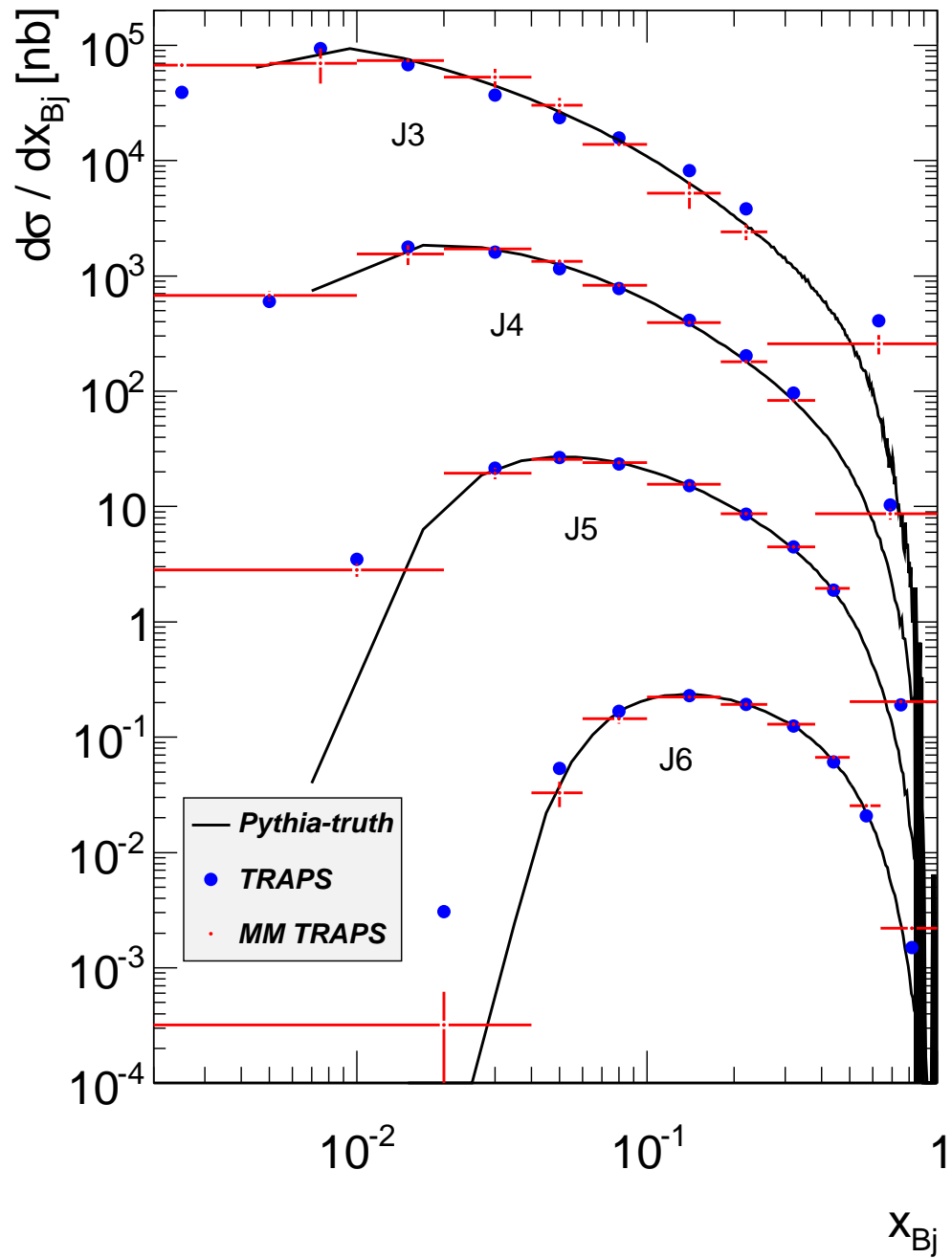


Figure 8.14: Inclusive differential cross-section with respect to  $x$  for PYTHIA samples J3-J6.



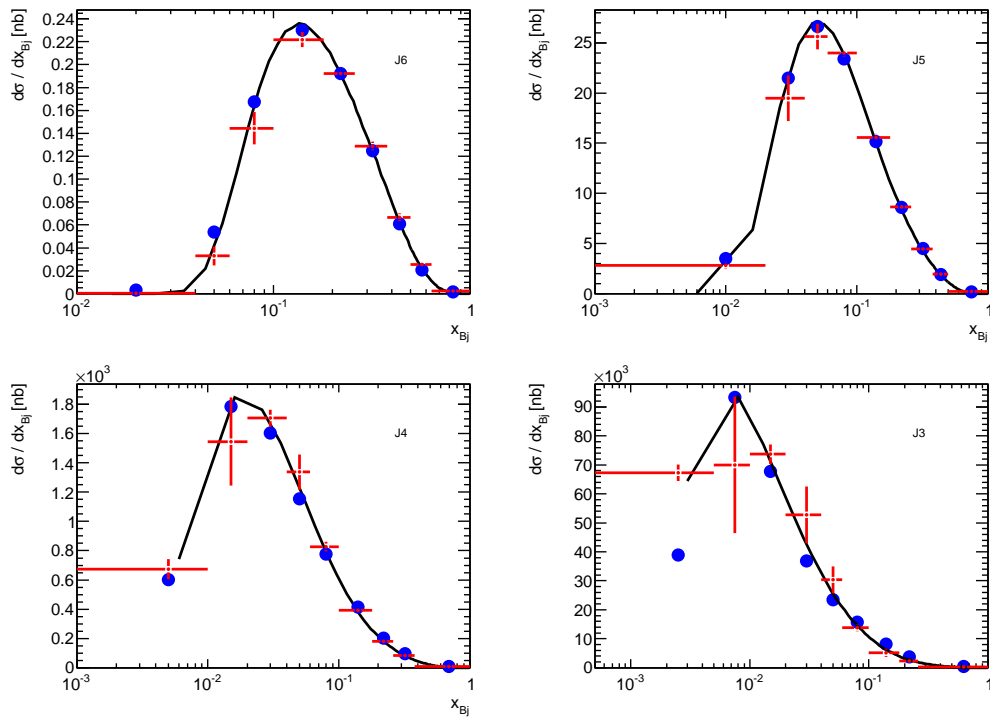


Figure 8.15: Inclusive differential cross-section with respect to  $x$  for PYTHIA samples J6-J3 with linear scale  $y$ -axis.

high  $x$  would be interacting with one of low  $x$  in order to put the event within a particular JX sample.

Information from figure 8.14 is reproduced in a different form in figure 8.16, in which the  $x$  bins with more than one entry in common are plotted against the mid-point  $\hat{p}_T$  of their JX sample. Lines of common  $x$  lie close to one another, showing signs of approximate scaling, although a full analysis would be needed to separate the kinematics from amplitude effects. So as to fully separate the lines of common  $x$ , each is multiplied by a factor of  $5^i$ , where the  $i$ -value is indicated.

Linearity is observed, which is essentially a result of the power law of the propagator for point particles. This is expected to be a negative fourth

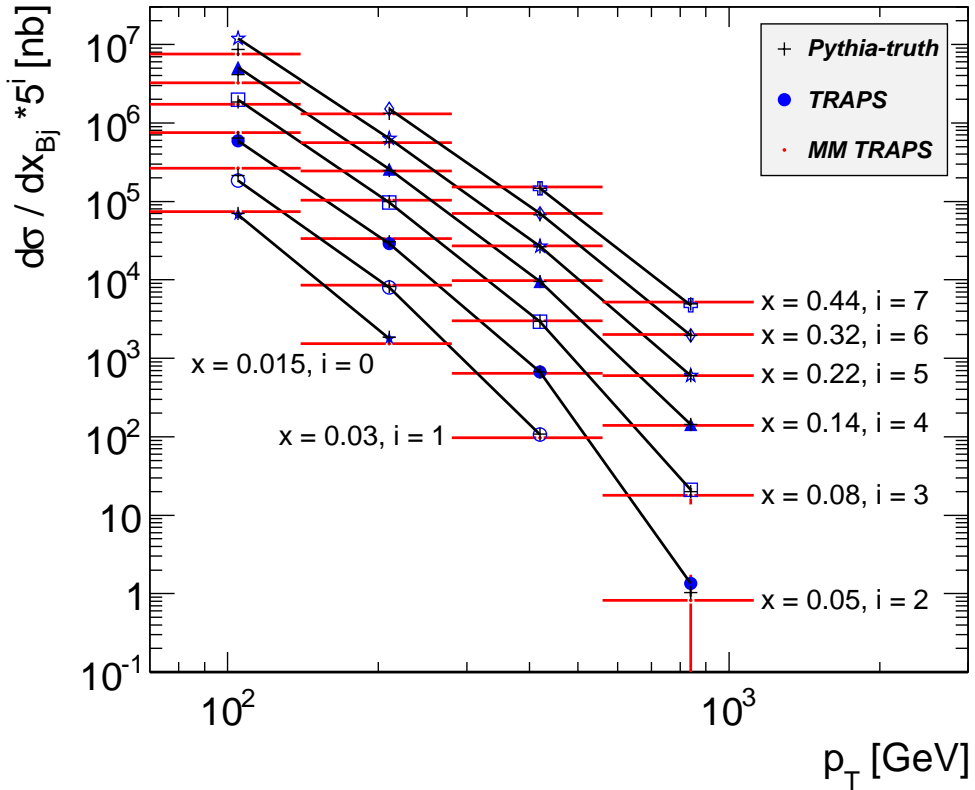


Figure 8.16: Inclusive differential cross-section with respect to  $x$  as a function of  $\hat{p}_T$ . The varying  $x$  points are separated by factors of  $5^i$ . The points on the  $\hat{p}_T$  axis are the mid-points of each JX sample.

power and the gradient here is calculated to be  $\sim -4.5$ .

## 8.6 Detector Scaling and Resolution

There are inevitable detector ‘smearing’ effects on any measurements taken by the ATLAS detector. Due to the limited spatial and energy resolution of the calorimetry, there will always be uncertainty associated with the detector.

Before any comparison with experimental data, TRAPS should be tested

on reconstructed MC, which takes into account detector effects, as described in section 2.5. TRAPS is already capable of running on reconstructed MC, and produces the same output variables as for generator level MC.

ATLAS is calibrated to the ‘EM scale’, which gives a good energy measurement for electromagnetic objects. But the detector has a different response to hadrons which must be accounted for by the ‘jet-energy scale’ (JES) [49]. This is a further calibration applied to the energy of jets and other hadronic objects. The recommended correction is a function of jet  $p_T$  and rapidity. The procedures for applying the latest JES within Athena, and for calculating the associated uncertainties are given in [67].

The ‘jet-energy resolution’ (JER) describes the calculated energy resolution of ATLAS in response to jets [68]. For jets with  $20 < p_T < 80\text{GeV}/c$  and in the rapidity range  $|Y| < 2.8$ , MC simulation describes the JER measured from data within 14% [68].

Only very preliminary checks on results from TRAPS running on reconstructed MC have been made, to ensure that the code is functioning correctly. Therefore no resolutions will be shown at this stage.

# Chapter 9

## Conclusions

A top-down dijet-finding algorithm has been conceived, written and tested against PYTHIA Monte Carlo QCD events. The Topological Reconstruction Algorithm for Parton Scatters, known as TRAPS, has been found to reconstruct with good resolution and very high acceptance, final state partons in high- $p_T$   $2 \rightarrow 2$  MC jet events. The most planar events are found to be those with the greatest  $\hat{p}_T$ .

A comparison of TRAPS with the performance of current leading jet-finders was carried out, using samples of hypothetical excited quark Monte Carlo events, generated at a range of resonant masses between 500 GeV and 2.5 TeV. TRAPS was found to be competitive with the anti- $k_T$  algorithm, and with the  $k_T$  algorithm configured to form a dijet state.

An exercise designed to apply TRAPS to the basic kinematic variable  $x$  within Monte Carlo was performed. In this test, a method was discussed to allow for small differences between PYTHIA and HERWIG Monte Carlo.

Although the objective to be entirely parameter-free was not entirely successful, allowing for the inclusion of the  $n$ -parameter and the underlying event region size, TRAPS does not define a fixed jet-size, which was the main

goal. In any case, both parameters passed the test of being largely insensitive to their assigned value.

As with any jet-finding method, it is left to the user of the algorithm to decide what event selection cuts are appropriate to their analysis. Otherwise, TRAPS has full acceptance for all events, in the sense that two TRAPS-objects are produced from any event input. However, selection of events should be carried out carefully, as acknowledged in the Data Analysis Model in the following section, so as to create a sample of high  $\hat{p}_T$  events. It is true that TRAPS occasionally produces some unphysical results with negative  $p_T$  due to underlying event subtraction, and as with every analysis the possibility of bias by removing such events must be studied.

The scope of the algorithm has always focused on the highest energy hadron-hadron collisions available from the LHC, and it is here where the best algorithm resolutions are found. However, overlap with the highest energy  $e^+e^-$  annihilation and DIS experimental fragmentation function results at around 100 GeV has proved within the reach of TRAPS.

It is shown that TRAPS is an infrared safe method, although this has not been proven mathematically. It is also computationally simple compared to other algorithms, implying a robust method and competitively fast performance.

TRAPS is already capable of running on the reconstruction of ATLAS-simulated Monte Carlo, although the review and testing of results are at an early stage and have not been included in this thesis. Various steps need to be taken, as discussed in the following section, to calibrate the results in preparation for data. Significant quantities of high-energy jet events have been recorded by ATLAS, and it is hoped that these will soon be analysed by the TRAPS algorithm.

## 9.1 The Future for TRAPS

It is unfortunate that the usual time-constraints restrict the advancement of the TRAPS algorithm, as there are many ideas for its progress. This section contains a brief summary of the most promising ideas for extension of the algorithm, an outline of the steps needed to proceed towards applying TRAPS to ATLAS data, and possible alternative methods and usage. Of these categories, the most immediate priority is to apply TRAPS to ATLAS data.

### Algorithm

- TRAPS already stores a vector of jet constituents, to be used in the ATLAS ‘good/bad/ugly’ test. This assesses the jets so as to identify those which include dead or faulty regions of the detector. These events can then be excluded as desired. One essential addition is the identification of good/bad/ugly *background* regions, thus background constituents must also be recorded. However, the event would not necessarily have to be abandoned due to a bad/ugly background. Since each TRAPS-object uses a mean average of *two* background areas, the problematic area could be ignored and the background subtraction performed using only one such region.
- The size of the search ellipse could be decreased in some events where the standard deviation is large. The ellipse as previously described is not necessarily optimum for some events, in which, for example two major deposits in a single semicylinder may not be at either the same  $\eta$  or  $\phi$ . This could be solved by the ‘tilted ellipse’, as illustrated in figure 9.1.

The standard deviations would be calculated in a frame rotated in the

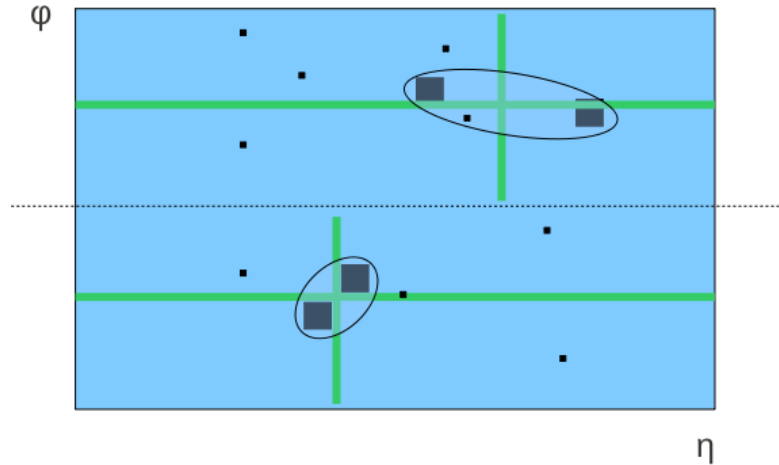


Figure 9.1: Diagram illustrating the un-implemented ‘tilted ellipse’ method.

$\eta$ - $\phi$  plane. The angle of this rotation is determined in much the same way as the acoplanarity calculation, as seen in step I of the TRAPS algorithm, except  $\eta$  and  $\phi$  are substituted for the  $x$  and  $y$  coordinates and the calculation is performed twice - once for each semicylinder. This should lead to smaller standard deviations and hence a smaller ellipse. The benefits are that there is a smaller likelihood of picking up background energy deposits, and that there is less chance of an overly-large background subtraction, associated with large ellipse areas.

- Other ideas include using the skew and kurtosis of the spread of  $p_T$  about the mean  $\eta - \phi$  positions in each semicylinder to perform jet separation in the case of three or more jet events. This may then lead to the tilted ellipse splitting into two circles in a seedless fashion, and naturally may also increase the number of parameters involved, but only, like  $n$ , as dimensionless measures of threshold.

### Generation

- Some more recent MC samples include the generation and reconstruction of pile-up, which was described in chapter 2. It is important to investigate the effect of pile-up that is already seen in ATLAS, on the algorithm resolutions. Up to  $\sim 20$  interactions per bunch crossing are observed to date. Note that the integrated incoherent background subtraction is expected to deal well with this background effect.
- Additional PDF sets could be modelled in MC, in order to test the algorithm's sensitivity to these variants, and perhaps find evidence in ATLAS data as to which is a better fit.

### Reconstruction

- Reconstructed MC must be studied in full. TRAPS is already capable of producing results from reconstructed data, however, work is required to apply the appropriate Jet Energy Scale (JES) as mentioned previously, and fully described in [49]. The associated errors are an essential part of this work.

### Experimental Data

- The minimum undertaking required in advance of applying TRAPS to ATLAS data would be to run on recent reconstructed MC and to implement the JES. TRAPS is already capable of running on ATLAS data, in ESD or RAW formats, or some DESD files, which contain the full calorimeter cell information.
- In addition, a full Data Analysis Model (DAM) must be prepared and implemented. This would involve selecting events from data which are



triggered with  $\sim 100\%$  efficiency, and similarly, cutting on MC samples so as to create an equivalent sample for direct comparison. A non-prescaled, high- $p_T$  Level-1 RoI trigger would be used. If higher level triggers are used to select data, i.e. those based on standard jet-finders, then these would be viewed as simply part of the trigger as a whole, and folded into the efficiency.

### Analysis

- One study would be a structure function measurement, although naturally this is not a straightforward process at the LHC. The approach would be to measure  $x_1$  and  $x_2$  separately and hence their correlations, which would lead to a crude effective  $(Q + \frac{4}{9}G)$  structure function.
- A fragmentation function analysis using both TRAPS and the anti- $k_T$  algorithm is already in progress. This work is being carried out by Tom Macey, and will be written up as his PhD thesis [69]. The total parton fragmentation will be calculated using various MC models. A *gluon* fragmentation function at  $\sim 100$  GeV could be measured for the first time. This would be done by subtracting quark fragmentation results, for example from a weighted mean of H1, ZEUS and LEP results, at around 100 GeV, from a *parton* fragmentation function measured from ATLAS data.

### Further Usage

- This thesis mentioned the potential application of the TRAPS algorithm in such measurements as  $x$  for structure functions, and parton fragmentation functions. It was also tested on a sample of excited quarks decaying to a two particle state. But the scope of TRAPS may be much

wider. It is suggested that TRAPS may be usefully employed in such diverse analyses as photon+jet, in which the photon may masquerade as a jet in the EM calorimeter, or in other exotic scenarios where decay to a 2-jet state is expected.

# Bibliography

- [1] CERN. *The CERN Large Hadron Collider: Accelerator and Experiments*.
- [2] F. D. Aaron et al. Charged Particle Production in High Q<sup>2</sup> Deep-Inelastic Scattering at HERA. *Phys. Lett.*, B654:148–159, 2007.
- [3] D. Traynor. *Hadronic Fragmentation Studies in Diffractive Deep Inelastic Scattering at HERA RAL-TH-2002-005*. PhD thesis, Queen Mary University of London, 2002.
- [4] J. Turney. *A Study of Identified Hadron Fragmentation in ep Collisions at HERA Using the H1 Detector RAL-TH-2002-004*. PhD thesis, Queen Mary University of London, 2002.
- [5] P. Dixon. *A Study of the Fragmentation of Quarks in ep Collisions at HERA RAL-TH-1998-016*. PhD thesis, Queen Mary University of London, 1998.
- [6] Georges Aad et al. Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector. *Eur. Phys. J.*, C71:1795, 2011.
- [7] I. Abt et al. The H1 detector at HERA. *Nucl. Instrum. Meth.*, A386:310–347, 1997.

- [8] ZEUS Collaboration. The ZEUS Detector. Status Report (unpublished). *DESY*, 1993.
- [9] The ATLAS Collaboration. *The ATLAS experiment at the CERN Large Hadron Collider*, 2007.
- [10] J. D. Bjorken and Emmanuel A. Paschos. Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon. *Phys. Rev.*, 185:1975–1982, 1969.
- [11] R.P. Feynman. Proceedings of the 3rd topical conference on high energy collision of hadrons. *Stony Brook, N. Y.*, 1969.
- [12] Georges Aad et al. Search for Quark Contact Interactions in Dijet Angular Distributions in pp Collisions at  $\sqrt{s} = 7$  TeV Measured with the ATLAS Detector. *Phys. Lett.*, B694:327–345, 2011.
- [13] Francis Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*.
- [14] B.R. Martin and G. Shaw. *Particle Physics*. John Wiley & Sons, second edition, 1997.
- [15] G.D. Coughlan and J.E. Dodd. *The ideas of particle physics - an introduction for scientists*. Cambridge University Press, second edition, 1994.
- [16] <http://en.wikipedia.org/wiki/Proton>.
- [17] K. Nakamura *et al.* Particle data group - 2010 review of particle physics. *J. Phys. G*, 37, 2010.
- [18] M. Wing. Standard model (lecture notes). *University of London*, 2007.

- [19] G. Aad *et al.* The ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *JINST*, 3:S08003, 2008.
- [20] R Bonciani *et al.* Nll resummation of the heavy-quark hadroproduction cross-section. *arXiv:hep-ph/9801375v2*, 1998.
- [21] The Durham HepData Project. *Online PDF plotting and calculation - <http://hepdata.cedar.ac.uk/pdf/pdf3.html>*.
- [22] R.D. Field and R.P. Feynman. A parametrization of the properties of quark jets. *Nuclear Physics B*, 136(1):1 – 76, 1978.
- [23] R.P. Feynman, R.D. Field, and G.C. Fox. Correlations among particles and jets produced with large transverse momenta. *Nuclear Physics B*, 128(1):1 – 65, 1977.
- [24] F. D. Aaron et al. Charged Particle Production in High Q<sup>2</sup> Deep-Inelastic Scattering at HERA. *Phys. Lett.*, B654:148–159, 2007.
- [25] Eram Rizvi. Methods and techniques in experimental particle physics. *University of London*, 2007.
- [26] S. Agostinelli *et al* GEANT 4 Collaboration. Geant 4 - a simulation toolkit. *NIM-A*, 506, 3:250 – 303, 2003.
- [27] T. Sjostrand *et al.* *Pythia 6.4 - Physics and Manual*, arXiv:hep-ph/0603175 2006.
- [28] M. Bahr *et al.* *Herwig++ - Physics and Manual*, arXiv:0803.0883v3 2008.
- [29] Tanju Gleisberg *et al.* *Sherpa manual*, [www.hepforge.org/archive/sherpa/howto-1.1.1.pdf](http://www.hepforge.org/archive/sherpa/howto-1.1.1.pdf) 2008.

- [30] Michelangelo L. Mangano *et al.* *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, CERN-TH/2002-129, FTN/T-2002/06 2006.
- [31] G. Gustafson. Dual Description of a Confined Color Field. *Phys. Lett.*, B175:453, 1986.
- [32] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to pythia 8.1. *Comput. Phys. Commun.*, 178(arXiv:0710.3820. CERN-LCGAPP-2007-04. LU TP 07-28. FERMILAB-PUB-07-512-CD-T):852 – 867. 27 p, Oct 2007.
- [33] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton Fragmentation and String Dynamics. *Phys. Rept.*, 97:31–145, 1983.
- [34] Richard D. Field and Stephen Wolfram. A QCD Model for  $e^+ e^-$  Annihilation. *Nucl. Phys.*, B213:65, 1983.
- [35] *The Rivet webpage: <http://projects.hepforge.org/rivet/>.*
- [36] *The ROOT User's Guide webpage: <http://root.cern.ch/drupal/content/users-guide>.*
- [37] *The AGILE webpage: <http://projects.hepforge.org/agile/>.*
- [38] The ATLAS Collaboration. *The ATLAS Workbook: <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/WorkBook>.*
- [39] *The Grid PP homepage - <http://www.gridpp.ac.uk/>.*
- [40] G. Aad *et al.* The ATLAS Collaboration. *Expected performance of the ATLAS experiment: detector, trigger and physics*. CERN, Geneva, 2009.

- [41] *The Data Summary website: <https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/DataSummary/2011/>.*
- [42] *The CERN public website: <http://public.web.cern.ch/public/en/LHC/Facts-en.html>.*
- [43] *The LHC website: <http://lhc.web.cern.ch/lhc/>.*
- [44] L Tompkins and B Heinemann. Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 7$  tev with the atlas detector. Technical Report ATLAS-COM-CONF-2011-007, CERN, Geneva, Jan 2011.
- [45] Mike Lamont. LHC Report: Steady as she goes. *CERN Bulletin*, 10th October 2011.
- [46] *The ATLAS public website: <http://atlas.ch>.*
- [47] J. Colas et al. Position resolution and particle identification with the ATLAS EM calorimeter. *Nucl. Instrum. Meth.*, A550:96–115, 2005.
- [48] M. Aharrouche et al. Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test- beam. *Nucl. Instrum. Meth.*, A568:601–623, 2006.
- [49] M Begel et al. Jet energy scale and its systematic uncertainty in proton-proton collisions at  $\sqrt{s}=7$  tev in atlas 2010 data. Technical Report ATLAS-COM-CONF-2011-053, CERN, Geneva, Mar 2011.
- [50] *The ATLAS trigger operations public results webpage - [https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults#2011\\_7\\_TeV](https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults#2011_7_TeV).*

- [51] R. Achenbach *et al.* The atlas level-1 calorimeter trigger. *JINST*, 3:P03001, 2008.
- [52] The ATLAS Collaboration. *ATLAS Level-1 Calorimeter Trigger: Timing Calibration with 2009 LHC Beam Splashes*, ATL-DAQ-PUB-2010-001 2010.
- [53] R. Achenbach *et al.* First data with the atlas level-1 calorimeter trigger. *ATL-DAQ-PROC-2008-006*, 2008.
- [54] *The ATLAS approved plots webpage* - <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsED>.
- [55] *The Atlantis homepage* - [www.cern.ch/atlantis](http://www.cern.ch/atlantis).
- [56] C. Ohm and T. Pauly. The atlas beam pick-up based timing system. *Nucl. Instr. and Meth. A*, *arXiv:0905.3648v1*, 2010.
- [57] C. Buttar *et al.* Standard Model Handles and Candles Working Group: Tools and Jets Summary Report. 2008.
- [58] Matteo Cacciari, Juan Rojo, and Gavin P. Salam. Quantifying the performance of jet definitions for kinematic reconstruction at the LHC. 2008.
- [59] Gavin P. Salam and Grégory Soyez. A practical seedless infrared-safe cone jet algorithm. *arXiv:0704.0292 [hep-ph]*, 2007.
- [60] Mario Campanelli. Jet physics with the first atlas data. Slides from QMUL Seminar, Mar 2011.
- [61] Matteo Cacciari, Gavin Salam, and Gregory Soyez. *The Fastjet webpage*: <http://www.lpthe.jussieu.fr/salam/fastjet/>.



- [62] Gavin P. Salam. Jet reconstruction with fastjet. Slides given at Jets in Proton-Proton and Heavy-Ion Collisions, Prague, Aug 2010.
- [63] L. Asquith et al. Performance of jet algorithms in the atlas detector. Technical Report ATL-PHYS-INT-2010-129, CERN, Geneva, Dec 2010.
- [64] P. Calafiura, C. G. Leggett, D. R. Quarrie, H. Ma, and S. Rajagopalan. The StoreGate: a Data Model for the Atlas Software Architecture. *Proceedings of Conference for Computing in High-Energy and Nuclear Physics, La Jolla, California, arXiv:cs/0306089*, 2003.
- [65] ATLAS Collaboration. ATLAS Monte Carlo Tunes for MC09. *CERN*, 2010.
- [66] *The ATLAS Metadata Interface (AMI) webpage: <http://ami.in2p3.fr/>.*
- [67] S Adomeit et al. Jet calibration and jes uncertainty at atlas - a user guide. Technical report, ATLAS Note, Geneva.
- [68] Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the atlas detector using proton-proton collisions at a center of mass energy  $\sqrt{s} = 7$  tev. Technical Report ATLAS-CONF-2010-054, CERN, Geneva, Jul 2010.
- [69] T. Macey. *Parton Fragmentation in pp Collisions at  $\sqrt{s} = 7$  TeV*. PhD thesis, Queen Mary University of London, Due 2013.